Universität Basel

Fakultät für
Psychologie

Unlocking the Potential:

Validity Evidence for a Comprehensive Developmental Assessment in Children and Adolescents

Inauguraldissertation zur Erlangung der Würde einer Doktorin der Philosophie

vorgelegt der Fakultät für Psychologie der Universität Basel von

Salome Dorothée Odermatt

aus Dallenwil, NW

Basel, 2023

Universität Basel

Fakultät für Psychologie

Genehmigt von der Fakultät für Psychologie auf Antrag von

Prof. Dr. Alexander Grob

Prof. Dr. Wenke Möhring

Datum des Doktoratsexamen: 08.12.2023

Dekan:in der Fakultät für Psychologie

Universität Basel

Fakultät für
Psychologie

Erklärung zur wissenschaftlichen Lauterkeit

Ich erkläre hiermit, dass ich die vorliegende Arbeit ohne die Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel selbstständig verfasst habe. Zu Hilfe genommene Quellen sind als solche gekennzeichnet. Die veröffentlichten oder zur Veröffentlichung in Zeitschriften eingereichten Manuskripte wurden in Zusammenarbeit mit den Koautoren erstellt und von keinem der Beteiligten an anderer Stelle publiziert, zur Publikation eingereicht oder einer anderen Prüfungsbehörde als Qualifikationsarbeit vorgelegt. Es handelt sich dabei um folgende Manuskripte:

1. **Odermatt, S. D.**, Möhring, W., Grieder, S., & Grob, A. (2022). Cognitive and developmental functions in autistic and non-autistic children and adolescents: Evidence from the Intelligence and Development Scales–2. *Journal of Intelligence, 10*(4), Article 4. https://doi.org/10.3390/jintelligence10040112

2. **Odermatt, S. D.**, Weidmann, R., Schweizer, F., & Grob, A. (2024). Academic performance through multiple lenses: Intelligence, conscientiousness, and achievement striving motivation as differential predictors of objective and subjective measures of academic achievement in two studies of adolescents. *Journal of Research in Personality*, Article 104461. https://doi.org/10.1016/j.jrp.2024.104461
   (Note: On the date of submission, the revised resubmission of the manuscript was currently under review. Updated reference with journal and DOI on January 20th, 2024.)

3. **Odermatt, S. D.**, Grieder, S., Schweizer, F., Bünger, A., & Grob, A. (2023). *The role of language proficiency in the assessment of cognitive and developmental functions in children: An analysis of the Intelligence and Development Scales–2.* Manuscript submitted for publication.

Spezifizierung des eigenen Forschungsbeitrags zu den Manuskripten:

1. Eigener Beitrag nach CRediT[1]:

   ☒ Conceptualization ☒ Data curation ☒ Formal Analysis
   ☐ Funding acquisition ☒ Investigation ☒ Methodology
   ☐ Project administration ☐ Resources ☒ Software
   ☐ Supervision ☒ Validation ☒ Visualization
   ☒ Writing – original draft
   ☒ Writing – review & editing

   Das Manuskript wurde in dieser Form bisher für keine anderen Qualifikationsarbeiten eingereicht.

2. Eigener Beitrag nach CRediT[1]:

   ☒ Conceptualization ☐ Data curation ☒ Formal Analysis
   ☐ Funding acquisition ☐ Investigation ☒ Methodology
   ☐ Project administration ☐ Resources ☐ Software
   ☐ Supervision ☐ Validation ☒ Visualization
   ☒ Writing – original draft
   ☒ Writing – review & editing

   Das Manuskript wurde in der bestehenden Form bisher für keine anderen Qualifikationsarbeiten eingereicht. Das Manuskript wurde ausgehend von einer früheren Version, welche Dr. Florine Schweizer für ihre Qualifikationsarbeit eingereicht hat, erarbeitet und in seiner Form substanziell verändert (d.h. Einbezug anderer Variablen und Stichproben, Durchführung neuer Analysen mit anderen statistischen Methoden, vollständige Überarbeitung des bestehenden Textes durch die Generierung neuer Textpassagen).

3. Eigener Beitrag nach CRediT[1]:

   ☒ Conceptualization ☐ Data curation ☒ Formal Analysis
   ☐ Funding acquisition ☒ Investigation ☒ Methodology
   ☐ Project administration ☐ Resources ☐ Software
   ☐ Supervision ☐ Validation ☒ Visualization
   ☒ Writing – original draft
   ☒ Writing – review & editing

   Das Manuskript wurde in dieser Form bisher für keine anderen Qualifikationsarbeiten eingereicht.

---

[1] https://casrai.org/credit/

Universität Basel

Fakultät für Psychologie

Open-Science Aspekte der Manuskripte:

1. Preregistration: ☐ ja ☒ nein

   Open-Access-Publikation: ☒ ja ☐ nein

   Open-Access-Data/Analyse: ☐ ja ☒ nein

2. Preregistration: ☐ ja ☒ nein

   Open-Access-Publikation[2]: ☐ ja ☐ nein

   Open-Access-Data[3]: ☐ ja ☒ nein

   Open-Access-Analyse[3]: ☒ ja ☐ nein

   URL der Analysen: https://osf.io/y3c5z/?view_only=ddeb7c5e62d74492af76c2969c4da1f4

3. Preregistration: ☐ ja ☒ nein

   Open-Access-Publikation[2]: ☐ ja ☐ nein

   Open-Access-Data[3]: ☐ ja ☒ nein

   Open-Access-Analyse[3]: ☒ ja ☐ nein

   URL der Analysen: https://osf.io/shrzk/?view_only=5a6f4444a2ff4c5d87885524c153ed86

Ort, Datum ………………………………………………….

Signatur ………………………………………………….

Vorname Nachname ………………………………………………….

---

[2] Zum Zeitpunkt der Einreichung der Dissertation noch nicht definiert.
[3] Da sich die Antworten in Bezug auf Open Access für die Daten und Analysen bei diesem Manuskript unterscheiden, werden diese beiden Bereiche getrennt dargestellt.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to the following people who supported me on my PhD journey:

- To Prof. Dr. Alexander Grob for your guidance and continuous support, and for giving me the opportunity to write this dissertation and to conduct my own research project by collecting data on autistic children and adolescents. Thank you for broadening my horizons with your overarching views and thoughts, and for providing me with many occasions where I could learn and grow.
- To Prof. Dr. Wenke Möhring, not only for being the second supervisor of this dissertation, but also for your constant encouragement and valuable advice in every respect. Your drive and dedication have been an inspiration. Thank you for your supportive nature that helped me follow my path on my doctoral journey.
- To Prof. Dr. Rui Mata for serving as chair on my dissertation committee.
- To my coauthors for your helpful feedback and thoughtful comments. I want to deeply thank Dr. Rebekka Weidmann, Dr. Silvia Grieder, Dr. Florine Schweizer, and Dr. Anette Bünger for the unwavering support you have shown me throughout my PhD and beyond. Thank you for your inspiring company in so many different ways and for enriching my life with your wisdom, kindness, and optimism.
- To current and former members of the PEP team and the “D-Treff” for the scientific discussions and personal exchanges. A heartfelt thank you goes to Leila T. Schächinger Tenés, Dr. Stephanie Klupp, and—especially—to Dr. Silvia Meyer for your open doors, for sharing ups and downs, and for motivating me to look for new ways when others did not work.
- To current and former members of the test development team for the time you have put into our project. I also want to express my gratitude to the master’s students, interns, and research assistants for their help during data collection. Moreover, I wish to thank all the children, adolescents, and parents who participated in our studies, without whom this dissertation would not have been possible.
- To my dear colleagues from the School Psychology Service Basel-Stadt for your interest in my work and belief in my abilities, which have been a great source of encouragement.
- To my partner, Lukas Jäger, for the endless support over the past years—not only emotionally, but also with the most delicious food creations—as well as for always motivating me and for being by my side for almost half of my life. Thank you for being my anchor whenever I need it.
- To my parents, Liliane and Werner Odermatt, and to my brother and his wife, Dr. med. Jonas Odermatt and Dr. med. Natalie Friedli, for your unconditional love, trust, and invaluable support throughout my doctoral journey and every day. Thank you for always believing in me.
- Last but not least, I want to thank my godchildren, Noélie and Yanna, and my little nephew, Lionel, for the joy and happiness you bring into my life. When I see you growing up and discovering the world, my “developmental-psychology heart” is fulfilled.

TABLE OF CONTENTS

ABSTRACT

In psychological research and practice, test procedures that assess children's and adolescents' developmental domains play a fundamental role as they serve a variety of important purposes. For example, practitioners such as school psychologists base their high-stakes diagnostic decisions and recommendations for future schooling support measures on the results of these instruments. To ensure accurate interpretations of scores obtained from such test procedures, their validity needs to be corroborated. At the same time, a holistic assessment of the core developmental domains (i.e., cognitive abilities, psychomotor skills, social-emotional skills, cultural skills, noncognitive personality characteristics) is integral to enabling an in-depth understanding of children's and adolescents' unique strengths and areas of development. In the pursuit of extending current knowledge on validity evidence and contributing to a comprehensive developmental assessment of children and adolescents, the current cumulative dissertation aims to provide insights on three validity aspects of test scores from the Intelligence and Development Scales (IDS) and the Intelligence and Development Scales–2 (IDS-2). Following this goal, three studies were conducted that analyzed (1) differential validity for a clinical subgroup of autistic individuals, (2) criterion validity for multi-informant academic achievement in typically developing individuals, and (3) the contribution of participant characteristics, in particular aspects of proficiency in the test language, to test performance, which may interfere with the validity of score interpretations in individuals at risk for linguistic disadvantages.

Samples consisting of children and adolescents aged between 5 and 20 years drawn from the IDS and IDS-2 standardization and validation studies and from additional validation studies were investigated with independent-samples *t* tests, structural equation modeling analyses, and hierarchical regression analyses. Results of the first study demonstrated the differential validity of test scores in psychomotor skills, language skills, and externally rated achievement motivation of the IDS-2 for autistic children and adolescents and showed that group differences in some developmental domains (e.g., social-emotional skills) may be dependent on factors such as age. The second study provided evidence on the criterion validity of IDS and IDS-2 intelligence scores for multi-informant academic achievement. Moreover, it revealed the incremental role of noncognitive personality characteristics, such as traits and motives, in predicting most of the objective and subjective academic achievement measures. Last, the third study uncovered the contribution of proficiency in the test language as a critical participant characteristic in the assessment of the IDS-2 domains. The findings suggested that the relative importance of different language aspects depends mainly on the verbal demands of the presented tasks.

In conclusion, the present work provides evidence on the validity of test scores from two comprehensive test batteries across different groups of children and adolescents and contexts. In addition, this dissertation highlights the need for a fine-grained view in the process of validating test procedures, offers conclusions on practical implications, and supports a holistic assessment of children's and adolescents' various developmental domains to ensure that the full potential of the individual is captured and unlocked.

1. Introduction

According to the constructivist view of development (Binet, 1909), children's and adolescents' developmental domains (i.e., cognitive abilities, psychomotor skills, social-emotional skills, cultural skills, noncognitive personality characteristics) are interrelated and involved in the active process of constructing their knowledge of the world. In this vein, a holistic assessment of children's and adolescents' development is required to gain a comprehensive understanding of their unique patterns of strengths and difficulties in and across various developmental domains and, hence, their full potential.

Moreover, results of individuals' performance on psychological tests lay the foundation for high-stakes diagnostic decisions, conclusions about future developmental progress, and the planning of support measures or educational placement (Flanagan & Harrison, 2012; Grob & Hagmann-von Arx, 2018b). Assessing children's and adolescents' developmental domains in a comprehensive way thus enables the provision of individualized support tailored to the needs of the individual.

However, for recommendations and interventions to be as effective and beneficial for the individuals' development as possible, the use of instruments that yield accurate and valid test scores is the core prerequisite. Hence, it is essential to examine psychological test procedures in terms of their fulfilment of testing standards and psychometric properties, such as validity (American Educational Research Association [AERA] et al., 2014). Validity is considered the foremost feature of test construction and evaluation as it addresses the extent to which conclusions drawn from scores align with evidence and theoretical foundations (AERA et al., 2014). For example, if an intelligence test is conducted to make predictions about children's future educational development, but the test shows no association with academic achievement, the interpretation of test scores for school-related decisions is invalid.

Validity can be further subdivided by its sources, including differential validity (i.e., whether test scores can differentiate between [clinical] subgroups; Schmidt-Atzert & Amelang, 2012) and criterion validity (i.e., whether test scores correlate with other criteria; Moosbrugger & Kelava, 2012). In addition, validating a test also means considering factors other than the measured construct, such as participant characteristics, that may contribute to test performance and therefore might compromise the validity of score interpretations (AERA et al., 2014). Yet, previous research on validity evidence for many tests has commonly investigated test scores from only a single domain, such as intelligence, leaving other crucial developmental domains out of consideration. In addition, previous studies have often relied on coarse-grained analyses, omitting potentially important aspects (e.g., distinct measures of academic achievement or language abilities) that may provide valuable insights for deriving implications for practice and research. Moreover, particularly for clinical subpopulations, the majority of previous validity evidence is limited and based on small sample sizes with constrained representativeness.

To address these gaps, the present cumulative dissertation aims to expand current evidence on the validity of assessments of developmental domains that are significant for children's and adolescents' development and hence often assessed in psychological practice. This work therefore follows an integrative approach and focuses on test batteries that are designed to assess multiple domains. In particular, this dissertation investigates the Intelligence and Development Scales–2 (IDS-2; Grob & Hagmann-von Arx, 2018a), which assess core developmental domains within a single test battery, and its precursor, the Intelligence and Development Scales (IDS; Grob, Meyer, & Hagmann-von Arx, 2009). In addition, the specific goal of this dissertation is to emphasize three aspects that relate to the validity of test scores across different groups of children and adolescents and different contexts. Study 1 examined differential validity of test scores from developmental domains for a clinical subgroup of autistic children and adolescents. Study 2 investigated criterion validity of intelligence test scores for multi-informant academic achievement measures in the school context by also considering noncognitive personality characteristics in typically developing adolescents. Study 3 analyzed the contribution of participant characteristics (i.e., aspects of proficiency in the test language) to test performance in developmental domains to evaluate the validity of test score interpretations in children at risk for linguistic disadvantages.

In the following, Section 2 provides an overview of the theoretical background of developmental domains, their assessment, and the concept of validity and a synthesis of previous research. Section 3 outlines the research questions. Section 4 provides information on the methods while Section 5 summarizes the results of the studies included in this dissertation. Section 6 presents a general discussion suggesting avenues for future studies and drawing conclusions for research, practice, and politics.

2. Theoretical Background

2.1 Developmental Domains

Developmental domains encompass cognitive abilities, areas of motor, social-emotional, and cultural skills, and noncognitive personality characteristics, which are crucial for a variety of life outcomes. Their paramount importance has been revealed in studies that demonstrated relationships with academic achievement (e.g., Cameron et al., 2016; Cortés Pascual et al., 2019; Romano et al., 2010; Roth et al., 2015) and occupational attainment (e.g., Bailey, 2007; Gross et al., 2009; Jones et al., 2015; Schmidt & Hunter, 2004), as well as with health, well-being, and longevity (e.g., Gottfredson & Deary, 2004; Ozer & Benet-Martínez, 2006; Reimann et al., 2020; Robinson et al., 2015).

2.1.1 Cognitive Abilities

Cognitive abilities comprise versatile skills and mental processes used to execute tasks related to learning, reasoning, comprehension, memorization, and perception (American Psychological Association [APA], 2015). Although cognitive abilities—such as intelligence and executive functions—are closely related (Berk, 2018), they represent separable constructs (e.g., Friedman et al., 2006), as they encompass different aspects of cognitive functioning, which are introduced in the following.

2.1.1.1 Intelligence

For over a century, human intelligence has been intensively examined in terms of its definition, structure, and measurement, making intelligence one of the most studied attributes in psychological research to date (Rost, 2009; Stern & Neubauer, 2016). Although many attempts to define intelligence in a generally accepted way have failed, a scientific consensus on the main components was reached in the 1990s by a group of 52 intelligence researchers (Gottfredson, 1997). They described intelligence as “a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience” (Gottfredson, 1997, p. 13).

In more than 100 years of psychometric intelligence research, numerous theories and models of the structure of intellectual functioning have been generated. Spearman (1904) was the first to lay the foundations with his *g* or *two-factor theory*. He posited that each intellectual task consists of a measure of general intelligence (*g*)—underlying all intellectual abilities—and a task-specific component (*S*) separate from *g.* Currently, the most influential theory in research and contemporary test construction is the Cattell–Horn–Carroll (CHC) model (McGrew, 1997, 2009; Schneider & McGrew, 2018). As the name implies, the CHC model is based on other intelligence theories, namely, (1) Cattell and Horn’s extension of the *Gf-Gc theory* (Cattell, 1941; Horn, 1991; Horn & Cattell, 1966), which incorporates narrow and broad abilities, such as Fluid Reasoning (Gf) and Comprehension-Knowledge (Gc),[1] but not an overall general intelligence factor *g*; and (2) Carroll's (1993) *three-stratum theory*, which separates intelligence into three strata consisting of narrow abilities, broad abilities, and a *g* factor, according to Spearman's (1904) *g*. Despite inconsistencies between the theories about the existence of *g*, the CHC model integrated the Cattell–Horn Gf-Gc and Carroll's three-stratum theory into one model. Following Carroll (1993), the CHC model is hierarchical in structure, with narrow abilities on Stratum I, broad abilities on Stratum II (e.g., Gf, Gc, Visual Processing [Gv], Processing Speed [Gs], Working Memory Capacity [Gwm], Auditory Processing [Ga], Learning Efficiency [Gl], Retrieval Fluency [Gr]), and in most cases a *g* factor at the top on Stratum III (Schneider & McGrew, 2018).

2.1.1.2 Executive Functions

In contrast to the long tradition of intelligence research, executive functions in their current understanding have only recently gained the attention of the scientific community (Miyake et al., 2000). Executive functions comprise top-down cognitive processes used for control and regulation in situations where it is not possible to count on automatic responses, instinct, or intuition (Diamond, 2013; Miyake & Friedman, 2012). These functions enable people, for instance, to concentrate and pay attention despite distractions (Diamond, 2013). In terms of the structure of executive functions, three main components have been identified, which are interrelated and at the same time differentiable (Miyake et al., 2000). These are (1) updating, which subsumes actively holding, monitoring, and manipulating verbal and visuospatial information; (2) inhibition, which entails intentionally controlling and overriding prepotent thoughts and reactions; and (3) shifting (or cognitive flexibility), which involves the ability to mentally switch between tasks or perspectives (Baddeley, 1998; Diamond, 2020; Miyake et al., 2000; Miyake & Friedman, 2012). These three core components are related to other constructs, such as attention (e.g., Drechsler, 2007), and provide the basis for higher-order executive functions (e.g., planning; Collins & Koechlin, 2012; Diamond, 2020; Lunt et al., 2012).

Cognitive abilities are crucial because they are closely related to acquiring and processing information and support the functioning of other developmental domains, such as motor skills, through bidirectional associations (e.g., Adolph & Joh, 2007; Klupp et al., 2023). Hence, further developmental domains, including motor skills, social-emotional skills, cultural skills, and noncognitive personality characteristics, also need to be considered to gain a comprehensive picture of children and adolescents.

2.1.2 Psychomotor Skills

Motor skills entail the individual's capacity to perform goal-directed movements (Burton & Rodgerson, 2001) that involve specific muscles of the body to execute activities such as walking, tying shoes, or writing and are therefore crucial for participation in daily life (e.g., Feder & Majnemer, 2007). Motor skills are also referred to as psychomotor skills because they comprise mental aspects including cognitive, perceptive, sensory, and motivational mechanisms (Grob & Hagmann-von Arx, 2018b; Singer & Bös, 1994). Moreover, these skills are often divided into different components, such as gross, fine, and visuomotor skills (Grob & Hagmann-von Arx, 2018a; Henderson et al., 2007). Gross motor skills rely on large muscle groups and are the "building blocks" needed for balance (e.g., standing on one leg), object control (e.g., throwing a ball), and locomotion (e.g., walking; Gallahue et al., 2012; Logan et al., 2018). In comparison, fine motor skills require small muscle movements of the hands and fingers that involve hand–eye coordination (e.g., grasping an object; Clark & Whitall, 1989; Strooband et al., 2020). Last, visuomotor

[1] Baltes (1987, 1990) referred to Gf and Gc as fluid mechanics (i.e., knowledge-independent basic cognitive operations) and crystallized pragmatics (i.e., the application of fluid mechanics), respectively.

skills include—besides the hand–eye coordination of fine motor skills—visual perception abilities (e.g., drawing geometric figures; Heubrock et al., 2004).

Especially for children and adolescents, psychomotor skills are often used and trained in social interactions (Leonard & Hill, 2014). The ability to move enables them to actively engage with their social environment, for example, in situations of joint attention with their parents or during play and team sports with their peers (Clearfield, 2011; Smyth & Anderson, 2000). Taking this interplay into account, social-emotional skills provide a further crucial domain of children's and adolescents' development.

2.1.3 Social-Emotional Skills

Humans are by nature social beings who begin to develop their social-emotional skills from early childhood through interactions with their environment to form secure relationships (Yates et al., 2008). Social-emotional skills are defined as the ability to "experience, regulate, and express emotions in socially and culturally appropriate ways" (Yates et al., 2008, p. 2) and encompass—besides the abilities of self-regulation and showing developmentally appropriate behavior—the subdomains emotional and social competence (Halle & Darling-Churchill, 2016). These two are closely intertwined as emotional competence, which involves abilities such as expressing, recognizing, understanding, and regulating emotions (Halle & Darling-Churchill, 2016; Petermann & Wiedebusch, 2016), represents a prerequisite for acting in a socially competent manner (Blair et al., 2004). On the other hand, social competence, described as the extent to which individuals are effective in social interactions (Rose-Krasnor, 1997), comprises an individual's ability to establish and maintain relationships and to adapt their behavior to meet the demands of different social environments (Fabes et al., 2006; Halle & Darling-Churchill, 2016; Han & Kemple, 2006). To attain social competence, prosocial behavior, social problem-solving and relationship skills, and empathy are needed (Denham, 2006; Janke, 2008).

In addition to social-emotional skills, cultural skills constitute an integral piece of the puzzle when considering the multidimensional development of children and adolescents. Research indicated important associations between social-emotional skills and cultural skills (Eisenberg et al., 2005). For instance, children and adolescents with more adaptive emotion regulation strategies show an advantage in mathematical competence (e.g., Kahl et al., 2021) and less language impairment (e.g., Fujiki et al., 2002). Therefore, cultural skills also need to be taken into account.

2.1.4 Cultural Skills

Cultural skills, also referred to as basic skills, subsume an individual's knowledge acquired through learning and training such as in school or at home including language skills, reading, writing, and mathematical skills (APA, 2015; Köller & Baumert, 2008). The acquisition of language skills, which encompass—among other aspects—an individual's ability to understand (i.e., receptive language ability) and to produce (i.e., expressive language ability) spoken words, represents an essential developmental task during the early childhood years (Kauschke, 2012; Weinert & Grimm, 2018). Moreover, language skills relate to the acquisition of literacy skills such as reading and writing (Melby-Lervåg et al., 2012). Whereas reading comprises the ability to decode and comprehend written texts (Hoover & Gough, 1990; Rost, 2001), writing skills include elements of spelling, grammar, and content (APA, 2015). Along with mathematical skills, literacy skills constitute one of the most important pillars learned in school, given their significance for participation in society (Bos et al., 2010). Finally, mathematical competence entails a wide range of skills acquired throughout the (educational) development such as knowledge of numbers, arithmetic, and shape and space (Dowker, 2019; Swiss-German Conference of Directors of Education, 2016) and can be defined as an individual's "insightful readiness to act appropriately in response to all kinds of mathematical challenges" (Niss & Højgaard, 2019, p. 12).

Research has demonstrated that, alongside cognitive abilities, individuals' noncognitive personality characteristics play a significant role in different measures of cultural skills (e.g., J. Meyer et al., 2019). Personality characteristics therefore complement the developmental domains of children and adolescents.

2.1.5 Noncognitive Personality Characteristics

According to Roberts and Wood's (2006) neo-socioanalytic theory, personality encompasses—in addition to cognitive aspects (i.e., abilities)—noncognitive characteristics, which subsume traits (e.g., conscientiousness), values and motives (e.g., achievement [striving] motivation), and narratives (e.g., narrative identity). Traits and motives are particularly crucial for children's and adolescents' development because of their important relationships to educational success (Lavrijsen et al., 2022; Mammadov, 2022). In contrast, narratives are thought to develop across adolescence and become more coherent later in life (Habermas & Bluck, 2000; Köber et al., 2015). These are therefore not included in this work.

Traits comprise persistent patterns of thoughts, feelings, and behaviors (Roberts & Wood, 2006) and are often assessed using the Big Five taxonomy (John et al., 2008). One of the Big Five traits is conscientiousness, which can be defined as the tendency to adhere to socially prescribed norms for goal-oriented behavior, impulse control, planning, and delaying rewards (Roberts et al., 2009) and entails "the readiness to do academic work" (Di Domenico & Fournier, 2015, p. 157). Hence, conscientious individuals are more likely to be described as dutiful, organized, and hardworking (Costa & McCrae, 1992; Roberts et al., 2009). On the other hand, values and motives refer to the qualities that are considered desirable to a person (Roberts & Wood, 2006) and have been grouped within the heterogeneous and multifaceted concept of motivation (Eccles & Wigfield, 2002). Thereby, achievement motivation represents an individual's endurance and perseverance on achievement-related tasks (Wigfield & Eccles, 2000) including the "willingness to do academic work" (Di Domenico & Fournier, 2015, p. 156). Finally, and more narrowly, its facet achievement striving (henceforth: *achievement striving motivation)* can be seen as an individual's need to succeed in life (Murray, 1938).

2.1.6 Summary

In sum, intelligence encompasses an individual's general mental ability to reason, solve problems, and learn quickly, whereas executive functions include an individual's set of control and regulatory mechanisms. Psychomotor skills can be seen as the individual's fundamental abilities to execute tasks of daily life and to explore their surroundings. On the other hand, to sustain connections with the social environment, social-emotional skills that refer to the individual's ability to read, understand, and regulate emotions and to show socially competent behavior are required. Cultural skills build on what has been understood and learned from previous experiences, for example, at home or at school. Finally, personality consists of different layers including noncognitive characteristics such as traits and motives. Thus, developmental domains comprise a multitude of areas and constructs that are of relevance not only for everyday life but also for the long-term development of children and adolescents. Moreover, these areas are interrelated and mutually reinforce each other. Hence, for a comprehensive assessment of human development—and an individual's full potential—it is necessary to take into account various developmental domains.

2.2 Assessment of Developmental Domains

As developmental domains represent latent constructs, they need to be operationalized with tasks or questions that assess the intended behavior or characteristic (Schmidt-Atzert & Amelang, 2012). These tasks and questions are bundled into standardized performance tests or rating scales. Test procedures are therefore based on theoretical models and concepts of the latent constructs (e.g., Schneider &

Flanagan, 2015). Although intelligence is the construct most frequently measured by practitioners, such as school psychologists (Benson et al., 2019), individuals express additional relevant domains, for instance, social-emotional and psychomotor skills. Even though such domains develop rapidly during childhood and adolescence (Siegler et al., 2021), they tend to be rather neglected in the assessment of children after the age of school enrollment. This might be partly explained by the lack of tests covering developmental domains other than intelligence beyond the pre-school years (Grob, Hagmann-von Arx, & Bodmer, 2009), which leads to the use of multiple test procedures or nonstandardized observational data when a broad evaluation is needed. Nevertheless, tests rely on different theoretical backgrounds, adopt different test administration methods, and are based on distinct characteristics of standardization samples. These are factors that must be considered when drawing conclusions from various test batteries. However, in recent years, two instruments that include multiple developmental domains within a single test have been developed, namely, the IDS (Grob, Meyer, & Hagmann-von Arx, 2009) and the IDS-2 (Grob & Hagmann-von Arx, 2018a). The IDS assesses cognitive (i.e., intelligence) and developmental (i.e., psychomotor skills, social-emotional skills, mathematics, language skills, achievement motivation) functions in children aged 5–10 years, whereas the IDS-2 measures cognitive (i.e., intelligence, executive functions) and developmental (i.e., psychomotor skills, social-emotional skills, basic skills, motivation and attitude) functions in children and adolescents aged 5–20 years. As with any test procedure, psychometric properties, such as validity, need to be examined to ensure that accurate and valid conclusions can be established on the basis of their test scores.

2.3 Validity Evidence for Developmental Assessment

Validity is defined as an evaluative judgment (Messick, 1995) concerning "the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests" (AERA et al., 2014, p. 11; see also Moosbrugger & Kelava, 2012; Schmidt-Atzert & Amelang, 2012). It can be structured into various sources including evidence based on *test content*, *response processes*, *internal structure*, and *relations to other variables* (AERA et al., 2014). Thereby, validity evidence should be established according to the specific purposes for which a test is intended to be used (AERA et al., 2014; International Test Commission, 2001).

Since in many assessment situations, the application of a test is based on the assumption that the construct being measured is associated with external variables, evidence of *relations to other variables* is particularly crucial for test validation (AERA et al., 2014). This source of validity evidence can be further divided into categorical and criterion variables. Categorical variables, such as group membership, "become relevant when the theory underlying a proposed test use suggests that group differences should be present or absent if a proposed test score interpretation is to be supported" (AERA et al., 2014, p. 16). This so-called *differential validity evidence* offers insights on whether test scores can distinguish between different subgroups, such as between clinical (e.g., autistic) and nonclinical (e.g., non-autistic) individuals (Schmidt-Atzert & Amelang, 2012). Evidence on differential validity is therefore particularly important for the interpretation of tests that are used in clinical practice.

In contrast, criterion variables are of relevance when the question is how accurately test scores predict variables outside of the test situation (AERA et al., 2014). The corresponding *criterion validity evidence* thus expresses "how well a test correlates with an established standard of comparison (i.e., a criterion)" (APA, 2015, p. 266) and can be further divided depending on whether the criterion is assessed in the future (*predictive validity evidence*) or at about the same time (*concurrent validity evidence*; AERA et al., 2014; Moosbrugger & Kelava, 2012). A third subcategory consists of *incremental validity evidence*, which provides insights on whether adding a new measure to an existing set of variables improves the prediction of a criterion (APA, 2015). Establishing criterion validity evidence is of paramount importance to demonstrate that conclusions based on test results are appropriate for predicting real-life outcomes, such as academic achievement.

In addition, validation requires the careful consideration and examination of measurement factors that may be related to test performance, including administration conditions, testing format, or participant characteristics, such as individuals' level of proficiency in the test language (AERA et al., 2014). Test language proficiency might interfere with performance on assessments when a test places high verbal demands on the participants (e.g., including complex verbal instructions; Cormier et al., 2022). Under such conditions the participants' true ability might be underestimated (Hagmann-von Arx et al., 2013), especially for individuals at risk for linguistic disadvantages, such as those with linguistically diverse backgrounds (e.g., multilingual individuals). Examining the contribution of participant characteristics to test performance is thus essential, as they could compromise the validity of score interpretations (AERA et al., 2014).[2]

However, determining the validity of test scores requires the ongoing examination of empirical evidence from previous findings and from new research, as "the validation process never ends" (AERA et al., 2014, p. 21). Therefore, the present dissertation aims to add validity evidence for assessments of developmental domains by focusing on three aspects: (1) differential validity for autistic individuals, (2) criterion validity for academic achievement, and (3) the contribution of participant characteristics (i.e., aspects of proficiency in the test language) for validity of test score interpretations. The following subsections review previous research on these aspects of validity and their gaps by concentrating on two tests in particular: the IDS and the IDS-2.

2.3.1 Evidence on Differential Validity With a Focus on Autistic Individuals

As the global prevalence of autism spectrum disorder (ASD) has risen to about 1–2% in recent years (e.g., Idring et al., 2015), it is now recognized as a frequent condition (also referred to as "autism epidemic"; Chiarotti & Venerosi, 2020). Consequently, autistic individuals represent a common clinical subgroup in psychological and medical practice. Deficits in social communication and interaction, such as impaired emotion recognition (Yeung, 2022) and regulation (Cai et al., 2018), represent one of the core characteristics—besides restricted repetitive behaviors—of the neurodevelopmental disorder ASD (American Psychiatric Association, 2013). Aside from the core symptoms, autistic individuals often experience difficulties in multiple developmental domains. For example, compared to non-autistic individuals, they show impairments in intelligence and executive functions (e.g., Demetriou et al., 2018; Maenner et al., 2020), obtain lower scores on tasks assessing psychomotor skills, such as gross and fine motor skills (e.g., Coll et al., 2020), and exhibit language deficits and delays (e.g., Kwok et al., 2015).

However, the assessment of developmental domains plays a crucial role in the diagnostic process of evaluating autistic individuals not only for reporting possible cognitive and language impairments according to the standards of the *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; *DSM-5*; American Psychiatric Association, 2013) and the *International Statistical Classification of Diseases and Related Health Problems* (11th ed.; *ICD-11*; World Health Organization, 2018) but also for obtaining valuable information for appropriate treatment and schooling measures (White et al., 2007). Therefore, gaining insights into autistic children's and adolescents' performance in developmental domains on widely used tests is essential to draw accurate conclusions based on test results.

---

[2] Although participant characteristics are also relevant to fairness in testing (AERA et al., 2014), this dissertation focuses specifically on aspects of proficiency in the test language for validity of test score interpretations.

Previous evidence on differential validity of the IDS-2 can be mainly found in the test's technical manual (Grob & Hagmann-von Arx, 2018b). Parallel to the IDS-2 standardization study, validation studies were conducted that compared children and adolescents who belonged to different subgroups with control samples. These comparisons included groups of participants with above-average intelligence (n = 62), intellectual disability (n = 70), attention-deficit/hyperactivity disorder (n = 65), motor problems (n = 120), reading and spelling disorders (n = 22), mathematical giftedness (n = 30), and ASD (n = 18). Yet, for some subgroups, comparisons were based on small, nonrepresentative samples, which reduces generalizability and the power to detect group differences. This is particularly the case for the ASD subgroup, as this sample comprised only 18 children and adolescents (M_{age} = 13.30 years, age range 8–17 years; 1 female participant). In addition, this sample was mainly diagnosed with Asperger's syndrome (n = 13) and not infantile autism, which may limit the representativeness of the spectrum of ASD.[3] Results showed that this group of autistic participants scored lower on the composites of psychomotor skills and social-emotional skills of the IDS-2 compared to controls (Grob & Hagmann-von Arx, 2018b). However, analyses at the level of subtests or intelligence group factors of the IDS-2 have been missing so far for this clinical group. With respect to the IDS, one study (Grob et al., 2013; see also: C. S. Meyer et al., 2009) found lower scores for autistic children (n = 38) on subtests of the intelligence (i.e., Selective Attention, Auditory Memory), psychomotor skills (i.e., Gross Motor Skills), social-emotional skills (i.e., Regulating Emotions, Understanding Social Situations, Socially Competent Behavior), and achievement motivation (i.e., Perseverance) domains of the IDS compared to controls. Nevertheless, this study exclusively included children with Asperger's syndrome.

To summarize, the first aim of the present dissertation was to fill this gap in the previous literature and to extend the current evidence on differential validity for a clinical subgroup, namely, autistic children and adolescents, on the IDS-2. To achieve this goal, further data were collected to obtain a larger sample of autistic individuals with a more representative distribution of sex and subtypes (i.e., including individuals with infantile autism and more girls). In addition, group differences at the level of subtests and intelligence group factors were investigated. Finally, and to explore age effects, previous research was extended by performing age-separated analyses for children and adolescents.

2.3.2 Evidence on Criterion Validity With a Focus on Academic Achievement

The extant literature has demonstrated the significance of intelligence in predicting numerous important life outcomes, including academic achievement (e.g., Kuncel et al., 2004). In line with meta-analytic results that indicated corrected correlations of ρ = .21 to .54 between intelligence and school grades (Richardson et al., 2012; Roth et al., 2015), intelligence, assessed with the IDS and the IDS-2, has been linked to academic achievement measures. Specifically, studies by Gut et al. (2012, 2013) and Gygi et al. (2017) showed that IDS intelligence predicted school grades (i.e., mathematics, language, science, grade point average [GPA]) longitudinally over 3 years in children (age range: 5–11 years). In addition, Gut et al. (2013) indicated that IDS intelligence was a concurrent predictor of parent-reported academic performance. For the IDS-2, concurrent associations of intelligence with grades and parent-reported

---

[3] Until the publication of the *DSM-5* (American Psychiatric Association, 2013) and the *ICD-11* (World Health Organization, 2018), autism was divided into subtypes including *Asperger's syndrome* (absence of general developmental delay, particularly in language skills and cognitive abilities; often clumsiness), *infantile autism* (developmental delay manifesting before the age of 3 years; often additional nonspecific problems), and *atypical autism* (developmental delay manifesting after the age of 3 years or criteria not fully met; often comorbidities such as language disorders; World Health Organization, 1990). Current diagnostic classification standards (*DSM-5*; *ICD-11*) follow a spectrum approach encompassing all autistic individuals of each former subtype because autistic individuals share common features (i.e., the core symptoms of deficits in social communication and interaction and restricted repetitive behaviors).

academic performance (i.e., language; geography and history [combined]; mathematics; biology, chemistry, and physics [combined]; environment) were reported in the technical manual of the IDS-2 (Grob & Hagmann-von Arx, 2018b). A study by Grieder et al. (2022) supported these results and found concurrent relationships between IDS-2 intelligence and school grades (i.e., language, mathematics, GPA) in participants aged 5–19 years.

However, aligned with a systemic point of view, the perspectives of all key actors in the school context (i.e., teachers, parents, *and* students) should be taken into account to gain a more comprehensive understanding of students' academic functioning. In addition to grades and parent-reported academic performance, students' self-reported academic performance offers insights into their learning progress and evaluation of competence (Marsh & Martin, 2011). Therefore, objective (i.e., grades) and subjective (i.e., parent-reported and self-reported academic performance) measures of academic achievement need to be considered. Yet, relations between intelligence and students' self-reported academic performance have not been investigated for the IDS and IDS-2 and have rarely been studied with other intelligence measures. Moreover, evidence for the predictive validity of IDS intelligence for academic achievement over a longer period of time and into adolescence is lacking. Analyses of the IDS-2 that particularly focus on adolescents are missing as well. Academic outcomes are especially important for this age group, as they pave the way for adolescents' further educational and occupational trajectories.

Even though intelligence is considered the most significant predictor of academic achievement (Mammadov, 2022), noncognitive personality characteristics, such as conscientiousness and achievement striving motivation, have also been shown to be crucial (e.g., Richardson et al., 2012; Steinmayr & Spinath, 2007). For example, for grades, meta-analyses reported corrected correlations of ρ = .23 to .50 with conscientiousness (Poropat, 2014; Richardson et al., 2012) and of ρ = .30 with subordinate achievement motivation (Robbins et al., 2004). Studies also indicated that conscientiousness (e.g., Mammadov, 2022) and achievement striving motivation (e.g., Steinmayr & Spinath, 2009) incrementally predict grades beyond intelligence. This empirical research also aligns with theoretical considerations, such as the neo-socioanalytic theory (Roberts & Wood, 2006), which emphasizes the importance of traits and motives, above abilities, for life outcomes. Yet, evidence on the associations between noncognitive personality characteristics and subjective performance ratings from parents and students is sparse. Furthermore, no study to date has taken a comprehensive approach and simultaneously investigated the relationships of intelligence and noncognitive personality characteristics with objective and subjective measures of academic achievement.

Therefore, the present dissertation aimed to contribute to current knowledge on the evidence on criterion validity of IDS and IDS-2 intelligence scores for multi-informant academic achievement in the school context (i.e., grades, parent-reported and self-reported academic performance) in typically developing adolescents. Moreover, previous research was extended by considering noncognitive personality characteristics (i.e., conscientiousness and achievement striving motivation) as concurrent and incremental predictors of academic achievement beyond IDS and IDS-2 intelligence.

2.3.3 Evidence on the Contribution of Participant Characteristics to Validity With a Focus on Aspects of Proficiency in the Test Language

Over the past decades, global migration rates have risen (United Nations, 2022) and therefore more individuals with linguistically diverse backgrounds undergo psychological assessment. As multilingual individuals often lack proficiency in the test language (e.g., Bialystok et al., 2010), the accurate and valid assessment of their developmental domains may be challenging. Specifically, these individuals might face difficulties in, for instance, comprehending verbal instructions or providing verbal answers during testing (Cormier et al., 2022; Weiss et al., 2006). These limitations mainly depend on the extent to which the test requires verbal interaction from the participant, for example, in test directions, response options, and task content (Cormier et al., 2011; Flanagan & Ortiz, 2001). Participants with linguistically diverse backgrounds might not be able to display their full potential (Hagmann-von Arx et al., 2013), which could have negative implications for their future educational and occupational development (Calero et al., 2013; Goldstein et al., 2015; Klingner et al., 2007; Sullivan, 2011). Hence, it is crucial to investigate how proficiency in the test language relates to performance on a test.

Previous evidence on the relations between aspects of proficiency in the test language (e.g., receptive and expressive language abilities, multilingualism) and test scores on the cognitive and developmental functions of the IDS-2 is limited. Only one study, by Schweizer et al. (2021), examined group differences between matched monolingual, simultaneously bilingual, and successively bilingual children and adolescents in the IDS-2 intelligence domain. Results revealed that successive bilinguals demonstrated lower mean values compared to monolinguals, and also to some extent compared to simultaneous bilinguals, in verbal-dependent intelligence scores (i.e., Verbal Reasoning and verbal Long-Term Memory, including corresponding subtests). Nonetheless, to date, evidence on the several other IDS-2 domains (e.g., executive functions, psychomotor skills, social-emotional skills, basic skills) beyond intelligence is lacking. For the IDS—and in line with the findings of the IDS-2—children with another native language than German scored lower than controls on an intelligence subtest measuring verbal long-term memory (Grob et al., 2013). Moreover, studies that also examined the developmental functions of the IDS mainly did not find group differences in the psychomotor skills domain, whereas inconsistent results emerged for the social-emotional skills and mathematics domains (Grob et al., 2013; Hagmann-von Arx et al., 2013). The few previous studies on the IDS and IDS-2 concluded that as the verbal demands of a task increased, group differences between individuals with and without another native language than German were more likely to be observed (Grob et al., 2013; Hagmann-von Arx et al., 2013; Schweizer et al., 2021).

Yet, previous research on the IDS and the IDS-2 followed the widespread approach of comparing different groups. But individuals within these groups may display varying levels of test language proficiency, placing them at different points along the language ability spectrum (Ortiz, 2019). Therefore, a dimensional and more fine-grained examination of the independent contribution of individuals' measurable language abilities is needed. This would help researchers formulate implications for psychological practice (Ortiz, 2019) and result in a better understanding of possible participant characteristics that might impede valid score interpretations. To my knowledge, only one recent study (Cormier et al., 2022) examined the effects of objectively measured language abilities in English on performance in a cognitive test battery (Woodcock–Johnson Tests of Cognitive Abilities, 4th ed.; Schrank et al., 2014). This study found associations with participants' receptive and expressive language abilities. However, evidence on the relations of receptive and expressive language abilities to test performance in other developmental domains is currently missing.

Thus, the last aim of the present dissertation was to close these gaps and add knowledge on participant characteristics that are critical to the valid interpretation of test results by examining the contribution of aspects of proficiency in the test language to performance in the cognitive *and* developmental functions of the IDS-2. Thereby, the current dissertation extends previous research by adopting a dimensional approach and by investigating the relative importance of multiple, distinct aspects (i.e., objectively measured receptive and expressive language abilities, and multilingualism).

3. Research Questions

The present dissertation aims to extend the current evidence on the validity of assessments of developmental domains in children and adolescents by taking an integrative approach (see Figure 1). Specifically, it focuses on three aspects that address the validity of test scores of the IDS and IDS-2 across different groups and contexts: evidence on differential validity for a clinical subgroup of individuals with ASD (Study 1); evidence on criterion validity—split into predictive, concurrent, and incremental validity—for objective and subjective academic achievement measures in the school context for typically developing individuals (Study 2); and, last, the contribution of participant characteristics, namely, aspects of proficiency in the test language, to test performance that may limit the validity of score interpretations in individuals at risk for linguistic disadvantages (Study 3). Note, Study 2 comprised two studies with different samples that are referred to here as Studies 2.1 and 2.2. The following research questions (RQs) were investigated in the three studies:

Differential Validity (Study 1). ***RQ 1a.*** Do autistic and non-autistic children and adolescents show mean-level differences in the IDS-2 cognitive and developmental functions? ***RQ 1b.*** Do autistic and non-autistic children (5–10 years) show the same pattern of mean-level differences in the IDS-2 cognitive and developmental functions as autistic and non-autistic adolescents (11–20 years)?

Criterion Validity (Study 2). ***RQ 2a.*** Is IDS intelligence a valid longitudinal predictor of multi-informant academic achievement in typically developing adolescents (Study 2.1)? ***RQ 2b.*** Is IDS-2 intelligence a valid concurrent predictor of multi-informant academic achievement in typically developing adolescents (Study 2.2)? ***RQ 2c.*** Are noncognitive personality characteristics (i.e., conscientiousness, achievement striving motivation) valid concurrent predictors of multi-informant academic achievement in typically developing adolescents (Studies 2.1 & 2.2)? ***RQ 2d.*** Are noncognitive personality characteristics (i.e., conscientiousness, achievement striving motivation) valid incremental predictors of multi-informant academic achievement in typically developing adolescents, beyond intelligence (Studies 2.1 & 2.2)?

Participant Characteristics (Study 3). ***RQ 3a.*** Does children's receptive language ability explain variance in the IDS-2 cognitive and developmental functions, beyond sex and socioeconomic status (SES)? ***RQ 3b.*** Does children's expressive language ability explain additional variance in the IDS-2 cognitive and developmental functions, beyond sex, SES, and receptive language ability? ***RQ 3c.*** Does multilingualism explain additional variance in the IDS-2 cognitive and developmental functions, beyond sex, SES, and receptive and expressive language abilities?

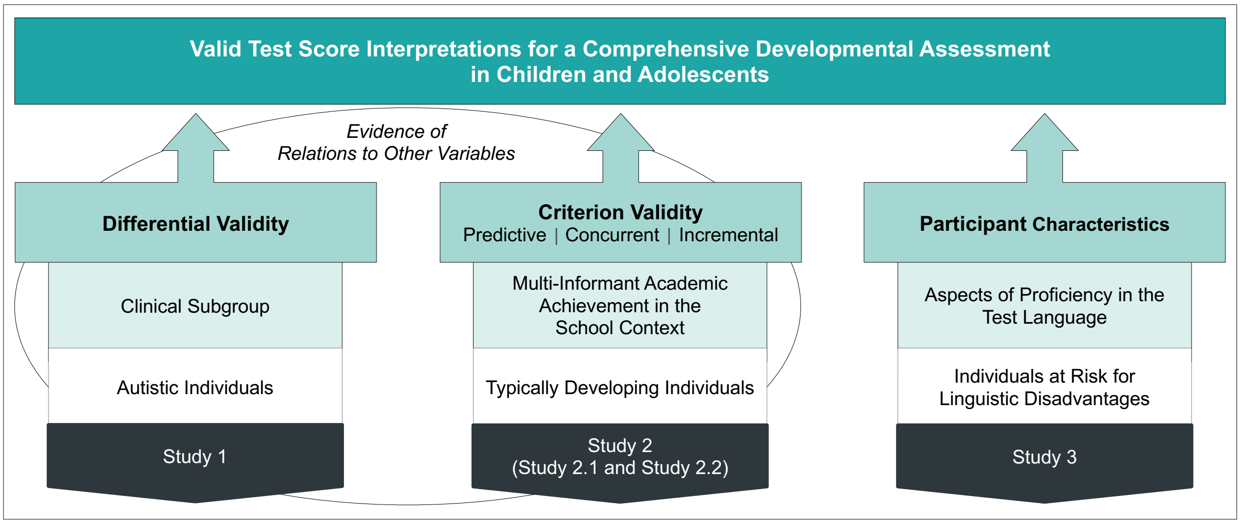


Figure 1. Dissertation concept. Study 1 (Odermatt, Möhring, Grieder, & Grob, 2022); Study 2 (Odermatt, Weidmann, Schweizer, & Grob, 2024); Study 3 (Odermatt, Grieder, Schweizer, Bünger, & Grob, 2023).

4. Methods

4.1 Samples

The samples employed in the three studies were collected as part of the *Test Development* research project in the Division of Developmental and Personality Psychology of the Department of Psychology at the University of Basel. All samples consisted of participants living in the German-speaking part of Switzerland, Germany, or Austria.

Study 1. The first study included data from a sample of children and adolescents aged 7–17 years diagnosed with ASD (n = 43; M_{age} = 12.30 years, SD = 3.08; 8 females) and a matched control sample of non-autistic children and adolescents aged 6–20 years (n = 43; M_{age} = 12.51 years, SD = 3.56; 8 females) obtained from the IDS-2 standardization and validation study and a further validation study. Autistic participants reported being diagnosed with infantile autism (n = 11), atypical autism (n = 6), or Asperger's syndrome (n = 24) or did not provide subtype information (n = 2).

Study 2. The second study was based on data from two independent samples of 766 typically developing individuals in total: Study 2.1's sample consisted of 301 children aged 5–10 years (M_{age} = 8.11 years, SD = 1.54) at Time 1 and adolescents aged 12–18 years (M_{age} = 15.45 years, SD = 1.52; 52% female) at Time 2. Study 2.2's sample consisted of 465 adolescents aged 12–18 years (M_{age} = 15.16 years, SD = 1.56; 53% female). The samples were drawn from the IDS (Study 2.1) and the IDS-2 (Study 2.2) standardization and validation studies.

Study 3. The third study was a subsample of the IDS-2 standardization and validation study consisting of 826 children aged 5–10 years (M_{age} = 8.06 years, SD = 1.66; 51% female) and including multilingual (n = 215) and monolingual (n = 611) individuals. Multilingual participants spoke German (the test language) and had at least one other language as their native language, whereas monolingual participants reported German as their single native language.

4.2 Measures and Procedure

In Studies 1 and 3, developmental domains were assessed with the IDS-2. In Study 2, intelligence was measured using the IDS (Study 2.1) and the IDS-2 (Study 2.2). Moreover, on the same questionnaires in Studies 2.1 and 2.2, conscientiousness was reported by parents, achievement striving motivation was indicated by the adolescent, and objective and subjective measures of academic achievement were assessed by parents and their adolescent child.

IDS and IDS-2. The IDS (Grob, Meyer, & Hagmann-von Arx, 2009) and the IDS-2 (Grob & Hagmann-von Arx, 2018a) are individually administered test batteries incorporating several age-standardized subtest scores and unit-weighted composites. An overview and description of the IDS and IDS-2 domains are presented in Tables 1 and 2, respectively (see pp. 21–23). The IDS was standardized between 2007 and 2008 in the German-speaking part of Switzerland, Germany, and Austria and measures cognitive (i.e., intelligence) and developmental (i.e., psychomotor skills, social-emotional skills, mathematics, language skills, achievement motivation) functions in children between the ages of 5 and 10 years using 19 subtests. The intelligence domain of the IDS is based on Spearman's (1904) *g* or two-factor theory and the concept of fluid mechanics (Baltes, 1987, 1990). The IDS-2 is the successor of the IDS and was standardized between 2015 and 2017 in the same regions. In recent years, a multitude of international language adaptations have been published (e.g., Dutch, English [UK], Italian, Polish; Grob et al., 2018, 2019, 2021, 2022) and further adaptations are underway in several other countries (e.g., Brazil, Denmark, Finland, France, Norway, Spain, Sweden, and the United States). The IDS-2 assesses cognitive (i.e., intelligence, executive functions) and developmental (i.e., psychomotor skills, social-emotional skills, basic skills, motivation and attitude) functions in children and adolescents between the ages of 5 and 20 years using 30 subtests. Among these is a measurement of participants' receptive and expressive language abilities. The intelligence domain of the IDS-2 is based on Spearman's (1904) *g* or two-factor theory and the CHC model (McGrew, 1997, 2009; Schneider & McGrew, 2018).

Self- and Parent-Reported Measures. In Studies 2.1 and 2.2, parents reported on their child's conscientiousness using the German Five Factors Questionnaire for Children [Fünf-Faktoren-Fragebogen für Kinder] (FFFK; Asendorpf, 1998). This questionnaire (α = .83 to .91; Asendorpf & Van Aken, 1999) contains eight items, each consisting of bipolar adjectives. Answers are given on a 5-point Likert scale (range: 1 to 5). In addition, adolescents rated their achievement striving motivation using the German Achievement Motivation Questionnaire for 7th- to 13th-Grade Students [Fragebogen zur Leistungsmotivation für Schüler der 7. bis 13. Klasse] (FLM 7-13; Petermann & Winkel, 2007). This questionnaire (α = .73; Petermann & Winkel, 2007) includes eight items evaluated on a 5-point Likert scale (range: 1 to 5). Regarding the academic achievement measures, parents reported their child's school grades in five subjects (i.e., language; geography and history [combined]; mathematics; biology, chemistry, and physics [combined]; environment) according to the school records (range: 1 to 6). Subjective performance ratings were assessed using a social comparison: Parents reported the academic performance of their child for each of the five subjects on a 5-point Likert scale (range: 1 to 5) by comparing their child's academic performance with the performance of their child's peers. Adolescents likewise compared their academic performance with the performance of their peers for each of the five subjects on a 5-point Likert scale (range: 1 to 5).

4.3 Statistical Analyses

Study 1, Study 2.2, and Study 3 were cross-sectional studies, whereas Study 2.1 followed a partial longitudinal design, as intelligence was assessed 7 years before the other variables. Multiple statistical approaches were used to enhance the evidence on validity of developmental domains. All analyses were conducted in R (R Core Team, 2022).

Study 1. In Study 1, the differential validity of scores from the cognitive and developmental functions of the IDS-2 was investigated by comparing autistic and non-autistic participants for mean-level differences. First, we matched the autistic sample and the non-autistic control sample on demographic characteristics (i.e., age, sex, SES). We then performed independent-samples *t* tests and accounted for multiple testing by adjusting the *p* values with Hommel's (1988) correction. We also conducted age-specific analyses for children (5–10 years) and adolescents (11–20 years). We interpreted mean-level differences as meaningful if the corrected *p* value was significant and at least a small effect size was detected. For the matching, we used the MatchIt package (Ho et al., 2011).

Study 2. In Study 2, we examined the criterion validity of IDS and IDS-2 intelligence scores for multi-informant academic achievement by considering noncognitive personality characteristics using structural equation modeling and hierarchical regressions. For theoretical considerations, we categorized the five school subjects into the following school-subject domains:[4] (1) *humanities*, consisting of (1a) language and (1b) geography and history; (2) *science*, consisting of (2a) mathematics and (2b) biology, chemistry, and physics; and (3) *environment* including the corresponding subject. To investigate predictive and concurrent validity for intelligence and noncognitive personality characteristics, we used structural equation modeling analyses. We computed parcels to model the latent factors intelligence, conscientiousness, and achievement striving motivation according to the item-to-construct balance

[4] Geography and history (1b) as well as biology, chemistry, and physics (2b) were treated as one subject each because the combined average grade of these individual subjects was asked for in the questionnaires.

technique (Little et al., 2002). We also compared the models of Studies 2.1 and 2.2 by analyzing the overlap of the 95% confidence intervals of the regression coefficients and by computing a χ^2 difference test of constrained and unconstrained multigroup structural equation models. We conducted analyses controlling for sex and SES in the models. To examine incremental validity for conscientiousness and achievement striving motivation, we used hierarchical regression analyses and added control variables (i.e., sex, SES) in Step 1, intelligence in Step 2, conscientiousness in Step 3, and achievement striving motivation in Step 4. We performed analyses for overall estimates and specific school-subject domains.

Study 3. In Study 3, we investigated whether participant characteristics in the form of different aspects of proficiency in the test language (i.e., receptive language ability, expressive language ability, multilingualism) contribute each to the explanation of additional variance in test scores from the cognitive and developmental functions of the IDS-2 using hierarchical regression analyses. Control variables (i.e., sex, SES) were entered in Step 1, children's receptive language ability in Step 2, their expressive language ability in Step 3, and the variable multilingualism in Step 4. We adjusted the *p* values with Hommel's (1988) correction to control for multiple testing.

5. Synopsis of Results

Study 1. The results of Study 1 revealed that the participants of the autistic sample showed significantly lower group mean values compared to the participants of the non-autistic control sample for the domains of psychomotor skills (i.e., composite, Gross Motor Skills, Fine Motor Skills) and language skills (i.e., composite, Phoneme Analysis, Language Receptive), and for the score of the evaluation of participation during testing of the developmental functions of the IDS-2. Effect sizes were in the medium-to-large range. The largest effect size was found for Gross Motor Skills. No significant group differences between autistic and non-autistic participants were detected in the other developmental domains or any cognitive functions of the IDS-2 after controlling for multiple testing. Separate analyses for children aged 5–10 years and adolescents aged 11–20 years revealed age-specific results: Autistic children showed significantly lower group mean values than non-autistic children in the composites of the cognitive functions, in three intelligence group factors (i.e., Auditory and Visuospatial Short-Term Memory, Verbal Reasoning), and in four intelligence subtests (i.e., [Rotated] Shape Memory, Naming Opposites, Story Recall). In the developmental functions, they scored lower in composites and subtests of the psychomotor skills, social-emotional skills, and basic skills domains. For adolescents, no significant differences were found between autistic and non-autistic adolescents in the cognitive and developmental functions of the IDS-2.

Study 2. In Studies 2.1 and 2.2, a largely similar pattern of results emerged. Intelligence was a longitudinal (IDS; Study 2.1) and cross-sectional (IDS-2; Study 2.2) predictor of objective and subjective measures of academic achievement. Conscientiousness was concurrently related to and incrementally explained variance beyond intelligence in grades and parent-reported academic performance, whereas in most analyses no associations were found with adolescents' self-reported academic performance. Achievement striving motivation was mostly concurrently linked to and incrementally explained variance beyond intelligence (and conscientiousness) in grades and subjective performance ratings. These results appeared for overall estimates as well as for specific school-subject domains. Comparisons of the structural equation models of Studies 2.1 and 2.2 showed no significant differences between their results. Controlling for sex and SES in the models yielded a largely similar pattern of results.

Study 3. The results of Study 3 showed that, after controlling for sex and SES, children's receptive language ability was significantly related to almost all scores of the cognitive and developmental functions of the IDS-2 (except for Socially Competent Behavior). Overall, children's expressive language ability accounted for little additional variance in the IDS-2 scores beyond receptive language ability. The highest amounts of additional explained variance emerged for the intelligence composites, the intelligence group factor Verbal Reasoning, including corresponding subtests Naming Categories and Naming Opposites, and for the basic skills subtests Logical-Mathematical Reasoning, Reading, and Spelling. Finally, multilingualism, included in the last step, explained variance beyond objectively measured language abilities only in subtests of Verbal Reasoning (i.e., Naming Opposites) and verbal Long-Term Memory (i.e., Story Recall).

6. General Discussion

The main goal of this dissertation was to extend current evidence on the assessment of core developmental domains in children and adolescents by examining three aspects that address the validity of test score interpretations across different groups and contexts. In the following, the evidence generated by this work on differential validity, criterion validity, and participant characteristics is summarized and discussed in light of previous findings. In addition, the overarching strengths and limitations of the current dissertation, as well as directions for future research efforts, are presented. Finally, a general conclusion is drawn.

6.1 Differential Validity

The first aim of this dissertation was to contribute to knowledge on differential validity for a clinical subgroup on the IDS-2. Specifically, it focused on whether theoretically or empirically derived group differences are present or absent in test score interpretations of children and adolescents with and without ASD. To accomplish this goal, the first step was to collect further data to obtain a larger and more representative sample of autistic individuals with respect to sex and subtypes. As the final sample consisted of n = 43 autistic individuals with different subtypes (i.e., also including infantile autism) and the male-to-female ratio corresponded to that of the autistic population (i.e., 4:1; Maenner et al., 2020), the current work meets this objective.

When autistic children and adolescents were then compared to a non-autistic control sample on developmental domains measured with the IDS-2, the findings suggest that autistic participants showed a lower performance in psychomotor skills (especially in the subtest Gross Motor Skills), language skills, and the evaluation of participation during testing of the developmental functions. This was in line with previous research (e.g., Coll et al., 2020; Kwok et al., 2015). On the other hand, the current results suggest similar performance for the domains of intelligence, executive functions, social-emotional skills, motivation and attitude, and the subtests Logical-Mathematical Reasoning, Reading, and Spelling. Hence, evidence for differential validity of the IDS-2 scores is provided for psychomotor skills, language skills, and achievement motivation evaluated by the test administrator (RQ 1a).

In particular, the lack of group differences in the social-emotional skills domain contrasts with previous studies (e.g., Cai et al., 2018), as one of the core symptoms of ASD is deficits in social communication and interaction (American Psychiatric Association, 2013; World Health Organization, 2018). A potential reason for this finding might be that in the IDS-2, social-emotional skills are measured through questions in which particularly explicit knowledge is requested and not through observing actual behavior in real-life situations. For these questions, autistic individuals might have been able to compensate for impairments in social-emotional skills with higher-order analytical strategies (Harms et al., 2010; Leung et al., 2022), as intelligence and social-emotional skills show particularly positive relations in individuals with ASD (Dyck et al., 2006; Salomone et al., 2019; Trevisan & Birmingham, 2016). This explanation is also supported by the current results of similar performance between autistic and non-autistic individuals in the intelligence and executive functions domains of the IDS-2, which implies that the autistic sample comprised participants of overall average cognitive abilities.

Another potential reason for the absent group differences in the social-emotional skills domain might be that autistic participants may have received interventions prior to the point of the assessment, especially for the deficits in the core characteristics of ASD. Support for this assumption comes from the age-separated analyses conducted in this work. For autistic children aged 5–10 years, evidence for differential validity of the IDS-2 scores is provided for cognitive abilities, psychomotor skills, basic skills (including language skills), achievement motivation evaluated by the test administrator, and, notably, for social-emotional skills. In contrast, the results do not suggest differential validity of IDS-2 test scores for autistic adolescents aged 11–20 years (RQ 1b). As ASD is often diagnosed in early childhood, allowing for timely interventions that are beneficial for the individual's development (Okoye et al., 2023), it might be that the autistic adolescents of this sample had already been provided with treatment measures. In contrast, the autistic children may have received their diagnosis only recently and hence had little or no previous support. It should be noted, however, that despite the larger sample size compared to previous studies, the power to detect small effects in the present analyses may still have been limited, especially given that the current effect sizes for social-emotional skills were moderate and comparable to previous meta-analytical findings (e.g., Yeung, 2022).

In summary, the present work suggests that differential validity of IDS-2 scores for a clinical subgroup of individuals with ASD is not necessarily to be found across all presumed developmental domains. Rather, it appears that it is especially the interpretation of test scores in terms of psychomotor skills (in particular Gross Motor Skills), language skills, and externally rated achievement motivation that provides information about the differentiation between autistic and non-autistic participants on the IDS-2. In addition, differential validity of test scores in social-emotional skills may depend on other factors, such as participants' intelligence, age, and, related to the latter, probably the therapeutic support they have received previously. Therefore, in practice, these are factors that should be considered in the clinical assessment of autistic individuals with the IDS-2. Moreover, the results suggest that clinicians should particularly pay attention to the domains of psychomotor and language skills of the IDS-2 as well as their evaluation of participants' achievement motivation during the diagnostic assessment of children and adolescents with ASD.

6.2 Criterion Validity

The second aim of this dissertation was to extend the evidence on criterion validity with its subcomponents predictive, concurrent, and incremental validity of IDS and IDS-2 intelligence and noncognitive personality characteristics for multi-informant academic achievement in typically developing adolescents. Specifically, the current work examined whether scores of intelligence, conscientiousness, and achievement striving motivation can predict and incrementally contribute to important perspectives available in the school context, namely, objective (grades) and subjective performance ratings from parents and students.

In terms of evidence on predictive validity, the current findings suggest that IDS intelligence was a longitudinal predictor of each of the three measures of academic achievement over 7 years (RQ 2a). Hence, intelligence, which mainly encompasses fluid aspects (Baltes, 1987, 1990), was able to predict school grades as well as parent-reported and self-reported academic performance over a long time span from childhood into adolescence. Looking at the evidence on concurrent validity, it appears that IDS-2 intelligence was a concurrent predictor of the three multi-informant measures of academic achievement (RQ 2b). Thus, intelligence, which comprises multiple abilities according to the CHC model (Schneider &

McGrew, 2018) and not exclusively fluid reasoning, was also cross-sectionally related to school grades and parent-reported and self-reported academic performance in adolescents.

These findings complement previous studies reporting associations of intelligence, measured with the IDS and the IDS-2, with grades (e.g., Grieder et al., 2022; Gygi et al., 2017) and parent-reported academic performance (Grob & Hagmann-von Arx, 2018b; Gut et al., 2013). Adding to this body of research, the present work provides evidence for criterion validity of intelligence for self-reported academic performance as a possible additional information source of academic achievement. This is important as it offers valuable insights into students' evaluation of their academic functioning in relation to their peers (Marsh & Martin, 2011) and therefore contributes to a more comprehensive understanding. It can thus be concluded that test scores, which assess the ability to solve problems, reason, and learn quickly (Gottfredson, 1997), play a crucial role in predicting current and future academic achievement as evaluated by *all* key players in the school context. As intelligence facilitates learning and comprehension (Di Fabio & Busoni, 2007), it stands to reason that students with higher intelligence are more likely to acquire more knowledge at school and to demonstrate more efficient problem-solving strategies at home, for example, when doing homework. This, in turn, might be reflected in higher grades and higher subjective performance ratings by parents and students themselves.

Following Roberts and Wood's (2006) neo-socioanalytic theory, the present work also examined concurrent and incremental validity for scores of noncognitive personality characteristics in predicting multi-informant academic achievement beyond intelligence. The current findings suggest that achievement striving motivation was a concurrent and incremental predictor of objective and subjective academic achievement measures, whereas conscientiousness was a concurrent and incremental predictor of grades and parent-reported—but largely not of self-reported—academic performance (RQ 2c and RQ 2d). Hence, beyond intelligence, also students' noncognitive personality characteristics predicted current academic achievement, which can be explained by the fact that they comprise the "will do" (Gottfredson, 2003, p. 369) for academic work. However, across all these examinations, only conscientiousness was mostly not able to predict and contribute incrementally to students' performance ratings. A possible explanation for this finding could be that individuals' conscientiousness increases over the life span (Roberts et al., 2006), so it may not yet be associated with students' self-reported academic performance at this age.

In summary, these results provide evidence on predictive (for the IDS) and concurrent (for the IDS-2) validity of intelligence for objective and subjective measures of academic achievement that encompass the major perspectives of players in the school environment. Hence, for the application in practice, these findings should give clinicians confidence in using the IDS and IDS-2 to generate predictions about individuals' academic achievement, even over the long term. Taking noncognitive personality characteristics into account, criterion validity may not be guaranteed for scores of conscientiousness predicting self-reported academic performance. Nonetheless, noncognitive personality characteristics should be considered in questions related to academic achievement as they appear to be less stable than intelligence (e.g., Roberts & Yoon, 2022) and thus could be targeted by interventions (Lazowski & Hulleman, 2016; Roberts et al., 2017; Stieger et al., 2021). Furthermore, in educational settings, subjective performance ratings from parents and students should be assessed by teachers alongside grades. Importantly, the current dissertation employed two independent studies to investigate criterion validity, enabling a cross-validation approach. The present findings can therefore be deemed robust and corroborated, as a largely similar pattern of results emerged in both studies.

6.3 Participant Characteristics

The last aim of the current dissertation was to expand previous knowledge about the contribution of participant characteristics to test performance and, therefore, to the valid interpretation of test results. Specifically, this work sought to disentangle the relative importance of aspects of proficiency in the test language, namely, objectively measured language abilities and multilingualism, for test scores in the cognitive *and* developmental functions of the IDS-2.

The current findings suggest that children's receptive language ability is an important contributor to test performance in almost all IDS-2 scores (RQ 3a). Thus, the ability to understand verbal instructions is required for performance on the cognitive and developmental functions of the IDS-2, as test directions are provided verbally. In addition, it became evident that children's expressive language ability—the ability to produce verbal information—contributes little to test performance on the IDS-2 beyond receptive language ability (RQ 3b). For both language abilities, the largest contributions appeared in the intelligence composites, Verbal Reasoning, including its subtests, and the basic skills subtests. Hence, according to the present findings, children's receptive and expressive language abilities are particularly crucial participant characteristics for test performance the higher the verbal demands of a task are in terms of complexity, length, and verbosity of instructions, type of response, and content. This was especially apparent for tasks assessing verbal and crystallized knowledge as elements of the construct to be measured, such as for Verbal Reasoning, which assesses the broad ability Gc of the CHC model (Schneider & McGrew, 2018). On the other hand, when the verbal component of a task is small and instructions are supported, for example, by gestures of the test administrator or pictorial illustrations, it is most likely that children's language abilities play a rather subordinate role in the performance on the IDS-2. Last, the present results suggest that multilingualism only contributes to test performance on intelligence subtests measuring Verbal Reasoning and verbal Long-Term Memory above children's receptive and expressive language abilities (RQ 3c). Thus, the results of the current work suggest that—when taking into account language abilities—other components of having a linguistically diverse background play a role solely in subtests that also capture "culturally-valued knowledge" (Schneider & McGrew, 2018, p. 114) and encompass content that may be partly socialization or culture specific.

These findings complement the conclusions of previous studies investigating group differences (Grob et al., 2013; Hagmann-von Arx et al., 2013; Schweizer et al., 2021). Yet, the present work goes beyond past work because it adopts a dimensional approach by considering the continuum of participants' language abilities. Furthermore, it disentangles the relative importance of different aspects of proficiency in the test language by also including objective measures of language abilities.

In summary, the findings of the present work suggest that aspects of children's proficiency in the test language contribute differently to test performance on cognitive and developmental functions of the IDS-2. In particular, their contribution depends largely on the verbal demands of the specific task and its assessment of verbal components, crystallized knowledge, and culture and socialization aspects. This finding has implications for practice. Clinicians should consider language proficiency particularly in verbal-dependent tasks in the assessment of cognitive and developmental functions with the IDS-2. Specifically, when testing individuals at risk for linguistic disadvantages, such as those with linguistically diverse backgrounds, their language abilities should be measured prior to the assessment of other domains. If sufficient proficiency in the test language cannot be guaranteed, a nonverbal test procedure should be used or, if not available, test results should be interpreted cautiously. The current results provide a novel perspective to our understanding of participant characteristics that may interfere with the validity of score interpretations.

6.4 Strengths, Limitations, and Call for Future Research

The present dissertation has both strengths and limitations. From a conceptual point of view, one of the main contributions of this work is the investigation of validity for the assessment of various developmental domains beyond cognitive abilities. By examining core cognitive and developmental functions, the present dissertation provides an integrative understanding concerning the validity of test score interpretations. It incorporated a holistic perspective on the assessment of children and adolescents and in doing so, revealed insights about various important domains simultaneously and went beyond previous research, which often emphasized single domains. On the other hand, a more fine-grained view was implemented to disentangle possible differential effects of age and level of analysis (i.e., composites, group factors, subtests; Study 1), distinct associations with multi-informant academic achievement (Study 2), and the contribution of specific aspects of proficiency in the test language (Study 3). Another strength of the present dissertation is its focus on the valid interpretation of test results for common participant groups and contexts in assessment settings. This is particularly important for autistic children and adolescents as well as for children and adolescents with linguistically diverse backgrounds, whose numbers have increased worldwide in recent years (Idring et al., 2015; United Nations, 2022). Moreover, the current work covered a wide age range across childhood and adolescence, allowing conclusions to be drawn about developmental domains in children and adolescents throughout the school years. Further, the comprehensive investigation of the various domains included in the IDS-2 using multiple samples promotes the validation of the IDS-2 and provides evidence on validity of test scores beyond previous studies that mostly examined the domain of intelligence.

From a methodological perspective, strengths of the present dissertation lie in the use of predominantly representative samples that are relatively large for the questions investigated compared to previous research and include not only subsamples from the standardization and validation studies of the IDS and IDS-2, but also further collected data. Additionally, both cross-sectional and longitudinal analyses were employed, and demographic variables (e.g., sex, SES) were accounted for as well as the alpha error inflation caused by multiple testing using Hommel's (1988) correction.

In addition to these strengths, the present dissertation encompasses three main limitations that await being addressed in future studies. First, the current work focused in particular on the investigation of one test battery, namely, the IDS-2, which is used for research and practice in German-speaking regions. However, according to guidelines on standards for psychological testing (AERA et al., 2014), it is necessary to provide evidence on validity specifically for each purpose for which a test is used. Therefore, investigation of the research questions explored in this dissertation for other tests and—particularly—for the additional language versions of the IDS-2 (Grob et al., 2018, 2019, 2021, 2022) is needed. Moreover, such analyses would also offer insights into possible similarities and differences concerning the validity of test score interpretations across different test procedures.

Second, although the present work has taken a broadly integrative approach, it has not been possible to examine the three aspects addressing validity consistently across all developmental domains. One reason for this is that some domains in the IDS-2 are not administered to all age groups. Related to this limitation, it was therefore also not feasible to include children and adolescents in all the conducted studies. Hence, future research is encouraged to consider these shortcomings and to continue following a holistic perspective on the validity evidence for developmental assessment in children and adolescents.

Last, the current dissertation exclusively focused on three aspects of validity evidence. Nevertheless, in addition to evidence based on *relations to other variables*, there are other sources, such as *test content*, *response processes*, and *internal structure* (AERA et al., 2014). Furthermore, besides participant characteristics, examiner characteristics (e.g., aspects of test administration and scoring) and test characteristics (e.g., aspects of test items) are listed as possible contributors that may impact test performance (AERA et al., 2014). Previous studies provided evidence for the aspect of *internal structure* and a subcomponent of *relations to other variables* (i.e., convergent validity evidence) for the IDS-2 (Bünger et al., 2021; Grieder et al., 2022, 2023; Grieder & Grob, 2020). Further information on validity evidence can be found in its technical manual (Grob & Hagmann-von Arx, 2018b), but there is a lack of research that specifically investigates these other sources and characteristics. In connection to this limitation, further research is required that continues the efforts of the current work by examining additional groups and criteria variables, such as other clinical subgroups (e.g., participants with speech disorders), longitudinal real-world outcomes (e.g., career success), and diverse participant characteristics (e.g., cultural background or degree of acculturation).

6.5 Conclusion

Three overarching conclusions can be drawn from the results of the present dissertation that are relevant to the assessment of developmental domains in children and adolescents. These also include further considerations that should be taken into account in research, practice, and politics.

First, as current findings suggest that validity of scores may differ for age groups, level of analysis, real-life criterion variables, or distinct aspects of participant characteristics, the present work highlights the importance of a fine-grained analysis when conducting validation studies. Only by considering the potential differential effects of specific aspects is it possible to disentangle their relevance to the validity of test scores and to derive concrete recommendations for practice. Second, and related to the previous point, practical implications can be drawn from the present findings that should be considered when assessing developmental domains in specific groups of children and adolescents and contexts. These can also be linked to each other: For example, if autistic children and adolescents exhibit limited language skills, their proficiency in the test language may also contribute to test performance. In such a case, clinicians should take into account the autistic individuals' language abilities particularly in tasks with high verbal demands, since autistic individuals may also be at risk for linguistic disadvantages. Finally, the current work highlights the importance of assessing children and adolescents holistically to gain a comprehensive understanding of the individuals' full potential. In doing so, specific strengths and difficulties in developmental domains, beyond cognitive abilities, can be identified and used for individualized support and treatment planning. It could also be a fruitful source for the personal development of children and adolescents and, moreover, for parents and educators, if the most comprehensive possible picture of the individual is drawn. Through a holistic assessment, it is possible to offer a fundamental and objective evaluation of the individual's developmental status, which can be linked to the ratings of parents, educators, and other caregivers. As time resources for psychological assessment become more and more limited due to increasing caseloads (Pastega & Riklin, 2023; Peter et al., 2023), this should be considered by policy and decision makers. Additional support should be provided to enable a holistic and valid assessment of children's and adolescents' developmental domains.

The present dissertation makes an important contribution by shedding light on three aspects of validity, extending previous validation findings. Nevertheless, continuing to build knowledge about the validity of test scores should remain an avenue for future research to ensure accurate and valid conclusions based on test results for high-stakes decisions, interventions, and educational placement across different groups of children and adolescents and contexts.

Table 1

Overview of the Cognitive and Developmental Functions of the IDS (Grob, Meyer, & Hagmann-von Arx, 2009)

Domain	Subtest	Description
Intelligence	Visual Perception[a]	Organize cards printed with lines of different lengths in consecutive order
	Selective Attention[a]	Cross out targets (ducks looking to the right with two orange characteristics) as quickly as possible
	Phonological Memory[a]	Repeat number and letter sequences
	Visuospatial Memory[a]	Remember geometric figures in a selection of figures
	Auditory Memory[a]	Listen to a story and recall it after a period of time
	Visual Reasoning[a]	Reproduce presented geometric figures with the help of rectangles and/or triangles
	Conceptual Reasoning[a]	Understand the concept of three pictures and choose two corresponding pictures from a selection of pictures
Psychomotor skills	Gross Motor Skills	Balance on a rope, catch and throw a ball, and jump sideways over a rope
	Fine Motor Skills	Quickly thread beads and cubes according to a picture
	Visuomotor Skills	Draw geometric figures
Social-emotional skills	Identifying Emotions	Recognize and name emotions of children in photos
	Regulating Emotions	Specify regulation strategies for the emotions of anger, fear, and grief
	Understanding Social Situations	Understand and explain social situations in two pictures
	Socially Competent Behavior	Name socially competent behavior according to a presented social situation
Basic skills	Mathematics	Solve logical-mathematical reasoning tasks
	Language Skills	
	Language Expressive	Form sentences from several words
	Language Receptive	Carry out instructions with wooden figures
Achievement motivation	Perseverance	The test administrator evaluates how persistently the participant works on the tasks
	Eagerness to Learn	The test administrator evaluates whether the participant takes pleasure in their performance

[a] Subtests included in the Full-Scale IQ.

Table 2

Overview of the Cognitive and Developmental Functions of the IDS-2 (Grob & Hagmann-von Arx, 2018a)

Domain	Group Factor	Subtest	Description
Intelligence	Visual Processing	Shape Design[a, b]	Reproduce geometric figures with the help of rectangles and triangles
		Washer Design[a]	Reproduce counter patterns according to a template
	Processing Speed	Parrots[a, b]	Cross out parrots with two orange features that look to the left from rows of different parrots
		Boxes[a]	Cross out groups of three or four boxes from rows of different groups of boxes
	Auditory Short-Term Memory	Digit and Letter Span[a, b]	Repeat number and letter sequences forward and backward
		Mixed Digit and Letter Span[a]	Repeat mixed number and letter sequences forward and backward
	Visuospatial Short-Term Memory	Shape Memory[a, b]	Remember figures and recognize them from a selection of figures and positions
		Rotated Shape Memory[a]	Remember figures and recognize them from a selection of rotated figures and positions
	Abstract Reasoning	Matrices: Completion[a, b, c]	Understand how a figure changes and transfer these changes to a continuing figure
		Matrices: Odd One Out[a]	Select from presented pictures the one that does not fit with the others
	Verbal Reasoning	Naming Categories[a, b, c]	Name categories for a group of pictures or words
		Naming Opposites[a]	Name opposites of presented words
	Long-Term Memory	Story Recall[a, b]	Listen to a semantically meaningful story and recall it after at least 20 min
		Picture Recall[a]	Look at a picture and recall key features and details after at least 20 min
Executive functions		Listing Words	List words based on categories or starting letters
		Divided Attention	Cross out parrots with two orange features that look to the left from different parrots and list animals at the same time
		Animal Colors	Say colors of animals as fast as possible
		Drawing Routes	Travel given routes as fast as possible once
Psychomotor skills		Gross Motor Skills	Balance on a rope, catch and throw a ball, and jump sideways over a rope
		Fine Motor Skills	Quickly screw nuts on and off bolts of different sizes and quickly thread beads of different sizes
		Visuomotor Skills	Move exactly between lines, draw figures, and reflect figures
Social-emotional skills		Identifying Emotions	Recognize and name emotions of children in photos
		Regulating Emotions	Specify regulation strategies for the emotions of anger, fear, and grief
		Socially Competent Behavior	Name socially competent behavior according to a presented social situation

Domain	Group Factor	Subtest	Description
Basic skills		Logical-Mathematical Reasoning	Solve logical-mathematical reasoning tasks
		Language Skills	
		Phoneme Analysis	Clap syllables, recognize rhymes, isolate on and off sounds, sound out words
		Phoneme–Grapheme Correspondence	Match phonemes and graphemes, recognize short and long vowels
		Language Expressive	Form sentences from several words
		Language Receptive	Carry out instructions
		Reading	
		Reading Words	Read real words
		Reading Pseudo Words	Read pseudo words
		Text Comprehension	Read and understand texts
		Spelling	Word dictation
Motivation and attitude		Conscientiousness	Rate statements regarding conscientiousness
		Achievement Motivation	Rate statements regarding achievement motivation
Participation during testing		Intelligence domain	The test administrator evaluates the participant's cooperation (achievement motivation) during testing of the intelligence domain
		Executive functions domain	The test administrator evaluates the participant's cooperation (achievement motivation) during testing of the executive functions domain
		Developmental functions domain	The test administrator evaluates the participant's cooperation (achievement motivation) during testing of the developmental functions domain

Note. Gross Motor Skills, Identifying Emotions, and Language Skills only for ages 5 to 10 years; Reading and Spelling only for ages 7 to 20 years; Conscientiousness and Achievement Motivation only for ages 11 to 20 years.

[a] Subtests included in the Profile IQ.

[b] Subtests included in the Full-Scale IQ.

[c] Subtests included in the Screening IQ.

7. References

Adolph, K. E., & Joh, A. S. (2007). Motor development: How infants get into the act. In A. Slater & M. Lewis (Eds.), *Introduction to infant development* (2nd ed., pp. 63–80). Oxford University Press.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.).

American Psychological Association. (2015). *APA dictionary of psychology* (2nd ed.).

Asendorpf, J. B. (1998). *Fünf-Faktoren-Fragebogen für Kinder [Five Factors Questionnaire for Children] (FFFK)*. Humboldt University, Institute for Psychology.

Asendorpf, J. B., & Van Aken, M. A. G. (1999). Resilient, overcontrolled, and undercontrolled personality prototypes in childhood: Replicability, predictive power, and the trait-type issue. *Journal of Personality and Social Psychology*, *77*(4), 815–832. https://doi.org/10.1037/0022-3514.77.4.815

Baddeley, A. (1998). Recent developments in working memory. *Current Opinion in Neurobiology*, *8*(2), 234–238. https://doi.org/10.1016/S0959-4388(98)80145-1

Bailey, C. E. (2007). Cognitive accuracy and intelligent executive function in the brain and in business. *Annals of the New York Academy of Sciences*, *1118*(1), 122–141. https://doi.org/10.1196/annals.1412.011

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology*, *23*(5), 611–626.

Baltes, P. B. (1990). Entwicklungspsychologie der Lebensspanne: Theoretische Leitsätze [Lifespan developmental psychology: Guiding theoretical principles]. *Psychologische Rundschau*, *41*, 1–24.

Benson, N. F., Floyd, R. G., Kranzler, J. H., Eckert, T. L., Fefer, S. A., & Morgan, G. B. (2019). Test use and assessment practices of school psychologists in the United States: Findings from the 2017 National Survey. *Journal of School Psychology*, *72*, 29–48. https://doi.org/10.1016/j.jsp.2018.12.004

Berk, L. E. (2018). *Development through the lifespan* (7th ed.). Pearson.

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *13*(4), 525–531. https://doi.org/10.1017/S1366728909990423

Binet, A. (1909). *Les idées modernes sur les enfants*. Flammarion.

Blair, K. A., Denham, S. A., Kochanoff, A., & Whipple, B. (2004). Playing it cool: Temperament, emotion regulation, and social behavior in preschoolers. *Journal of School Psychology*, *42*(6), 419–443. https://doi.org/10.1016/j.jsp.2004.10.002

Bos, W., Klieme, E., & Köller, O. (2010). *Schulische Lerngelegenheiten und Kompetenzentwicklung [School learning opportunities and competence development]*. Waxmann.

Bünger, A., Grieder, S., Schweizer, F., & Grob, A. (2021). The comparability of intelligence test results: Group- and individual-level comparisons of seven intelligence tests. *Journal of School Psychology*, *88*, 101–117. https://doi.org/10.1016/j.jsp.2021.09.002

Burton, A. W., & Rodgerson, R. W. (2001). New perspectives on the assessment of movement skills and motor abilities. *Adapted Physical Activity Quarterly*, *18*(4), 347–365. https://doi.org/10.1123/apaq.18.4.347

Cai, R. Y., Richdale, A. L., Uljarević, M., Dissanayake, C., & Samson, A. C. (2018). Emotion regulation in autism spectrum disorder: Where we are and where we need to go. *Autism Research*, *11*(7), 962–978. https://doi.org/10.1002/aur.1968

Calero, M. D., Fernández-Parra, A., López-Rubio, S., Carles, R., Mata, S., Vives, M. del C., Navarro, E., & Márquez, J. (2013). Variables involved in personal, social and school adjustment in a sample of preschool-aged children from different cultural backgrounds. *European Journal of Psychology of Education*, *28*(1), 133–155. https://doi.org/10.1007/s10212-012-0107-8

Cameron, C. E., Cottone, E. A., Murrah, W. M., & Grissmer, D. W. (2016). How are motor skills linked to children's school performance and academic achievement? *Child Development Perspectives*, *10*(2), 93–98. https://doi.org/10.1111/cdep.12168

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

Cattell, R. B. (1941). Some theoretical issues in adult intelligence testing. *Psychological Bulletin*, *38*, 592.

Chiarotti, F., & Venerosi, A. (2020). Epidemiology of autism spectrum disorders: A review of worldwide prevalence estimates since 2014. *Brain Sciences*, *10*(5), Article 274. https://doi.org/10.3390/brainsci10050274

Clark, J. E., & Whitall, J. (1989). What is motor development? The lessons of history. *Quest*, *41*(3), 183–202. https://doi.org/10.1080/00336297.1989.10483969

Clearfield, M. W. (2011). Learning to walk changes infants' social interactions. *Infant Behavior and Development*, *34*(1), 15–25. https://doi.org/10.1016/j.infbeh.2010.04.008

Coll, S.-M., Foster, N. E. V., Meilleur, A., Brambati, S. M., & Hyde, K. L. (2020). Sensorimotor skills in autism spectrum disorder: A meta-analysis. *Research in Autism Spectrum Disorders*, *76*, Article 101570. https://doi.org/10.1016/j.rasd.2020.101570

Collins, A., & Koechlin, E. (2012). Reasoning, learning, and creativity: Frontal lobe function and human decision-making. *PLoS Biology*, *10*(3), Article e1001293. https://doi.org/10.1371/journal.pbio.1001293

Cormier, D. C., Bulut, O., McGrew, K. S., & Kennedy, K. (2022). Linguistic influences on cognitive test performance: Examinee characteristics are more important than test characteristics. *Journal of Intelligence*, *10*(1), Article 1. https://doi.org/10.3390/jintelligence10010008

Cormier, D. C., McGrew, K. S., & Evans, J. J. (2011). Quantifying the "degree of linguistic demand" in spoken intelligence test directions. *Journal of Psychoeducational Assessment*, *29*(6), 515–533. https://doi.org/10.1177/0734282911405962

Cortés Pascual, A., Moyano Muñoz, N., & Quílez Robres, A. (2019). The relationship between executive functions and academic performance in primary education: Review and meta-analysis. *Frontiers in Psychology*, *10*. https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01582

Costa, P. T., & McCrae, R. R. (1992). *Neo Personality Inventory-Revised (NEO PI-R)*. Psychological Assessment Resources.

Demetriou, E. A., Lampit, A., Quintana, D. S., Naismith, S. L., Song, Y. J. C., Pye, J. E., Hickie, I., & Guastella, A. J. (2018). Autism spectrum disorders: A meta-analysis of executive function. *Molecular Psychiatry*, *23*(5), Article 5. https://doi.org/10.1038/mp.2017.75

Denham, S. A. (2006). Social-emotional competence as support for school readiness: What is it and how do we assess it? *Early Education and Development*, *17*(1), 57–89. https://doi.org/10.1207/s15566935eed1701_4

Di Domenico, S. I., & Fournier, M. A. (2015). Able, ready, and willing: Examining the additive and interactive effects of intelligence, conscientiousness, and autonomous motivation on undergraduate academic performance. *Learning and Individual Differences*, *40*, 156–162.

https://doi.org/10.1016/j.lindif.2015.03.016

Di Fabio, A., & Busoni, L. (2007). Fluid intelligence, personality traits and scholastic success: Empirical evidence in a sample of Italian high school students. *Personality and Individual Differences*, *43*(8), 2095–2104. https://doi.org/10.1016/j.paid.2007.06.025

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Diamond, A. (2020). Executive functions. In *Handbook of Clinical Neurology* (Vol. 173, pp. 225–240). Elsevier. https://doi.org/10.1016/B978-0-444-64150-2.00020-4

Dowker, A. (2019). *Individual differences in arithmetic: Implications for psychology, neuroscience, and education*. Routledge.

Drechsler, R. (2007). Exekutive Funktionen [Executive functions]. *Zeitschrift für Neuropsychologie*, *18*(3), 233–248. https://doi.org/10.1024/1016-264X.18.3.233

Dyck, M. J., Piek, J. P., Hay, D., Smith, L., & Hallmayer, J. (2006). Are abilities abnormally interdependent in children with autism? *Journal of Clinical Child & Adolescent Psychology*, *35*(1), 20–33. https://doi.org/10.1207/s15374424jccp3501_3

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*(1), 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Eisenberg, N., Sadovsky, A., & Spinrad, T. L. (2005). Associations of emotion-related regulation with language skills, emotion knowledge, and academic outcomes. *New Directions for Child and Adolescent Development*, *2005*(109), 109–118. https://doi.org/10.1002/cd.143

Fabes, R. A., Gaertner, B. M., & Popp, T. K. (2006). Getting along with others: Social competence in early childhood. In K. McCartney & D. Phillips (Eds.), *Blackwell handbook of early childhood development* (pp. 297–316). Blackwell. https://doi.org/10.1002/9780470757703.ch15

Feder, K. P., & Majnemer, A. (2007). Handwriting development, competency, and intervention. *Developmental Medicine & Child Neurology*, *49*(4), 312–317. https://doi.org/10.1111/j.1469-8749.2007.00312.x

Flanagan, D. P., & Harrison, P. L. (2012). *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed.). Guilford Press.

Flanagan, D. P., & Ortiz, S. O. (2001). *Essentials of cross-battery assessment*. Wiley.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*(2), 172–179. https://doi.org/10.1111/j.1467-9280.2006.01681.x

Fujiki, M., Brinton, B., & Clarke, D. (2002). Emotion regulation in children with specific language impairment. *Language, Speech, and Hearing Services in Schools*, *33*(2), 102–111. https://doi.org/10.1044/0161-1461(2002/008)

Gallahue, D. L., Ozmun, J. C., & Goodway, J. D. (2012). *Understanding motor development: Infants, children, adolescents, adults* (7th ed.). McGraw-Hill.

Goldstein, S., Princiotta, D., & Naglieri, J. A. (2015). *Handbook of intelligence. Evolutionary theory, historical perspective, and current concepts*. Springer.

Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, *24*(1), 13–23. https://doi.org/10.1016/S0160-2896(97)90011-8

Gottfredson, L. S. (2003). Dissecting practical intelligence theory: Its claims and evidence. *Intelligence*, *31*(4), 343–397. https://doi.org/10.1016/S0160-2896(02)00085-5

Gottfredson, L. S., & Deary, I. J. (2004). Intelligence predicts health and longevity, but why? *Current*

Directions in Psychological Science, *13*(1), 1–4. https://doi.org/10.1111/j.0963-7214.2004.01301001.x

Grieder, S., Bünger, A., Odermatt, S. D., Schweizer, F., & Grob, A. (2022). Limited internal comparability of general intelligence composites: Impact on external validity, possible predictors, and practical remedies. *Assessment*, *29*(6), 1172–1189. https://doi.org/10.1177/10731911211005171

Grieder, S., & Grob, A. (2020). Exploratory factor analyses of the Intelligence and Development Scales–2: Implications for theory and practice. *Assessment*, *27*(8), 1853–1869. https://doi.org/10.1177/1073191119845051

Grieder, S., Visser, L., Timmerman, M. E., & Grob, A. (2023). Measurement invariance of the Intelligence and Development Scales–2 across language versions, gender, and age. *European Journal of Psychological Assessment*. Advance online publication. https://doi.org/10.1027/1015-5759/a000771

Grob, A., & Hagmann-von Arx, P. (2018a). *Intelligence and Development Scales–2 (IDS-2). Intelligenz- und Entwicklungsskalen für Kinder und Jugendliche [Intelligence and Development Scales for Children and Adolescents]*. Hogrefe.

Grob, A., & Hagmann-von Arx, P. (2018b). *Intelligence and Development Scales–2 (IDS-2). Intelligenz- und Entwicklungsskalen für Kinder und Jugendliche. Manual zu Theorie, Interpretation und Gütekriterien [Intelligence and Development Scales for Children and Adolescents. Manual on theory, interpretation and psychometric criteria]*. Hogrefe.

Grob, A., Hagmann-von Arx, P., Barnett, A., Stuart, N., & Vanzan, S. (2021). *Intelligence and Development Scales–2 (IDS-2). Intelligence and Development Scales for Children and Adolescents*. Hogrefe.

Grob, A., Hagmann-von Arx, P., Ferri, R., Rea, M., & Casagrande, M. (2022). *Intelligence and Development Scales–2. Scale di intelligenza e sviluppo per bambini e adolescenti [Intelligence and Development Scales for Children and Adolescents]*. Hogrefe.

Grob, A., Hagmann-von Arx, P. H., & Bodmer, N. M. (2009). Entwicklungsdiagnostik [Developmental diagnostics]. In S. Schneider & J. Margraf (Eds.), *Lehrbuch der Verhaltenstherapie: Störungen im Kindes- und Jugendalter [Textbook of behavior therapy: Disorders in childhood and adolescence]* (Vol. 3, pp. 145–158). Springer. https://doi.org/10.1007/978-3-540-79545-2_9

Grob, A., Hagmann-von Arx, P., Jaworowska, A., Matczak, A., & Fecenec, D. (2019). *Intelligence and Development Scales–2. Inteligencji i Rozwoju dla Dzieci i Młodzieży [Intelligence and Development Scales for Children and Adolescents]*. Hogrefe.

Grob, A., Hagmann-von Arx, P., Ruiter, S. A. J., Timmerman, M. E., & Visser, L. (2018). *Intelligence and Development Scales–2 (IDS-2). Intelligentie- en ontwikkelingsschalen voor kinderen en jongeren. [Intelligence and Development Scales for Children and Adolescents]*. Hogrefe.

Grob, A., Meyer, C. S., & Hagmann-von Arx, P. (2009). *Intelligence and Development Scales (IDS)*. Huber.

Grob, A., Meyer, C. S., & Hagmann-von Arx, P. (2013). *Intelligence and Development Scales (IDS): Intelligenz- und Entwicklungsskalen für Kinder von 5–10 Jahren [Intelligence and Development Scales for Children 5–10 years old]*. Hans Huber.

Gross, J., Hudson, C., & Price, D. (2009). *The long term costs of numeracy difficulties*. Every Child a Chance Trust/ KPMG Foundation.

Gut, J., Reimann, G., & Grob, A. (2012). Kognitive, sprachliche, mathematische und sozial-emotionale Kompetenzen als Prädiktoren späterer schulischer Leistungen: Können die Leistungen eines

Kindes in den IDS dessen Schulleistungen drei Jahre später vorhersagen? [Cognitive, language, mathematical, and social-emotional skills as predictors of later academic achievement: Can a child's performance in the IDS predict their school performance three years later?] *Zeitschrift für Pädagogische Psychologie*, *26*(3), 213–220. https://doi.org/10.1024/1010-0652/a000070

Gut, J., Reimann, G., & Grob, A. (2013). A contextualized view on long-term predictors of academic performance. *Journal of Educational Psychology*, *105*(2), 436–443. https://doi.org/10.1037/a0031503

Gygi, J. T., Hagmann-von Arx, P., Schweizer, F., & Grob, A. (2017). The predictive validity of four intelligence tests for school grades: A small sample longitudinal study. *Frontiers in Psychology*, *8*, 1–9. https://doi.org/10.3389/fpsyg.2017.00375

Habermas, T., & Bluck, S. (2000). Getting a life: The emergence of the life story in adolescence. *Psychological Bulletin*, *126*(5), 748–769. https://doi.org/10.1037/0033-2909.126.5.748

Hagmann-von Arx, P., Petermann, F., & Grob, A. (2013). Konvergente und diskriminante Validität der WISC-IV und der Intelligence and Development Scales (IDS) bei Kindern mit Migrationshintergrund [Convergent and discriminant validity of the WISC-IV and the Intelligence and Development Scales (IDS) in children with a migration background]. *Diagnostica*, *59*(4), 170–182. https://doi.org/10.1026/0012-1924/a000091

Halle, T. G., & Darling-Churchill, K. E. (2016). Review of measures of social and emotional development. *Journal of Applied Developmental Psychology*, *45*, 8–18. https://doi.org/10.1016/j.appdev.2016.02.003

Han, H. S., & Kemple, K. M. (2006). Components of social competence and strategies of support: Considering what to teach and how. *Early Childhood Education Journal*, *34*(3), 241–246. https://doi.org/10.1007/s10643-006-0139-2

Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in autism spectrum disorders: A review of behavioral and neuroimaging studies. *Neuropsychology Review*, *20*(3), 290–322. https://doi.org/10.1007/s11065-010-9138-6

Henderson, S. E., Sugden, D. A., & Barnett, A. L. (2007). *Movement Assessment Battery for Children-2: MABC-2*. Pearson Assessment.

Heubrock, D., Eberl, I., & Petermann, F. (2004). *Abzeichentest für Kinder (ATK) [Drawing test for children]*. Hogrefe.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, *42*(8), 1–62. https://doi.org/10.18637/jss.v042.i08

Hommel, G. (1988). A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika*, *75*, 383–386.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, *2*(2), 127–160. https://doi.org/10.1007/BF00401799

Horn, J. L. (1991). Measurement of intellectual capabilities: A review of theory. In K. S. McGrew, J. Werder, & R. W. Woodcock (Eds.), *WJ-R technical manual* (pp. 197–232). Riverside.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, *57*(5), 253–270. https://doi.org/10.1037/h0023816

Idring, S., Lundberg, M., Sturm, H., Dalman, C., Gumpert, C., Rai, D., Lee, B. K., & Magnusson, C. (2015). Changes in prevalence of autism spectrum disorders in 2001–2011: Findings from the

Stockholm Youth Cohort. *Journal of Autism and Developmental Disorders*, *45*(6), 1766–1773. https://doi.org/10.1007/s10803-014-2336-y

International Test Commission. (2001). International guidelines for test use. *International Journal of Testing*, *1*(2), 93–114. https://doi.org/10.1207/S15327574IJT0102_1

Janke, B. (2008). Emotionswissen und Sozialkompetenz von Kindern im Alter von drei bis zehn Jahren [Emotion knowledge and social competence of children aged three to ten years]. *Zeitschrift zu Theorie und Praxis erziehungswissenschaftlicher Forschung*, *22*, 127–144.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy. In O. P. John, R. W. Robins, & A. P. Lawrence (Eds.), *Handbook of personality: Theory and research* (pp. 114–158). Guilford Press.

Jones, D. E., Greenberg, M., & Crowley, M. (2015). Early social-emotional functioning and public health: The relationship between kindergarten social competence and future wellness. *American Journal of Public Health*, *105*(11), 2283–2290. https://doi.org/10.2105/AJPH.2015.302630

Kahl, T., Grob, A., & Möhring, W. (2021). Does emotion regulation compensate deficits in various executive functions in children's and adolescents' mathematical achievement? *Learning and Individual Differences*, *89*, Article 102034. https://doi.org/10.1016/j.lindif.2021.102034

Kauschke, C. (2012). *Kindlicher Spracherwerb im Deutschen: Verläufe, Forschungsmethoden, Erklärungsansätze [Children's language acquisition in German: Trajectories, research methods, explanatory approaches]*. Walter de Gruyter.

Klingner, J. K., Blanchett, W. J., & Harry, B. (2007). Race, culture, and developmental disabilities. In S. L. Odom, R. H. Horner, M. E. Snell, & J. Blacher (Eds.), *Handbook of developmental disabilities* (pp. 55–75). Guilford Press.

Klupp, S., Grob, A., & Möhring, W. (2023). Aerobic fitness and fine motor skills are related to switching and updating in typically developing children. *Psychological Research*, *87*(5), 1401–1416. https://doi.org/10.1007/s00426-022-01749-w

Köber, C., Schmiedek, F., & Habermas, T. (2015). Characterizing lifespan development of three aspects of coherence in life narratives: A cohort-sequential study. *Developmental Psychology*, *51*(2), 260–275. https://doi.org/10.1037/a0038668

Köller, O., & Baumert, J. (2008). Entwicklung schulischer Leistungen [Development of scholastic performance]. In R. Oerter & L. Montada (Eds.), *Entwicklungspsychologie [Developmental psychology]* (Vol. 6, pp. 735–768). Beltz. https://pure.mpg.de/pubman/faces/ViewItemOverviewPage.jsp?itemId=item_2102046

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology*, 148–161.

Kwok, E. Y. L., Brown, H. M., Smyth, R. E., & Oram Cardy, J. (2015). Meta-analysis of receptive and expressive language skills in autism spectrum disorder. *Research in Autism Spectrum Disorders*, *9*, 202–222. https://doi.org/10.1016/j.rasd.2014.10.008

Lavrijsen, J., Vansteenkiste, M., Boncquet, M., & Verschueren, K. (2022). Does motivation predict changes in academic achievement beyond intelligence and personality? A multitheoretical perspective. *Journal of Educational Psychology*, *114*(4), 772–790. https://doi.org/10.1037/edu0000666

Lazowski, R. A., & Hulleman, C. S. (2016). Motivation interventions in education: A meta-analytic review. *Review of Educational Research*, *86*(2), 602–640. https://doi.org/10.3102/0034654315617832

Leonard, H. C., & Hill, E. L. (2014). Review: The impact of motor development on typical and atypical social cognition and language: A systematic review. *Child and Adolescent Mental Health*, *19*(3), 163–170. https://doi.org/10.1111/camh.12055

Leung, F. Y. N., Sin, J., Dawson, C., Ong, J. H., Zhao, C., Veić, A., & Liu, F. (2022). Emotion recognition across visual and auditory modalities in autism spectrum disorder: A systematic review and meta-analysis. *Developmental Review*, *63*, Article 101000. https://doi.org/10.1016/j.dr.2021.101000

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 151–173. https://doi.org/10.1207/S15328007SEM0902_1

Logan, S. W., Ross, S. M., Chee, K., Stodden, D. F., & Robinson, L. E. (2018). Fundamental motor skills: A systematic review of terminology. *Journal of Sports Sciences*, *36*(7), 781–796. https://doi.org/10.1080/02640414.2017.1340660

Lunt, L., Bramham, J., Morris, R. G., Bullock, P. R., Selway, R. P., Xenitidis, K., & David, A. S. (2012). Prefrontal cortex dysfunction and 'jumping to conclusions': Bias or deficit? *Journal of Neuropsychology*, *6*(1), 65–78. https://doi.org/10.1111/j.1748-6653.2011.02005.x

Maenner, M. J., Shaw, K. A., Baio, J., Washington, A., Patrick, M., DiRienzo, M., Christensen, D. L., Wiggins, L. D., Pettygrove, S., Andrews, J. G., Lopez, M., Hudson, A., Baroud, T., Schwenk, Y., White, T., Rosenberg, C. R., Lee, L.-C., Harrington, R. A., Huston, M., … Dietz, P. M. (2020). Prevalence of autism spectrum disorder among children aged 8 years—Autism and developmental disabilities monitoring network, 11 sites, United States, 2016. *MMWR Surveillance Summaries*, *69*(4), 1–12. https://doi.org/10.15585/mmwr.ss6904a1

Mammadov, S. (2022). Big Five personality traits and academic performance: A meta-analysis. *Journal of Personality*, *90*(2), 222–255. https://doi.org/10.1111/jopy.12663

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, *81*(1), 59–77. https://doi.org/10.1348/000709910X503501

McGrew, K. S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive Gf-Gc framework. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 151–119). Guilford Press.

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), 1–10. https://doi.org/10.1016/j.intell.2008.08.004

Melby-Lervåg, M., Lyster, S.-A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, *138*(2), 322–352. https://doi.org/10.1037/a0026744

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741–749. https://doi.org/10.1037/0003-066X.50.9.741

Meyer, C. S., Hagmann-von Arx, P., & Grob, A. (2009). Die Intelligence and Development Scale Sozial-Emotionale Kompetenz (IDS-SEK): Psychometrische Eigenschaften eines Tests zur Erfassung sozial-emotionaler Fähigkeiten [The Intelligence and Development Scale Social-Emotional Competence (IDS-SEK): Psychometric properties of a test to assess social-emotional skills].

Diagnostica, *55*(4), 234–244. https://doi.org/10.1026/0012-1924.55.4.234

Meyer, J., Fleckenstein, J., Retelsdorf, J., & Köller, O. (2019). The relationship of personality traits and different measures of domain-specific achievement in upper secondary education. *Learning and Individual Differences*, *69*, 45–59. https://doi.org/10.1016/j.lindif.2018.11.005

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Moosbrugger, H., & Kelava, A. (2012). Qualitätsanforderungen an einen psychologischen Test [Quality requirements for a psychological test]. In H. Moosbrugger & A. Kelava (Eds.), *Testtheorie und Fragebogenkonstruktion [Test theory and questionnaire construction]* (2nd ed., pp. 7–26). Springer.

Murray, H. (1938). *Explorations in personality*. Oxford University Press.

Niss, M., & Højgaard, T. (2019). Mathematical competencies revisited. *Educational Studies in Mathematics*, *102*(1), 9–28. https://doi.org/10.1007/s10649-019-09903-9

Okoye, C., Obialo-Ibeawuchi, C. M., Obajeun, O. A., Sarwar, S., Tawfik, C., Waleed, M. S., Wasim, A. U., Mohamoud, I., Afolayan, A. Y., & Mbaezue, R. N. (2023). Early diagnosis of autism spectrum disorder: A review and analysis of the risks and benefits. *Cureus*, *15*(8), Article e43226. https://doi.org/10.7759/cureus.43226

Ortiz, S. O. (2019). On the measurement of cognitive abilities in English learners. *Contemporary School Psychology*, *23*(1), 68–86. https://doi.org/10.1007/s40688-018-0208-8

Ozer, D. J., & Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, *57*(1), 401–421. https://doi.org/10.1146/annurev.psych.57.102904.190127

Pastega, N., & Riklin, F. (2023, September 30). Immer mehr Kinder “abgeklärt” – “Nicht alle sind gleich ein Therapiefall” [More and more children “tested”—"Not all are a therapy case"]. *Tages-Anzeiger*. https://www.tagesanzeiger.ch/immer-mehr-kinder-abgeklaert-nicht-alle-sind-gleich-ein-therapiefall-410906981524

Peter, C., Tuch, A., & Schuler, D. (2023). *Psychische Gesundheit—Erhebung Herbst 2022. Wie geht es der Bevölkerung in der Schweiz? Sucht sie sich bei psychischen Problemen Hilfe? [Mental health—Fall 2022 survey. How is the population in Switzerland doing? Do they seek help for mental health problems?]* (Obsan report 03/2023). Schweizerisches Gesundheitsobservatorium.

Petermann, F., & Wiedebusch, S. (2016). *Emotionale Kompetenz bei Kindern [Emotional competence in children]*. Hogrefe.

Petermann, F., & Winkel, S. (2007). *Fragebogen zur Leistungsmotivation für Schüler der 7. bis 13. Klasse [Achievement Motivation Questionnaire for 7th- to 13th-Grade Students] (FLM 7-13)*. Harcourt Test Services.

Poropat, A. E. (2014). A meta-analysis of adult-rated child personality and academic performance in primary education. *British Journal of Educational Psychology*, *84*(2), 239–252. https://doi.org/10.1111/bjep.12019

R Core Team. (2022). *R: A language and environment for statistical computing* (Version 2022.12.0+353)

[Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Reimann, Z., Miller, J. R., Dahle, K. M., Hooper, A. P., Young, A. M., Goates, M. C., Magnusson, B. M., & Crandall, A. (2020). Executive functions and health behaviors associated with the leading causes of death in the United States: A systematic review. *Journal of Health Psychology*, *25*(2), 186–196. https://doi.org/10.1177/1359105318800829

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, *138*(2), 353–387. https://doi.org/10.1037/a0026838

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, *130*(2), 261–288. https://doi.org/10.1037/0033-2909.130.2.261

Roberts, B. W., Hill, P. L., & Davis, J. P. (2017). How to change conscientiousness: The sociogenomic trait intervention model. *Personality Disorders: Theory, Research, and Treatment*, *8*, 199–205. https://doi.org/10.1037/per0000242

Roberts, B. W., Jackson, J. J., Fayard, J. V., Edmonds, G., & Meints, J. (2009). Conscientiousness. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 369–381). Guilford Press.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *132*(1), 1–25. https://doi.org/10.1037/0033-2909.132.1.1

Roberts, B. W., & Wood, D. (2006). Personality development in the context of the neo-socioanalytic model of personality. In D. Mroczek & T. Little (Eds.), *Handbook of personality development* (pp. 11–39). Erlbaum.

Roberts, B. W., & Yoon, H. J. (2022). Personality psychology. *Annual Review of Psychology*, *73*(1), 489–516. https://doi.org/10.1146/annurev-psych-020821-114927

Robinson, L. E., Stodden, D. F., Barnett, L. M., Lopes, V. P., Logan, S. W., Rodrigues, L. P., & D'Hondt, E. (2015). Motor competence and its effect on positive developmental trajectories of health. *Sports Medicine*, *45*(9), 1273–1284. https://doi.org/10.1007/s40279-015-0351-6

Romano, E., Babchishin, L., Pagani, L. S., & Kohen, D. (2010). School readiness and later achievement: Replication and extension using a nationwide Canadian survey. *Developmental Psychology*, *46*(5), 995–1007. https://doi.org/10.1037/a0018880

Rose-Krasnor, L. (1997). The nature of social competence: A theoretical review. *Social Development*, *6*(1), 111–135. https://doi.org/10.1111/j.1467-9507.1997.tb00097.x

Rost, D. H. (2001). Leseverständnis [Reading comprehension]. In D. H. Rost (Ed.), *Handwörterbuch Pädagogische Psychologie [Handbook of educational psychology]* (pp. 449–456). Beltz.

Rost, D. H. (2009). *Intelligenz: Fakten und Mythen. [Intelligence: Facts and myths]*. 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. https://doi.org/10.1016/j.intell.2015.09.002

Salomone, E., Bulgarelli, D., Thommen, E., Rossini, E., & Molina, P. (2019). Role of age and IQ in emotion understanding in autism spectrum disorder: Implications for educational interventions. *European Journal of Special Needs Education*, *34*(3), 383–392. https://doi.org/10.1080/08856257.2018.1451292

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, *86*(1), 162–173. https://doi.org/10.1037/0022-3514.86.1.162

Schmidt-Atzert, L., & Amelang, M. (2012). *Psychologische Diagnostik [Psychological diagnostics]* (5th ed.). Springer.

Schneider, W. J., & Flanagan, D. P. (2015). The relationship between theories of intelligence and intelligence tests. In S. Goldstein, D. Princiotta, & J. A. Naglieri (Eds.), *Handbook of intelligence: Evolutionary theory, historical perspective, and current concepts* (pp. 317–340). Springer. https://doi.org/10.1007/978-1-4939-1562-0_21

Schneider, W. J., & McGrew, K. S. (2018). The Cattell–Horn–Carroll theory of cognitive abilities. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 73–163). Guilford Press.

Schrank, F. A., McGrew, K. S., & Mather, N. (2014). *Woodcock—Johnson IV Tests of Cognitive Abilities*. Riverside Publishing.

Schweizer, F., Grieder, S., Bünger, A., & Grob, A. (2021). Vergleich von Intelligenztestleistungen bei monolingualen und bilingualen Kindern und Jugendlichen in den Intelligence and Development Scales–2 (IDS-2) [Comparison of intelligence test performance of monolingual and bilingual children and adolescents on the Intelligence and Development Scales–2 (IDS-2)]. *Diagnostica*, *67*(1), 36–46. https://doi.org/10.1026/0012-1924/a000260

Siegler, R., Saffran, J. R., Gershoff, E. T., & Eisenberg, N. (2021). *Entwicklungspsychologie im Kindes- und Jugendalter: Deutsche Auflage unter Mitarbeit von Sabina Pauen [Developmental psychology in childhood and adolescence: German edition with the collaboration of Sabina Pauen]* (5th ed.). Springer. https://doi.org/10.1007/978-3-662-62772-3

Singer, R., & Bös, K. (1994). Motorische Entwicklung: Gegenstandbereich und Entwicklungseinflüsse [Motor development: Domain and developmental influences]. In J. Baur, K. Bös, & R. Singer (Eds.), *Motorische Entwicklung. Ein Handbuch [Motor development. A manual]* (pp. 15–26). Hofmann.

Smyth, M. M., & Anderson, H. I. (2000). Coping with clumsiness in the school playground: Social and physical play in children with coordination impairments. *British Journal of Developmental Psychology*, *18*(3), 389–413. https://doi.org/10.1348/026151000165760

Spearman, C. (1904). “General Intelligence,” objectively determined and measured. *The American Journal of Psychology*, *15*(2), 201–292. https://doi.org/10.2307/1412107

Steinmayr, R., & Spinath, B. (2007). Predicting school achievement from motivation and personality. *Zeitschrift für Pädagogische Psychologie*, *21*(3/4), 207–216. https://doi.org/10.1024/1010-0652.21.3.207

Steinmayr, R., & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences*, *19*(1), 80–90. https://doi.org/10.1016/j.lindif.2008.05.004

Stern, E., & Neubauer, A. (2016). Intelligenz: Kein Mythos, sondern Realität [Intelligence: Not a myth but reality]. *Psychologische Rundschau*, *67*(1), 15–27. https://doi.org/10.1026/0033-3042/a000290

Stieger, M., Flückiger, C., Rüegger, D., Kowatsch, T., Roberts, B. W., & Allemand, M. (2021). Changing personality traits with the help of a digital personality change intervention. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(8), Article e2017548118. https://doi.org/10.1073/pnas.2017548118

Strooband, K. F. B., Rosnay, M. de, Okely, A. D., & Veldman, S. L. C. (2020). Systematic review and meta-analyses: Motor skill interventions to improve fine motor development in children aged birth to 6 years. *Journal of Developmental & Behavioral Pediatrics*, *41*(4), 319–331. https://doi.org/10.1097/DBP.0000000000000779

Sullivan, A. L. (2011). Disproportionality in special education identification and placement of English language learners. *Exceptional Children*, *77*(3), 317–334. https://doi.org/10.1177/001440291107700304

Swiss-German Conference of Directors of Education. (2016). *Lehrplan 21 [Curriculum 21]*. https://v-fe.lehrplan.ch/container/V_FE_DE_Gesamtausgabe.pdf

Trevisan, D. A., & Birmingham, E. (2016). Are emotion recognition abilities related to everyday social functioning in ASD? A meta-analysis. *Research in Autism Spectrum Disorders*, *32*, 24–42. https://doi.org/10.1016/j.rasd.2016.08.004

United Nations. (2022). *World migration report 2022*. https://worldmigrationreport.iom.int/wmr-2022-interactive/

Weinert, S., & Grimm, H. (2018). Sprachentwicklung [Language development]. In W. Schneider & U. Lindenberger (Eds.), *Entwicklungspsychologie [Developmental psychology]* (8th ed., pp. 517–550). Beltz.

Weiss, L. G., Harris, J. G., Prifitera, A., Courville, T., Rolfhus, E., Saklofske, D. H., & Holdnack, J. A. (2006). WISC-IV interpretation in societal context. The essentials and beyond. In L. G. Weiss, D. H. Saklofske, A. Prifitera, & J. A. Holdnack (Eds.), *WISC-IV: Advanced clinical interpretation* (pp. 1–49). Elsevier.

White, S. W., Scahill, L., Klin, A., Koenig, K., & Volkmar, F. R. (2007). Educational placements and service use patterns of individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *37*(8), 1403–1412. https://doi.org/10.1007/s10803-006-0281-0

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

World Health Organization. (1990). *The international statistical classification of diseases and related health problems* (10th ed.).

World Health Organization. (2018). *International classification of diseases for mortality and morbidity statistics* (11th rev.). https://icd.who.int/browse11/l-m/en

Yates, T., Ostrosky, M. M., Cheatham, G. A., Fettig, A., Shaffer, L., & Santos, R. M. (2008). *Research synthesis on screening and assessing social-emotional competence*. The Center on the Social Emotional Foundations for Early Learning. http://csefel.vanderbilt.edu/documents/rs_screening_assessment.pdf

Yeung, M. K. (2022). A systematic review and meta-analysis of facial emotion recognition in autism spectrum disorder: The specificity of deficits and the role of task characteristics. *Neuroscience & Biobehavioral Reviews*, *133*, Article 104518. https://doi.org/10.1016/j.neubiorev.2021.104518

APPENDIX A: STUDY 1

Odermatt, S. D., Möhring, W., Grieder, S., & Grob, A. (2022). Cognitive and developmental functions in autistic and non-autistic children and adolescents: Evidence from the Intelligence and Development Scales–2. *Journal of Intelligence, 10*(4), Article 4. https://doi.org/10.3390/jintelligence10040112

Please note that this is the author's version of a work that was accepted for publication in *Journal of Intelligence*. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. References are formatted according to the guidelines of the journal.
The final article is available via its DOI: https://doi.org/10.3390/jintelligence10040112

Cognitive and Developmental Functions in Autistic and Non-Autistic Children and Adolescents: Evidence from the Intelligence and Development Scales–2

Salome D. Odermatt[1], Wenke Möhring[1,2], Silvia Grieder[1], and Alexander Grob[1]

[1] Department of Psychology, University of Basel, Switzerland
[2] Department of Educational Psychology and Health Psychology, University of Education Schwäbisch Gmünd, Germany

Author Note

The authors declare no conflict of interest, with one exception: A.G. is recipient of royalties for the Intelligence and Development Scales–2 (IDS-2). This research received no external funding. The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee Northwest and Central Switzerland (protocol code: PB_2016-01836 and date of approval: 2 May 2019). Informed consent was obtained from all subjects involved in the study. The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy issues and property rights. Conceptualization, S.D.O., W.M., S.G. and A.G.; methodology, S.D.O., W.M. and S.G.; software, S.D.O.; validation, S.D.O., W.M. and S.G.; formal analysis, S.D.O.; investigation, S.D.O.; resources, A.G.; data curation, S.D.O.; writing—original draft preparation, S.D.O.; writing—review and editing, S.D.O., W.M., S.G. and A.G.; visualization, S.D.O.; supervision, A.G.; project administration, A.G.; funding acquisition, A.G. All authors have read and agreed to the published version of the manuscript. We would like to thank all the children and adolescents for their participation and all the test administrators, especially Miriam Weibel, for their support during data collection. We thank Anita Todd for copyediting the manuscript.

Correspondence concerning this article should be addressed to Salome D. Odermatt, Department of Psychology, University of Basel, Missionsstrasse 62, 4055 Basel, Switzerland. Email: salome.odermatt@unibas.ch

Abstract

Autistic individuals often show impairments in cognitive and developmental domains beyond the core symptoms of lower social communication skills and restricted repetitive behaviors. Consequently, the assessment of cognitive and developmental functions constitutes an essential part of the diagnostic evaluation. Yet, evidence on differential validity from intelligence and developmental tests, which are commonly used with autistic individuals, varies widely. In the current study, we investigated the cognitive (i.e., intelligence, executive functions) and developmental (i.e., psychomotor skills, social-emotional skills, basic skills, motivation and attitude, participation during testing) functions of autistic and non-autistic children and adolescents using the Intelligence and Development Scales–2 (IDS-2). We compared 43 autistic (M_{age} = 12.30 years) with 43 non-autistic (M_{age} = 12.51 years) participants who were matched for age, sex, and maternal education. Autistic participants showed significantly lower mean values in psychomotor skills, language skills, and the evaluation of participation during testing of the developmental functions compared to the control sample. Our findings highlight that autistic individuals show impairments particularly in motor and language skills using the IDS-2, which therefore merit consideration in autism treatment in addition to the core symptoms and the individuals' intellectual functioning. Moreover, our findings indicate that particularly motor skills might be rather neglected in autism diagnosis and may be worthy of receiving more attention. Nonsignificant group differences in social-emotional skills could have been due to compensatory effects of average cognitive abilities in our autistic sample.

Keywords: autism spectrum disorder; cognitive functions; developmental functions; Intelligence and Development Scales–2; children and adolescents

1. Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by difficulties in social communication and interaction accompanied by restricted repetitive behaviors, activities, and interests (American Psychiatric Association 2013). The worldwide prevalence of ASD has increased in recent years to approximately 1–2% (Idring et al. 2015; Maenner et al. 2020) and ASD is now considered a comparatively frequent condition (Happé and Frith 2020). Autistic individuals often experience difficulties beyond the core symptoms, such as impairments in cognitive and developmental domains, which in turn predict long-term development (e.g., Howlin and Moss 2012). Information about each individual's cognitive and developmental abilities is particularly important when it comes to making decisions about access to social services, the selection of appropriate treatment programs, and educational placement (White et al. 2007). Moreover, the amount of provided support is oftentimes determined on the basis of a cognitive assessment (Bowen 2014). According to the criteria of the *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; American Psychiatric Association 2013) and the *International Statistical Classification of Diseases and Related Health Problems* (11th ed.; World Health Organization 2018), clinicians have to report potential difficulties such as intellectual and language impairments in the diagnostic evaluation. Therefore, assessments with intelligence and developmental test batteries—in addition to autism-specific test procedures—represent a core part of the diagnostic process for autistic children and adolescents.

Yet, current tests for children and adolescents mainly allow the assessment of only single characteristics, such as intelligence, at a time and test batteries including multiple cognitive and developmental functions are missing so far. Consequently, when information about several domains or a broad assessment in a diagnostic evaluation is needed, clinicians often have to use various tests. This can be challenging, as the theoretical background and test administration differ widely among tests and dealing with these differences requires resources from the clinician. Moreover, tests build upon different characteristics of standardization samples and thus show less comparable scaled scores. The Intelligence and Development Scales–2 (IDS-2; Grob and Hagmann-von Arx 2018a) is a standardized test battery that assesses cognitive (i.e., intelligence and executive functions) and developmental (i.e., psychomotor skills, social-emotional skills, basic skills, motivation and attitude, and participation during testing) functions in 5- to 20-year-olds. The IDS-2 thus provides a comprehensive picture of an individual's strengths and difficulties with a single test battery across a wide age range from childhood to adolescence. In addition, the IDS-2 contains clear instructions and structured tasks, and many subtests use a closed-response format, which is particularly important for autistic children because of frequent structural language difficulties (Boucher 2012), making it suitable for administration with autistic individuals. Since the publication of the IDS-2 in 2018, it has often been used in psychological and medical practice in German-speaking countries. Further international adaptations for several other languages are currently in progress or have recently been published (e.g., Dutch, English, Italian, Polish; Grob et al. 2018, 2019, 2021, 2022). In the present study, we aimed to compare autistic children and adolescents to a matched non-autistic control sample on cognitive and developmental functions to study the differential validity of test scores from the IDS-2. By doing so, we can assess whether the IDS-2 is able to distinguish between clinical subgroups and typically developing individuals (Schmidt-Atzert and Amelang 2012).

Although general intellectual functioning varies substantially among autistic individuals, the latest report from the Centers for Disease Control and Prevention showed that almost 60% of autistic children are classified in the below-average intelligence range (IQ < 85), with about half of these children meeting criteria for intellectual disability (IQ ≤ 70; Maenner et al. 2020). Autistic individuals typically display uneven cognitive profiles, with relative strengths in nonverbal domains (e.g., Coolican et al. 2008; Grondhuis et al.

2018) and in tasks assessing abstract reasoning and visuospatial abilities (Charman et al. 2011; Nader et al. 2016), such as a well-documented peak in the Block Design subtest of the Wechsler Intelligence Scales (e.g., Muth et al. 2014). In contrast, relative weaknesses have been demonstrated in verbal domains, particularly in the Comprehension subtest of the Wechsler Intelligence Scales (e.g., Oliveras-Rentas et al. 2012), and in processing speed and working memory tasks[1] (Mayes and Calhoun 2003a; Nader et al. 2016; Oliveras-Rentas et al. 2012).

Autistic individuals often experience further cognitive difficulties on measures assessing executive functions (e.g., Hill 2004). Executive functions include a set of mental top-down regulation and control mechanisms (Miyake and Friedman 2012). In the theory of executive dysfunction, it is assumed that impairments in executive functions are responsible for some of the autism symptoms (Pennington and Ozonoff 1996), such as repetitive behavior (e.g., de Vries and Geurts 2012; Yerys et al. 2009). Demetriou et al. (2018) reported in the largest meta-analysis to date (235 studies) that autistic individuals showed moderate impairments in executive functions, both overall and in subdomains such as cognitive flexibility, fluency, planning, and inhibition,—which are also assessed with the IDS-2 (see Table S1 in the Supplement for an overview)—compared to non-autistic individuals.

Moreover, previous research showed significant impairments in autistic individuals' motor abilities, beginning in early childhood with deficits in the acquisition of motor milestones, such as later independent walking (e.g., Manicolo et al. 2019), and delays in gross and fine motor skills, for example, diminished object manipulation activity (Libertus et al. 2014; Provost et al. 2007). In a recent meta-analysis of 139 studies with samples of autistic children, adolescents, and young adults, their overall motor ability as well as gross and fine motor skills were strongly impaired in comparison to non-autistic peers (Coll et al. 2020). In line with this result, several studies found that autistic children, compared to non-autistic samples, scored lower on subscales (i.e., manual dexterity, ball skills, and balance) of the Movement Assessment Battery for Children–2 (M-ABC-2; Petermann 2008), which is a test of motor development that contains tasks similar to those in the IDS-2 psychomotor skills domain (Liu and Breslin 2013; Manicolo et al. 2019; Siaperas et al. 2012).

Further, research has shown that lower motor skills of autistic children were significantly associated with poorer social communication skills (MacDonald et al. 2013b). It has been suggested that motor problems might even precede social and communication deficits in autistic individuals because they may limit social participation and interaction with peers during play and may interfere with effective and timely movements, such as turning the head or pointing to something, that are particularly important for joint attention (Bhat et al. 2011). Impairments in social communication and interaction, such as difficulties in social-emotional reciprocity and nonverbal communicative behaviors, as well as in developing, maintaining, and understanding relationships constitute a core diagnostic characteristic of ASD (American Psychiatric Association 2013; World Health Organization 2018). These impairments are reflected in less accurate emotion recognition in human faces, with increased response times (Leung et al. 2022; Yeung 2022), more maladaptive emotion regulation strategies (Cai et al. 2018), including more reliance on others to regulate their emotions (Cibralic et al. 2019), and fewer socially competent behaviors (e.g., Meyer et al. 2009) compared to non-autistic individuals.

Additionally, language difficulties commonly co-occur with autism (Kjellmer et al. 2018). Some autistic individuals do not acquire verbal language at all (Brignell et al. 2018). Among those who develop language, delays often begin in infancy with retardations in the production of first words and in early language comprehension (e.g., Luyster et al. 2007; Mitchell et al. 2006). Moreover, across the preschool years, autistic children exhibit difficulties in phonological awareness skills (e.g., identifying syllables or onset-rimes), with slower development than their non-autistic peers (Dynia et al. 2019). Regarding language production and comprehension (i.e., expressive and receptive language skills, respectively), some studies indicated an atypical pattern, with better expressive and poorer receptive language skills in autistic individuals (e.g., Hudry et al. 2010). However, a meta-analysis examining 74 studies reported that autistic children and adolescents had scores that were approximately 1.5 standard deviations lower in receptive *as well as* expressive language abilities compared to non-autistic samples (Kwok et al. 2015).

In terms of academic skills, research indicated that autistic students demonstrate variable performance (Keen et al. 2016). Specifically, in previous studies, autistic individuals showed similar basic word-reading skills, such as word recognition, compared to non-autistic peers, but they tended to have difficulties in reading comprehension (for a meta-analysis: Brown et al. 2013). Autistic individuals with higher (vs. lower) reading skills also seemed to demonstrate better writing abilities (Zajic et al. 2020). Studies predominantly indicated deficits in text generation abilities for autistic individuals, while overall intact or slightly impaired spelling skills were reported (Finnegan and Accardo 2018; Mayes and Calhoun 2003a, 2003b). Similarly, the majority of autistic individuals exhibited average competencies in mathematics, such as mathematical problem solving, compared to non-autistic peers or to the norm population in previous research (Chiang and Lin 2007; Titeca et al. 2017; Troyb et al. 2014).

Concerning motivation and attitude, a recent meta-analysis reported that autistic individuals displayed significantly lower levels of conscientiousness than non-autistic individuals (Lodi-Smith et al. 2019). In contrast, less is known regarding achievement motivation in autistic individuals. A few studies reported that autistic individuals encountered problems with self-regulation (e.g., Jahromi et al. 2012; Konstantareas and Stewart 2006) and displayed higher interest in mathematics while simultaneously showing more fear of failure and lower mastery goals (Georgiou et al. 2018). Moreover, autistic children tended to exhibit impaired engagement (Keen 2009), especially in assessment situations where they frequently demonstrated off-task behaviors (Akshoomoff 2006) and a lack of willingness to complete tasks (Mandelbaum et al. 2006).

Previous research has rarely used the IDS-2 in order to test autistic individuals. The only study so far reported in the technical manual of the IDS-2 (Grob and Hagmann-von Arx 2018b) built upon a small sample of autistic children and adolescents (N = 18; M_{age} = 13 years 4 months, age range 8–17 years; 17 males and 1 female). Findings showed significantly lower group mean values for autistic children and adolescents compared to non-autistic peers in the composite score of social-emotional skills (d = 0.62) and the composite score of psychomotor skills (d = 1.01) of the IDS-2. No differences were found in the composite scores of other domains. However, evidence of possible differences at the level of subtests is currently lacking, as analyses on this level have not been performed. Moreover, the study included mainly children and adolescents with Asperger's syndrome (n = 13) and no participants with previously diagnosed infantile autism. Given the small sample size, which may have diminished the power to find group differences, and the biased distribution of sex and subtype, it remains unknown to what extent these results can be generalized.

Building on this theoretical background, we pursued two goals for the present study: First, we aimed to extend previous research on various cognitive and developmental functions in autistic children and adolescents using a single test procedure and based on the norms of a large and representative standardization sample. By doing so, our findings will provide a comparable and comprehensive view of participants' performance in relevant domains. Second, we aimed to add knowledge regarding the differential validity evidence for test scores of the IDS-2 in autistic individuals, as psychological test procedures need to be examined in terms of their scientific quality in order to draw appropriate conclusions based on their test results. Given that previous research had some limitations (Grob and Hagmann-von Arx 2018b), we attempted to overcome these shortcomings by assessing a larger sample, including a more representative mapping of sex and subtypes, and performing analyses at the level of subtests, which have not yet been investigated in this population. We therefore examined possible mean-level differences between a large sample of autistic children and adolescents and a control sample of non-autistic children and adolescents matched by age, sex, and maternal education in the cognitive and developmental functions measured by the IDS-2. We included maternal education as a proxy for socioeconomic status (SES) to control for the fact that more autistic children and adolescents come from families with higher SES than from other SES groups (Thomas et al. 2012; Van Meter et al. 2010).

Taking into consideration the presented literature, we hypothesized that autistic children and adolescents would score lower than the control sample of non-autistic children in the following IDS-2 domains as displayed in Table 1, while we assumed that autistic children and adolescents' scores would be similar to those of the control sample in the other IDS-2 domains (see Table 1 for a summary).

2. Materials and Methods

2.1 Participants and Procedure

Forty-three autistic children and adolescents (M_{age} = 12 years 4 months, age range 7–17 years; 35 males and 8 females) were recruited during (*n* = 18) or after (*n* = 25) the IDS-2 standardization and validation study with the help of local child and adolescent psychiatric services and hospitals, privately practicing psychiatrists and psychotherapists who are experts in autism diagnoses, and associations for autistic individuals. All included children and adolescents were diagnosed with ASD (infantile autism: *n* = 11, atypical autism: *n* = 6, Asperger's syndrome: *n* = 24, not specified: *n* = 2) but were not selected on the basis of specific subtypes. Participants had received the diagnosis on average 4.08 years (*SD* = 2.61) prior to their participation in the present study. The ratio of males to females corresponded to the distribution of approximately four males to one female diagnosed with ASD in the population (Maenner et al. 2020).

A control sample of 43 non-autistic children and adolescents (M_{age} = 12 years 6 months, age range 6–20 years; 35 males and 8 females) was drawn from the German standardization and validation sample of the IDS-2 (*N* = 2030; M_{age} = 12 years 3 months, age range 5–20 years; 977 males and 1053 females). The control sample was matched by age, sex, and maternal education (as a proxy for SES) and did not differ regarding demographic characteristics from the sample of autistic children and adolescents (see Table 2). Non-autistic children and adolescents were recruited from kindergartens and schools.

All participants were individually tested using the IDS-2 by psychologists or trained psychology students. For the administration of the IDS-2 with autistic children and adolescents, we received input from psychiatrists and psychotherapists who specialize in autism. Test administration lasted approximately 4 h and was split into two sessions no longer than 1 week apart upon a participant's request. Participants were tested either at their homes or in a laboratory at the university. The local ethics committee (Ethics Committee Northwest and Central Switzerland) provided approval and the study was conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from participants and/or their parents.

2.2 Instrument

A detailed description of the IDS-2 (Grob and Hagmann-von Arx 2018a) can be found in the Supplemental Material (Table S1). Psychometric properties have been demonstrated in several studies for the standardization sample (Grieder and Grob 2020; Grob and Hagmann-von Arx 2018b). Demographic characteristics were assessed through a parental interview at the beginning of the first test session.

2.3 Statistical Analyses

Analyses were conducted with R (R Core Team 2021). To obtain a non-autistic sample that would be comparable to the autistic sample with respect to demographic characteristics, we performed a matching procedure using the MatchIt package (Ho et al. 2011). We matched the two samples by age (nearest; continuous), sex (exact; 0 = male, 1 = female), and maternal education (nearest; 1 = compulsory school, 2 = apprenticeship, 3 = high school, 4 = higher vocational education, 5 = university degree, 6 = other, 7 = unknown). We calculated independent-samples *t* tests to investigate mean-level differences between the autistic sample and the non-autistic sample in cognitive and developmental domains using standardized scores (M = 100, SD = 15, for Profile IQ, Full-Scale IQ, Screening IQ, and the seven intelligence group factors; M = 10, SD = 3, for other composite scores and subtests). To reduce the alpha error inflation caused by multiple testing, *p* values were adjusted with Hommel's (1988) correction by including *p* values from all tests simultaneously. Effect sizes were computed (Cohen 1988) and interpreted in accordance with common practice (Cohen's *d*; small effect: $d \geq 0.20$, medium effect: $d \geq 0.50$, large effect: $d \geq 0.80$). A post-hoc power analysis using G*Power (Faul et al. 2007) revealed that with $\alpha = .05$ and power = .80, small effects ($d = 0.30$) could be detected in the present sample (note that this is without accounting for multiple testing). Differences were interpreted as meaningful if they were significant after Hommel's correction and showed at least a small effect size. In addition, we reported reliabilities for all IDS-2 scores, consisting of Cronbach's alpha for homogeneous subtests; reliabilities calculated according to a formula of Lienert and Raatz (1998) for composite scores, which are based on intercorrelations and reliabilities of those subtests or tasks that are included in the corresponding score; or retest reliabilities reported in the technical manual of the IDS-2 (Grob and Hagmann-von Arx 2018b) for subtests that contain a single score or consist of heterogeneous tasks.

3. Results

Reliabilities, descriptive statistics, and results of the independent-samples *t* tests[2] are presented in Table 3 for the cognitive functions and in Table 4 for the developmental functions. Reliabilities were high for composite scores and high-to-satisfactory for subtests in both samples.

3.1 Cognitive Functions

Figure 1 displays the means and standard deviations in the cognitive functions of the IDS-2 for the autistic and non-autistic samples. Before controlling for multiple testing, we found significant group differences for the intelligence composite scores: Profile IQ, $t(77) = 1.96$, $p = .027$, and Screening IQ, $t(82) = 1.80$, $p = .038$, with small effect sizes (d = 0.44 and 0.39, respectively), indicating lower scores for the autistic sample than the control sample. Furthermore, we observed group differences for the intelligence group factors: Auditory Short-Term Memory, $t(79) = 2.12$, $p = .019$, and Visuospatial Short-Term Memory, $t(79) = 2.70$, $p = .004$, with small-to-medium effect sizes (d = 0.47 and 0.60, respectively), and the corresponding subtests Mixed Digit and Letter Span, $t(79) = 2.51$, $p = .007$, and Rotated Shape Memory, $t(79) = 2.78$, $p = .003$, with medium effect sizes (d = 0.56 and 0.62, respectively), such that the autistic participants showed lower mean values than the control sample. Moreover, the autistic participants had significantly lower mean values in the executive functions composite score, $t(71) = 2.27$, $p = .013$, and the subtests Listing Words, $t(73) = 2.38$, $p = .010$, Divided Attention, $t(71) = 2.13$, $p = .019$, and Animal Colors, $t(72) = 1.70$, $p = .047$. Effect sizes were in the small-to-medium range (d = 0.40 to 0.55). We found no differences between autistic and non-autistic participants in the Full-Scale IQ, $t(81) = 1.58$, $p = .059$, in the intelligence group factors Visual Processing, $t(80) = 1.46$, $p = .148$, Processing Speed, $t(80) = 1.15$, $p = .126$, Abstract Reasoning, $t(80) = 0.62$, $p = .539$, Verbal Reasoning, $t(81) = 1.48$, $p = .071$, and Long-Term

Memory, $t(79) = 1.57$, $p = .060$, including corresponding intelligence subtests, and in the executive functions subtest Drawing Routes, $t(74) = 0.88$, $p = .192$.

However, after controlling for multiple testing, the significant differences in intelligence and executive functions fell above the Hommel-corrected *p*-value threshold (see Table 3).

3.2 Developmental Functions

Figure 2 shows the means and standard deviations in the developmental functions of the IDS-2 for the autistic and non-autistic samples. Before controlling for multiple testing, results indicate that autistic participants scored significantly lower than non-autistic participants in psychomotor skills [composite score, $t(81) = 4.60$, $p < .001$; Gross Motor Skills, $t(32) = 5.30$, $p < .001$; Fine Motor Skills, $t(79) = 3.20$, $p < .001$; Visuomotor Skills, $t(81) = 3.01$, $p = .002$] with medium-to-large effect sizes ($d = 0.66$ to 1.82). We found a similar group difference for participants' social-emotional skills [composite score, $t(82) = 2.71$, $p = .004$; Identifying Emotions, $t(32) = 2.07$, $p = .023$; Regulating Emotions, $t(82) = 2.37$, $p = .010$; Socially Competent Behavior, $t(80) = 2.29$, $p = .012$] with medium effect sizes ($d = 0.51$ to 0.71), and in language skills [composite score, $t(28) = 4.11$, $p < .001$; Phoneme Analysis, $t(29) = 3.75$, $p < .001$; Language Expressive, $t(28) = 3.31$, $p = .001$; Language Receptive, $t(29) = 4.52$, $p < .001$] with large effect sizes ($d =$ 1.22 to 1.63). Furthermore, autistic participants showed significantly lower group mean values than the control sample for the evaluation of participation during the test session of intelligence, $t(81) = 2.68$, $p =$ .004, executive functions, $t(71) = 2.13$, $p = .018$, and developmental functions, $t(79) = 3.30$, $p < .001$, with medium effect sizes ($d = 0.50$ to 0.73). We found no differences in the subtests Logical-Mathematical Reasoning, $t(81) = 1.44$, $p = .153$, Reading, $t(74) = 1.35$, $p = .182$, Spelling, $t(65) = 1.26$, $p = .212$, and in the motivation and attitude domain [composite score, $t(46) = 0.11$, $p = .458$; Conscientiousness, $t(45) =$ 0.06, $p = .477$; Achievement Motivation, $t(47) = -0.07$, $p = .528$], indicating similar performance in autistic and non-autistic participants.

After controlling for multiple testing, significant group differences remained for the composite score of psychomotor skills ($p_H < .001$) and subtests Gross Motor Skills ($p_H < .001$) and Fine Motor Skills ($p_H =$ .046). Moreover, the composite score of language skills remained significant ($p_H = .008$) as well as Phoneme Analysis ($p_H = .019$) and Language Receptive ($p_H = .003$) tasks. Finally, the evaluation of participation during testing of the developmental functions remained significant ($p_H = .035$; see Table 4).[3]

3.3 Post-Hoc Analyses

To assess for age-related differences between children and adolescents, we further performed post-hoc analyses separately for children aged 5–10 years ($n = 17$) and adolescents aged 11–20 years (n $= 26$). After Hommel's (1988) correction, autistic children scored significantly lower than non-autistic children in the composite scores of the cognitive functions, the intelligence group factors, Auditory Short-Term Memory, Visuospatial Short-Term Memory, and Verbal Reasoning (including the corresponding subtests) as well as in psychomotor skills, social-emotional skills, and basic skills of the developmental functions (see Tables S3 and S4 in the Supplemental Material for results). We found no significant group differences between autistic and non-autistic adolescents for the cognitive and developmental functions of the IDS-2 after controlling for multiple testing (see Tables S5 and S6 in the Supplemental Material).

4. Discussion

In the present study, we compared autistic children and adolescents to a matched control sample on six cognitive and developmental functions assessed with the IDS-2. Our results provide evidence for differential validity for the IDS-2 test scores in psychomotor skills, language skills, and in the evaluation of participation during testing of the developmental functions, with autistic children and adolescents scoring lower than non-autistic participants in these domains. No group differences were detected in the other domains after controlling for multiple testing. Overall, our findings provide an overview of important cognitive and developmental functions in autistic children and adolescents using a single comprehensive and standardized test battery.

In line with our hypotheses, we found similar performance in autistic and non-autistic participants for the intelligence group factors Visual Processing and Abstract Reasoning, which corresponds to studies reporting relative strengths for autistic individuals in nonverbal domains (e.g., Grondhuis et al. 2018) and in subtests measuring fluid reasoning and visuospatial abilities (Charman et al. 2011; Nader et al. 2016). Specifically, the Shape Design subtest, which is part of the Visual Processing group factor of the IDS-2, requires participants to reproduce presented geometric figures with rectangles and triangles. This task is similar to the Block Design subtest of the Wechsler Intelligence Scales, for which autistic individuals oftentimes show at least comparable performance to non-autistic controls (e.g., Muth et al. 2014).

However, in contrast to our hypotheses and previous research (e.g., Demetriou et al. 2018), no significant group differences emerged for the other cognitive functions scores of the IDS-2 after correcting for multiple testing, even though effect sizes were in the small-to-medium range. This finding suggests that our autistic sample included participants with overall average cognitive abilities. One explanation for this result could be that about half of our autistic participants had been diagnosed with Asperger's syndrome, which is known for impairments in social interaction and restricted interests, but without deficits in cognitive development (10th ed.; World Health Organization 2016). Moreover, when assessing age-related differences in a set of post-hoc analyses, we found that autistic adolescents scored similarly to non-autistic adolescents in the IDS-2, while autistic children obtained significantly lower scores in several domains of the IDS-2 compared to non-autistic children. In particular, group differences between autistic and non-autistic children remained significant after controlling for multiple testing in the composite scores of the intelligence and executive functions domains as well as in the intelligence group factors Verbal Reasoning and Auditory and Visuospatial Short-Term Memory. These results are in line with previous research reporting weaknesses of autistic children in verbal domains (e.g., Oliveras-Rentas et al. 2012) and in working memory tasks (e.g., Mayes and Calhoun 2003a) as the IDS-2 Auditory and Visuospatial Short-Term Memory group factors also include tasks measuring working memory (i.e., [Mixed] Digit and Letter Span—backwards and Rotated Shape Memory; see Table S1 in the Supplement). In addition, autistic children scored lower on motor and language skills, and importantly, also on social-emotional skills. Interestingly, we did not find any differences between autistic and non-autistic participants when focusing on adolescents only. One reason for this finding could be that autistic adolescents have already received support and intervention in crucial developmental areas, whereas the included autistic children may have been recently diagnosed with autism and thus have had little or no treatment to that point. However, it should be noted that these results are based on small sample sizes. Thus, future studies should use larger age-specific samples to investigate developmental effects across childhood and adolescence and simultaneously control for previous interventions.

Autistic participants had significant impairments in overall psychomotor skills as well as lower scores in gross and fine motor skills in the IDS-2 compared to the non-autistic participants. This finding is in line with results of a previous meta-analysis (Coll et al. 2020) and studies using the M-ABC-2 to assess motor abilities (e.g., Manicolo et al. 2019). Motor skills are particularly important for carrying out everyday tasks (e.g., grasping a glass) and performing activities of daily living (MacDonald et al. 2013a), as well as for participating in activities at school or in the community (Oliveira et al. 2021). It has been suggested that one reason for these motor differences may be that autistic individuals encounter problems in the translation of sensory inputs into movements (Hannant et al. 2016). Moreover, structural and functional alterations in motor cortex regions of the brain (Mostofsky et al. 2007; Nebel et al. 2014) and in the cerebellum (Fatemi et al. 2012; Mostofsky et al. 2009) have been detected for autistic individuals, which might explain some of the motor impairments. The strong group difference we observed in gross motor skills, representing the largest effect in our study, is in accordance with previous research (Coll et al. 2020) and may be associated with the high prevalence of autistic individuals exhibiting hypotonia (51%) or motor apraxia (34%; Ming et al. 2007). Hence, autistic individuals tend to experience difficulties especially in movements that require activation of muscles in the entire body including balance, arm movements, and coordination. However, as this subtest is administered only to 5- to 10-year-olds in the IDS-2 and correlational research has shown that autistic children's motor skills improve with age (Coll et al. 2020), future longitudinal studies are needed to study possible developmental effects. Although it is not compulsory to report potential difficulties in motor skills as part of the diagnostic criteria of ASD, our findings support the importance of assessing psychomotor abilities during the diagnostic evaluation of children and adolescents at increased likelihood of ASD, as they might be crucial for treatment programs (Bhat et al. 2011; Colombo-Dougovito and Block 2019).

As stated in previous studies, we found that autistic children scored lower in language skills, such as in phoneme analysis (Dynia et al. 2019) and receptive language tasks (Kwok et al. 2015), compared to the non-autistic participants. However, we detected no significant group differences after correcting for multiple testing in expressive language tasks. Although a previous meta-analysis showed equally impaired receptive and expressive language skills in autistic individuals (Kwok et al. 2015), our finding is in line with other studies that also indicated an atypical language pattern of autistic individuals with an advantage in expressive over receptive language skills (e.g., Hudry et al. 2010). One reason for this result might be that we used a direct measurement of language skills in our study. Previous research also found this pattern when using a similar test procedure but did not detect any expressive language advantages when using caregiver reports (Ellis Weismer et al. 2010). Given that having better language production than comprehension skills is contrary to what is generally anticipated in typically developing peers, researchers even suggested that this pattern may be unique to autism (e.g., Volden et al. 2011) and therefore could be used for differential diagnosis (Mitchell et al. 2011) and specific interventions (Hudry et al. 2010). Nevertheless, as the expressive and receptive language tasks are conducted only with 5- to 10-year-olds in the IDS-2 and previous studies have reported a decrease in the expressive–receptive discrepancy in older autistic individuals (Kwok et al. 2015; Volden et al. 2011), it could also be that our result was driven by age effects. Because of the diagnostic and therapeutic potential of this finding, future studies should continue to examine this potential discrepancy between expressive and receptive language in autistic individuals across development.

Additionally, we found no significant group differences in tasks measuring phoneme-grapheme correspondence, which is consistent with our finding that autistic participants also scored similarly to the non-autistic control group in the reading and spelling subtests in our study. This result might be explained by the fact that knowledge of letter–sound correspondence is a prerequisite for the development of literacy skills (Carnine et al. 2010) and therefore needs to be intact for average reading and spelling skills. The finding that our autistic participants showed no differences in the basic skills logical-mathematical reasoning, reading, and spelling compared to non-autistic peers is in line with other studies (e.g., Brown et al. 2013; Chiang and Lin 2007). One reason may refer to the fact that most of the autistic participants in our study attended inclusive educational settings. The enrollment in integrative settings can have a positive impact on autistic individuals' academic skills as individualized education plans in mainstream programs focus more on academic enhancement than in specialized settings which place more emphasis on life competencies and developmental domains (Kurth and Mastergeorge 2010).

Contrary to previous research (e.g., Cai et al. 2018; Yeung 2022), we found no significant group differences for social-emotional skills after correcting for multiple testing. One explanation for this result could be that the tasks assessing social-emotional skills in the IDS-2 mainly measure explicit knowledge, such as naming socially competent behavior in hypothetical social situations, rather than actual behavior in real-life situations. Since we did not observe any group differences in the cognitive functions of the IDS-2 either, it might be that autistic participants could compensate for difficulties in social-emotional skills with higher-level analytical strategies (Harms et al. 2010; Leung et al. 2022). This would be in line with studies reporting that intelligence is positively associated with social-emotional skills (Jones et al. 2011), especially in autistic individuals (Dyck et al. 2006; Salomone et al. 2019; Trevisan and Birmingham 2016). We found further evidence for this assumption in supplementary analyses where we matched the non-autistic control sample by age, sex, and Full-Scale IQ and obtained lower effect sizes for the social-emotional skills composite score as well as for the subtests Identifying Emotions and Regulating Emotions compared to the effect sizes obtained by matching the samples by age, sex, and maternal education (see Table S2 in the Supplemental Material). In addition, time limits in testing procedures might explain part of the nonsignificant group differences in social-emotional skills. Nagy et al. (2021) found impairments only when time limits for responding were applied, and the present tasks assessing social-emotional skills did not have any time restrictions. However, it is important to note that although meta-analyses and reviews show significant deficits in social-emotional abilities of autistic individuals (e.g., Cai et al. 2018; Yeung 2022), several previous studies were also not able to detect impairments in emotion recognition and regulation (e.g., Jones et al. 2011; Mazefsky et al. 2014; Rosset et al. 2008) or reported difficulties only for certain emotions, for example, for negative emotions (e.g., Shanok et al. 2019). To clarify the interplay between explicit knowledge and social-emotional skills in the IDS-2, future research should use multiple methods to assess social-emotional skills and compare the autistic participants' performance in the IDS-2 with the behavior they demonstrate in real-life social interactions using observational measures. Even though the group differences in the social-emotional skills of the IDS-2 were no longer significant after correcting for multiple testing, it is crucial to mention that effect sizes were within a medium range and comparable to those in a previous meta-analysis (Yeung 2022) which at least tends to indicate differential validity of test scores from the social-emotional skills domain of the IDS-2.

A strength of our study is that we assessed the cognitive and developmental functions using a standardized test procedure with good psychometric properties. Moreover, we used a single test battery based on one standardization sample for the assessment of a broad range of cognitive and developmental domains. In addition, our sample covered a wide age range and was representative of the autistic population, in that the male:female ratio was approximately 4:1 (Maenner et al. 2020), different subtypes were included, and children and adolescents exhibited known comorbid conditions (Leyfer et al. 2006; Salazar et al. 2015). We also consider it a strength that we included participants with intellectual functioning below 70, which represents an understudied subpopulation in autism research (Russell et al. 2019). In addition, by selecting the control sample through a matching procedure, we could control for possible confounding influences of age, sex, and SES.

The present study also has limitations that need to be considered and addressed in future research. First, we relied on diagnostic evaluations carried out by clinical services and experienced psychiatrists and psychotherapists and hence could not consider the standardization and comparability of the diagnoses. Second, we had no information regarding symptom severity or previous treatment programs and could therefore not control for these factors. Third, analyses were conducted at the group level, which limits generalizability to individuals. Finally, although the sample size was larger than in previous studies, an even larger sample of children and adolescents would further increase the power to detect small effects in future studies.

5. Conclusions

In sum, our findings suggest that in particular, motor and language skills as well as achievement motivation rated by the test administrator were impaired in autistic children and adolescents in the IDS-2 compared to non-autistic participants, which provides evidence for differential validity for these domains of the IDS-2. The largest difference was found in gross motor skills. We therefore advise that therapists working with autistic children should gain knowledge in the area of motor and language therapeutic intervention. Speech-language pathologists as well as psychomotor therapists should obtain autism-specific knowledge, so that autistic children with limited motor and language skills receive appropriate therapeutic support regardless of the background of the therapist. Arguably, with optimal training, autistic participants may also perform tasks in the psychomotor and language domains with greater engagement, which, in turn, could have a positive impact on the long-term development of their motor and language abilities. In conclusion, our results highlight important domains beyond the core symptoms of ASD that need to be considered in future research, educational contexts, and clinical assessment and that seem particularly critical for interventions.

Footnotes

1. According to current models of intelligence (Schneider and McGrew 2018) and executive functions (Miyake et al. 2000), working memory can be understood as a component of intelligence or executive functions. Because working memory is included in the intelligence domain in the IDS-2, we subsumed working memory under the realm of intelligence.
2. Although the sample size met the robustness criteria for using independent-samples *t* tests (Eid et al. 2017), we also examined the variables regarding normal distribution and variance homogeneity. Analyses using the Shapiro–Wilk test showed that 12 of the 55 dependent variables may not fulfill the normality assumption. Therefore, we additionally calculated Mann–Whitney *U* tests for these variables. The results remained largely the same with two exceptions: First, the mean difference in the subtest Identifying Emotions was no longer significant before controlling for multiple testing. Second, the mean difference in the composite score of language skills was no longer significant after controlling for multiple testing. Furthermore, we found that the Levene's test was significant for fewer than 10 of the dependent variables, indicating unequal variances. Thus, Welch's *t* tests were additionally performed. The results were identical to those obtained from the independent-samples *t* tests.
3. To control for effects of intelligence, we repeated the independent-samples *t* tests for the developmental functions with a non-autistic control sample matched by age, sex, and intelligence (Full-Scale IQ). The pattern of results remained largely the same, showing lower group mean values for the autistic participants than for the control sample in the domains psychomotor skills, social-emotional skills, language skills, and participation during testing (see Table S2 in the Supplemental Material for full results). These differences hold when correcting for multiple testing in the domain of psychomotor skills. These post-hoc analyses underscore the robustness of our findings.

6. References

Akshoomoff, Natacha. 2006. Use of the Mullen Scales of Early Learning for the Assessment of Young Children with Autism Spectrum Disorders. *Child Neuropsychology* 12: 269–77. https://doi.org/10.1080/09297040500473714.

American Psychiatric Association. 2013. *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed. Washington, DC: American Psychiatric Association.

Bhat, Anjana N., Rebecca J. Landa, and James C. (Cole) Galloway. 2011. Current Perspectives on Motor Functioning in Infants, Children, and Adults with Autism Spectrum Disorders. *Physical Therapy* 91: 1116–29. https://doi.org/10.2522/ptj.20100294.

Boucher, Jill. 2012. Research Review: Structural Language in Autistic Spectrum Disorder—Characteristics and Causes. *Journal of Child Psychology and Psychiatry* 53: 219–33. https://doi.org/10.1111/j.1469-7610.2011.02508.x.

Bowen, Sonya E. 2014. *Autism Spectrum Disorders (ASD): State of the States of Services and Supports for People with ASD*. Washington, DC: L & M Policy Research.

Brignell, Amanda, Angela T. Morgan, Susan Woolfenden, Felicity Klopper, Tamara May, Vanessa Sarkozy, and Katrina Williams. 2018. A Systematic Review and Meta-Analysis of the Prognosis of Language Outcomes for Individuals with Autism Spectrum Disorder. *Autism & Developmental Language Impairments* 3: 1–19. https://doi.org/10.1177/2396941518767610.

Brown, Heather M., Janis Oram-Cardy, and Andrew Johnson. 2013. A Meta-Analysis of the Reading Comprehension Skills of Individuals on the Autism Spectrum. *Journal of Autism and Developmental Disorders* 43: 932–55. https://doi.org/10.1007/s10803-012-1638-1.

Cai, Ru Ying, Amanda L. Richdale, Mirko Uljarević, Cheryl Dissanayake, and Andrea C. Samson. 2018. Emotion Regulation in Autism Spectrum Disorder: Where We Are and Where We Need to Go. *Autism Research* 11: 962–78. https://doi.org/10.1002/aur.1968.

Carnine, Douglas W., Jerry Silbert, Edward J. Kame'enui, and Sara G. Tarver. 2010. *Direct Instruction Reading*, 5th ed. Columbus: Merrill.

Charman, Tony, Andrew Pickles, Emily Simonoff, Susie Chandler, Tom Loucas, and Gillian Baird. 2011. IQ in Children with Autism Spectrum Disorders: Data from the Special Needs and Autism Project (SNAP). *Psychological Medicine* 41: 619–27. https://doi.org/10.1017/S0033291710000991.

Chiang, Hsu-Min, and Yueh-Hsien Lin. 2007. Mathematical Ability of Students with Asperger Syndrome and High-Functioning Autism: A Review of Literature. *Autism* 11: 547–56. https://doi.org/10.1177/1362361307083259.

Cibralic, Sara, Jane Kohlhoff, Nancy Wallace, Catherine McMahon, and Valsamma Eapen. 2019. A Systematic Review of Emotion Regulation in Children with Autism Spectrum Disorder. *Research in Autism Spectrum Disorders* 68: 101422. https://doi.org/10.1016/j.rasd.2019.101422.

Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale: Erlbaum.

Coll, Sarah-Maude, Nicholas E. V. Foster, Alexa Meilleur, Simona M. Brambati, and Krista L. Hyde. 2020. Sensorimotor Skills in Autism Spectrum Disorder: A Meta-Analysis. *Research in Autism Spectrum Disorders* 76: 101570. https://doi.org/10.1016/j.rasd.2020.101570.

Colombo-Dougovito, Andrew M., and Martin E. Block. 2019. Fundamental Motor Skill Interventions for Children and Adolescents on the Autism Spectrum: A Literature Review. *Review Journal of Autism and Developmental Disorders* 6: 159–71. https://doi.org/10.1007/s40489-019-00161-2.

Coolican, Jamesie, Susan E. Bryson, and Lonnie Zwaigenbaum. 2008. Brief Report: Data on the Stanford–Binet Intelligence Scales (5th Ed.) in Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders* 38: 190–97. https://doi.org/10.1007/s10803-007-0368-2.

de Vries, Marieke, and Hilde M. Geurts. 2012. Cognitive Flexibility in ASD; Task Switching with Emotional Faces. *Journal of Autism and Developmental Disorders* 42: 2558–68. https://doi.org/10.1007/s10803-012-1512-1.

Demetriou, Eleni A., Amit Lampit, Daniel S. Quintana, Sharon L. Naismith, Yun J.C. Song, Julia E. Pye, Ian Hickie, and Adam J. Guastella. 2018. Autism Spectrum Disorders: A Meta-Analysis of Executive Function. *Molecular Psychiatry* 23: 1198–204. https://doi.org/10.1038/mp.2017.75.

Dyck, Murray J., Jan P. Piek, David Hay, Leigh Smith, and Joachim Hallmayer. 2006. Are Abilities Abnormally Interdependent in Children With Autism? *Journal of Clinical Child & Adolescent Psychology* 35: 20–33. https://doi.org/10.1207/s15374424jccp3501_3.

Dynia, Jaclyn M., Allison Bean, Laura M. Justice, and Joan N. Kaderavek. 2019. Phonological Awareness Emergence in Preschool Children with Autism Spectrum Disorder. *Autism & Developmental Language Impairments* 4: 1–15. https://doi.org/10.1177/2396941518822453.

Eid, Michael, Mario Gollwitzer, and Manfred Schmitt. 2017. *Statistik und Forschungsmethoden [Statistics and Research Methods]*, 5th ed. Weinheim: Beltz.

Ellis Weismer, Susan, Catherine Lord, and Amy Esler. 2010. Early Language Patterns of Toddlers on the Autism Spectrum Compared to Toddlers with Developmental Delay. *Journal of Autism and Developmental Disorders* 40: 1259–73. https://doi.org/10.1007/s10803-010-0983-1.

Fatemi, S. Hossein, Kimberly A. Aldinger, Paul Ashwood, Margaret L. Bauman, Charles D. Blaha, Gene J. Blatt, Abha Chauhan, Ved Chauhan, Stephen R. Dager, and Price E. Dickson, and et al. 2012. Consensus Paper: Pathological Role of the Cerebellum in Autism. *The Cerebellum* 11: 777–807. https://doi.org/10.1007/s12311-012-0355-9.

Faul, Franz, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2007. G*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences. *Behavior Research Methods* 39: 175–91. https://doi.org/10.3758/BF03193146.

Finnegan, Elizabeth, and Amy L. Accardo. 2018. Written Expression in Individuals with Autism Spectrum Disorder: A Meta-Analysis. *Journal of Autism and Developmental Disorders* 48: 868–82. https://doi.org/10.1007/s10803-017-3385-9.

Georgiou, Alexandra, Spyridon-Georgios Soulis, and Danai Rapti. 2018. Motivation in Mathematics of High Functioning Students With Autism Spectrum Disorder (ASD). *Journal of Psychology Research* 8: 96–106. https://doi.org/10.17265/2159-5542/2018.03.002.

Grieder, Silvia, and Alexander Grob. 2020. Exploratory Factor Analyses of the Intelligence and Development Scales–2: Implications for Theory and Practice. *Assessment* 27: 1853–69. https://doi.org/10.1177/1073191119845051.

Grob, Alexander, and Priska Hagmann-von Arx. 2018a. *Intelligence and Development Scales–2 (IDS-2). Intelligenz- und Entwicklungsskalen für Kinder und Jugendliche [Intelligence and Development Scales for Children and Adolescents]*. Bern: Hogrefe.

Grob, Alexander, and Priska Hagmann-von Arx. 2018b. *Intelligence and Development Scales–2 (IDS-2). Intelligenz- und Entwicklungsskalen für Kinder und Jugendliche. Manual zu Theorie, Interpretation und Gütekriterien [Intelligence and Development Scales for Children and Adolescents. Manual on Theory, Interpretation and Psychometric Criteria]*. Bern: Hogrefe.

Grob, Alexander, Priska Hagmann-von Arx, Aleksandra Jaworowska, Anna Matczak, and Diana Fecenec. 2019. *Intelligence and Development Scales-2. Inteligencji i Rozwoju Dla Dzieci i Młodzieży [Intelligence and Development Scales for Children and Adolescents]*. Warsaw: Hogrefe.

Grob, Alexander, Priska Hagmann-von Arx, Anna Barnett, Nichola Stuart, and Serena Vanzan. 2021. *Intelligence and Development Scales-2 (IDS-2). Intelligence and Development Scales for Children and Adolescents*. Oxford: Hogrefe.

Grob, Alexander, Priska Hagmann-von Arx, Rosa Ferri, Monica Rea, and Maria Casagrande. 2022. *Intelligence and Development Scales-2. Scale Di Intelligenza e Sviluppo per Bambini e Adolescenti [Intelligence and Development Scales for Children and Adolescents]*. Florence: Hogrefe.

Grob, Alexander, Priska Hagmann-von Arx, Selma A. J. Ruiter, Marieke E. Timmerman, and Linda Visser. 2018. *Intelligence and Development Scales–2 (IDS-2). Intelligentie- En Ontwikkelingsschalen Voor Kinderen En Jongeren. [Intelligence and Development Scales for Children and Adolescents]*. Amsterdam: Hogrefe.

Grondhuis, Sabrina N., Luc Lecavalier, L. Eugene Arnold, Benjamin L. Handen, Lawrence Scahill, Christopher J. McDougle, and Michael G. Aman. 2018. Differences in Verbal and Nonverbal IQ Test Scores in Children with Autism Spectrum Disorder. *Research in Autism Spectrum Disorders* 49: 47–55. https://doi.org/10.1016/j.rasd.2018.02.001.

Hannant, Penelope, Teresa Tavassoli, and Sarah Cassidy. 2016. The Role of Sensorimotor Difficulties in Autism Spectrum Conditions. *Frontiers in Neurology* 7: 124.

Happé, Francesca, and Uta Frith. 2020. Annual Research Review: Looking Back to Look Forward—Changes in the Concept of Autism and Implications for Future Research. *Journal of Child Psychology and Psychiatry* 61: 218–32. https://doi.org/10.1111/jcpp.13176.

Harms, Madeline B., Alex Martin, and Gregory L. Wallace. 2010. Facial Emotion Recognition in Autism Spectrum Disorders: A Review of Behavioral and Neuroimaging Studies. *Neuropsychology Review* 20: 290–322. https://doi.org/10.1007/s11065-010-9138-6.

Hill, Elisabeth L. 2004. Executive Dysfunction in Autism. *Trends in Cognitive Sciences* 8: 26–32. https://doi.org/10.1016/j.tics.2003.11.003.

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2011. MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software* 42: 1–62. https://doi.org/10.18637/jss.v042.i08.

Hommel, Gerhard. 1988. A Stagewise Rejective Multiple Test Procedure Based on a Modified Bonferroni Test. *Biometrika* 75: 383–86.

Howlin, Patricia, and Philippa Moss. 2012. Adults with Autism Spectrum Disorders. *The Canadian Journal of Psychiatry* 57: 275–83. https://doi.org/10.1177/070674371205700502.

Hudry, Kristelle, Kathy Leadbitter, Kathryn Temple, Vicky Slonims, Helen McConachie, Catherine Aldred, Patricia Howlin, Tony Charman, and the PACT Consortium. 2010. Preschoolers with Autism Show Greater Impairment in Receptive Compared with Expressive Language Abilities. *International Journal of Language & Communication Disorders* 45: 681–90. https://doi.org/10.3109/13682820903461493.

Idring, Selma, Michael Lundberg, Harald Sturm, Christina Dalman, Clara Gumpert, Dheeraj Rai, Brian K. Lee, and Cecilia Magnusson. 2015. Changes in Prevalence of Autism Spectrum Disorders in 2001–2011: Findings from the Stockholm Youth Cohort. *Journal of Autism and Developmental Disorders* 45: 1766–73. https://doi.org/10.1007/s10803-014-2336-y.

Jahromi, Laudan B., Shantel E. Meek, and Sharman Ober-Reynolds. 2012. Emotion Regulation in the Context of Frustration in Children with High Functioning Autism and Their Typical Peers. *Journal of Child Psychology and Psychiatry* 53: 1250–58. https://doi.org/10.1111/j.1469-7610.2012.02560.x.

Jones, Catherine R. G., Andrew Pickles, Milena Falcaro, Anita J. S. Marsden, Francesca Happé, Sophie K. Scott, Disa Sauter, Jenifer Tregay, Rebecca J Phillips, Gillian Baird, and et al. 2011. A Multimodal Approach to Emotion Recognition Ability in Autism Spectrum Disorders. *Journal of Child Psychology and Psychiatry* 52: 275–85. https://doi.org/10.1111/j.1469-7610.2010.02328.x.

Keen, Deb. 2009. Engagement of Children with Autism in Learning. *Australasian Journal of Special Education* 33: 130–40. https://doi.org/10.1375/ajse.33.2.130.

Keen, Deb, Amanda Webster, and Greta Ridley. 2016. How Well Are Children with Autism Spectrum Disorder Doing Academically at School? An Overview of the Literature. *Autism* 20: 276–94. https://doi.org/10.1177/1362361315580962.

Kjellmer, Liselotte, Elisabeth Fernell, Christopher Gillberg, and Fritjof Norrelgen. 2018. Speech and Language Profiles in 4- to 6-Year-Old Children with Early Diagnosis of Autism Spectrum Disorder without Intellectual Disability. *Neuropsychiatric Disease and Treatment* 14: 2415–27. https://doi.org/10.2147/NDT.S171971.

Konstantareas, M. Mary, and Kelly Stewart. 2006. Affect Regulation and Temperament in Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders* 36: 143–54. https://doi.org/10.1007/s10803-005-0051-4.

Kurth, Jennifer A., and Ann M. Mastergeorge. 2010. Academic and Cognitive Profiles of Students with Autism: Implications for Classroom Practice and Placement. *International Journal of Special Education* 25: 8–14.

Kwok, Elaine Y. L., Heather M. Brown, Rachael E. Smyth, and Janis Oram Cardy. 2015. Meta-Analysis of Receptive and Expressive Language Skills in Autism Spectrum Disorder. *Research in Autism Spectrum Disorders* 9: 202–22. https://doi.org/10.1016/j.rasd.2014.10.008.

Leung, Florence Yik Nam, Jacqueline Sin, Caitlin Dawson, Jia Hoong Ong, Chen Zhao, Anamarija Veić, and Fang Liu. 2022. Emotion Recognition across Visual and Auditory Modalities in Autism Spectrum Disorder: A Systematic Review and Meta-Analysis. *Developmental Review* 63: 101000. https://doi.org/10.1016/j.dr.2021.101000.

Leyfer, Ovsanna T., Susan E. Folstein, Susan Bacalman, Naomi O. Davis, Elena Dinh, Jubel Morgan, Helen Tager-Flusberg, and Janet E. Lainhart. 2006. Comorbid Psychiatric Disorders in Children with Autism: Interview Development and Rates of Disorders. *Journal of Autism and Developmental Disorders* 36: 849–61. https://doi.org/10.1007/s10803-006-0123-0.

Libertus, Klaus, Kelly A. Sheperd, Samuel W. Ross, and Rebecca J. Landa. 2014. Limited Fine Motor and Grasping Skills in 6-Month-Old Infants at High Risk for Autism. *Child Development* 85: 2218–31. https://doi.org/10.1111/cdev.12262.

Lienert, Gustav A., and Ulrich Raatz. 1998. *Testaufbau und Testanalyse [Test Design and Test Analysis]*. Weinheim: Beltz.

Liu, Ting, and Casey M. Breslin. 2013. Fine and Gross Motor Performance of the MABC-2 by Children with Autism Spectrum Disorder and Typically Developing Children. *Research in Autism Spectrum Disorders* 7: 1244–49. https://doi.org/10.1016/j.rasd.2013.07.002.

Lodi-Smith, Jennifer, Jonathan D. Rodgers, Sara A. Cunningham, Christopher Lopata, and Marcus L. Thomeer. 2019. Meta-Analysis of Big Five Personality Traits in Autism Spectrum Disorder. *Autism* 23: 556–65. https://doi.org/10.1177/1362361318766571.

Luyster, Rhiannon, Kristina Lopez, and Catherine Lord. 2007. Characterizing Communicative Development in Children Referred for Autism Spectrum Disorders Using the MacArthur-Bates Communicative Development Inventory (CDI). *Journal of Child Language* 34: 623–54. https://doi.org/10.1017/S0305000907008094.

MacDonald, Megan, Catherine Lord, and Dale A. Ulrich. 2013a. The Relationship of Motor Skills and Adaptive Behavior Skills in Young Children with Autism Spectrum Disorders. *Research in Autism Spectrum Disorders* 7: 1383–90. https://doi.org/10.1016/j.rasd.2013.07.020.

MacDonald, Megan, Catherine Lord, and Dale A. Ulrich. 2013b. The Relationship of Motor Skills and Social Communicative Skills in School-Aged Children with Autism Spectrum Disorder. *Adapted Physical Activity Quarterly* 30: 271–82. https://doi.org/10.1123/apaq.30.3.271.

Maenner, Matthew J., Kelly A. Shaw, Jon Baio, Anita Washington, Mary Patrick, Monica DiRienzo, Deborah L. Christensen, Lisa D. Wiggins, Sydney Pettygrove, Jennifer G. Andrews, and et al. 2020. Prevalence of Autism Spectrum Disorder among Children Aged 8 Years—Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2016. *MMWR Surveillance Summaries* 69: 1–12. https://doi.org/10.15585/mmwr.ss6904a1.

Mandelbaum, David E., Michael Stevens, Eric Rosenberg, Max Wiznitzer, Mitchell Steinschneider, Saul R. Korey, Pauline Filipek, Isabelle Rapin, and Saul R. Korey. 2006. Sensorimotor Performance in School-Age Children with Autism, Developmental Language Disorder, or Low IQ. *Developmental Medicine & Child Neurology* 48: 33–39. https://doi.org/10.1017/S0012162206000089.

Manicolo, Olivia, Mark Brotzmann, Priska Hagmann-von Arx, Alexander Grob, and Peter Weber. 2019. Gait in Children with Infantile/Atypical Autism: Age-Dependent Decrease in Gait Variability and Associations with Motor Skills. *European Journal of Paediatric Neurology* 23: 117–25. https://doi.org/10.1016/j.ejpn.2018.09.011.

Mayes, Susan Dickerson, and Susan L. Calhoun. 2003a. Analysis of WISC-III, Stanford-Binet:IV, and Academic Achievement Test Scores in Children with Autism. *Journal of Autism and Developmental Disorders* 33: 329–41.

Mayes, Susan Dickerson, and Susan L. Calhoun. 2003b. Ability Profiles in Children with Autism: Influence of Age and IQ. *Autism* 7: 65–80. https://doi.org/10.1177/1362361303007001006.

Mazefsky, Carla A., Xenia Borue, Taylor N. Day, and Nancy J. Minshew. 2014. Emotion Regulation Patterns in Adolescents with High-Functioning Autism Spectrum Disorder: Comparison to Typically Developing Adolescents and Association with Psychiatric Symptoms. *Autism Research* 7: 344–54. https://doi.org/10.1002/aur.1366.

Meyer, Christine Sandra, Priska Hagmann-von Arx, and Alexander Grob. 2009. Die Intelligence and Development Scale Sozial-Emotionale Kompetenz (IDS-SEK): Psychometrische Eigenschaften eines Tests zur Erfassung sozial-emotionaler Fähigkeiten [The Intelligence and Development Scale Social-Emotional Competence (IDS-SEK): Psychometric properties of a test to assess social-emotional skills]. *Diagnostica* 55: 234–44. https://doi.org/10.1026/0012-1924.55.4.234.

Ming, Xue, Michael Brimacombe, and George C. Wagner. 2007. Prevalence of Motor Impairment in Autism Spectrum Disorders. *Brain and Development* 29: 565–70. https://doi.org/10.1016/j.braindev.2007.03.002.

Mitchell, Shelley, Janis Oram Cardy, and Lonnie Zwaigenbaum. 2011. Differentiating Autism Spectrum Disorder from Other Developmental Delays In The First Two Years Of Life. *Developmental Disabilities Research Reviews* 17: 130–40. https://doi.org/10.1002/ddrr.1107.

Mitchell, Shelley, Jessica Brian, Lonnie Zwaigenbaum, Wendy Roberts, Peter Szatmari, Isabel Smith, and Susan Bryson. 2006. Early Language and Communication Development of Infants Later Diagnosed with Autism Spectrum Disorder. *Journal of Developmental & Behavioral Pediatrics* 27: S69–S78.

Miyake, Akira, and Naomi P. Friedman. 2012. The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current Directions in Psychological Science* 21: 8–14. https://doi.org/10.1177/0963721411429458.

Miyake, Akira, Naomi P. Friedman, Michael J. Emerson, Alexander H. Witzki, Amy Howerter, and Tor D. Wager. 2000. The Unity and Diversity of Executive Functions and Their Contributions to Complex 'Frontal Lobe' Tasks: A Latent Variable Analysis. *Cognitive Psychology* 41: 49–100. https://doi.org/10.1006/cogp.1999.0734.

Mostofsky, Stewart H., Melanie P. Burgess, and Jennifer C. Gidley Larson. 2007. Increased Motor Cortex White Matter Volume Predicts Motor Impairment in Autism. *Brain* 130: 2117–22. https://doi.org/10.1093/brain/awm129.

Mostofsky, Stewart H., Stephanie K. Powell, Daniel J. Simmonds, Melissa C. Goldberg, Brian Caffo, and James J. Pekar. 2009. Decreased Connectivity and Cerebellar Activity in Autism during Motor Task Performance. *Brain* 132: 2413–25. https://doi.org/10.1093/brain/awp088.

Muth, Anne, Johannes Hönekopp, and Christine M. Falter. 2014. Visuo-Spatial Performance in Autism: A Meta-Analysis. *Journal of Autism and Developmental Disorders* 44: 3245–63. https://doi.org/10.1007/s10803-014-2188-5.

Nader, Anne-Marie, Valérie Courchesne, Michelle Dawson, and Isabelle Soulières. 2016. Does WISC-IV Underestimate the Intelligence of Autistic Children? *Journal of Autism and Developmental Disorders* 46: 1582–89. https://doi.org/10.1007/s10803-014-2270-z.

Nagy, Emese, Louise Prentice, and Tess Wakeling. 2021. Atypical Facial Emotion Recognition in Children with Autism Spectrum Disorders: Exploratory Analysis on the Role of Task Demands. *Perception* 50: 819–33. https://doi.org/10.1177/03010066211038154.

Nebel, Mary Beth, Suresh E. Joel, John Muschelli, Anita D. Barber, Brian S. Caffo, James J. Pekar, and Stewart H. Mostofsky. 2014. Disruption of Functional Organization within the Primary Motor Cortex in Children with Autism. *Human Brain Mapping* 35: 567–80. https://doi.org/10.1002/hbm.22188.

Oliveira, Katherine Simone Caires, Déborah Ebert Fontes, Egmar Longo, Hércules Ribeiro Leite, and Ana Cristina Resende Camargos. 2021. Motor Skills Are Associated with Participation of Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders* 21: 1–10. https://doi.org/10.1007/s10803-021-05318-9.

Oliveras-Rentas, Rafael E., Lauren Kenworthy, Richard B. Roberson, Alex Martin, and Gregory L. Wallace. 2012. WISC-IV Profile in High-Functioning Autism Spectrum Disorders: Impaired Processing Speed Is Associated with Increased Autism Communication Symptoms and Decreased Adaptive Communication Abilities. *Journal of Autism and Developmental Disorders* 42: 655–64. https://doi.org/10.1007/s10803-011-1289-7.

Pennington, Bruce F., and Sally Ozonoff. 1996. Executive Functions and Developmental Psychopathology. *Journal of Child Psychology and Psychiatry* 37: 51–87. https://doi.org/10.1111/j.1469-7610.1996.tb01380.x.

Petermann, Franz. 2008. *Movement Assessment Battery for Children*, 2nd ed. Frankfurt: Pearson Assessment.

Provost, Beth, Brian R. Lopez, and Sandra Heimerl. 2007. A Comparison of Motor Delays in Young Children: Autism Spectrum Disorder, Developmental Delay, and Developmental Concerns. *Journal of Autism and Developmental Disorders* 37: 321–28. https://doi.org/10.1007/s10803-006-0170-6.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing (Version 4.0.3) [Computer Software]*. Vienna: R Foundation for Statistical Computing. 2021. Available online: https://www.R-project.org/ (accessed on 13 February 2021).

Rosset, Delphine B., Cécilie Rondan, David Da Fonseca, Andreia Santos, Brigitte Assouline, and Christine Deruelle. 2008. Typical Emotion Processing for Cartoon but Not for Real Faces in Children with Autistic Spectrum Disorders. *Journal of Autism and Developmental Disorders* 38: 919–25. https://doi.org/10.1007/s10803-007-0465-2.

Russell, Ginny, William Mandy, Daisy Elliott, Rhianna White, Tom Pittwood, and Tamsin Ford. 2019. Selection Bias on Intellectual Ability in Autism Research: A Cross-Sectional Review and Meta-Analysis. *Molecular Autism* 10: 9. https://doi.org/10.1186/s13229-019-0260-x.

Salazar, Fernando, Gillian Baird, Susie Chandler, Evelin Tseng, Tony O'sullivan, Patricia Howlin, Andrew Pickles, and Emily Simonoff. 2015. Co-Occurring Psychiatric Disorders in Preschool and Elementary School-Aged Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders* 45: 2283–94. https://doi.org/10.1007/s10803-015-2361-5.

Salomone, Erica, Daniela Bulgarelli, Evelyne Thommen, Emanuelle Rossini, and Paola Molina. 2019. Role of Age and IQ in Emotion Understanding in Autism Spectrum Disorder: Implications for Educational Interventions. *European Journal of Special Needs Education* 34: 383–92. https://doi.org/10.1080/08856257.2018.1451292.

Schmidt-Atzert, Lothar, and Manfred Amelang. 2012. *Psychologische Diagnostik [Psychological Assessment]*, 5th ed. Berlin: Springer.

Schneider, W. Joel, and Kevin S. McGrew. 2018. The Cattell–Horn–Carroll Theory of Cognitive Abilities. In *Contemporary Intellectual Assessment: Theories, Tests, and Issues*, 4th ed. Edited by Dawn P. Flanagan and Erin M. McDonough. New York: Guilford Press, pp. 73–163.

Shanok, Nathaniel A., Nancy Aaron Jones, and Nikola N. Lucas. 2019. The Nature of Facial Emotion Recognition Impairments in Children on the Autism Spectrum. *Child Psychiatry & Human Development* 50: 661–67. https://doi.org/10.1007/s10578-019-00870-z.

Siaperas, Panagiotis, Howard A. Ring, Catherine J. McAllister, Sheila Henderson, Anna Barnett, Peter Watson, and Anthony J. Holland. 2012. Atypical Movement Performance and Sensory Integration in Asperger's Syndrome. *Journal of Autism and Developmental Disorders* 42: 718–25. https://doi.org/10.1007/s10803-011-1301-2.

Thomas, Pauline, Walter Zahorodny, Bo Peng, Soyeon Kim, Nisha Jani, William Halperin, and Michael Brimacombe. 2012. The Association of Autism Diagnosis with Socioeconomic Status. *Autism* 16: 201–13. https://doi.org/10.1177/1362361311413397.

Titeca, Daisy, Herbert Roeyers, and Annemie Desoete. 2017. Early Numerical Competencies in 4- and 5-Year-Old Children with Autism Spectrum. *Focus on Autism and Other Developmental Disabilities* 32: 279–92.

Trevisan, Dominic A., and Elina Birmingham. 2016. Are Emotion Recognition Abilities Related to Everyday Social Functioning in ASD? A Meta-Analysis. *Research in Autism Spectrum Disorders* 32: 24–42. https://doi.org/10.1016/j.rasd.2016.08.004.

Troyb, Eva, Alyssa Orinstein, Katherine Tyson, Molly Helt, Inge-Marie Eigsti, Michael Stevens, and Deborah Fein. 2014. Academic Abilities in Children and Adolescents with a History of Autism

Spectrum Disorders Who Have Achieved Optimal Outcomes. *Autism* 18: 233–43. https://doi.org/10.1177/1362361312473519.

Van Meter, Karla C., Lasse E. Christiansen, Lora D. Delwiche, Rahman Azari, Tim E. Carpenter, and Irva Hertz-Picciotto. 2010. Geographic Distribution of Autism in California: A Retrospective Birth Cohort Analysis. *Autism Research* 3: 19–29. https://doi.org/10.1002/aur.110.

Volden, Joanne, Isabel M. Smith, Peter Szatmari, Susan Bryson, Eric Fombonne, Pat Mirenda, Wendy Roberts, Tracy Vaillancourt, Charlotte Waddell, Lonnie Zwaigenbaum, and et al. 2011. Using the Preschool Language Scale, Fourth Edition to Characterize Language in Preschoolers with Autism Spectrum Disorders. *American Journal of Speech-Language Pathology* 20: 200–208. https://doi.org/10.1044/1058-0360(2011/10-0035).

White, Susan Williams, Lawrence Scahill, Ami Klin, Kathleen Koenig, and Fred R. Volkmar. 2007. Educational Placements and Service Use Patterns of Individuals with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders* 37: 1403–12. https://doi.org/10.1007/s10803-006-0281-0.

World Health Organization. 2016. *International Statistical Classification of Diseases and Related Health Problems (10th Rev.)*. Geneva: World Health Organization. Available online: https://icd.who.int/browse10/2019/en (accessed on 5 November 2022).

World Health Organization. 2018. *International Classification of Diseases for Mortality and Morbidity Statistics (11th Rev.)*. Geneva: World Health Organization. Available online: https://icd.who.int/browse11/l-m/en (accessed on 6 January 2022).

Yerys, Benjamin E., Gregory L. Wallace, Jennifer L. Sokoloff, Devon A. Shook, Joette D. James, and Lauren Kenworthy. 2009. Attention Deficit/Hyperactivity Disorder Symptoms Moderate Cognition and Behavior in Children with Autism Spectrum Disorders. *Autism Research* 2: 322–33. https://doi.org/10.1002/aur.103.

Yeung, Michael K. 2022. A Systematic Review and Meta-Analysis of Facial Emotion Recognition in Autism Spectrum Disorder: The Specificity of Deficits and the Role of Task Characteristics. *Neuroscience & Biobehavioral Reviews* 133: 104518. https://doi.org/10.1016/j.neubiorev.2021.104518.

Zajic, Matthew C., Emily J. Solari, Ryan P. Grimm, Nancy S. McIntyre, and Peter C. Mundy. 2020. Relationships between Reading Profiles and Narrative Writing Abilities in School-Age Children with Autism Spectrum Disorder. *Reading and Writing* 33: 1531–56. https://doi.org/10.1007/s11145-020-10015-7.

Table 1
Summary of our Hypotheses

Domain	Assumed Differences in Performance between Autistic and Non-Autistic Participants	Assumed Similar Performance in Autistic and Non-Autistic Participants
	Variable	Variable
Intelligence	Composite scores (Profile IQ, Full-Scale IQ, Screening IQ)	Visual Processing
	Processing Speed	Shape Design
	Parrots	Washer Design
	Boxes	Abstract Reasoning
	Auditory Short-Term Memory	Matrices: Completion
	Digit and Letter Span	Matrices: Odd One Out
	Mixed Digit and Letter Span	
	Visuospatial Short-Term Memory	
	Shape Memory	
	Rotated Shape Memory	
	Verbal Reasoning	
	Naming Categories	
	Naming Opposites	
	Long-Term Memory	
	Story Recall	
	Picture Recall	
Executive functions	Composite score	
	Listing Words	
	Divided Attention	
	Animal Colors	
	Drawing Routes	
Psychomotor skills	Composite score	
	Gross Motor Skills	
	Fine Motor Skills	
	Visuomotor Skills	
Social-emotional skills	Composite score	
	Identifying Emotions	
	Regulating Emotions	
	Socially Competent Behavior	
Basic skills	Language skills	Composite score
	Phoneme Analysis	Logical-Mathematical Reasoning
	Phoneme–Grapheme Correspondence	Reading
	Language Expressive	Reading Words
	Language Receptive	Reading Pseudo Words
	Text Comprehension	Spelling
Motivation and attitude	Composite score	
	Conscientiousness	
	Achievement Motivation	
Participation during testing	intelligence	
	executive functions	
	developmental functions	

Note. Differences in performance between autistic and non-autistic participants are interpreted as meaningful if the *p* value is significant after Hommel's correction and the effect size is at least small.

Table 2

Demographic Characteristics of Autistic and Non-Autistic Children and Adolescents

Characteristic	Autistic Sample n = 43		Non-Autistic Sample n = 43		χ^2	p
	n	%	n	%		
Sex					11.33	1.000
Female	8	19	8	19		
Male	35	81	35	81		
Maternal education					14.24	1.000
No postsecondary education	23	54	23	54		
Compulsory school	1	2	2	5		
Apprenticeship	16	37	15	35		
High school	1	2	1	2		
Higher vocational education	5	12	5	12		
Postsecondary education (university degree)	19	44	19	44		
Other	0	0	0	0		
Unknown	1	2	1	2		
Participants' current education					7.00	1.000
Kindergarten	0	0	1	2		
Elementary school	14	33	20	47		
Secondary school	10	23	11	26		
School for special education	8	19	1	2		
High school	6	14	5	12		
Apprenticeship	3	7	4	9		
University	0	0	1	2		
None	2	5	0	0		
Intelligence level					11.09	1.000
<70	9	21	1	2		
70–84	7	16	6	14		
85–99	8	19	16	37		
100–114	11	26	14	33		
≥115	6	14	6	14		
Comorbid condition					12.15	1.000
Visual impairment	6	14	8	19		
Motor problems	4	9	0	0		
Speech problems	4	9	1	2		
Dyslexia	2	5	2	5		
Dyscalculia	0	0	2	5		
AD(H)D	4	9	4	9		
Depression	1	2	1	2		
Medical problems	10	23	2	5		
Ethnicity					10.04	1.000
German-speaking country	38	88	39	91		
Other European country	4	9	2	5		
Non-European country	1	2	1	2		
Unknown	0	0	1	2		
Native language					10.99	1.000
Monolingual German	32	74	35	81		
Bilingual	6	14	6	14		
Other language than German	5	12	2	5		

Note. Samples were matched for age, sex, and maternal education (as a proxy for socioeconomic status). Autistic sample: M_{age} = 12.3 years, SD = 3.08; non-autistic sample: M_{age} = 12.51 years, SD = 3.56. Paired-sample t test for age: t = 0.34, p = .733. χ^2 test for sex (0 = male, 1 = female), maternal education (0 = no postsecondary education, 1 = postsecondary education), participants' current education (0 = no special education, 1 = special education), intelligence level (0 = average, 1 = below/above average), comorbid condition (0 = no, 1 = yes), ethnicity (0 = German-speaking country, 1 = other), and native language (0 = monolingual, 1 = not monolingual). AD(H)D = attention deficit/hyperactivity disorder or attention deficit disorder.

Table 3

Reliabilities, Means, Standard Deviations, and t tests of the Cognitive Functions From the Intelligence and Development Scales–2 for Autistic and Non-Autistic Children and Adolescents

Variable	Autistic Sample $n = 43$				Non-Autistic Sample $n = 43$				*t*	*df*	*p*	p_H	*d*
	Rel	*M*	*SD*	Range	Rel	*M*	*SD*	Range					
Profile IQ[b]	.99	90.16	19.98	55–131	.99	97.68	13.82	61–121	1.96	77	.027	.406	0.44
Full-Scale IQ[b]	.99	91.58	20.77	55–129	.98	97.63	13.60	63–120	1.58	81	.059	.627	0.35
Screening IQ[b]	.98	93.54	19.35	55–125	.98	100.58	16.45	61–134	1.80	82	.038	.490	0.39
Visual Processing[b]	.99	97.03	21.50	55–129	.97	102.67	12.50	80–129	1.46	80	.148	.809	0.32
Processing Speed[b]	.98	95.58	20.47	55–143	.98	100.19	15.57	56–126	1.15	80	.126	.758	0.25
Auditory Short-Term Memory[b]	.97	90.77	16.95	55–139	.96	97.76	12.54	64–121	2.12	79	.019	.352	0.47
Visuospatial Short-Term Memory[b]	.97	88.92	15.12	55–118	.94	96.79	10.89	77–118	2.70	79	.004	.161	0.60
Abstract Reasoning[b]	.98	95.10	20.67	55–141	.97	97.55	14.93	63–122	0.62	80	.539	.846	0.14
Verbal Reasoning[b]	.99	94.32	19.29	58–126	.97	99.98	15.29	61–131	1.48	81	.071	.674	0.33
Long-Term Memory[b]	.97	88.08	15.82	55–113	.97	93.64	16.06	58–137	1.57	79	.060	.627	0.35
Shape Design[a]	.95	9.62	4.24	1–16	.89	10.65	2.55	7–16	1.36	83	.176	.846	0.30
Washer Design[a]	.94	9.54	3.83	1–17	.92	10.57	2.78	4–19	1.41	81	.162	.811	0.31
Parrots[a]	.92	9.15	4.28	1–19	.91	9.74	2.94	1–17	0.75	82	.228	.846	0.16
Boxes[a]	.93	9.28	3.29	1–16	.90	10.40	3.36	2–17	1.54	80	.064	.642	0.34
Digit and Letter Span[a]	.90	9.02	3.67	1–18	.82	9.95	2.17	5–14	1.42	83	.079	.693	0.31
Mixed Digit and Letter Span[a]	.86	8.49	3.19	1–18	.84	10.21	2.99	1–17	2.51	79	.007	.231	0.56
Shape Memory[a]	.88	8.41	2.99	1–14	.78	9.28	2.36	5–16	1.47	82	.072	.674	0.32
Rotated Shape Memory[a]	.90	8.36	3.06	2–17	.82	10.12	2.63	6–18	2.78	79	.003	.132	0.62
Matrices: Completion[a]	.93	9.32	3.63	3–18	.90	10.70	3.07	4–17	1.89	82	.063	.628	0.41
Matrices: Odd One Out[a]	.93	9.72	3.82	2–18	.86	9.14	2.93	2–15	−0.78	80	.440	.846	0.17
Naming Categories[a]	.95	9.38	3.92	1–16	.92	10.33	3.49	2–18	1.17	83	.122	.755	0.25
Naming Opposites[a]	.92	9.41	3.49	1–16	.87	10.50	2.90	3–19	1.54	81	.063	.633	0.34
Story Recall[a]	.93	8.40	3.56	1–14	.88	9.42	3.17	1–16	1.39	82	.084	.693	0.30
Picture Recall[a]	.85	8.05	2.73	3–14	.88	8.90	3.32	3–18	1.27	80	.104	.726	0.28
Executive functions composite score[b]	.97	8.84	2.20	4–13	.96	9.95	1.98	6–15	2.27	71	.013	.317	0.53
Listing Words[c]	.75	7.89	3.13	1–14	.75	9.55	2.93	4–17	2.38	73	.010	.292	0.55
Divided Attention[b]	.92	8.76	2.93	4–15	.90	10.11	2.52	5–17	2.13	71	.019	.352	0.50
Animal Colors[c]	.72	8.09	3.55	1–14	.72	9.47	3.44	3–19	1.70	72	.047	.547	0.40
Drawing Routes[b]	.96	9.91	2.60	5–15	.94	10.41	2.41	5–15	0.88	74	.192	.846	0.20

Note. Samples were matched for age, sex, and maternal education (as a proxy for socioeconomic status). p_H indicates *p* values adjusted with Hommel's (1988) correction. Please note that after this correction, none of the comparisons were significant. Rel indicate reliabilities. The following reliabilities are reported: [a]Cronbach's alpha, [b]reliability calculated according to a formula by Lienert and Raatz (1998), or [c]retest reliability.

Table 4

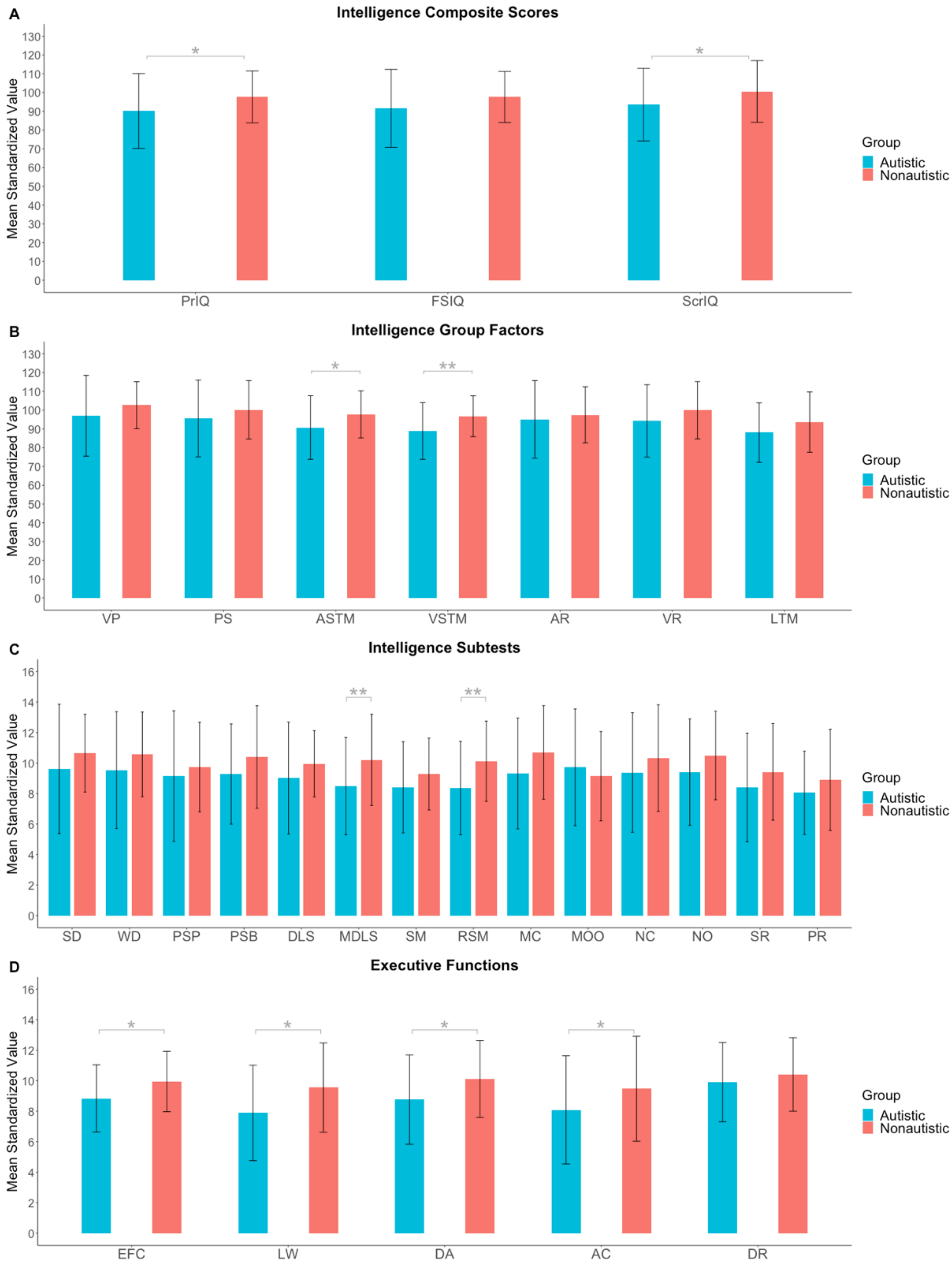
Reliabilities, Means, Standard Deviations, and t tests of the Developmental Functions From the Intelligence and Development Scales–2 for Autistic and Non-Autistic Children and Adolescents

Variable	Autistic Sample *n* = 43				Non-Autistic Sample *n* = 43				*t*	*df*	*p*	p_H	*d*
	Rel	*M*	*SD*	Range	Rel	*M*	*SD*	Range					
Psychomotor skills composite score[b]	.98	8.49	2.22	4–12	.95	10.43	1.57	7–15	4.60	81	< .001	**< .001**	**1.01**
Gross Motor Skills[a]	.72	5.29	3.57	1–11	.77	11.35	3.08	5–15	5.30	32	< .001	**< .001**	**1.82**
Fine Motor Skills[b]	.96	8.65	3.03	2–14	.96	10.59	2.39	4–16	3.20	79	< .001	**.046**	**0.71**
Visuomotor Skills[b]	.95	8.79	1.97	4–13	.87	10.01	1.73	7–13	3.01	81	.002	.077	0.66
Social-emotional skills composite score[b]	.96	8.32	2.86	1–13	.95	9.82	2.15	5–13	2.71	82	.004	.154	0.59
Identifying Emotions[c]	.85	7.71	4.27	1–12	.85	10.24	2.68	4–12	2.07	32	.023	.388	0.71
Regulating Emotions[c]	.78	8.40	3.39	1–13	.78	10.00	2.75	4–15	2.37	82	.010	.292	0.52
Socially Competent Behavior[c]	.71	8.32	3.17	1–15	.71	9.81	2.70	5–15	2.29	80	.012	.311	0.51
Basic skills composite score[b]	.99	9.77	2.53	2–14	.99	9.92	2.15	5–13	0.29	74	.772	.846	0.07
Logical-Mathematical Reasoning[a]	.98	9.07	4.21	1–17	.96	10.24	3.07	3–16	1.44	81	.153	.809	0.32
Language Skills[b]	.98	7.12	2.65	3–14	.98	10.50	1.86	6–13	4.11	28	< .001	**.008**	**1.51**
Phoneme Analysis[a]	.92	6.07	3.00	1–13	.97	10.24	3.13	3–15	3.75	29	< .001	**.019**	**1.35**
Phoneme–Grapheme Correspondence[a]	.86	9.07	4.20	1–15	.96	10.00	2.18	5–13	0.79	29	.217	.846	0.29
Language Expressive[a]	.81	7.08	2.96	1–14	.89	10.41	2.55	5–15	3.31	28	.001	.058	1.22
Language Receptive[a]	.85	6.71	3.34	2–14	.81	11.35	2.37	7–16	4.52	29	< .001	**.003**	**1.63**
Reading[b]	.98	8.76	3.33	1–15	.95	9.65	2.40	5–14	1.35	74	.182	.846	0.31
Reading Words[c]	.79	9.34	3.41	2–16	.79	9.47	2.56	5–14	0.20	76	.846	.846	0.04
Reading Pseudo Words[c]	.67	8.89	2.89	2–14	.67	9.82	2.48	4–14	1.52	75	.132	.792	0.35
Text Comprehension[a]	.69	9.40	5.02	1–16	.69	10.37	2.79	4–16	1.00	68	.160	.809	0.24
Spelling[a]	.88	8.89	3.19	3–15	.88	9.79	2.66	4–15	1.26	65	.212	.846	0.31
Motivation and attitude composite score[b]	.96	10.56	3.24	6–17	.96	10.65	2.78	6–19	0.11	46	.458	.846	0.03
Conscientiousness[a]	.82	10.21	3.27	6–18	.79	10.26	2.85	6–19	0.06	45	.477	.846	0.02
Achievement Motivation[a]	.87	11.12	3.96	4–19	.86	11.04	3.11	6–19	−0.07	47	.528	.846	0.02
Participation during testing, intelligence[a]	.93	8.19	3.15	1–16	.93	10.17	3.57	1–16	2.68	81	.004	.169	0.59
Participation during testing, executive functions[a]	.89	8.76	2.75	1–16	.91	10.17	2.89	4–16	2.13	71	.018	.351	0.50
Participation during testing, developmental functions[a]	.95	8.33	3.31	1–16	.92	10.66	3.02	5–16	3.30	79	< .001	**.035**	**0.73**

Note. Samples were matched for age, sex, and maternal education (as a proxy for socioeconomic status). p_H indicates *p* values adjusted with Hommel's (1988) correction. Significant results after accounting for multiple testing (Hommel correction) are presented in bold. Rel indicate reliabilities. The following reliabilities are reported: [a]Cronbach's alpha, [b]reliability calculated according to a formula by Lienert and Raatz (1998), or [c]retest reliability.

Figure 1

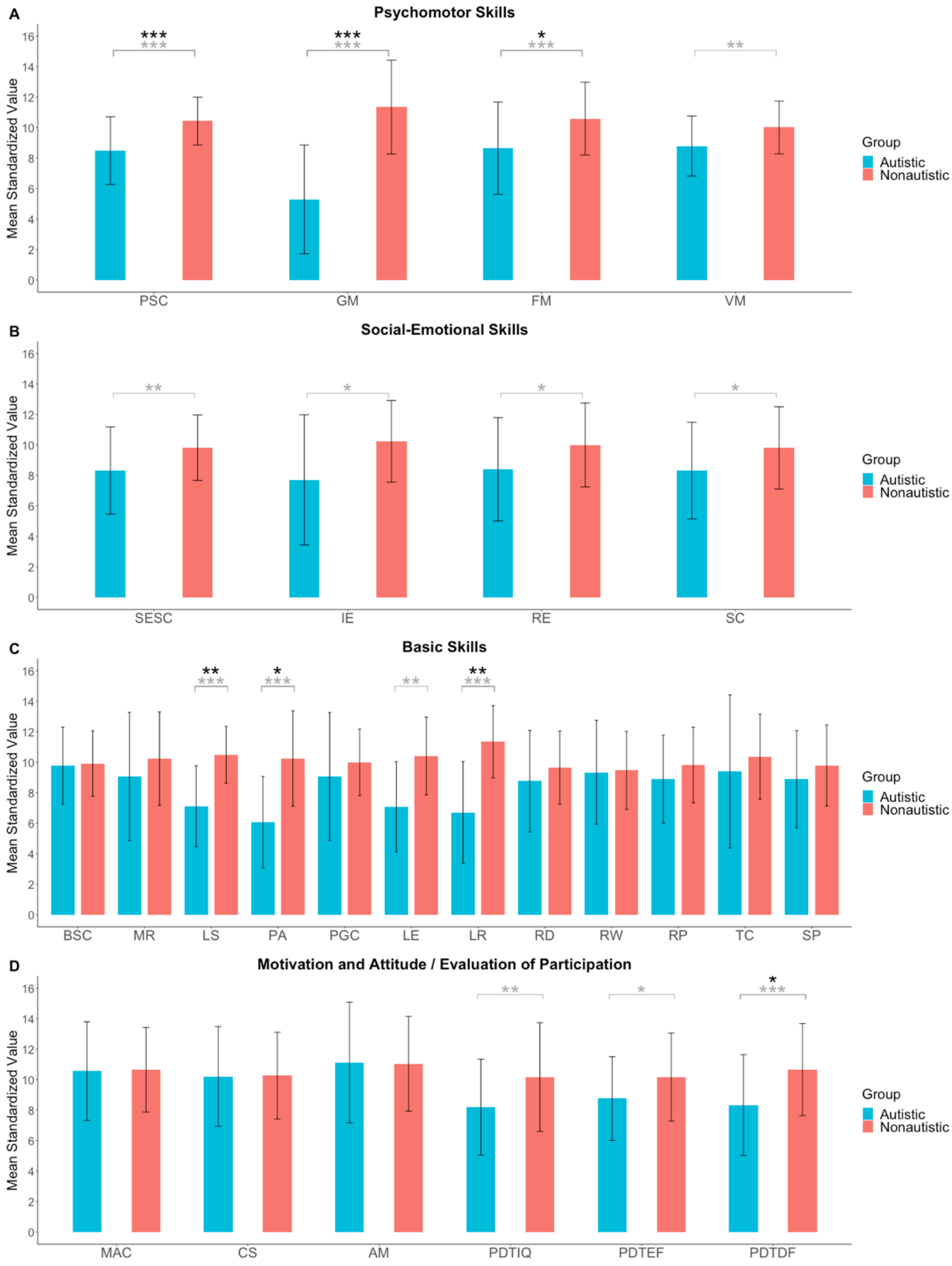
Means and Standard Deviations of the Cognitive Functions From the Intelligence and Development Scales–2 for Autistic and Non-Autistic Children and Adolescents



Note. Means and standard deviations are reported for (**A**) intelligence composite scores, (**B**) intelligence group factors, (**C**) intelligence subtests, and (**D**) executive functions composite score and subtests of the Intelligence and Development Scales–2 for autistic and non-autistic children and adolescents. Asterisks in grey indicate *p* values not adjusted with Hommel's (1988) correction. Asterisks in black indicate *p* values adjusted according to Hommel (1988). Please note that after this correction, none of the comparisons were significant and therefore, no black asterisks are included in the present graphs. PrIQ = Profile IQ; FSIQ = Full-Scale IQ; ScrIQ = Screening IQ; VP = Visual Processing; PS = Processing Speed; ASTM = Auditory Short-Term Memory; VSTM = Visuospatial Short-Term Memory; AR = Abstract Reasoning; VR = Verbal Reasoning; LTM = Long-Term Memory; SD = Shape Design; WD = Washer Design; PSP = Parrots; PSB = Boxes; DLS = Digit and Letter Span; MDLS = Mixed Digit and Letter Span; SM = Shape Memory; RSM = Rotated Shape Memory; MC = Matrices: Completion; MOO = Matrices: Odd One Out; NC = Naming Categories; NO = Naming Opposites; SR = Story Recall; PR = Picture Recall; EFC = Executive functions composite score; LW = Listing Words; DA = Divided Attention; AC = Animal Colors; DR = Drawing Routes.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Figure 2

Means and Standard Deviations of the Developmental Functions From the Intelligence and Development Scales–2 for Autistic and Non-Autistic Children and Adolescents



Note. Means and standard deviations are reported for (**A**) psychomotor skills composite score and subtests, (**B**) social-emotional skills composite score and subtests, (**C**) basic skills composite score and subtests, and (**D**) motivation and attitude composite score and subtests as well as for the evaluation of participation during testing of the Intelligence and Development Scales–2 for autistic and non-autistic children and adolescents. Asterisks in grey indicate *p* values not adjusted with Hommel's (1988) correction. Asterisks in black indicate *p* values adjusted according to Hommel (1988). PSC = Psychomotor skills composite score; GM = Gross Motor Skills; FM = Fine Motor Skills; VM = Visuomotor Skills; SESC = Social-emotional skills composite score; IE = Identifying Emotions; RE = Regulating Emotions; SC = Socially Competent Behavior; BSC = Basic skills composite score; MR = Logical-Mathematical Reasoning; LS = Language Skills; PA = Phoneme Analysis; PGC = Phoneme–Grapheme Correspondence; LE = Language Expressive; LR = Language Receptive; RD = Reading; RW = Reading Words; RP = Reading Pseudo Words; TC = Text Comprehension; SP = Spelling; MAC = Motivation and attitude composite score; CS = Conscientiousness; AM = Achievement Motivation; PDTIQ = Participation during testing, intelligence; PDTEF = Participation during testing, executive functions; PDTDF = Participation during testing, developmental functions.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Supplementary Material

Table S1

Description of the Composites, Group Factors, and Subtests of the Intelligence and Development Scales–2

Composite	Group factor	Subtest/Task	No. of items/subtests	Description	Measurement of
Profile IQ (PrIQ)			14	Including the subtests SD, WD, PSP, PSB, DLS, MDLS, SM, RSM, MC, MOO, NC, NO, SR, PR	
Full-Scale IQ (FSIQ)			7	Including one subtest from every group factor (i.e., SD, PSP, DLS, SM, MC, NC, SR)	
Screening IQ (ScrIQ)			2	Including one subtest from the group factor AR (i.e., MC) and one subtest from the group factor VR (i.e., NC)	
	Visual Processing (VP)	Shape Design (SD)	20	Reproduce geometric figures with the help of rectangles and triangles	Visual processing
		Washer Design (WD)	2–4[a]	Reproduce counter patterns according to a template	Visual processing
	Processing Speed (PS)	Parrots (PSP)	56–180[a]	Cross out parrots with two orange features that look to the left from rows of different parrots	Processing speed
		Boxes (PSB)	104–180[a]	Cross out groups of three or four boxes from rows of different groups of boxes	Processing speed
	Auditory Short-Term Memory (ASTM)	Digit and Letter Span (DLS)	40	Repeat number and letter sequences forward and backward	Forward: Short-term memory Backward: Working memory
		Mixed Digit and Letter Span (MDLS)	36	Repeat mixed number and letter sequences forward and backward	Forward: Short-term memory Backward: Working memory
	Visuospatial Short-Term Memory (VSTM)	Shape Memory (SM)	23	Remember figures and recognize them from a selection of figures and positions	Short-term memory
		Rotated Shape Memory (RSM)	23	Remember figures and recognize them from a selection of rotated figures and positions	Working memory
	Abstract Reasoning (AR)	Matrices: Completion (MC)	35	Understand how a figure changes and transfer these changes to a continuing figure	Fluid reasoning
		Matrices: Odd One Out (MOO)	31	Select from presented pictures the one that does not fit with the others	Fluid reasoning
	Verbal Reasoning (VR)	Naming Categories (NC)	34	Name categories for a group of pictures or words	Comprehension-knowledge
		Naming Opposites (NO)	34	Name opposites of presented words	Comprehension-knowledge
	Long-Term Memory (LTM)	Story Recall (SR)	19–32[a]	Listen to a semantically meaningful story and recall it after at least 20 min	Long-term memory
		Picture Recall (PR)	11–21[a]	Look at a picture and recall key features and details after at least 20 min	Long-term memory
Executive functions composite score (EFC)			4	Including the subtests LW, DA, AC, DR	
		Listing Words (LW)	2-4[a]	List words based on categories or starting letters	Word fluency
		Divided Attention (DA)	50-100[a]	Cross out parrots with two orange features that look to the left from different parrots and list animals at the same time	Cognitive flexibility
		Animal Colors (AC)	3	Say colors of animals as fast as possible	Inhibition
		Drawing Routes (DR)	14	Travel given routes as fast as possible once	Planning

Composite	Group factor	Subtest/Task	No. of items/subtests	Description	Measurement of
Psychomotor skills composite score (PSC)			3	Including the subtests GM, FM, VM	
		Gross Motor Skills (GM)	3	Balance on a rope, catch and throw a ball, and jump sideways over a rope	Gross motor skills
		Fine Motor Skills (FM)	6	Quickly screw nuts on and off bolts of different sizes and quickly thread beads of different sizes	Fine motor skills
		Visuomotor Skills (VM)	12	Move exactly between lines, draw figures, and reflect figures	Visuomotor skills
Social-emotional skills composite score (SESC)			3	Including the subtests IE, RE, SC	
		Identifying Emotions (IE)	10	Recognize and name emotions of children in photos	Emotion recognition
		Regulating Emotions (RE)	6-9[a]	Specify regulation strategies for the emotions of anger, fear, and grief	Emotion regulation
		Socially Competent Behavior (SC)	6-9[a]	Name socially competent behavior according to a presented social situation	Socially competent behavior
Basic skills composite score (BSC)			4	Including the subtests MR, LS, RD, SP	
		Logical-Mathematical Reasoning (MR)	64	Solve logical-mathematical reasoning tasks	Mathematical skills
		Language Skills (LS)	4	Including the tasks PA, PGC, LE, LR	Language skills
		Phoneme Analysis (PA)	56	Clap syllables, recognize rhymes, isolate on and off sounds, sound out words	Phonological awareness
		Phoneme–Grapheme Correspondence (PGC)	42	Match phonemes and graphemes, recognize short and long vowels	Knowledge of letter–sound correspondence
		Language Expressive (LE)	12	Form sentences from several words	Language production
		Language Receptive (LR)	15	Carry out instructions	Language comprehension
		Reading (RD)	3	Including RW, RP, TC	Reading
		Reading Words (RW)	60	Read real words	Word recognition
		Reading Pseudo Words (RP)	60	Read pseudo words	Synthetic reading
		Text Comprehension (TC)	14-19[a]	Read and understand texts	Reading comprehension
		Spelling (SP)	40-60[a]	Word dictation	Spelling skills
Motivation and attitude composite score (MAC)			2	Including the subtests CS, AM	
		Conscientiousness (CS)	18	Rate statements regarding conscientiousness	Conscientiousness
		Achievement Motivation (AM)	18	Rate statements regarding achievement motivation	Achievement motivation
		Participation during testing, intelligence (PDTIQ)	12	The test administrator evaluates the participant's cooperation during testing of the intelligence domain	Achievement motivation
		Participation during testing, executive functions (PDTEF)	12	The test administrator evaluates the participant's cooperation during testing of the executive functions domain	Achievement motivation
		Participation during testing, developmental functions (PDTDF)	12	The test administrator evaluates the participant's cooperation during testing of the developmental functions domain	Achievement motivation

Note. Gross Motor Skills, Identifying Emotions, and Language Skills only for ages 5 to 10 years; Reading and Spelling only for ages 7 to 20 years; Conscientiousness and Achievement Motivation only for ages 11 to 20 years. [a]Depending on age and skill.

Table S2

Means, Standard Deviations, and t tests of the Developmental Functions From the Intelligence and Development Scales–2 for Autistic and Non-Autistic Children and Adolescents Matched by Age, Sex, and Intelligence

Variable	Autistic Sample *n* = 43			Non-Autistic Sample *n* = 43			*t*	*df*	*p*	p_H	*d*
	M	*SD*	Range	*M*	*SD*	Range					
Psychomotor skills composite score	8.49	2.22	4–12	10.15	2.03	6–15	3.58	82	< .001	**.007**	**0.78**
Gross Motor Skills	5.29	3.57	1–11	10.31	3.09	4–17	4.31	31	< .001	**.002**	**1.50**
Fine Motor Skills	8.65	3.03	2–14	10.19	2.81	2–17	2.41	81	.009	.166	0.53
Visuomotor Skills	8.79	1.97	4–13	10.16	1.81	7–15	3.29	80	< .001	**.018**	**0.73**
Social-emotional skills composite score	8.32	2.86	1–13	9.70	2.19	4–15	2.50	83	.007	.135	0.54
Identifying Emotions	7.71	4.27	1–12	9.41	3.18	1–12	1.32	32	.098	.884	0.45
Regulating Emotions	8.40	3.39	1–13	9.35	3.01	1–15	1.36	83	.089	.884	0.30
Socially Competent Behavior	8.32	3.17	1–15	9.98	2.92	6–18	2.45	80	.008	.155	0.54
Basic skills composite score	9.77	2.53	2–14	9.01	3.09	1–13	-1.15	75	.254	.999	0.26
Logical-Mathematical Reasoning	9.07	4.21	1–17	9.63	3.86	1–15	0.63	82	.531	.999	0.14
Language Skills	7.12	2.65	3–14	9.01	3.44	2–14	1.65	28	.055	.662	0.61
Phoneme Analysis	6.07	3.00	1–13	9.18	4.05	1–16	2.38	29	.012	.218	0.86
Phoneme–Grapheme Correspondence	9.07	4.20	1–15	9.41	4.05	1–17	0.23	29	.410	.999	0.08
Language Expressive	7.08	2.96	1–14	9.47	4.85	1–16	1.57	28	.064	.762	0.58
Language Receptive	6.71	3.34	2–14	8.00	3.76	1–14	1.00	29	.164	.999	0.36
Reading	8.76	3.33	1–15	8.76	2.95	1–14	0.00	71	.999	.999	0.00
Reading Words	9.34	3.41	2–16	8.46	3.00	1–14	-1.19	73	.238	.999	0.27
Reading Pseudo Words	8.89	2.89	2–14	8.89	3.06	1–16	0.00	72	.999	.999	0.00
Text Comprehension	9.40	5.02	1–16	10.21	3.62	2–16	0.72	62	.236	.999	0.18
Spelling	8.89	3.19	3–15	8.29	3.76	1–14	-0.67	60	.507	.999	0.17
Motivation and attitude composite score	10.56	3.24	6–17	10.20	2.57	5–19	-0.45	51	.674	.999	0.13
Conscientiousness	10.21	3.27	6–18	10.25	2.49	6–19	0.05	50	.479	.999	0.01
Achievement Motivation	11.12	3.96	4–19	10.14	3.25	1–19	-0.99	52	.836	.999	0.27
Participation during testing, intelligence	8.19	3.15	1–16	9.74	3.05	2–16	2.31	83	.012	.210	0.50
Participation during testing, executive functions	8.76	2.75	1–16	9.53	2.83	2–15	1.21	75	.116	.955	0.28
Participation during testing, developmental functions	8.33	3.31	1–16	9.85	3.08	3–17	2.14	80	.018	.315	0.47

Note. Samples were matched for age, sex, and intelligence (Full-Scale IQ). p_H indicates *p* values adjusted with Hommel's (1988) correction. Significant results after accounting for multiple testing (Hommel correction) are presented in bold.

Table S3

Means, Standard Deviations, and t tests of the Cognitive Functions From the Intelligence and Development Scales–2 for Autistic and Non-Autistic Children (Aged 5–10 Years)

Variable	Autistic Children $n = 17$			Non-Autistic Children $n = 17$			t	df	p	p_H	d
	M	*SD*	Range	*M*	*SD*	Range					
Profile IQ	77.69	20.08	55–114	100.93	13.49	74–118	3.64	26	< .001	**.017**	**1.38**
Full-Scale IQ	79.53	20.55	55–119	99.65	14.99	74–120	3.19	30	.002	**.033**	**1.13**
Screening IQ	83.12	18.47	55–122	101.59	16.54	70–134	3.03	31	.002	**.037**	**1.06**
Visual Processing	84.21	27.23	55–129	102.62	11.65	80–117	2.46	28	.020	.115	0.90
Processing Speed	83.53	20.75	55–131	101.62	19.10	56–125	2.53	29	.009	.070	0.91
Auditory Short-Term Memory	81.15	12.53	61–97	96.06	15.13	64–120	2.84	27	.004	**.047**	**1.06**
Visuospatial Short-Term Memory	81.64	15.99	55–105	99.56	11.42	77–117	3.57	28	< .001	**.019**	**1.31**
Abstract Reasoning	84.80	21.49	55–141	98.75	13.78	73–119	2.17	29	.039	.155	0.78
Verbal Reasoning	85.00	17.74	61–123	102.75	14.01	78–131	3.10	29	.002	**.036**	**1.12**
Long-Term Memory	85.50	16.26	55–107	97.62	13.84	67–125	2.21	28	.018	.107	0.81
Shape Design	7.00	5.10	1–16	10.47	2.60	7–16	2.49	31	.019	.111	0.87
Washer Design	8.07	4.67	1–14	10.75	2.70	7–15	1.98	29	.058	.231	0.71
Parrots	6.56	4.02	1–16	9.47	3.68	1–15	2.17	31	.019	.113	0.76
Boxes	7.47	3.76	1–15	11.06	3.62	4–17	2.71	29	.006	.056	0.97
Digit and Letter Span	7.19	3.12	1–14	9.76	2.33	5–14	2.70	31	.006	.056	0.94
Mixed Digit and Letter Span	7.15	2.61	2–10	9.81	3.71	1–16	2.18	27	.019	.115	0.81
Shape Memory	7.38	3.44	1–13	10.29	2.52	5–16	2.79	31	.004	**.049**	**0.97**
Rotated Shape Memory	7.29	2.52	2–11	10.06	1.73	6–12	3.55	28	< .001	**.019**	**1.30**
Matrices: Completion	7.44	3.81	3–18	10.71	3.20	4–17	2.67	31	.012	.092	0.93
Matrices: Odd One Out	8.20	3.78	2–16	9.56	2.73	5–15	1.16	29	.257	.257	0.42
Naming Categories	7.75	3.38	1–16	10.65	3.79	2–18	2.31	31	.014	.096	0.81
Naming Opposites	7.73	3.56	1–13	11.00	2.07	6–14	3.15	29	.002	**.034**	**1.13**
Story Recall	7.15	3.30	1–12	10.53	2.60	4–16	3.28	31	.001	**.029**	**1.14**
Picture Recall	8.57	2.59	4–14	9.31	2.94	5–14	0.73	28	.236	.257	0.27
Executive functions composite score	7.38	1.66	4–9	10.12	2.35	6–15	2.93	21	.004	**.047**	**1.28**
Listing Words	7.00	2.28	4–11	10.14	3.25	5–16	2.72	23	.006	.058	1.09
Divided Attention	7.06	2.01	4–10	10.13	2.93	5–14	2.77	22	.006	.055	1.17
Animal Colors	6.30	3.33	1–12	8.80	3.80	3–15	1.69	23	.052	.210	0.69
Drawing Routes	9.00	2.97	5–14	11.11	1.86	7–14	2.17	23	.020	.115	0.88

Note. Samples were matched for age, sex, and maternal education (as a proxy for socioeconomic status). p_H indicates p values adjusted with Hommel's (1988) correction. Significant results after accounting for multiple testing (Hommel correction) are presented in bold.

Table S4

Means, Standard Deviations, and t tests of the Developmental Functions From the Intelligence and Development Scales–2 for Autistic and Non-Autistic Children (Aged 5–10 Years)

Variable	Autistic Children n = 17			Non-Autistic Children n = 17			t	df	p	p_H	d
	M	SD	Range	M	SD	Range					
Psychomotor skills composite score	7.40	2.42	4–11	10.63	1.58	8–13	4.52	30	< .001	**.002**	**1.60**
Gross Motor Skills	5.29	3.57	1–11	11.35	3.08	5–15	5.30	32	< .001	**< .001**	**1.82**
Fine Motor Skills	8.25	3.10	3–13	10.50	2.73	4–15	2.18	30	.019	.113	0.77
Visuomotor Skills	7.88	2.27	4–12	10.00	1.78	7–13	2.95	30	.003	**.041**	**1.04**
Social-emotional skills composite score	7.70	3.02	2–11	10.67	1.99	6–13	3.38	32	.001	**.024**	**1.16**
Identifying Emotions	7.71	4.27	1–12	10.24	2.68	4–12	2.07	32	.023	.117	0.71
Regulating Emotions	7.53	3.61	1–13	10.94	2.86	5–15	3.06	32	.002	**.036**	**1.05**
Socially Competent Behavior	7.88	3.24	1–12	10.82	2.90	5–15	2.79	32	.004	**.048**	**0.96**
Basic skills composite score	6.99	2.24	2–10	9.91	2.21	5–13	3.19	24	.004	**.047**	**1.32**
Logical-Mathematical Reasoning	5.80	3.05	1–13	9.82	3.36	3–15	3.53	30	.001	**.030**	**1.25**
Language Skills	7.12	2.65	3–14	10.50	1.86	6–13	4.11	28	< .001	**.007**	**1.51**
Phoneme Analysis	6.07	3.00	1–13	10.24	3.13	3–15	3.75	29	< .001	**.013**	**1.35**
Phoneme–Grapheme Correspondence	9.07	4.20	1–15	10.00	2.18	5–13	0.79	29	.217	.257	0.29
Language Expressive	7.08	2.96	1–14	10.41	2.55	5–15	3.31	28	.001	**.029**	**1.22**
Language Receptive	6.71	3.34	2–14	11.35	2.37	7–16	4.52	29	< .001	**.002**	**1.63**
Reading	5.86	3.44	1–12	9.78	2.80	5–14	3.32	26	.003	**.039**	**1.26**
Reading Words	7.07	3.29	2–14	10.33	2.66	6–14	2.94	27	.007	.060	1.09
Reading Pseudo Words	6.93	2.76	2–11	10.00	2.51	4–14	3.14	27	.004	**.047**	**1.17**
Text Comprehension	5.50	5.35	1–15	10.92	2.64	5–14	3.14	22	.002	**.037**	**1.28**
Spelling	5.71	1.98	3–8	9.87	2.72	5–15	3.60	20	.002	**.034**	**1.65**
Motivation and attitude composite score	-	-	-	-	-	-	-	-	-	-	-
Conscientiousness	-	-	-	-	-	-	-	-	-	-	-
Achievement Motivation	-	-	-	-	-	-	-	-	-	-	-
Participation during testing, intelligence	6.69	2.52	2–10	10.12	4.27	1–16	2.79	31	.005	.050	0.97
Participation during testing, executive functions	8.09	1.30	7–10	10.43	3.44	7–16	2.13	23	.022	.115	0.86
Participation during testing, developmental functions	6.59	2.98	1–10	10.57	3.59	5–16	3.38	29	.001	**.025**	**1.22**

Note. Samples were matched for age, sex, and maternal education (as a proxy for socioeconomic status). p_H indicates p values adjusted with Hommel's (1988) correction. Significant results after accounting for multiple testing (Hommel correction) are presented in bold. Conscientiousness and Achievement Motivation only for ages 11 to 20 years.

Table S5

Means, Standard Deviations, and t tests of the Cognitive Functions From the Intelligence and Development Scales–2 for Autistic and Non-Autistic Adolescents (Aged 11–20 Years)

Variable	Autistic Adolescents n = 26			Non-Autistic Adolescents n = 26			t	df	p	p_H	d
	M	SD	Range	M	SD	Range					
Profile IQ	96.64	16.91	55–131	95.81	13.91	61–121	-0.19	49	.576	.926	0.05
Full-Scale IQ	98.80	17.60	55–129	96.31	12.75	63–116	-0.58	49	.718	.926	0.16
Screening IQ	100.20	17.10	55–125	99.92	16.68	61–125	-0.06	49	.523	.926	0.02
Visual Processing	103.92	13.93	77–123	102.69	13.22	84–129	-0.33	50	.745	.926	0.09
Processing Speed	102.80	16.88	70–143	99.31	13.29	73–126	-0.82	49	.793	.926	0.23
Auditory Short-Term Memory	95.58	17.01	55–139	98.81	10.84	74–121	0.82	50	.209	.926	0.23
Visuospatial Short-Term Memory	93.00	13.24	64–118	95.08	10.41	78–118	0.62	49	.268	.926	0.17
Abstract Reasoning	101.28	17.85	59–137	96.81	15.82	63–122	-0.95	49	.348	.926	0.27
Verbal Reasoning	99.69	18.35	58–126	98.27	16.05	61–128	-0.30	50	.616	.926	0.08
Long-Term Memory	89.52	15.71	55–113	91.19	17.08	58–137	0.36	49	.359	.926	0.10
Shape Design	11.23	2.60	7–16	10.77	2.57	7–16	-0.64	50	.522	.926	0.18
Washer Design	10.38	3.05	4–17	10.46	2.87	4–19	0.09	50	.926	.926	0.03
Parrots	10.80	3.63	4–19	9.92	2.42	6–17	-1.02	49	.844	.926	0.29
Boxes	10.36	2.46	5–16	10.00	3.20	2–15	-0.45	49	.672	.926	0.13
Digit and Letter Span	10.15	3.57	1–18	10.08	2.10	5–14	-0.09	50	.538	.926	0.03
Mixed Digit and Letter Span	9.15	3.29	1–18	10.46	2.50	6–17	1.61	50	.057	.926	0.45
Shape Memory	9.08	2.52	4–14	8.62	2.04	5–13	-0.73	49	.764	.926	0.20
Rotated Shape Memory	8.96	3.21	3–17	10.15	3.09	6–18	1.35	49	.091	.926	0.38
Matrices: Completion	10.52	3.00	4–16	10.69	3.04	4–16	0.20	49	.840	.926	0.06
Matrices: Odd One Out	10.64	3.62	3–18	8.88	3.06	2–14	-1.87	49	.067	.926	0.52
Naming Categories	10.38	3.95	1–16	10.12	3.34	4–16	-0.27	50	.604	.926	0.07
Naming Opposites	10.38	3.11	4–16	10.19	3.31	3–19	-0.22	50	.585	.926	0.06
Story Recall	9.20	3.55	1–14	8.69	3.34	1–15	-0.53	49	.699	.926	0.15
Picture Recall	7.77	2.80	3–13	8.65	3.57	3–18	0.99	50	.162	.926	0.28
Executive functions composite score	9.31	2.17	6–13	9.84	1.76	6–14	0.96	48	.171	.926	0.27
Listing Words	8.27	3.39	1–14	9.21	2.73	4–17	1.07	48	.144	.926	0.30
Divided Attention	9.40	3.00	4–15	10.10	2.30	7–17	0.93	47	.180	.926	0.26
Animal Colors	8.83	3.43	2–14	9.88	3.22	6–19	1.10	47	.138	.926	0.31
Drawing Routes	10.29	2.39	5–15	10.02	2.63	5–15	-0.38	49	.648	.926	0.11

Note. Samples were matched for age, sex, and maternal education (as a proxy for socioeconomic status). p_H indicates p values adjusted with Hommel's (1988) correction. Please note that after this correction, none of the comparisons were significant.

Table S6

Means, Standard Deviations, and t tests of the Developmental Functions From the Intelligence and Development Scales–2 for Autistic and Non-Autistic Adolescents (Aged 11–20 Years)

Variable	Autistic Adolescents *n* = 26			Non-Autistic Adolescents *n* = 26			*t*	*df*	*p*	p_H	*d*
	M	*SD*	Range	*M*	*SD*	Range					
Psychomotor skills composite score	9.12	1.87	4–12	10.29	1.58	7–15	2.42	49	.010	.456	0.68
Gross Motor Skills	-	-	-	-	-	-	-	-	-	-	-
Fine Motor Skills	8.90	3.01	2–14	10.65	2.19	8–16	2.31	47	.013	.567	0.66
Visuomotor Skills	9.35	1.55	6–13	10.02	1.74	7–13	1.46	49	.075	.926	0.41
Social-emotional skills composite score	8.74	2.72	1–13	9.24	2.10	5–13	0.73	48	.235	.926	0.21
Identifying Emotions	-	-	-	-	-	-	-	-	-	-	-
Regulating Emotions	9.00	3.16	1–13	9.36	2.53	4–13	0.44	48	.329	.926	0.13
Socially Competent Behavior	8.65	3.16	2–15	9.12	2.37	5–14	0.58	46	.281	.926	0.17
Basic skills composite score	10.77	1.78	8–14	9.93	2.16	5–13	-1.49	48	.142	.926	0.42
Logical-Mathematical Reasoning	10.96	3.61	1–17	10.52	2.90	4–16	-0.48	49	.633	.926	0.13
Language Skills	-	-	-	-	-	-	-	-	-	-	-
Phoneme Analysis	-	-	-	-	-	-	-	-	-	-	-
Phoneme–Grapheme Correspondence	-	-	-	-	-	-	-	-	-	-	-
Language Expressive	-	-	-	-	-	-	-	-	-	-	-
Language Receptive	-	-	-	-	-	-	-	-	-	-	-
Reading	10.40	1.84	8–15	9.58	2.19	5–13	-1.41	46	.167	.926	0.41
Reading Words	10.67	2.76	5–16	8.96	2.41	5–13	-2.31	47	.025	.893	0.66
Reading Pseudo Words	10.09	2.29	6–14	9.72	2.51	4–14	-0.53	46	.601	.926	0.15
Text Comprehension	11.43	3.47	3–16	10.09	2.87	4–16	-1.43	44	.921	.926	0.42
Spelling	9.95	2.80	6–15	9.75	2.67	4–14	-0.25	43	.806	.926	0.07
Motivation and attitude composite score	10.56	3.24	6–17	10.65	2.78	6–19	0.11	46	.458	.926	0.03
Conscientiousness	10.21	3.27	6–18	10.26	2.85	6–19	0.06	45	.477	.926	0.02
Achievement Motivation	11.12	3.96	4–19	11.04	3.11	6–19	-0.07	47	.528	.926	0.02
Participation during testing, intelligence	9.12	3.18	1–16	10.21	3.08	4–15	1.23	48	.112	.926	0.35
Participation during testing, executive functions	9.04	3.16	1–16	10.00	2.56	4–14	1.14	46	.129	.926	0.33
Participation during testing, developmental functions	9.46	3.06	1–16	10.71	2.71	5–16	1.52	48	.068	.926	0.43

Note. Samples were matched for age, sex, and maternal education (as a proxy for socioeconomic status). p_H indicates *p* values adjusted with Hommel's (1988) correction. Please note that after this correction, none of the comparisons were significant. Gross Motor Skills, Identifying Emotions, and Language Skills only for ages 5 to 10 years.

Supplemental References

Hommel, G. (1988). A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika*, *75*, 383–386.

APPENDIX B: STUDY 2

Odermatt, S. D., Weidmann, R., Schweizer, F., & Grob, A. (2024). Academic performance through multiple lenses: Intelligence, conscientiousness, and achievement striving motivation as differential predictors of objective and subjective measures of academic achievement in two studies of adolescents. *Journal of Research in Personality*, Article 104461. https://doi.org/10.1016/j.jrp.2024.104461
(Note: On the date of submission, the revised resubmission of the manuscript was currently under review. Updated reference with journal and DOI on January 20th, 2024.)

Note that Study 2 comprised two studies with different samples that are referred to as Studies 2.1 and 2.2 in the dissertation and as Studies 1 and 2 in the manuscript, respectively.

Please note that this is the author's version of a work that was accepted for publication in *Journal of Research in Personality*. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article.
The final article is available via its DOI: https://doi.org/10.1016/j.jrp.2024.104461

Academic Performance Through Multiple Lenses: Intelligence, Conscientiousness, and Achievement Striving Motivation as Differential Predictors of Objective and Subjective Measures of Academic Achievement in Two Studies of Adolescents

Salome D. Odermatt[1], Rebekka Weidmann[2], Florine Schweizer[1], and Alexander Grob[1]

[1] Department of Psychology, University of Basel, Switzerland
[2] Department of Psychology, Michigan State University, United States

Author Note

The authors declare no conflict of interest, with one exception: AG is a recipient of the royalties for the Intelligence and Development Scales (IDS) and the Intelligence and Development Scales–2 (IDS-2). This work was supported by a Postdoc.Mobility Fellowship (P400PS_186724) from the Swiss National Science Foundation granted to Rebekka Weidmann. The funding source was not involved in the study design, data collection, analysis, interpretation, writing, or in the decision to submit this article for publication. We thank all participants and their parents for participation and all test administrators for their assistance. Further, we would like to thank Priska Hagmann-von Arx and Jasmin T. Gygi for their valuable comments on an earlier version of this manuscript and Anita Todd for editing the manuscript. The data presented in the two studies are available on request from the corresponding author. The data and study material are not publicly available because of property rights restrictions. The analysis code can be accessed at https://osf.io/y3c5z/?view_only=ddeb7c5e62d74492af76c2969c4da1f4.

Correspondence concerning this article should be addressed to Salome D. Odermatt, Department of Psychology, University of Basel, Missionsstrasse 62, 4055 Basel, Switzerland. Email: salome.odermatt@unibas.ch

Abstract

In a two-sample investigation of 766 adolescents, we examined the associations and incremental validity of test-based intelligence, parent-rated conscientiousness, and self-rated achievement striving motivation with objective (grades) and subjective (parent-reported and self-reported academic performance) measures of academic achievement. The two studies yielded largely similar results. Intelligence was related to objective and subjective performance ratings. Conscientiousness showed associations and explained variance beyond intelligence in grades and parent-reported but mostly not in self-reported academic performance. Achievement striving motivation was largely related to and explained variance beyond intelligence and conscientiousness in grades and subjective performance ratings. Findings indicate that traits and motives predict objective and subjective academic achievement measures incrementally to abilities. Differential relationships for conscientiousness depended on the informant of academic performance.

Keywords: intelligence; conscientiousness; achievement striving motivation; school grades; subjective performance ratings; adolescents

1. Introduction

For adolescents, school represents a major part of their everyday lives, where they experience various learning settings, acquire important skills, and create the foundation for their future careers. Especially during this age span, academic achievement measures are of particular importance as they set the path for later educational and professional development. Intelligence has been identified in previous literature as the most significant predictor of overall academic achievement (e.g., Mammadov, 2022), including school grades (e.g., Roth et al., 2015) and standardized achievement tests (e.g., Deary et al., 2007). Although intelligence explains a large share of the variance in academic performance, a substantial proportion remains unaccounted for. According to Roberts and Wood's (2006) neo-socioanalytic theory, other personality characteristics besides abilities are critical for important life outcomes, namely, traits, motives, and narratives. In the context of academic performance, particularly traits and motives, such as conscientiousness and achievement striving motivation,[1] have been identified as meaningful predictors of academic achievement (e.g., Richardson et al., 2012; Steinmayr & Spinath, 2007) as they are described as the "will do" (Gottfredson, 2003, p. 369) for academic work. Moreover, research has suggested that conscientiousness (e.g., Mammadov, 2022) and achievement striving motivation (e.g., Steinmayr & Spinath, 2009) show incremental contributions beyond intelligence to variance in grades. However, studies investigating the relations of traits and motives with subjective measures of academic achievement, that is, parent-reported and self-reported academic performance, in addition to grades, are scarce or lacking in adolescents.

The aim of the present work[2] was therefore to examine the relationships of students' intelligence, parent-rated conscientiousness, and self-rated achievement striving motivation with grades as well as with subjective performance ratings from parents and students[3]—who offer additional perspectives on students' academic achievement. Moreover, incremental validity was assessed. To achieve these goals, two independently recruited samples of adolescents were used, which allowed for direct cross-validation of the results.

1.1 Intelligence and Academic Achievement Measures

Intelligence has been described as a general mental capacity for reasoning, problem solving, planning, and learning quickly (Gottfredson, 1997) and represents the individual's "can do" (Gottfredson, 2002, p. 37) or "*ability* to do academic work" (Di Domenico & Fournier, 2015, p. 156). Previous literature has demonstrated the significance of intelligence for academic achievement, as it explains up to a quarter of the variance in academic achievement, and has suggested an average correlation of r = .50 between intelligence and academic performance (Gustafsson & Undheim, 1996; Kuncel et al., 2004). Meta-analyses of cross-sectional and longitudinal studies have reported corrected correlations between intelligence and school grades of ρ = .54 for the elementary to secondary school level (N = 105,185; Roth et al., 2015) and ρ = .21 for the postsecondary school level (N = 7,820; Richardson et al., 2012).

In comparison, research on the relationship between students' intelligence and parent-reported academic performance is scarce. However, a few studies have suggested that children's intelligence is linked to parental factors, such as parents' perceptions, beliefs, and expectations of their children's academic competence (e.g., Phillipson, 2010; Phillipson & Phillipson, 2012), as parents observe how their children use their cognitive ability, such as their reasoning and planning skills, in everyday tasks (e.g., when solving homework problems). Gut et al. (2013), for instance, found very high associations (β = .56) between intelligence and parents' performance ratings of their 5- to 7-year-old children.

Students' intelligence also appears to be related to their self-reported academic performance. Studies that assessed intelligence and students' global performance ratings—among other constructs—

found correlations ranging from moderate to very large (e.g., Chamorro-Premuzic et al., 2010; Lotz et al., 2018). Students are likely to evaluate their academic performance on the basis of their intelligence because their learning skills directly impact their ability to complete school-related tasks. To our knowledge, only one study has primarily addressed the relationship between intelligence and self-reported academic performance. Spinath and Spinath (2005) reported in a cross-sectional study that the strength of the correlation between intelligence and self-reported academic ability increased from first to fourth grade and reached a moderate association (r = .23) at the end of elementary school. However, longitudinal evidence on the association between students' intelligence and their self-reported academic performance in adolescents is lacking. Building on these findings, we expected that intelligence would be positively related to objective and subjective measures of academic achievement (i.e., grades, parent-reported and self-reported academic performance) in our two studies (Hypothesis 1).

1.2 Conscientiousness and Academic Achievement Measures

In the neo-socioanalytic framework (Roberts & Wood, 2006), the Big Five traits represent the broadest personality domains. Among these traits, conscientiousness encompasses the extent to which individuals are self-disciplined, dutiful, and achievement oriented (Costa & McCrae, 1992) and has been described as "the *readiness* to do academic work" (Di Domenico & Fournier, 2015, p. 157). More conscientious individuals tend to be more able to delay gratification, plan ahead, be goal directed, and control their impulses (Roberts et al., 2009). Empirically, conscientiousness is considered the strongest predictor of academic achievement among the Big Five traits and an incremental predictor beyond intelligence (Mammadov, 2022). Meta-analyses have indicated corrected correlation coefficients of ρ = .23 (N = 27,875; postsecondary school level; Richardson et al., 2012), ρ = .27 (N = 413,074; elementary to postsecondary school level; Mammadov, 2022), and ρ = .50 (N = 5,706; elementary school level; Poropat, 2014) for conscientiousness and average grades.

Yet, less is known regarding the association between students' conscientiousness and their parent-reported academic performance. There is evidence of intercorrelations between students' conscientiousness and the effort they invest in their homework (Göllner, Damian, et al., 2017; Trautwein & Lüdtke, 2007) and studying for exams (Bleidorn, 2012). As parents directly observe the level of effort their children put into academic-related activities and their children's study habits, adolescents' conscientiousness consequently might be related to parents' performance ratings of their children. However, to our knowledge, the link between conscientiousness and parent-reported academic performance has not yet been directly examined.

In terms of self-reported academic performance, it can be assumed that part of adolescents' performance ratings is built on their personality traits, in particular their level of conscientiousness. In other words, it is likely that more conscientious adolescents also evaluate themselves as being more academically competent. This might be partly because more conscientious individuals stay more focused on tasks even in stressful situations (Saklofske et al., 2012), use self-regulated learning strategies such as time-management (Bidjerano & Dai, 2007), and have more likely experienced being good in school (e.g., getting good grades on report cards) in the past. Thus, their conscientiousness might be linked to their self-reported academic performance. In line with this reasoning, previous research has suggested a positive moderate relationship between conscientiousness and students' self-reported academic performance (e.g., Wilmot & Ones, 2019). Marsh et al. (2006) found in a study of adolescents that conscientiousness was positively linked to multiple variables of self-reported performance ratings, showing the highest association (of moderate strength) with self-reported academic performance in mathematics (r = .26). Another study using data from the Program for International Student Assessment also demonstrated positive but somewhat lower correlations between conscientiousness and different school subjects in ninth and 10th graders, reporting the highest relationship with self-reported performance ratings in French (r = .20; Spengler et al., 2016). On the basis of this research, we expected that conscientiousness would be positively associated with objective and subjective measures of academic achievement (i.e., grades, parent-reported and self-reported academic performance; Hypothesis 2.1) and, in addition, would be an incremental predictor for grades beyond intelligence (Hypothesis 2.2) in our two studies. Moreover, we sought to explore the incremental validity of conscientiousness for subjective measures of academic achievement (Research Question 1): Does conscientiousness explain additional variance in parent-reported and self-reported academic performance beyond intelligence?

1.3 Achievement Striving Motivation and Academic Achievement Measures

A third domain of the neo-socioanalytic theory (Roberts & Wood, 2006) encompasses values and motives. Values and motives represent all the characteristics that are desirable to a person (Roberts & Wood, 2006) and have been subsumed under the broad and heterogeneous concept of motivation (Eccles & Wigfield, 2002). For academic achievement, especially the facet achievement striving motivation—a person's need to be successful in life (Murray, 1938)—seems to be a relevant factor as it contains the "*willingness* to do academic work" (Di Domenico & Fournier, 2015, p. 156). A meta-analysis reported corrected correlations of subordinate achievement motivation of ρ = .30 with school grades for the postsecondary school level (N = 9,330; Robbins et al., 2004). Likewise, a more recent study of secondary school students showed that achievement motives explained some variance in school grades (2.72%), even after intelligence, the Big Five traits (including conscientiousness), and other motivational processes were controlled (Lavrijsen et al., 2022). Achievement striving motivation, more specifically, was also associated with grade point average (GPA; β = .23–.36) in a sample of adolescents (Steinmayr & Spinath, 2007, 2009).

Regarding parent-reported academic performance, various theories, such as the self-determination theory (Deci & Ryan, 1985) and the expectancy–value theory (Wigfield & Eccles, 2000), postulate the importance of parents' beliefs for their children's achievement motivation. For example, within the framework of self-determination theory (Deci & Ryan, 1985), it has been proposed that parents' engagement, including the communication of reasonable expectations and rules, may satisfy students' needs for autonomy, competence, and relatedness, which in turn may enhance students' motivation. Support for this notion comes from Pomerantz et al. (2007), who argued that parents' perceptions might promote or diminish their children's motivation when parents share their beliefs during everyday interactions, for instance, while assisting with homework. However, empirical research supporting this assumption is scarce and limited to parents' influence in predicting motivation (e.g., Dinkelmann & Buff, 2016). Nevertheless, given that any kind of development reflects the output of reciprocal processes between individuals and their environmental systems (Bronfenbrenner, 1979), it can be assumed that the relationship between adolescents' achievement striving motivation and parents' performance ratings of their children operates bidirectionally.

Adolescents' achievement striving motivation might also be linked to their self-reported academic performance, as students with higher levels of achievement motivation engage with more persistence and intensity in learning activities (Wigfield et al., 2015) and may therefore rate themselves as more academically competent. Empirically, a study of seventh graders revealed a large correlation (r = .35) between achievement motivation and self-reported performance ratings (Lavrijsen et al., 2022). Likewise, focus, accuracy, and persistence—core aspects of achievement striving motivation—were overall moderately to very strongly associated (β = .21–.47) with self-reported academic performance in a study of adolescents (Kulakow, 2020). Earlier literature found slightly lower correlations (r = .11–.21) between achievement striving motivation and self-reported performance ratings in different subjects (Steinmayr & Spinath, 2007, 2009). However, only a handful of studies have examined the direct association between students' achievement striving motivation and their self-reported academic performance so far. Drawing from this literature, we expected that achievement striving motivation would show positive associations with objective and subjective measures of academic achievement (i.e., grades, parent-reported and self-reported academic performance; Hypothesis 3.1) and, additionally, would be an incremental predictor for grades beyond intelligence and conscientiousness (Hypothesis 3.2) in our two studies. Moreover, we aimed to explore the incremental validity of achievement striving motivation for subjective measures of academic achievement (Research Question 2): Does achievement striving motivation explain additional variance in parent-reported and self-reported academic performance beyond intelligence and conscientiousness?

1.4 The Present Work

Parent-reported and self-reported academic performance provide different perspectives on students' academic functioning. Parents' reports of academic performance are based on implicit evaluations, judgments, and expectations about their child's academic capabilities drawn from daily observations of their child's behavior outside of school, feedback from teachers, and comparisons with other children's performance (Frischknecht et al., 2014). Especially in the case of scholastic difficulties, parental performance ratings of their children play an important role because parents are the primary observers of developmental deficiencies, and because their involvement leads to higher acceptance of adopted measures (e.g., long-term educational decisions; Schrader, 2006). For example, if parents consider their child to be very competent, but the school grades do not reflect this, it might be difficult for parents to accept and support assistance for their child. On the other hand, self-reported academic performance reflects a person's evaluations of their competence within a specific academic domain based on feedback, performance experiences, and social comparison (Marsh & Martin, 2011). Self-evaluations become crucial for students' academic development as they begin to reflect on their own capabilities. Moreover, these evaluations appear to be reciprocally related to other measures of academic achievement with both skill-development and self-enhancement effects (Guay et al., 2003; Marsh et al., 2005). In general, subjective performance ratings from parents and students are likely to be based on actual school grades but depend, in addition, on the child's and parents' awareness and knowledge of the performance of the child as well as of other students in the class as we assessed them in the sense of a social comparison based on relative performance rankings in the present work.

Thus, including assessments from all key players in the school context (i.e., teachers, parents, and students) results in a more valid picture of students' academic achievement (Morris & Bronfenbrenner, 2006) and supports, from a systemic point of view, the successful promotion of students' learning needs. For example, when questions arise regarding students' academic performance, diagnosticians such as school psychologists usually consider all perspectives in the diagnostic process to connect school and home environments. Parents, as experts on their adolescents' home life, can provide important insights into students' academic behavior outside of school, and students can add information about their own learning progress. An isolated view of school grades is questionable because school is a distinct setting compared to the home environment, with its distractions, differing social demands, and performance situations. Therefore, a comprehensive assessment of students' academic achievement by including multiple perspectives is needed.

Furthermore, conscientiousness and achievement striving motivation seem to be somewhat less stable compared to intelligence (e.g., Roberts & Yoon, 2022) and might therefore be used for interventions (Lazowski & Hulleman, 2016; Roberts et al., 2017; Stieger et al., 2021). Thus, it is of great interest to investigate their associations with academic achievement measures. Because our goal was to deepen the understanding regarding the associations of conscientiousness and achievement striving motivation with academic performance from multiple perspectives, we went beyond objective performance ratings, such as grades, and solicited subjective performance ratings from parents and students. Moreover, we add insights regarding the incremental validity of conscientiousness and achievement striving motivation beyond intelligence for the prediction of objective and, importantly, subjective academic achievement measures. We also contribute to previous findings by broadening the scope to different school-subject domains (i.e., humanities and science). Specifically, we investigated the associations and incremental contributions of students' intelligence, parent-rated conscientiousness, and self-rated achievement striving motivation with school grades and parent-reported and self-reported academic performance in two studies considering overall academic achievement estimates and achievement in specific school-subject domains. The two studies included samples of adolescents in the same age range and identical measures regarding conscientiousness, achievement striving motivation, and academic achievement as well as similar intelligence test procedures. Thus, we were able to validate our results across two similar samples with identical analytical strategies.

2. Methods

2.1 Procedure

Both studies were approved by the legally responsible Ethics Committee (BLINDED) and performed in accordance with the 1964 Helsinki declaration [Study 1: 221/10, Study 2: 2015-009]. Informed consent was obtained from parents and their children. We recruited participants in the German-speaking part of Switzerland, Germany, and Austria in 2007 (Study 1) and 2015 (Study 2) by distributing flyers with study information at different schools without targeting specific classrooms. Parents and their children could register to participate in the studies by sending back a registration form free of charge. Participants were then contacted to schedule an appointment for the intelligence testing. To assess intelligence, participants were individually tested with standardized test procedures by psychologists or trained undergraduate psychology students. The testing of the intelligence domain took about 45 min in both studies and was conducted in a laboratory at the university or at the participants' homes. Two questionnaires were sent to the families subsequently (in Study 2) or 7 years later (in Study 1). Parents rated the child's conscientiousness, reported the academic performance of their child, and indicated the child's school grades and the highest completed degree of the child's mother in a questionnaire. Adolescents provided information about their achievement striving motivation and reported their academic performance in another questionnaire.

2.2 Participants

2.2.1 Study 1

Study 1 consisted of 301 students who participated at both time points (Time 1: 2007, Time 2: 2014). Students who were assessed only once were excluded from the analyses. Participants ranged in age from 5 to 10 years during intelligence testing at Time 1 (M = 8.11 years, SD = 1.54) and 12 to 18 years during questionnaire collection at Time 2 (M = 15.45 years, SD = 1.52; 52% female). Seventy-six percent of the participants' mothers had completed formal job training or compulsory education (i.e., no postsecondary education), whereas 24% had earned a university degree (i.e., postsecondary education).

This is comparable with census data of the corresponding population in recent years (e.g., 2014–2021: 22.7–29.2% of Swiss females had completed postsecondary education; Swiss Federal Statistical Office, 2022). Adolescents attended the first (*n* = 139) or second (*n* = 152) secondary school level, or other school programs (*n* = 6). Twenty-two students had skipped at least one grade, and 28 had to repeat one grade. About 9% of the adolescents had a diagnosed learning disorder (dyscalculia: *n* = 9; dyslexia: *n* = 17; language development disorder: *n* = 2).

2.2.2 Study 2

Study 2 consisted of 465 students who participated in the study between 2015 and 2017. Participants ranged in age from 12 to 18 years (*M* = 15.16 years, *SD* = 1.56; 53% female) during intelligence testing and questionnaire collection. Sixty-six percent of the participants' mothers had completed formal job training or compulsory education (i.e., no postsecondary education), whereas 34% had earned a university degree (i.e., postsecondary education). This is slightly higher than in the corresponding population in recent years (e.g., 2014–2021: 22.7–29.2% of Swiss females had completed postsecondary education; Swiss Federal Statistical Office, 2022). Adolescents attended the first (*n* = 169) or second (*n* = 157) secondary school level, or other school programs (*n* = 21). Thirteen students had skipped at least one grade, and 24 had to repeat one grade. About 6% of the adolescents had a diagnosed learning disorder (dyscalculia: *n* = 4; dyslexia: *n* = 15; language development disorder: *n* = 7).

2.3 Measures

2.3.1 Intelligence

Intelligence was assessed using the Intelligence and Development Scales (IDS; Grob et al., 2009) in Study 1 and the Intelligence and Development Scales–2 (IDS-2; Grob & Hagmann-von Arx, 2018a) in Study 2. Both test batteries measure general intelligence (*M* = 100, *SD* = 15) and contain comparable tasks, as the IDS-2 is based on the IDS (see Table S1 for a description of the IDS and Table S2 for a description of the IDS-2 intelligence subtests). The intelligence domains of both test batteries are largely grounded in Spearman's (1904) *g* or two-factor theory. In addition, the IDS intelligence domain is based on the theory of fluid mechanics by Baltes (1990), focusing on fluid intelligence aspects. The IDS-2 is additionally based on the Cattell–Horn–Carroll (CHC) model (McGrew, 1997, 2009; Schneider & McGrew, 2018), which also includes scholastic abilities, such as quantitative knowledge or reading and writing, within the intelligence domain. However, in contrast to the CHC model, the IDS-2 does not assess scholastic skills within the intelligence domain.

Both tests are often used in psychological research and practice in German-speaking countries and several international adaptations with other languages have been conducted (e.g., a Dutch version; Grob et al., 2018). Psychometric properties reported for the standardization studies of the IDS and the IDS-2 show excellent internal consistencies (IDS: α = .92; IDS-2: α = .97) and test–retest reliabilities (IDS: *r* = .83; IDS-2: *r* = .89) for general intelligence (Grob et al., 2009; Grob & Hagmann-von Arx, 2018b). Several studies have demonstrated predictive validity evidence of intelligence assessed with the IDS (β = .21–.32; Gut et al., 2012, 2013) and concurrent validity evidence of intelligence assessed with the IDS-2 (β = .38; Grieder et al., 2022) for school grades. Moreover, both test batteries show differential validity (e.g., Hagmann-von Arx et al., 2008; Odermatt et al., 2022) and convergent validity with other intelligence tests often used in German-speaking countries, such as with the *Wechsler Intelligence Scale for Children–Fourth Edition* (IDS: *r* = .69; IDS-2: *r* = .67; Grob et al., 2009; Grob & Hagmann-von Arx, 2018b; Petermann & Petermann, 2011).[4]

2.3.2 Conscientiousness

In both studies, conscientiousness was measured with a parental questionnaire using the German Five Factors Questionnaire for Children [Fünf-Faktoren-Fragebogen für Kinder] (FFFK; Asendorpf, 1998). A previous meta-analysis indicated that adult reports on conscientiousness demonstrated higher validity than self-reports in predicting academic performance (Poropat, 2014). Parents evaluated their child's conscientiousness on eight items. The items consist of bipolar adjectives and were answered on a 5-point Likert scale ranging from 1 to 5, introduced with the following instructions: "Here you find trait words. They are pairs of opposites that describe how children can be. We are interested in your opinion about how your child compares to other children their age." An example of an item is "My child is: 1 (sloppy) vs. 5 (conscientious)." Studies have indicated good internal consistencies (range: α = .83 to .91; Asendorpf & Van Aken, 1999) and validity (Asendorpf & Van Aken, 2003) for the FFFK.

2.3.3 Achievement Striving Motivation

Achievement striving motivation was assessed in both studies with an adolescent questionnaire using the German Achievement Motivation Questionnaire for 7th- to 13th-Grade Students [Fragebogen zur Leistungsmotivation für Schüler der 7. bis 13. Klasse] (Petermann & Winkel, 2007). Adolescents rated their achievement striving motivation on eight items on a 5-point Likert scale ranging from 1 (*totally disagree*) to 5 (*totally agree*). An example of an item is "At school, I want to be the best." Analyses of the standardization sample demonstrated acceptable internal consistency (α = .73) and test–retest reliability (r = .72) after 6 weeks as well as construct and criterion validity for achievement striving motivation (Petermann & Winkel, 2007).

2.3.4 School Grades

School grades were obtained in both studies using parent reports of their children's school grades in five subjects (i.e., language; geography and history; mathematics; biology, chemistry, and physics; environment) from the school records. These five subjects cover the core school subjects for students in this age group. In the Swiss school system, students' grades range from 1 (*lowest*) to 6 (*highest*), with grades 4 through 6 representing passing grades. Grades of German (ranging from 6 [*lowest*] to 1 [*highest*]) and Austrian (ranging from 5 [*lowest*] to 1 [*highest*]) students were converted to the Swiss grading scale.[5]

2.3.5 Parent-Reported Academic Performance

Parent-reported academic performance was assessed in both studies with the parental questionnaire. In line with previous research (e.g., Frome & Eccles, 1998; Gut et al., 2013), parents were asked: "How would you rate your child's school performance? Please compare your child's performance with their peers' performance." They indicated their performance ratings of their children for each of the five subjects on a 5-point Likert scale ranging from 1 (*much worse*) to 5 (*much better*).

2.3.6 Self-Reported Academic Performance

Self-reported academic performance was assessed in both studies with the adolescent questionnaire. Similar to the assessment of the parent-reported academic performance, adolescents were asked: "How would you rate your school performance? Please compare your performance with the performance of your peers." They indicated their performance ratings for each of the five subjects on a 5-point Likert scale ranging from 1 (*much worse*) to 5 (*much better*).

2.4 Statistical Analyses

We performed identical analyses for both studies to establish comparability in the studies' results. We calculated descriptive statistics and bivariate correlations for the demographic characteristics and study variables. Reliability was computed using omega reliabilities. Effect sizes were interpreted according to Funder and Ozer (2019; small effect: $r \geq .10$, medium effect: $r \geq .20$, large effect: $r \geq .30$, very large effect: $r \geq .40$). All analyses were performed in R (R Core Team, 2022).

Information about school grades and parent-reported and self-reported academic performance in the five subjects were grouped into three school-subject domains for theoretical reasons. First, the school-subject domain *humanities* comprised the two aggregated subjects *language* and *geography and history*, as they focus on aspects of human experience, include social facets, and help promote intercultural competences. Second, the school-subject domain *science* was composed of the two averaged subjects *mathematics* and *biology, chemistry, and physics*, as they are part of the natural sciences and share the same methods and principles, such as relying on experiments. Third, the last domain, *environment*, focuses on the acquisition of skills in four different areas: (1) nature and technology, (2) economy, work, and household, (3) spaces, times, and societies, and (4) ethics, religions, and community. This subject aims to improve students' ability to understand fundamental knowledge about the world (Swiss-German Conference of Directors of Education, 2016). Accordingly, as this subject contains elements from both the *humanities* and *science* school-subject domains, we decided to treat it as its own subject domain.

To investigate whether different personality characteristics contributed to the variance in academic achievement measures beyond intelligence (and conscientiousness), we conducted hierarchical regression analyses (cf., Hypotheses 2.2 and 3.2; Research Questions 1 and 2). For that purpose, we included the control variables—sex and socioeconomic status (SES; represented as maternal educational attainment)—in Step 1, intelligence in Step 2, conscientiousness in Step 3, and achievement striving motivation in Step 4 as predictors. We used the standardized scores (M = 100, SD = 15) for intelligence and the raw sum scores for conscientiousness and achievement striving motivation. To further test our Hypotheses 1, 2.1, and 3.1, we performed structural equation modeling analyses to investigate the relationships between the constructs while correcting for measurement error. For these models, we generated three parcels per latent factor following the item-to-construct balance technique (Little et al., 2002) to create the factors for intelligence, conscientiousness, and achievement striving motivation (a detailed description of the statistical approach for the structural equation modeling analyses is provided in the Supplement). We used the standardized scores (M = 10, SD = 3) for the intelligence subtests and the raw scores for the conscientiousness and achievement striving motivation items. We employed the lavaan package to conduct the structural equation modeling analyses (Rosseel, 2012).

3. Results

3.1 Descriptive Statistics and Bivariate Correlations

Table 1 provides an overview of the descriptive statistics of the study variables of Studies 1 and 2. Omega reliabilities of the measures ranged from .71 to .91 (except for intelligence in Study 1, which had a reliability of .54). Distributions of the academic achievement measures are presented in the Supplement (see Figures S1 to S3). Table 2 shows the bivariate correlations between the sample characteristics and the study variables. Patterns of correlations were largely similar in the two studies.

3.2 Overall Estimates

Hierarchical regression analyses to investigate the incremental validity of conscientiousness and achievement striving motivation for the overall estimates of academic achievement measures are presented in Table 3 for both studies. After controlling for sex and SES, analyses demonstrated positive associations between intelligence and each academic achievement measure in Step 2, explaining between 4% and 14% of additional variance. In line with Hypothesis 2.2, conscientiousness, entered in Step 3, was also significantly associated with grades, explaining between 8% and 10% of additional variance beyond intelligence. Concerning Research Question 1, conscientiousness was also significantly associated with subjective measures of academic achievement, explaining between 5% and 12% of additional variance, with one exception: No relation was found with adolescents' self-reported academic performance in Study 1. In accordance with Hypothesis 3.2, achievement striving motivation, included in Step 4, was related to grades, explaining between 2% and 7% of additional variance beyond intelligence and conscientiousness. Regarding Research Question 2, achievement striving motivation revealed positive associations with subjective measures of academic achievement, except for self-reported academic performance in Study 2. This final step explained between 1% and 7% of additional variance in subjective performance ratings.

Structural equation models[6] of both studies for the overall estimates of the academic achievement measures are displayed in Figure 1.[7] The overall models meet the criteria of a satisfactory to good model fit (Study 1: comparative fit index [CFI] = .943, root-mean-square error of approximation [RMSEA] = .051, standardized root-mean-square residual [SRMR] = .061; Study 2: CFI = .968, RMSEA = .037, SRMR = .053). In line with Hypothesis 1, intelligence was related to GPA (95% confidence interval [CI] [0.03, 0.21]; 95% CI [0.07, 0.19]), parent-reported academic performance (95% CI [0.08, 0.33]; 95% CI [0.12, 0.25]), and self-reported academic performance (95% CI [0.01, 0.23]; 95% CI [0.04, 0.14]). In accordance with Hypothesis 2.1, conscientiousness was associated with GPA (95% CI [0.11, 0.30]; 95% CI [0.08, 0.24]) and parent-reported academic performance (95% CI [0.03, 0.24]; 95% CI [0.07, 0.28]), but no associations emerged between conscientiousness and adolescents' self-reported academic performance in either study (95% CI [-0.05, 0.12]; 95% CI [-0.07, 0.09]). Finally, confirming Hypothesis 3.1, achievement striving motivation was related to GPA (95% CI [0.12, 0.33]; 95% CI [0.14, 0.33]), parent-reported academic performance (95% CI [0.10, 0.39]; 95% CI [0.16, 0.42]), and self-reported academic performance (95% CI [0.14, 0.45]; 95% CI [0.24, 0.47]). To contrast the models of the two studies, we compared the 95% CIs of the regression coefficients of the structural equation models. All 95% CIs showed moderate to high overlap according to the definition of Cumming (2012). Hence, no significant differences were detected between the regression coefficients of the two studies. In addition, the χ^2 difference test showed no significant difference between the multigroup structural equation models with and without constrained regression coefficients ($\Delta\chi^2 = 2.490$, $\Delta df = 6$, $p = .870$).

3.3 School-Subject Domains

3.3.1 Humanities

Table S3 presents the hierarchical regression analyses for the school-subject domain *humanities* for both studies. Intelligence was significantly associated with each academic achievement measure after accounting for sex and SES, explaining between 2% and 11% of additional variance. In Step 3, conscientiousness also showed significant relations with GPA as well as with parent-reported academic performance, but no relationship emerged with self-reported academic performance in *humanities* in Study 1. Between 3% and 12% of additional variance was explained beyond intelligence. Achievement striving motivation, entered in Step 4, was associated with each academic achievement measure, explaining between 2% and 9% of additional variance beyond intelligence and conscientiousness, except for the nonsignificant association with parent-reported academic performance in *humanities* in Study 2.

Figure S4 displays the structural equation models of both studies for the academic achievement measures in *humanities.* The overall models meet the criteria of a good model fit (Study 1: CFI = .953, RMSEA = .050, SRMR = .049; Study 2: CFI = .983, RMSEA = .030, SRMR = .045). Similar to the overall estimates, intelligence was related to GPA (95% CI [0.04, 0.26]; 95% CI [0.10, 0.24]), parent-reported academic performance (95% CI [0.07, 0.53]; 95% CI [0.14, 0.32]), and self-reported academic performance (95% CI [0.02, 0.37]; 95% CI [0.06, 0.26]) in *humanities*. Conscientiousness was associated with GPA

(95% CI [0.12, 0.32]; 95% CI [0.08, 0.26]) and parent-reported academic performance (95% CI [0.07, 0.33]; 95% CI [0.09, 0.39]) but not with adolescents' self-reported academic performance (95% CI [-0.12, 0.14]; 95% CI [-0.09, 0.20]) in *humanities*. Achievement striving motivation was related to GPA (95% CI [0.07, 0.27]; 95% CI [0.09, 0.30]), parent-reported academic performance (95% CI [0.07, 0.42]; 95% CI [0.04, 0.37]), and self-reported academic performance (95% CI [0.13, 0.45]; 95% CI [0.24, 0.48]) in *humanities*.

3.3.2 Science

Table S4 displays the hierarchical regression analyses for the school-subject domain *science* for both studies. Results revealed positive associations between intelligence and each academic achievement measure when controlling for sex and SES. Between 2% and 17% of additional variance was explained. Conscientiousness, entered in Step 3, also showed positive relations with each academic achievement measure, explaining between 1% and 11% of additional variance beyond intelligence, except for self-reported academic performance in *science* in Study 1. Achievement striving motivation, added in the final step, was significantly associated with each academic achievement measure beyond intelligence and conscientiousness. Additional accounted variance ranged from 1% to 7%.

Figure S5 shows the structural equation models of both studies for the academic achievement measures in *science*. The overall models meet the criteria of a good model fit (Study 1: CFI = .959, RMSEA = .050, SRMR = .051; Study 2: CFI = .966, RMSEA = .045, SRMR = .047). Similar to the overall estimates and the academic achievement measures in *humanities*, intelligence was related to GPA (95% CI [0.07, 0.32]; 95% CI [0.08, 0.21]), parent-reported academic performance (95% CI [0.14, 0.50]; 95% CI [0.16, 0.35]), and self-reported academic performance (95% CI [0.05, 0.39]; 95% CI [0.08, 0.23]) in *science*. Conscientiousness was associated with GPA (95% CI [0.12, 0.40]; 95% CI [0.06, 0.26]) and parent-reported academic performance (95% CI [0.03, 0.39]; 95% CI [0.03, 0.35]) but not with adolescents' self-reported academic performance (95% CI [-0.17, 0.23]; 95% CI [-0.15, 0.16]) in *science*. Achievement striving motivation was related to GPA (95% CI [0.18, 0.49]; 95% CI [0.20, 0.47]), parent-reported academic performance (95% CI [0.19, 0.74]; 95% CI [0.28, 0.68]), and self-reported academic performance (95% CI [0.37, 0.94]; 95% CI [0.41, 0.80]) in *science*.

3.3.3 Summary

The patterns of results for *humanities* and *science* were almost identical to the overall estimates reported earlier. Conscientiousness contributed to additional explained variance beyond intelligence in grades (Hypothesis 2.2) and parent-reported but not self-reported academic performance in *humanities* and *science* in Study 1 (Research Question 1). Achievement striving motivation explained additional variance beyond intelligence and conscientiousness in grades (Hypothesis 3.2) and subjective performance ratings, except for parent-reported academic performance in *humanities* in Study 2 (Research Question 2). As they were for the overall estimates, both Hypotheses 1 and 3.1 were confirmed for the specific school-subject domains, but Hypothesis 2.1 was only partly supported, as conscientiousness was not related to self-reported academic performance in *humanities* or *science*.

3.4 Post-Hoc Analyses

3.4.1 Socioeconomic Status (SES)

To control for effects of SES, we further conducted post-hoc analyses including the manifest variable maternal educational attainment as a predictor of the academic achievement measures in the models. In both studies, model fit deteriorated slightly (Study 1: CFI = .928, RMSEA = .055, SRMR = .065; Study 2: CFI = .962, RMSEA = .038, SRMR = .057). All regressions remained significant except for the association between students' intelligence and their self-reported academic performance in Study 1. Maternal educational attainment was not related to any of the academic achievement measures in Study

1, while in Study 2 a significant association with parent-reported academic performance was detected in that participants whose mothers had completed postsecondary education were rated by their parents as more academically competent than participants of mothers without postsecondary education (see Figure S8 in the Supplement for full results).

3.4.2 Sex

We also performed post-hoc analyses controlling for effects of sex in the structural equation models. In both studies, model fit deteriorated with the manifest variable sex as a predictor of the academic achievement measures in the models (Study 1: CFI = .919, RMSEA = .059, SRMR = .070; Study 2: CFI = .948, RMSEA = .044, SRMR = .064). All regressions remained significant. Sex was related to parent-reported academic performance in Studies 1 and 2 and to self-reported academic performance in Study 2 in that parents rated their sons to be more academically competent than their daughters and boys exhibited higher self-ratings of academic performance than girls (see Figure S9 in the Supplement for full results).

3.4.3 Summary

The observed results of the structural equation models showed minimal variation when controlling for SES or sex. Nevertheless, these results should be interpreted with caution because of the lower model fit.

4. Discussion

In two studies of adolescents (total N = 766) aged 12–18 years, we found a largely identical pattern of results for overall estimates and specific school-subject domains: Students' intelligence was associated with grades and subjective performance ratings. Conscientiousness was related to and incrementally explained additional variance beyond intelligence in grades and parent-reported academic performance, while mostly no associations were evident between conscientiousness and adolescents' self-reported academic performance. Achievement striving motivation was in most analyses associated with and explained additional variance beyond intelligence and conscientiousness in grades and subjective performance ratings. The findings suggest that traits and motives are incremental predictors for objective and subjective measures of academic achievement beyond abilities. Moreover, traits, such as conscientiousness, seem to be differentially related depending on the informant and, hence, the lens through which academic performance is evaluated.

4.1 Personality Characteristics and School Grades

Intelligence, conscientiousness, and achievement striving motivation showed significant associations and independent contributions with school grades in both studies. In line with previous research conducted with the IDS (Gut et al., 2012, 2013) and the IDS-2 (Grieder et al., 2022), intelligence was related to school grades longitudinally (Study 1) and cross-sectionally (Study 2) and explained between 2% and 12% of additional variance. Intelligence eases comprehension and learning in school (Di Fabio & Busoni, 2007) because it reflects the capability to reason, plan, solve problems, and learn quickly (Gottfredson, 1997). Students with higher intelligence can therefore acquire more knowledge and can use their better learning and problem-solving skills in scholastic tasks. Intelligence had a weaker association with school grades in both studies compared to studies where academic achievement was measured with standardized achievement tests (e.g., r = .81 between intelligence and national school test scores; Deary et al., 2007). In contrast to standardized achievement tests, grades assess academic achievement over a longer time period and comprise composites of several performances—not only exams but also, for example, extended assignments that require long-term organization and perseverance to succeed. Thus, other characteristics such as conscientiousness and achievement striving motivation may be also linked to grades.

In accordance with this reasoning and findings in the literature, conscientiousness (e.g., Mammadov, 2022) and achievement striving motivation (e.g., Steinmayr & Spinath, 2009) were associated with school grades and incrementally explained 7% to 12% and 2% to 7% of the variance beyond intelligence (and conscientiousness), respectively. Hence, not only the individual's "can do" (Gottfredson, 2002, p. 37) is important for grades but also the "will do" (Gottfredson, 2003, p. 369). Conscientious students, who are more self-disciplined, dutiful, and achievement oriented (Costa & McCrae, 1992), are thus more likely to get higher grades as these characteristics are crucial to be successful in school. This might be explained by the fact that conscientious students report more regular engagement in study routines and less absenteeism in class (Wilmot & Ones, 2019) and show fewer counterproductive academic behaviors at school, such as cheating, breaking rules, or working with low effort (Cuadrado et al., 2021). It might therefore be that students' conscientiousness-related behavior affects teachers' ratings and, accordingly, their school grades. Conscientiousness might also be associated with school grades because it is linked to motivation for goal-directed and persistent performance, which is most effectively realized in a predictable and well-structured environment with clear expectations and predefined goals, such as in educational settings (Wilmot & Ones, 2019). Moreover, students who demonstrated high levels of achievement striving motivation, describing themselves as ambitious, hardworking, and perseverant (Petermann & Winkel, 2007), achieved higher school grades. One reason may be that as motivation influences behavior (Spinath et al., 2006), highly motivated students might intend to achieve more than others (Petermann & Winkel, 2007). Therefore, they might embrace challenges in achievement situations, such as speaking up and contributing their ideas in the classroom. This active participation may be acknowledged by teachers and then be reflected, for instance, in a higher oral grade.

4.2 Personality Characteristics and Subjective Performance Ratings

Intelligence, conscientiousness, and achievement striving motivation were significantly related and incrementally contributed to the prediction of subjective performance ratings from parents and students in both studies, except for the nonsignificant associations between conscientiousness and adolescents' self-reported academic performance.

4.2.1 Personality Characteristics and Parent-Reported Academic Performance

In line with previous research (Gut et al., 2013), our results show similar associations between intelligence and parent-reported academic performance, as those found for children aged 5–7 years, which suggests that parents continue to have close interactions with their children into adolescence. Moreover, intelligence explained between 2% and 17% of variance in parent-reported academic performance. It can be assumed that how parents rate the academic competence of their children is mainly the result of adolescents' previous performance in school and behaviors in daily situations at home (Frischknecht et al., 2014; Pomerantz & Dong, 2006). For example, if students struggle with homework or studying, they will most likely seek help from their parents first. Hence, parents might observe how able their children are to perform achievement-related tasks and what learning or problem-solving strategies they use.

Moreover, conscientiousness was also associated with parent-reported academic performance and explained between 1% and 9% of variance beyond intelligence. This result supports previous research findings that students with higher conscientiousness studied and did their homework with more effort (Bleidorn, 2012; Trautwein & Lüdtke, 2007), which in turn might lead to parents perceiving their children as more academically competent. Last, achievement striving motivation was also related to parent-reported academic performance and explained between 1% and 6% of additional variance over and above intelligence and conscientiousness. This finding supports theoretical considerations (e.g., self-determination theory; Deci & Ryan, 1985) and previous studies that underline the importance of parental beliefs in the prediction of achievement motivation (e.g., Dinkelmann & Buff, 2016). However, our result suggests that the link between adolescents' achievement motivation and parent-reported academic performance could also operate in the other direction: It might be that students high on achievement striving motivation may invest more time in doing their homework or studying for exams than less-motivated peers. This may be recognized by their parents, which could lead them to rate their child as academically competent.

4.2.2 Personality Characteristics and Self-Reported Academic Performance

With respect to self-reported academic performance, associations were observed with students' intelligence which correspond to previous findings (e.g., Lotz et al., 2018), explaining between 2% and 6% of variance. Students may be aware of their cognitive ability because it affects their competence to do homework, study for exams, or complete assignments which might then be reflected in their self-reported academic performance. In addition, students, particularly during adolescence, may compare their learning abilities to those of their peers. In Study 1, we demonstrated for the first time that intelligence predicts self-reported academic performance in a longitudinal design over 7 years, which supports the idea that the effects of intelligence are stable and long lasting.

Moreover, consistent with previous literature (e.g., Steinmayr & Spinath, 2009), achievement striving motivation was also related to self-reported academic performance and explained between 5% and 9% of variance beyond intelligence and conscientiousness. One explanation might be that students high on achievement motivation may take more responsibility for their learning while showing persistence in studying activities and engaging with more intensity (Wigfield et al., 2015). Furthermore, some theories, such as the expectancy–value theory (Wigfield & Eccles, 2000), also consider performance ratings to be self-perceptions of academic competence, reflecting a motivational construct (i.e., expectancies to succeed), which may also partially explain the associations we found. Nevertheless, as we assessed cognitive-evaluative performance ratings (e.g., "I am good at history") and not affective evaluations (e.g., "I like science"; Möller & Trautwein, 2015), which would have been closer to the concept of motivation, we might have captured a different construct.

In contrast to previous studies (e.g., Marsh et al., 2006; Wilmot & Ones, 2019), conscientiousness was not related to adolescents' self-reported academic performance in most of our analyses. We attribute slight differences between the results of the hierarchical regression and structural equation model analyses to the fact that structural equation modeling corrects the estimated associations for measurement error variance. Hence, the finding suggests that how parents think about how conscientious their children are, seems not to be connected to how adolescents rate their own academic performance. This is interesting when contrasted with the result that conscientiousness showed associations with parent-reported academic performance.

There are several potential reasons for these differential findings between parent- and self-reported academic performance. First, observer reports of personality and academic performance by parents may provide unique information that differs from what adolescents report about themselves because parents may tap into aspects of their children's personality and academic performance that the adolescents might not report, disclose, or be aware of. Parents may use a different set of comparison and compare, for example, their child's academic performance with that of siblings or with their own past school experiences whereas adolescents may compare their performance with that of their classmates. Second, it could also be that students focus on other characteristics such as intelligence and motivation when evaluating their academic performance. For this reason, the student perspective should also be considered in the assessment of academic achievement, as adolescents do not necessarily rely on the same sources of information as their parents. Third, another potential reason might be that parents' and children's ratings of conscientiousness during adolescence do not perfectly align (Göllner, Roberts, et al., 2017). Parental evaluations of conscientiousness might be contingent on factors other than those considered by the child, which might drive the differences in effects. Fourth, individuals tend to become more conscientious over the life span (Roberts et al., 2006), suggesting that conscientiousness might not yet be related to adolescents' self-reported academic performance at this stage in life as studies found associations between conscientiousness and self-reported academic ability in university students (Noftle & Robins, 2007). Last, however, that an association of conscientiousness with parental performance ratings but not with adolescents' self-reported academic performance emerged, could also be due to shared method variance or measurement artefacts, such as the halo effect, where a positive impression in one domain (i.e., parent-rated conscientiousness) can lead to a positive impression in another (i.e., parent-reported academic performance).

4.3 Strengths, Limitations, and Future Directions

A major strength of the present work is that we included multiple perspectives of academic performance and took into account both objective *and* subjective measures of academic achievement, thereby revealing insights into differential associations with personality characteristics and presenting a more comprehensive picture than previous research. Moreover, we considered not only overall estimates of academic achievement but also specific school-subject domains and could show that our findings hold for humanities and science. Furthermore, the intelligence tests used in both studies do not assess school-related skills (e.g., quantitative knowledge) within the intelligence domain and therefore provide purer measures of intelligence. Additionally, we controlled for important sociodemographic confounders, such as sex and SES (e.g., Liu et al., 2020). Finally, we performed the same analyses with two independent samples, thus providing a direct replication of the main results, which underscores the robustness of our findings.

Our work also has limitations that should be recognized in future studies. First, adolescents attended different school tracks within school levels, which is due to the national school systems. In German-speaking countries, there exist several school tracks for specific age groups that are divided according to students' performance. Unfortunately, we did not collect information about specific schools and classes that the students attended. Thus, we could not control for the probably partly nested nature of our data in the analyses and, hence, for influences of students' reference group on the academic achievement measures, such as the big-fish–little-pond effect (for a review, see Marsh & Seaton, 2015). Future studies should gather more information about students' current school situation and control for clustering of students within schools and classrooms. Second, the predictors had unequal and, in some cases, only marginally sufficient reliability. Concerning the low reliability of intelligence in Study 1, however, age-separated analyses showed higher reliability for older children and the mean and standard deviation were distributed nearly the same as in the general population (see Table 1). Third, associations were almost completely cross-sectional and therefore no conclusions on causal relationships can be made. In future studies, growth curve designs using longitudinal data or experimental studies should be employed to investigate the directionality of the effects. Fourth, our findings may be generalizable exclusively to Western countries, as our samples included participants from so-called WEIRD (Western, educated, industrialized, rich, and democratic) nations (Muthukrishna et al., 2020) as well as only to the two intelligence tests employed, as intelligence test scores can differ greatly across tests on an individual level (e.g., Bünger et al., 2021; Hagmann-von Arx et al., 2018). Future studies that take into account the cultural context and use different test batteries are required. Fifth, as conscientiousness was solely rated by parents and achievement striving motivation only rated by adolescents in our studies, we cannot rule out the possibility that the results are driven by measurement-related factors (e.g., shared method variance). Hence, both self- and parental evaluations of traits and motives should be included in future work. Finally, since few previous studies have examined the associations of traits and motives with performance ratings from parents and students, research should further investigate subjective measures of academic achievement, in addition to grades, and also examine subjective performance ratings from teachers for a better understanding of academic achievement.

4.4 Implications

Implications for theory, research, and practice can be drawn from our findings. In line with Roberts and Wood's (2006) neo-socioanalytic theory, our results suggest that abilities as well as traits and motives constitute overall important aspects for both objective and subjective measures of academic achievement. Specifically and for most cases, traits and motives show incremental validity beyond intelligence. This implies that it is crucial to pursue a holistic understanding of personality by considering multiple characteristics and abilities of the individual.

Moreover, as relations of personality characteristics with objective and subjective measures of academic achievement varied depending on whether academic performance was reported by parents or by the students themselves, the goal should also be to consider the informant when applying interventions. Our findings suggest that for promoting self-reported academic performance, conscientiousness may play a subordinate role compared to achievement striving motivation, whereas for parent-reported academic performance both conscientiousness and achievement striving motivation might be beneficial. It has been suggested that conscientious behavior and habits such as punctuality or focus on a task can be practiced (Dumfart & Neubauer, 2016). Similarly, motivation depends on, for example, learning contexts, situational factors, and instructional characteristics (Midgley et al., 1995). However, specific intervention studies are needed that examine whether increases in conscientiousness and achievement striving motivation can also improve subjective performance ratings.

Finally, as our findings indicated that both objective and subjective measures of academic achievement are of relevance, multiple perspectives of academic performance should be assessed in educational settings. Hence, we encourage teachers to regularly assess parents' and students' performance ratings in addition to assigning school grades, especially when working with adolescents. For example, when discussing learning reports with parents and students, their performance ratings should be elicited and considered when making decisions. Evaluations from all key stakeholders in the school context (i.e., teachers as experts on adolescents at school, parents as experts on adolescents at home, and adolescents themselves as experts on their own learning) will provide a more comprehensive picture when identifying students' strengths and difficulties. Teachers should be more aware of potential perceptual discrepancies and pay attention to the multiple perspectives available in the school context.

Footnotes

1. According to personality (e.g., Costa & McCrae, 1992) and motivation (e.g., Eccles & Wigfield, 2002) conceptualizations, achievement striving represents a facet of conscientiousness as well as of achievement motivation. However, the way achievement striving was measured in the present work likely reflects part of the concept of achievement motivation, which may also explain why conscientiousness and achievement striving were not, or only weakly, correlated in the two studies (see descriptive statistics in Results section). To make it clear that achievement striving refers to motivation in the present work, we use the term "achievement striving motivation."
2. The two studies and our hypotheses were not preregistered. The data and study material presented here are not publicly available owing to property rights restrictions. The first author was involved in: Conceptualization; Methodology; Formal analysis; Writing – original draft; Writing – review & editing; Visualization. The second author was involved in: Conceptualization; Methodology; Writing – review & editing. The third author was involved in: Conceptualization; Writing – original draft; Writing – review & editing. The last author was involved in: Conceptualization; Funding acquisition; Project administration; Resources; Writing – review & editing; Supervision.
3. In the two studies, parent-reported and self-reported academic performance are measured as a social comparison relative to the performance of the children's peers in five school subjects.
4. It has to be noted that the IDS-2 includes three intelligence composite scores (a Screening IQ with two subtests, a Full-Scale IQ with seven subtests, and a Profile IQ with 14 subtests). As the Screening IQ is intended only for a rough initial assessment and the Profile IQ is primarily administered to calculate intelligence factors, the Full-Scale IQ was used in Study 2 because it provides a reliable and valid estimate of a general intelligence score that is most similar to the intelligence composite score obtained with the IDS in Study 1.
5. The following formula, employed by Swiss universities, was used to convert international grades into the Swiss grading scale (6 = maximum grade in Swiss grading scale; 2 = conversion factor; Grade_{max} = maximum grade in the foreign grading scale; Grade_{min} = lowest passing grade in the foreign grading scale; Grade = grade that is intended to be converted):

$$\text{Converted grade} = 6 - 2 \cdot \left(\frac{\text{Grade}_{\text{max}} - \text{Grade}}{\text{Grade}_{\text{max}} - \text{Grade}_{\text{min}}}\right)$$

6. We also performed the structural equation models of Studies 1 and 2 without any restricted covariances and received a similar pattern of results with one exception: Intelligence was no longer significantly related to self-reported academic performance in Study 1 (see Figure S6 in the Supplement). Moreover, we additionally conducted the structural equation modeling analyses with the five individual school subjects (i.e., language; geography and history; mathematics; biology, chemistry, and physics; environment) as indicators of the latent factors of the academic achievement measures and obtained comparable results (see Figure S7 in the Supplement).
7. We restricted the covariances between intelligence and achievement striving motivation (r = .10, 95% CI [-0.03, 0.14]; r = -.02, 95% CI [-0.14, 0.11]) as well as between GPA and self-reported academic performance (r = .27, 95% CI [0.00, 0.05]; r = .12, 95% CI [-0.01, 0.03]) to zero (see dashed lines in Figure 1) as these correlations were nonsignificant in both studies.

Highlights

- Included subjective performance ratings additional to grades
- Consisted of two studies of adolescents (total of *N* = 766)
- Intelligence, conscientiousness, achievement striving motivation were predictors
- Conscientiousness was not related to self-reported academic performance

5. References

Asendorpf, J. B. (1998). *Fünf-Faktoren-Fragebogen für Kinder [Five Factors Questionnaire for Children] (FFFK)*. Humboldt University, Institute for Psychology.

Asendorpf, J. B., & Van Aken, M. A. G. (1999). Resilient, overcontrolled, and undercontrolled personality prototypes in childhood: Replicability, predictive power, and the trait-type issue. *Journal of Personality and Social Psychology*, *77*(4), 815–832. https://doi.org/10.1037/0022-3514.77.4.815

Asendorpf, J. B., & Van Aken, M. A. G. (2003). Validity of Big Five personality judgments in childhood: A 9 year longitudinal study. *European Journal of Personality*, *17*(1), 1–17. https://doi.org/10.1002/per.460

Baltes, P. B. (1990). Entwicklungspsychologie der Lebensspanne: Theoretische Leitsätze [Lifespan developmental psychology: Guiding theoretical principles]. *Psychologische Rundschau*, *41*, 1–24.

Bidjerano, T., & Dai, D. Y. (2007). The relationship between the big-five model of personality and self-regulated learning strategies. *Learning and Individual Differences*, *17*(1), 69–81. https://doi.org/10.1016/j.lindif.2007.02.001

Bleidorn, W. (2012). Hitting the road to adulthood: Short-term personality development during a major life transition. *Personality and Social Psychology Bulletin*, *38*(12), 1594–1608. https://doi.org/10.1177/0146167212456707

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press.

Bünger, A., Grieder, S., Schweizer, F., & Grob, A. (2021). The comparability of intelligence test results: Group- and individual-level comparisons of seven intelligence tests. *Journal of School Psychology*, *88*, 101–117. https://doi.org/10.1016/j.jsp.2021.09.002

Chamorro-Premuzic, T., Harlaar, N., Greven, C. U., & Plomin, R. (2010). More than just IQ: A longitudinal examination of self-perceived abilities as predictors of academic performance in a large sample of UK twins. *Intelligence*, *38*(4), 385–392. https://doi.org/10.1016/j.intell.2010.05.002

Costa, P. T., & McCrae, R. R. (1992). *Neo Personality Inventory-Revised (NEO PI-R)*. Psychological Assessment Resources.

Cuadrado, D., Salgado, J. F., & Moscoso, S. (2021). Personality, intelligence, and counterproductive academic behaviors: A meta-analysis. *Journal of Personality and Social Psychology*, *120*(2), 504–537. https://doi.org/10.1037/pspp0000285

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. Routledge.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*(1), 13–21. https://doi.org/10.1016/j.intell.2006.02.001

Deci, E. L., & Ryan, R. M. (1985). *Motivation and self-determination in human behavior*. Plenum.

Di Domenico, S. I., & Fournier, M. A. (2015). Able, ready, and willing: Examining the additive and interactive effects of intelligence, conscientiousness, and autonomous motivation on undergraduate academic performance. *Learning and Individual Differences*, *40*, 156–162. https://doi.org/10.1016/j.lindif.2015.03.016

Di Fabio, A., & Busoni, L. (2007). Fluid intelligence, personality traits and scholastic success: Empirical evidence in a sample of Italian high school students. *Personality and Individual Differences*, *43*(8), 2095–2104. https://doi.org/10.1016/j.paid.2007.06.025

Dinkelmann, I., & Buff, A. (2016). Children’s and parents’ perceptions of parental support and their effects on children’s achievement motivation and achievement in mathematics. A longitudinal predictive mediation model. *Learning and Individual Differences*, *50*, 122–132. https://doi.org/10.1016/j.lindif.2016.06.029

Dumfart, B., & Neubauer, A. C. (2016). Conscientiousness is the most powerful noncognitive predictor of school achievement in adolescents. *Journal of Individual Differences*, *37*(1), 8–15. https://doi.org/10.1027/1614-0001/a000182

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*(1), 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Frischknecht, M.-C., Reimann, G., Gut, J., Ledermann, T., & Grob, A. (2014). Wie genau können Mütter die Mathematik- und Sprachleistungen ihrer Kinder einschätzen? [How accurately can mothers assess their children’s math and language performance?]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *46*(2), 67–78. https://doi.org/10.1026/0049-8637/a000101

Frome, P. M., & Eccles, J. S. (1998). Parents’ influence on children’s achievement-related perceptions. *Journal of Personality and Social Psychology*, *74*(2), 435–452. https://doi.org/10.1037/0022-3514.74.2.435

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168. https://doi.org/10.1177/2515245919847202

Göllner, R., Damian, R. I., Rose, N., Spengler, M., Trautwein, U., Nagengast, B., & Roberts, B. W. (2017). Is doing your homework associated with becoming more conscientious? *Journal of Research in Personality*, *71*, 1–12. https://doi.org/10.1016/j.jrp.2017.08.007

Göllner, R., Roberts, B. W., Damian, R. I., Lüdtke, O., Jonkmann, K., & Trautwein, U. (2017). Whose “storm and stress” is it? Parent and child reports of personality development in the transition to early adolescence. *Journal of Personality*, *85*(3), 376–387. https://doi.org/10.1111/jopy.12246

Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, *24*(1), 79–132. https://doi.org/10.1016/S0160-2896(97)90014-3

Gottfredson, L. S. (2002). Where and why g matters: Not a mystery. *Human Performance*, *15*(1–2), 25–46. https://doi.org/10.1080/08959285.2002.9668082

Gottfredson, L. S. (2003). Dissecting practical intelligence theory: Its claims and evidence. *Intelligence*, *31*(4), 343–397. https://doi.org/10.1016/S0160-2896(02)00085-5

Grieder, S., Bünger, A., Odermatt, S. D., Schweizer, F., & Grob, A. (2022). Limited internal comparability of general intelligence composites: Impact on external validity, possible predictors, and practical remedies. *Assessment*, *29*(6), 1172–1189. https://doi.org/10.1177/10731911211005171

Grob, A., & Hagmann-von Arx, P. (2018a). *Intelligence and Development Scales–2 (IDS-2). Intelligenz- und Entwicklungsskalen für Kinder und Jugendliche [Intelligence and Development Scales for Children and Adolescents]*. Hogrefe.

Grob, A., & Hagmann-von Arx, P. (2018b). *Intelligence and Development Scales–2 (IDS-2). Intelligenz- und Entwicklungsskalen für Kinder und Jugendliche. Manual zu Theorie, Interpretation und Gütekriterien [Intelligence and Development Scales for Children and Adolescents. Manual on theory, interpretation and psychometric criteria]*. Hogrefe.

Grob, A., Hagmann-von Arx, P., Ruiter, S. A. J., Timmerman, M. E., & Visser, L. (2018). *Intelligence and Development Scales–2 (IDS-2). Intelligentie- en ontwikkelingsschalen voor kinderen en* *jongeren. [Intelligence and Development Scales for Children and Adolescents]*. Hogrefe.

Grob, A., Meyer, C. S., & Hagmann-von Arx, P. (2009). *Intelligence and Development Scales (IDS)*. Huber.

Guay, F., Marsh, H. W., & Boivin, M. (2003). Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology*, *95*(1), 124–136. https://doi.org/10.1037/0022-0663.95.1.124

Gustafsson, J.-E., & Undheim, J. O. (1996). Individual differences in cognitive functions. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 186–242). Prentice Hall International.

Gut, J., Reimann, G., & Grob, A. (2012). Kognitive, sprachliche, mathematische und sozial-emotionale Kompetenzen als Prädiktoren späterer schulischer Leistungen: Können die Leistungen eines Kindes in den IDS dessen Schulleistungen drei Jahre später vorhersagen? [Cognitive, language, mathematical, and social-emotional skills as predictors of later academic achievement: Can a child's performance in the IDS predict their school performance three years later?] *Zeitschrift für Pädagogische Psychologie*, *26*(3), 213–220. https://doi.org/10.1024/1010-0652/a000070

Gut, J., Reimann, G., & Grob, A. (2013). A contextualized view on long-term predictors of academic performance. *Journal of Educational Psychology*, *105*(2), 436–443. https://doi.org/10.1037/a0031503

Hagmann-von Arx, P., Lemola, S., & Grob, A. (2018). Does IQ = IQ? Comparability of intelligence test scores in typically developing children. *Assessment*, *25*(6), 691–701. https://doi.org/10.1177/1073191116662911

Hagmann-von Arx, P., Meyer, C. S., & Grob, A. (2008). Assessing intellectual giftedness with the WISC-IV and the IDS. *Journal of Psychology*, *216*(3), 172–179. https://doi.org/10.1027/0044-3409.216.3.172

Kulakow, S. (2020). Academic self-concept and achievement motivation among adolescent students in different learning environments: Does competence-support matter? *Learning and Motivation*, *70*, Article 101632. https://doi.org/10.1016/j.lmot.2020.101632

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology*, *86*(1), 148–161. https://doi.org/10.1037/0022-3514.86.1.148

Lavrijsen, J., Vansteenkiste, M., Boncquet, M., & Verschueren, K. (2022). Does motivation predict changes in academic achievement beyond intelligence and personality? A multitheoretical perspective. *Journal of Educational Psychology*, *114*(4), 772–790. https://doi.org/10.1037/edu0000666

Lazowski, R. A., & Hulleman, C. S. (2016). Motivation interventions in education: A meta-analytic review. *Review of Educational Research*, *86*(2), 602–640. https://doi.org/10.3102/0034654315617832

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 151–173. https://doi.org/10.1207/S15328007SEM0902_1

Liu, J., Peng, P., & Luo, L. (2020). The relation between family socioeconomic status and academic achievement in China: A meta-analysis. *Educational Psychology Review*, *32*(1), 49–76. https://doi.org/10.1007/s10648-019-09494-0

Lotz, C., Schneider, R., & Sparfeldt, J. R. (2018). Differential relevance of intelligence and motivation for grades and competence tests in mathematics. *Learning and Individual Differences*, *65*, 30–40.

https://doi.org/10.1016/j.lindif.2018.03.005
Mammadov, S. (2022). Big Five personality traits and academic performance: A meta-analysis. *Journal of Personality*, *90*(2), 222–255. https://doi.org/10.1111/jopy.12663
Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, *81*(1), 59–77. https://doi.org/10.1348/000709910X503501
Marsh, H. W., & Seaton, M. (2015). The big-fish–little-pond effect, competence self-perceptions, and relativity: Substantive advances and methodological innovation. In A. J. Elliot (Ed.), *Advances in motivation science* (Vol. 2, pp. 127–184). Elsevier. https://doi.org/10.1016/bs.adms.2015.05.002
Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*(2), 397–416. https://doi.org/10.1111/j.1467-8624.2005.00853.x
Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Integration of multidimensional self-concept and core personality constructs: Construct validation and relations to well-being and achievement. *Journal of Personality*, *74*(2), 403–456. https://doi.org/10.1111/j.1467-6494.2005.00380.x
McGrew, K. S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive Gf-Gc framework. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 151–119). Guilford Press.
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), 1–10. https://doi.org/10.1016/j.intell.2008.08.004
Midgley, C., Anderman, E., & Hicks, L. (1995). Differences between elementary and middle school teachers and students: A goal theory approach. *The Journal of Early Adolescence*, *15*(1), 90–113. https://doi.org/10.1177/0272431695015001006
Möller, J., & Trautwein, U. (2015). Selbstkonzept [Self-concept]. In E. Wild & J. Möller (Eds.), *Pädagogische Psychologie* (2nd ed.). Springer.
Morris, B. U., & Bronfenbrenner, U. (2006). The bioecological model of human development. In R. M. Lerner & W. Damon (Eds.), *Handbook of child psychology: Theoretical models of human development* (Vol. 6, pp. 793–828). Wiley.
Murray, H. (1938). *Explorations in personality*. Oxford University Press.
Muthukrishna, M., Bell, A. V., Henrich, J., Curtin, C. M., Gedranovich, A., McInerney, J., & Thue, B. (2020). Beyond Western, educated, industrial, rich, and democratic (WEIRD) psychology: Measuring and mapping scales of cultural and psychological distance. *Psychological Science*, *31*(6), 678–701. https://doi.org/10.1177/0956797620916782
Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology*, *93*(1), 116–130. https://doi.org/10.1037/0022-3514.93.1.116
Odermatt, S. D., Möhring, W., Grieder, S., & Grob, A. (2022). Cognitive and developmental functions in autistic and non-autistic children and adolescents: Evidence from the Intelligence and Development Scales–2. *Journal of Intelligence*, *10*(4), Article 4. https://doi.org/10.3390/jintelligence10040112
Petermann, F., & Petermann, U. (2011). *Wechsler Intelligence Scale for Children–Fourth Edition (WISC-*

IV). German Adaptation. Pearson.

Petermann, F., & Winkel, S. (2007). *Fragebogen zur Leistungsmotivation für Schüler der 7. bis 13. Klasse [Achievement Motivation Questionnaire for 7th- to 13th-Grade Students] (FLM 7-13)*. Harcourt Test Services.

Phillipson, S. (2010). Modeling parental role in academic achievement: Comparing high ability to low and average ability students. *Talent Development & Excellence*, *2*, 83–103.

Phillipson, S., & Phillipson, S. N. (2012). Children's cognitive ability and their academic achievement: The mediation effects of parental expectations. *Asia Pacific Education Review*, *13*(3), 495–508. https://doi.org/10.1007/s12564-011-9198-1

Pomerantz, E. M., & Dong, W. (2006). Effects of mothers' perceptions of children's competence: The moderating role of mothers' theories of competence. *Developmental Psychology*, *42*(5), 950–961. https://doi.org/10.1037/0012-1649.42.5.950

Pomerantz, E. M., Moorman, E. A., & Litwack, S. D. (2007). The how, whom, and why of parents' involvement in children's academic lives: More is not always better. *Review of Educational Research*, *77*(3), 373–410. https://doi.org/10.3102/003465430305567

Poropat, A. E. (2014). A meta-analysis of adult-rated child personality and academic performance in primary education. *British Journal of Educational Psychology*, *84*(2), 239–252. https://doi.org/10.1111/bjep.12019

R Core Team. (2022). *R: A language and environment for statistical computing (Version 2022.12.0+353) [Computer software]*. R Foundation for Statistical Computing. https://www.R-project.org/

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, *138*(2), 353–387. https://doi.org/10.1037/a0026838

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, *130*(2), 261–288. https://doi.org/10.1037/0033-2909.130.2.261

Roberts, B. W., Hill, P. L., & Davis, J. P. (2017). How to change conscientiousness: The sociogenomic trait intervention model. *Personality Disorders: Theory, Research, and Treatment*, *8*, 199–205. https://doi.org/10.1037/per0000242

Roberts, B. W., Jackson, J. J., Fayard, J. V., Edmonds, G., & Meints, J. (2009). Conscientiousness. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 369–381). Guilford Press.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *132*(1), 1–25. https://doi.org/10.1037/0033-2909.132.1.1

Roberts, B. W., & Wood, D. (2006). Personality development in the context of the neo-socioanalytic model of personality. In D. Mroczek & T. Little (Eds.), *Handbook of personality development* (pp. 11–39). Erlbaum.

Roberts, B. W., & Yoon, H. J. (2022). Personality psychology. *Annual Review of Psychology*, *73*(1), 489–516. https://doi.org/10.1146/annurev-psych-020821-114927

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36.

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

Saklofske, D. H., Austin, E. J., Mastoras, S. M., Beaton, L., & Osborne, S. E. (2012). Relationships of personality, affect, emotional intelligence and coping with student stress and academic success: Different patterns of association for stress and success. *Learning and Individual Differences*, *22*(2), 251–257. https://doi.org/10.1016/j.lindif.2011.02.010

Schneider, W. J., & McGrew, K. S. (2018). The Cattell–Horn–Carroll theory of cognitive abilities. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 73–163). Guilford Press.

Schrader, F.-W. (2006). Diagnostische Kompetenzen von Eltern und Lehrern [Diagnostic competencies of parents and teachers]. In D. H. Rost (Ed.), *Handwörterbuch Pädagogische Psychologie* (3rd ed., pp. 95–100). Psychologie Verlag Union.

Spearman, C. (1904). “General intelligence,” objectively determined and measured. *American Journal of Psychology*, *15*, 201–292.

Spengler, M., Brunner, M., Martin, R., & Lüdtke, O. (2016). The role of personality in predicting (change in) students’ academic success across four years of secondary school. *European Journal of Psychological Assessment*, *32*(1), 95–103. https://doi.org/10.1027/1015-5759/a000330

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*(2), 190–204. https://doi.org/10.1016/j.cogdev.2005.01.001

Spinath, B., Spinath, F. M., Harlaar, N., & Plomin, R. (2006). Predicting school achievement from general cognitive ability, self-perceived ability, and intrinsic value. *Intelligence*, *34*(4), 363–374. https://doi.org/10.1016/j.intell.2005.11.004

Steinmayr, R., & Spinath, B. (2007). Predicting school achievement from motivation and personality. *Zeitschrift für Pädagogische Psychologie*, *21*(3/4), 207–216. https://doi.org/10.1024/1010-0652.21.3.207

Steinmayr, R., & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences*, *19*(1), 80–90. https://doi.org/10.1016/j.lindif.2008.05.004

Stieger, M., Flückiger, C., Rüegger, D., Kowatsch, T., Roberts, B. W., & Allemand, M. (2021). Changing personality traits with the help of a digital personality change intervention. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(8), Article e2017548118. https://doi.org/10.1073/pnas.2017548118

Swiss Federal Statistical Office. (2022). *Bildungsstand der Wohnbevölkerung nach Alter und Geschlecht [Educational attainment of the resident population by age and gender]*. https://www.bfs.admin.ch/bfs/de/home/statistiken/kataloge-datenbanken/tabellen.assetdetail.22024469.html

Swiss-German Conference of Directors of Education. (2016). *Lehrplan 21 [Curriculum 21]*. https://v-fe.lehrplan.ch/container/V_FE_DE_Gesamtausgabe.pdf

Trautwein, U., & Lüdtke, O. (2007). Students’ self-reported effort and time on homework in six school subjects: Between-students differences and within-student variation. *Journal of Educational Psychology*, *99*(2), 432–444. https://doi.org/10.1037/0022-0663.99.2.432

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Wigfield, A., Eccles, J. S., Fredricks, J. A., Simpkins, S., Roeser, R. W., & Schiefele, U. (2015). Development of achievement motivation and engagement. In M. E. Lamb & R. M. Lerner (Eds.), *Handbook of child psychology and developmental science: Socioemotional processes* (pp. 657–700). Wiley.

Wilmot, M. P., & Ones, D. S. (2019). A century of research on conscientiousness at work. *Proceedings of the National Academy of Sciences*, *116*(46), 23004–23010. https://doi.org/10.1073/pnas.1908430116

Table 1

Descriptive Statistics of Variables from Study 1 (n = 301) and Study 2 (n = 465)

Variable	Study 1				Study 2			
	M	*SD*	Range	ω	*M*	*SD*	Range	ω
Grade point average	4.87	0.48	3.50–6.00	.83	4.95	0.46	3.33–6.00	.88
Performance-based intelligence	101.14	13.58	64–135	.54	101.17	14.86	55–141	.70
Parent-reported conscientiousness	29.93	5.37	6–40	.84	28.33	6.61	9–40	.91
Parent-reported academic performance	3.47	0.58	1.60–5.00	.71	3.41	0.58	1.40–5.00	.77
Self-reported achievement striving motivation	25.86	5.47	9–39	.79	26.37	5.57	12–39	.82
Self-reported academic performance	3.44	0.57	2.00–5.00	.72	3.41	0.52	2.00–5.00	.60

Note. Omega reliabilities (ω) of intelligence from Study 1 for each age group separately: 5-year-olds: sample size too small; 6-year-olds: ω = .67; 7-year-olds: ω = .46; 8-year-olds: ω = .80; 9-year-olds: ω = .66; 10-year-olds: ω = .70.

Table 2

Intercorrelations of Variables and Demographic Characteristics from Study 1 (n = 301) and Study 2 (n = 465)

	Variable	1	2	3	4	5	6	7	8
1	Grade point average		.37***	.45***	.62***	.25***	.41***	.15	.10
2	Performance-based intelligence	.25***		.20**	.42***	-.04	.23**	.13*	.18***
3	Parent-reported conscientiousness	.31***	.01		.44***	.16*	.24**	.19**	-.03
4	Parent-reported academic performance	.54***	.34***	.16**		.31***	.65***	-.03	.23**
5	Self-reported achievement striving motivation	.28***	.08	.07	.30***		.34***	-.01	-.04
6	Self-reported academic performance	.42***	.23***	.04	.51***	.32***		-.08	.13
7	Sex	.10	.00	.24***	-.13*	-.24***	-.14*		.00
8	Maternal educational attainment	.09	.27***	-.08	.21***	.01	.11	.00	

Note. Correlations below the diagonal represent coefficients for Study 1 and correlations above the diagonal represent coefficients for Study 2.
Sex: 0 = male, 1 = female; maternal educational attainment: 0 = no postsecondary education, 1 = postsecondary education.
*p < .05. **p < .01. ***p < .001.

Table 3

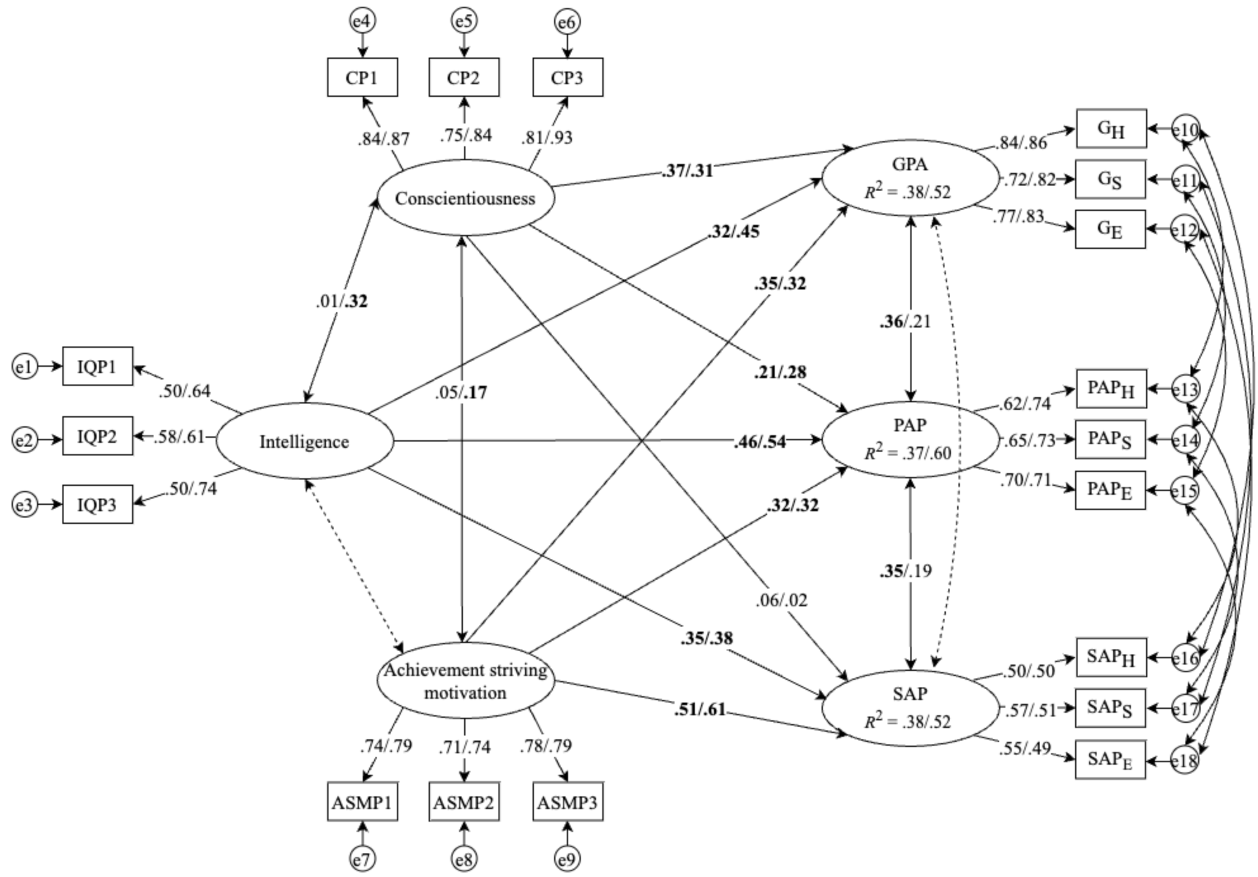
Hierarchical Regression Analyses of Study 1 (n = 301) and Study 2 (n = 465) Focusing on Overall Estimates

Variable	Grade point average					Parent-reported academic performance					Self-reported academic performance				
	β	*B*	95% CI for *B*		*p*	β	*B*	95% CI for *B*		*p*	β	*B*	95% CI for *B*		*p*
			LL	*UL*				*LL*	*UL*				*LL*	*UL*	
Step 1															
Sex	.10/.14	0.10/0.13	-0.02/-0.01	0.21/0.26	.113/.064	-.12/-.04	-0.14/-0.04	-0.28/-0.21	-0.01/0.12	.031/.612	**-.14**/-.08	-0.16/-0.08	-0.29/-0.23	-0.03/0.07	.014/.271
MEA	.09/.10	0.10/0.09	-0.04/-0.05	0.24/0.23	.147/.183	**.21/.23**	0.28/0.26	0.13/0.08	0.44/0.44	<.001/.006	.11/.14	0.15/0.14	-0.01/-0.01	0.30/0.30	.062/.068
R^2			.02/.03					.06/.06					.03/.02		
Step 2															
Sex	.10/.10	0.10/0.09	-0.02/-0.03	0.22/0.22	.115/.150	**-.13**/-.08	-0.15/-0.08	-0.28/-0.24	-0.02/0.07	.020/.299	**-.16**/-.10	-0.18/-0.10	-0.31/-0.25	-0.05/0.05	.009/.175
MEA	.04/.02	0.04/0.02	-0.10/-0.12	0.19/0.16	.562/.768	**.14**/.15	0.20/0.17	0.04/-0.01	0.36/0.34	.015/.061	.07/.11	0.10/0.11	-0.06/-0.04	0.27/0.27	.222/.157
Intelligence	**.21/.36**	0.01/0.01	0.00/0.01	0.01/0.02	.002/<.001	**.29/.39**	0.01/0.01	0.01/0.01	0.02/0.02	<.001/<.001	**.18/.23**	0.01/0.01	0.00/0.00	0.01/0.01	.003/.003
R^2			.06/.15					.14/.20					.07/.07		
ΔR^2			.04/.12					.08/.14					.04/.05		
Step 3															
Sex	.02/.04	0.02/0.04	-0.10/-0.09	0.14/0.16	.757/.572	**-.19**/-.12	-0.22/-0.12	-0.35/-0.28	-0.09/0.03	<.001/.110	**-.17**/-.08	-0.20/-0.07	-0.33/-0.27	-0.06/0.12	.006/.464
MEA	.06/.05	0.07/0.05	-0.07/-0.08	0.21/0.18	.341/.440	**.16/.17**	0.22/0.19	0.06/0.02	0.37/0.35	.006/.029	.08/.09	0.11/0.09	-0.06/-0.11	0.27/0.28	.198/.382
Intelligence	**.19/.32**	0.01/0.01	0.00/0.01	0.01/0.01	.003/< .001	**.28/.35**	0.01/0.01	0.01/0.01	0.02/0.02	<.001/<.001	**.18/.24**	0.01/0.01	0.00/0.00	0.01/0.02	.003/.025
Conscientiousness	**.30/.31**	0.03/0.02	0.02/0.01	0.04/0.03	<.001/.001	**.23/.31**	0.02/0.03	0.01/0.01	0.04/0.04	<.001/<.001	.06/**.33**	0.01/0.02	-0.01/0.01	0.02/0.04	.354/.002
R^2			.14/.25					.19/.29					.07/.19		
ΔR^2			.08/.10					.05/.09					.00/.12		
Step 4															
Sex	.09/.04	0.09/0.04	-0.03/-0.09	0.21/0.16	.161/.543	**-.14**/-.11	-0.16/-0.11	-0.30/-0.27	-0.03/0.04	.020/.152	-.09/-.10	-0.11/-0.10	-0.25/-0.30	0.03/0.10	.130/.335
MEA	.06/.06	0.07/0.06	-0.06/-0.07	0.21/0.19	.301/.399	**.16/.16**	0.21/0.17	0.06/0.01	0.37/0.34	.006/.040	.07/.08	0.10/0.08	-0.06/-0.12	0.26/0.28	.217/.419
Intelligence	**.19/.31**	0.01/0.01	0.00/0.01	0.01/0.01	.003/<.001	**.28/.34**	0.01/0.01	0.01/0.01	0.02/0.02	<.001/<.001	.14/.26	0.01/0.01	0.00/0.00	0.01/0.02	.019/.017
Conscientiousness	**.25/.29**	0.02/0.02	0.01/0.01	0.03/0.03	<.001/<.001	**.19/.29**	0.02/0.02	0.01/0.01	0.03/0.04	<.001/<.001	.03/**.26**	0.00/0.02	-0.01/0.00	0.02/0.04	.667/.022
Achievement striving motivation	**.27/.22**	0.02/0.02	0.01/0.01	0.03/0.03	<.001/.002	**.23/.17**	0.02/0.02	0.01/0.00	0.04/0.03	<.001/.025	**.28**/.19	0.03/0.02	0.02/-0.00	0.04/0.04	<.001/.087
R^2			.21/.27					.25/.30					.14/.21		
ΔR^2			.07/.02					.06/.01					.07/.02		

Note. The standardized coefficients for Study 1 are shown left of the slash, those for Study 2 right of the slash. Sex: 0 = male, 1 = female; MEA = maternal educational attainment: 0 = no postsecondary education, 1 = postsecondary education; CI = confidence interval; LL = lower limit; UL = upper limit. Standardized coefficients in bold are significant ($p < .05$).

Figure 1

Structural Equation Model of Study 1 (n = 301) and Study 2 (n = 465) Focusing on Overall Estimates (Standardized Coefficients for Study 1 Are Shown Left of the Slash, Those for Study 2 Right of the Slash)



Note. Model fit was χ^2 = 203.186, *p* < .001, *df* = 113, comparative fit index (CFI) = .943, root-mean-square error of approximation (RMSEA) = .051, standardized root-mean-square residual (SRMR) = .061 for Study 1 and χ^2 = 182.655, *p* < .001, *df* = 113, CFI = .968, RMSEA = .037, SRMR = .053 for Study 2. Nonsignificant covariances between grade point average (GPA) and self-reported academic performance (SAP) and between intelligence and achievement striving motivation were restricted to zero (see dashed lines). Correlations of residuals (e) are shown in Table S7. Significant coefficients with a *p* value below .05 are depicted in bold. In Study 1, intelligence was measured 7 years before the other constructs (Time 1); all other constructs were assessed at the same time (Time 2). In Study 2, all constructs were assessed at the same time (Time 1). IQP1 = intelligence parcel 1; IQP2 = intelligence parcel 2; IQP3 = intelligence parcel 3; CP1 = conscientiousness parcel 1; CP2 = conscientiousness parcel 2; CP3 = conscientiousness parcel 3; ASMP1 = achievement striving motivation parcel 1; ASMP2 = achievement striving motivation parcel 2; ASMP3 = achievement striving motivation parcel 3; G_H = grade in humanities; G_S = grade in science; G_E = grade in environment; PAP = parent-reported academic performance; PAP_H = PAP in humanities; PAP_S = PAP in science; PAP_E = PAP in environment; SAP = self-reported academic performance; SAP_H = SAP in humanities; SAP_S = SAP in science; SAP_E = SAP in environment.

Supplementary Material

Supplemental Methods: Structural Equation Modeling

We used the comparative fit index (CFI), the root-mean-square error of approximation (RMSEA), and the standardized root-mean-square residual (SRMR) for model evaluation. According to Hu and Bentler (1999) and van de Schoot et al. (2012), CFIs above .95, RMSEAs below .06, and SRMRs below .08 were considered a good model fit and CFIs above .90 as satisfactory. To deal with missing data that resulted from parents or adolescents not completing the questionnaires or test sessions where not all intelligence subtests were administered, we used the full-information maximum likelihood approach.

The latent factor grade point average (GPA) was created by the grades in *humanities*, *science*, and *environment*. In addition, the latent parent-reported academic performance factor and the self-reported academic performance factor were generated by the performance ratings from parents and students in the three school-subject domains, respectively. For the *humanities* domain, we used the grades in *language* and *geography and history* to model the latent factor GPA in *humanities* as well as the performance ratings from parents and students in the two subjects to create the parent-reported and self-reported academic performance factor in *humanities*, respectively. For the *science* domain, we used the grades in *mathematics* and *biology, chemistry, and physics* to model the latent factor GPA in *science* as well as the performance ratings from parents and students in the two subjects to create the parent-reported and self-reported academic performance factor in *science*, respectively.

We allowed the residuals of the corresponding school-subject domains and subjects of the grades, performance ratings by parents, and performance ratings by adolescents to covary as the wording of the items was similar and they represented the same subject domains and subjects. Nonsignificant covariances between latent factors, when found for both studies, were adjusted to zero in the structural equation models to simplify the models and reduce model complexity according to the principle of parsimony. Descriptive statistics and intercorrelations of indicator variables are shown in Tables S5 and S6 for both studies.

To compare the regression coefficients of the models of Studies 1 and 2 focusing on overall estimates and thus including all school-subject domains, we used two approaches: First, we calculated the 95% confidence intervals of the regression coefficients for both models. If the 95% confidence intervals showed at least a moderate overlap, which was defined as half the length of the average margin of error (i.e., half the length of the confidence interval; Cumming, 2012), we concluded that the coefficients of the two studies were not significantly different. Second, we compared a multigroup structural equation model without any constraints with a multigroup structural equation model with constrained regression coefficients across the two studies (except for the intelligence parameters, as different intelligence test batteries were used in the two studies) by calculating a χ^2 difference test. If the χ^2 difference test was nonsignificant, we concluded that there were no differences between the two studies that were significantly different from each other.

Supplemental Results

Table S1

Description of the Intelligence Subtests of the Intelligence and Development Scales (Grob et al., 2009)

Subtest	No. of items	Task description
Visual Perception	7	Organize cards printed with lines of different lengths in consecutive order
Selective Attention	225	Cross out targets (ducks looking to the right with two orange characteristics) as quickly as possible
Phonological Memory	12	Repeat number and letter sequences
Visuospatial Memory	11	Remember geometric figures in a selection of figures
Auditory Memory	24	Listen to a story and recall it after a period of time
Visual Reasoning	10	Reproduce presented geometric figures with the help of rectangles and/or triangles
Conceptual Reasoning	12	Understand the concept of three pictures and choose two corresponding pictures from a selection of pictures

Table S2

Description of the Intelligence Subtests of the Intelligence and Development Scales–2 (Grob & Hagmann-von Arx, 2018)

Subtest	No. of items	Task description
Shape Design	20	Reproduce presented geometric figures with the help of rectangles and/or triangles
Parrots	56–180[a]	Cross out targets (parrots looking to the left with two orange characteristics) as quickly as possible
Digit and Letter Span	40	Repeat number and letter sequences
Shape Memory	23	Remember geometric figures in a selection of figures
Matrices: Completion	35	Understand how a figure changes and transfer these changes to a continuing figure
Naming Categories	34	Name categories for a group of pictures or words
Story Recall	19–32[a]	Listen to a story and recall it after a period of time

Note. These intelligence subtests belong to the Full-Scale IQ of the Intelligence and Development Scales–2.
[a] Depending on age.

Table S3

Hierarchical Regression Analyses of Study 1 (n = 301) and Study 2 (n = 465) Focusing on the School-Subject Domain Humanities

Variable	Grade point average					Parent-reported academic performance					Self-reported academic performance				
	β	*B*	95% CI for *B*		*p*	β	*B*	95% CI for *B*		*p*	β	*B*	95% CI for *B*		*p*
			LL	*UL*				*LL*	*UL*				*LL*	*UL*	
Step 1															
Sex	**.17/.22**	0.14/0.20	0.03/0.06	0.25/0.33	.013/.005	.00/.10	0.00/0.13	-0.16/-0.06	0.16/0.31	.967/.187	-.05/-.01	-0.06/-0.01	-0.22/-0.15	0.09/0.12	.410/.865
MEA	.05/.09	0.05/0.08	-0.08/-0.06	0.18/0.22	.439/.250	.09/.13	0.14/0.18	-0.05/-0.01	0.32/0.37	.150/.068	.08/.05	0.12/0.06	-0.06/-0.08	0.29/0.20	.202/.414
R^2			.03/.06					.01/.03					.01/.00		
Step 2															
Sex	**.17/.20**	0.14/0.18	0.03/0.05	0.25/0.31	.013/.007	.00/.07	0.00/0.09	-0.16/-0.10	0.17/0.27	.965/.360	-.07/-.02	-0.08/-0.03	-0.24/-0.17	0.07/0.11	.292/.682
MEA	.02/.02	0.02/0.02	-0.11/-0.12	0.15/0.15	.792/.811	.07/.08	0.12/0.11	-0.08/-0.08	0.32/0.30	.250/.260	.07/.02	0.10/0.03	-0.09/-0.11	0.29/0.18	.314/.655
Intelligence	**.20/.34**	0.01/0.01	0.00/0.01	0.01/0.02	.005/<.001	**.15/.27**	0.01/0.01	0.00/0.01	0.01/0.02	.020/<.001	**.18/.15**	0.01/0.01	0.00/0.00	0.01/0.01	.006/.010
R^2			.07/.17					.03/.10					.05/.02		
ΔR^2			.04/.11					.02/.07					.04/.02		
Step 3															
Sex	.08/.13	0.06/0.12	-0.05/-0.01	0.17/0.25	.262/.067	-.05/.02	-0.07/0.02	-0.23/-0.16	0.10/0.21	.421/.823	-.08/-.04	-0.10/-0.05	-0.27/-0.24	0.06/0.14	.207/.573
MEA	.04/.04	0.04/0.04	-0.08/-0.09	0.16/0.17	.526/.555	.09/.09	0.14/0.12	-0.06/-0.06	0.34/0.31	.162/.196	.07/-.02	0.11/-0.02	-0.08/-0.22	0.30/0.18	.271/.836
Intelligence	**.18/.30**	0.01/0.01	0.00/0.00	0.01/0.01	.009/<.001	**.14/.22**	0.01/0.01	0.00/0.00	0.01/0.02	.022/.003	**.18**/.13	0.01/0.01	0.00/-0.00	0.01/0.01	.007/.099
Conscientiousness	**.35/.31**	0.03/0.02	0.02/0.01	0.04/0.03	<.001/<.001	**.20/.30**	0.03/0.03	0.01/0.02	0.04/0.04	.001/<.001	.06/**.18**	0.01/0.02	-0.01/0.00	0.02/0.03	.327/.019
R^2			.19/.26					.07/.17					.05/.05		
ΔR^2			.12/.09					.04/.07					.00/.03		
Step 4															
Sex	**.14**/.13	0.11/0.11	0.00/-0.01	0.22/0.24	.042/.079	-.01/.02	-0.02/0.02	-0.19/-0.17	0.15/0.21	.826/.836	-.01/-.02	-0.01/-0.03	-0.17/-0.21	0.16/0.15	.931/.750
MEA	.04/.06	0.03/0.05	-0.09/-0.08	0.15/0.18	.589/.447	.09/.11	0.14/0.14	-0.06/-0.05	0.33/0.33	.166/.148	.06/.01	0.09/0.01	-0.10/-0.18	0.28/0.20	.347/.918
Intelligence	**.17/.29**	0.01/0.01	0.00/0.00	0.01/0.01	.010/<.001	**.15/.21**	0.01/0.01	0.00/0.00	0.01/0.02	.015/.006	**.16**/.12	0.01/0.01	0.00/-0.00	0.01/0.01	.013/.105
Conscientiousness	**.30/.30**	0.02/0.02	0.01/0.01	0.03/0.03	<.001/<.001	**.18/.28**	0.02/0.03	0.01/0.01	0.04/0.04	.005/<.001	.02/.14	0.00/0.01	-0.01/-0.00	0.02/0.03	.717/.061
Achievement striving motivation	**.24/.19**	0.02/0.02	0.01/0.00	0.03/0.03	<.001/.007	**.14**/.12	0.02/0.01	0.00/-0.00	0.03/0.03	.025/.091	**.23/.32**	0.03/0.04	0.01/0.02	0.04/0.05	<.001/<.001
R^2			.24/.28					.09/.17					.10/.14		
ΔR^2			.05/.02					.02/.00					.05/.09		

Note. The standardized coefficients for Study 1 are shown left of the slash, those for Study 2 right of the slash. Sex: 0 = male, 1 = female; MEA = maternal educational attainment: 0 = no postsecondary education, 1 = postsecondary education; CI = confidence interval; LL = lower limit; UL = upper limit. Standardized coefficients in bold are significant ($p < .05$).

Table S4

Hierarchical Regression Analyses of Study 1 (n = 301) and Study 2 (n = 465) Focusing on the School-Subject Domain Science

Variable	Grade point average					Parent-reported academic performance					Self-reported academic performance				
	β	*B*	95% CI for *B*		*p*	β	*B*	95% CI for *B*		*p*	β	*B*	95% CI for *B*		*p*
			LL	*UL*				*LL*	*UL*				*LL*	*UL*	
Step 1															
Sex	-.03/.05	-0.03/0.05	-0.20/-0.11	0.13/0.22	.694/.527	**-.16**/-.08	-0.26/-0.11	-0.43/-0.31	-0.08/0.08	.005/.262	**-.24/-.23**	-0.36/-0.34	-0.54/-0.50	-0.18/-0.18	<.001/<.001
MEA	.08/.09	0.10/0.09	-0.09/-0.08	0.29/0.26	.284/.285	**.12/.21**	0.23/0.30	0.02/0.10	0.44/0.51	.033/.004	**.15**/.07	0.26/0.10	0.06/-0.06	0.47/0.27	.013/.226
R^2			.01/.01					.04/.05					.08/.06		
Step 2															
Sex	-.04/.02	-0.04/0.02	-0.21/-0.14	0.13/0.18	.616/.802	**-.16**/-.13	-0.25/-0.17	-0.43/-0.36	-0.07/0.01	.007/.062	**-.24/-.26**	-0.36/-0.38	-0.55/-0.54	-0.18/-0.23	<.001/<.001
MEA	.02/.01	0.02/0.02	-0.18/-0.15	0.22/0.18	.821/.851	.05/.12	0.09/0.17	-0.13/-0.02	0.31/0.36	.412/.087	.11/.02	0.20/0.02	-0.02/-0.14	0.42/0.19	.080/.764
Intelligence	**.18/.33**	0.01/0.01	0.00/0.01	0.01/0.02	.025/<.001	**.25/.42**	0.01/0.02	0.01/0.01	0.02/0.03	<.001/<.001	**.16/.25**	0.01/0.01	0.00/0.01	0.02/0.02	.016/<.001
R^2			.03/.11					.10/.22					.10/.12		
ΔR^2			.02/.10					.06/.17					.02/.06		
Step 3															
Sex	-.10/-.06	-0.11/-0.07	-0.28/-0.22	0.06/0.09	.189/.401	**-.19/-.18**	-0.30/-0.25	-0.49/-0.43	-0.12/-0.07	.001/.008	**-.26/-.25**	-0.39/-0.34	-0.58/-0.54	-0.20/-0.14	<.001/<.001
MEA	.04/.05	0.05/0.06	-0.15/-0.11	0.24/0.22	.627/.496	.06/.13	0.11/0.18	-0.11/-0.01	0.33/0.37	.327/.060	.12/.02	0.21/0.02	-0.01/-0.18	0.43/0.23	.065/.815
Intelligence	**.17/.30**	0.01/0.01	0.00/0.01	0.01/0.02	.029/<.001	**.25/.38**	0.01/0.02	0.01/0.01	0.02/0.02	<.001/<.001	**.15/.30**	0.01/0.01	0.00/0.01	0.02/0.02	.017/<.001
Conscientiousness	**.27/.31**	0.03/0.02	0.01/0.01	0.04/0.04	<.001/<.001	**.13/.24**	0.02/0.03	0.00/0.01	0.04/0.04	.030/<.001	.08/.15	0.01/0.02	-0.01/0.00	0.03/0.03	.201/.045
R^2			.10/.22					.11/.27					.11/.16		
ΔR^2			.07/.11					.01/.05					.01/.04		
Step 4															
Sex	-.05/-.07	-0.06/-0.08	-0.22/-0.23	0.11/0.08	.508/.327	**-.14/-.18**	-0.23/-0.24	-0.42/-0.43	-0.03/-0.06	.022/.011	**-.19/-.23**	-0.29/-0.32	-0.48/-0.52	-0.10/-0.12	.004/.002
MEA	.03/.06	0.04/0.07	-0.15/-0.09	0.22/0.22	.679/.410	.05/**.14**	0.09/0.20	-0.13/0.01	0.31/0.39	.413/.043	.12/.02	0.21/0.03	-0.01/-0.17	0.42/0.24	.061/.758
Intelligence	**.18/.29**	0.01/0.01	0.00/0.01	0.01/0.02	.018/<.001	**.27/.38**	0.02/0.02	0.01/0.01	0.02/0.03	<.001/<.001	**.14/.32**	0.01/0.02	0.00/0.01	0.01/0.02	.028/<.001
Conscientiousness	**.22/.27**	0.02/0.02	0.01/0.01	0.04/0.03	.004/<.001	.11/**.22**	0.02/0.02	-0.00/0.01	0.03/0.04	.076/.002	.06/.12	0.01/0.01	-0.01/-0.00	0.03/0.03	.349/.122
Achievement striving motivation	**.25/.25**	0.03/0.02	0.01/0.01	0.04/0.04	<.001/.001	**.16/.18**	0.02/0.02	0.01/0.01	0.04/0.04	.009/.009	**.28/.25**	0.04/0.03	0.02/0.01	0.06/0.05	<.001/<.001
R^2			.16/.25					.15/.28					.18/.21		
ΔR^2			.06/.03					.04/.01					.07/.05		

Note. The standardized coefficients for Study 1 are shown left of the slash, those for Study 2 right of the slash. Sex: 0 = male, 1 = female; MEA = maternal educational attainment: 0 = no postsecondary education, 1 = postsecondary education; CI = confidence interval; LL = lower limit; UL = upper limit. Standardized coefficients in bold are significant ($p < .05$).

Table S5

Descriptive Statistics of Indicators From Study 1 (n = 301) and Study 2 (n = 465)

Latent factor	Indicator	No. of items/subtests	Study 1				Study 2			
			M	*SD*	Range	ω	*M*	*SD*	Range	ω
Performance-based intelligence	IQP1	3	10.17	1.84	5.33–15.67	.75	10.32	2.24	2.00–16.00	.65
	IQP2	2	10.27	2.05	4.50–15.00	.24	10.35	2.45	3.50–17.50	.44
	IQP3	2	9.99	2.18	4.00–15.00	.32	10.71	2.54	1.50–19.00	.59
Parent-reported conscientiousness	CP1	3	3.46	0.76	1.33–5.00	.73	3.29	0.91	1.00–5.00	.84
	CP2	3	3.96	0.69	1.67–5.00	.68	3.68	0.83	1.00–5.00	.77
	CP3	2	3.89	0.81	1.00–5.00	.70	3.71	0.99	1.00–5.00	.86
Self-reported achievement striving motivation	ASMP1	3	3.23	0.74	1.00–5.00	.51	3.37	0.71	1.33–5.00	.57
	ASMP2	3	3.42	0.78	1.00–5.00	.57	3.39	0.79	1.00–5.00	.53
	ASMP3	2	2.98	0.98	1.00–5.00	.64	3.04	0.99	1.00–5.00	.66
Grade point average	G_H	2	4.88	0.42	3.75–6.00	.71	4.96	0.46	3.50–6.00	.78
	G_S	2	4.85	0.57	3.50–6.00	.72	4.96	0.51	3.50–6.00	.76
	G_E	1	5.05	0.53	3.50–6.00	–	5.04	0.61	3.25–6.00	–
Parent-reported academic performance	PAP_H	2	3.47	0.68	1.50–5.00	.41	3.30	0.68	1.50–5.00	.66
	PAP_S	2	3.50	0.78	1.50–5.00	.44	3.47	0.72	1.00–5.00	.67
	PAP_E	1	3.54	0.69	2.00–5.00	–	3.50	0.70	2.00–5.00	–
Self-reported academic performance	SAP_H	2	3.42	0.64	2.00–5.00	.47	3.35	0.66	1.50–5.00	.47
	SAP_S	2	3.40	0.77	1.50–5.00	.48	3.41	0.76	1.00–5.00	.54
	SAP_E	1	3.66	0.76	2.00–5.00	–	3.50	0.75	2.00–5.00	–

Note. For G_E, PAP_E, and SAP_E, we did not compute omega reliabilities (ω) as only one item per indicator was used. IQP1 = intelligence parcel 1; IQP2 = intelligence parcel 2; IQP3 = intelligence parcel 3; CP1 = conscientiousness parcel 1; CP2 = conscientiousness parcel 2; CP3 = conscientiousness parcel 3; ASMP1 = achievement striving motivation parcel 1; ASMP2 = achievement striving motivation parcel 2; ASMP3 = achievement striving motivation parcel 3; G_H = grade in humanities; G_S = grade in science; G_E = grade in environment; PAP_H = parent-reported academic performance in humanities; PAP_S = parent-reported academic performance in science; PAP_E = parent-reported academic performance in environment; SAP_H = self-reported academic performance in humanities; SAP_S = self-reported academic performance in science; SAP_E = self-reported academic performance in environment.

Table S6

Intercorrelations of Indicators From Study 1 (n = 301) and Study 2 (n = 465)

	Indicator	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	IQP1		.41***	.47***	.02	.22**	.13	-.05	.06	-.06	.20**	.22**	.20	.18*	.35***	.20*	.03	.28***	.02
2	IQP2	.31***		.43***	.06	.20**	.17*	-.08	-.03	-.03	.26***	.23**	.27*	.22**	.31***	.19*	.09	.07	.04
3	IQP3	.28***	.27***		.10	.30***	.18*	-.07	-.02	.00	.34***	.34***	.32**	.34***	.35***	.27***	.21***	.14**	.03
4	CP1	.08	.02	-.03		.71***	.82***	.12	.15*	.13*	.32***	.29***	.30**	.31***	.24***	.27***	.11	.08	.04
5	CP2	.10	.03	-.02	.61***		.78***	.05	.12	.13*	.42***	.41***	.37***	.44***	.39***	.32***	.19**	.14*	.11
6	CP3	.01	-.05	-.03	.69***	.59***		.07	.11	.10	.38***	.33***	.28**	.36***	.30***	.26***	.15*	.11	.07
7	ASMP1	.01	.04	-.02	.03	.15**	-.07		.59***	.63***	.20**	.22**	.15	.12	.13*	.18**	.25***	.25***	.16**
8	ASMP2	.16**	.12*	.05	.06	.11	-.08	.54***		.59***	.06	.19**	.16	.12	.26***	.19**	.22***	.34***	.11
9	ASMP3	-.01	-.01	.09	.08	.07	-.05	.59***	.54***		.25***	.29***	.26*	.20**	.23***	.27***	.21***	.33***	.21***
10	G_H	.13	.21**	.15*	.27***	.41***	.25***	.17*	.28***	.26***		.68***	.77***	.55***	.34***	.40***	.35***	.17*	.09
11	G_S	.26**	.11	.02	.26***	.36***	.15*	.16*	.36***	.23**	.60***		.70***	.32***	.56***	.34***	.22**	.46***	.13
12	G_E	.00	.21*	.12	.16	.27**	.18*	.18*	.27**	.23**	.69***	.60***		.41***	.43***	.50***	.15	.17	.34*
13	PAP_H	.08	.17**	.18**	.17**	.16**	.11	.05	.22***	.21***	.52***	.26***	.39***		.49***	.57***	.46***	.12	.19*
14	PAP_S	.23***	.18**	.18**	.09	.14*	.05	.06	.31***	.18**	.33***	.56***	.32***	.38***		.51***	.21**	.62***	.20*
15	PAP_E	.11	.23***	.12	.10	.17**	.08	.12	.28***	.17*	.41***	.37***	.55***	.45***	.47***		.28***	.23**	.34***
16	SAP_H	.08	.16*	.20**	.09	.04	-.02	.19**	.23***	.24***	.34***	-.01	.38***	.52***	.10	.26***		.15**	.32***
17	SAP_S	.19**	.14*	.05	.06	.06	-.02	.22***	.35***	.31***	.21**	.56***	.21*	.16*	.64***	.29***	.15*		.34***
18	SAP_E	.04	.12	.07	.01	.06	-.10	.06	.06	.08	.19*	.19	.35***	.26**	.21**	.39***	.43***	.44***	

Note. Correlations below the diagonal represent coefficients for Study 1 and correlations above the diagonal represent coefficients for Study 2. IQP1 = intelligence parcel 1; IQP2 = intelligence parcel 2; IQP3 = intelligence parcel 3; CP1 = conscientiousness parcel 1; CP2 = conscientiousness parcel 2; CP3 = conscientiousness parcel 3; ASMP1 = achievement striving motivation parcel 1; ASMP2 = achievement striving motivation parcel 2; ASMP3 = achievement striving motivation parcel 3; G_H = grade in humanities; G_S = grade in science; G_E = grade in environment; PAP_H = parent-reported academic performance in humanities; PAP_S = parent-reported academic performance in science; PAP_E = parent-reported academic performance in environment; SAP_H = self-reported academic performance in humanities; SAP_S = self-reported academic performance in science; SAP_E = self-reported academic performance in environment.

*$p < .05$. **$p < .01$. ***$p < .001$.

Table S7

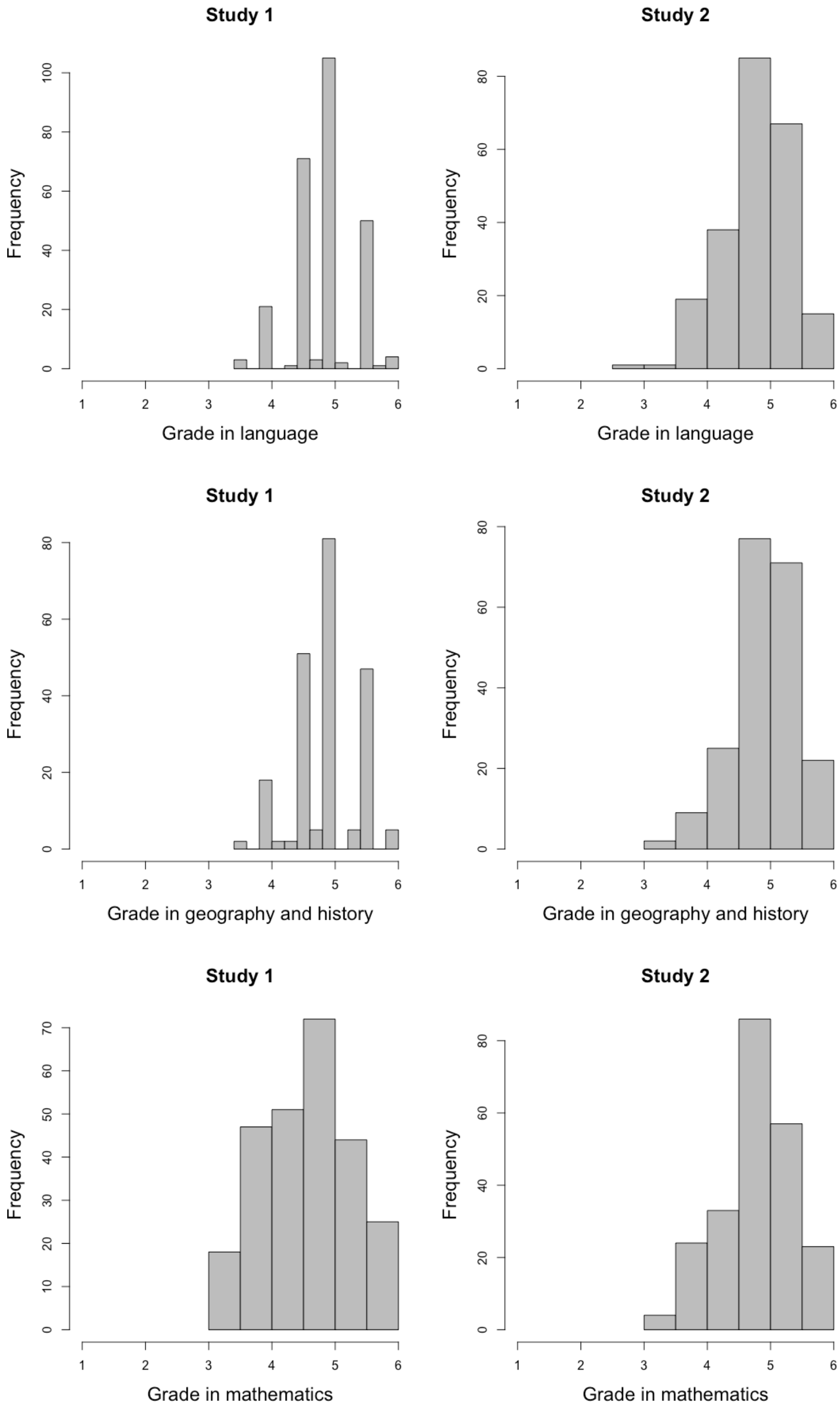
Correlations of the Residuals From the Structural Equation Models from Study 1 (n = 301) and Study 2 (n = 465) Focusing on Overall Estimates (see Figure 1)

Residual	Study 1		Study 2	
	r	*p*	*r*	*p*
e10 ⟷ e13	.43	< .001	.49	< .001
e10 ⟷ e16	.36	< .001	.36	< .01
e11 ⟷ e14	.48	< .001	.55	< .001
e11 ⟷ e17	.58	< .001	.57	< .001
e12 ⟷ e15	.50	< .001	.47	< .01
e12 ⟷ e18	.33	< .01	.57	< .001
e13 ⟷ e16	.52	< .001	.45	< .001
e14 ⟷ e17	.69	< .001	.72	< .001
e15 ⟷ e18	.29	< .01	.25	< .05

Note. e10 is the residual of G_H, e11 is the residual of G_S, e12 is the residual of G_E, e13 is the residual of PAP_H, e14 is the residual of PAP_S, e15 is the residual of PAP_E, e16 is the residual of SAP_H, e17 is the residual of SAP_S, and e18 is the residual of SAP_E. G_H = grade in humanities; G_S = grade in science; G_E = grade in environment; PAP_H = parent-reported academic performance in humanities; PAP_S = parent-reported academic performance in science; PAP_E = parent-reported academic performance in environment; SAP_H = self-reported academic performance in humanities; SAP_S = self-reported academic performance in science; SAP_E = self-reported academic performance in environment.

Figure S1

Histogram of Grades in Different Subjects of Study 1 (n = 301) and Study 2 (n = 465)



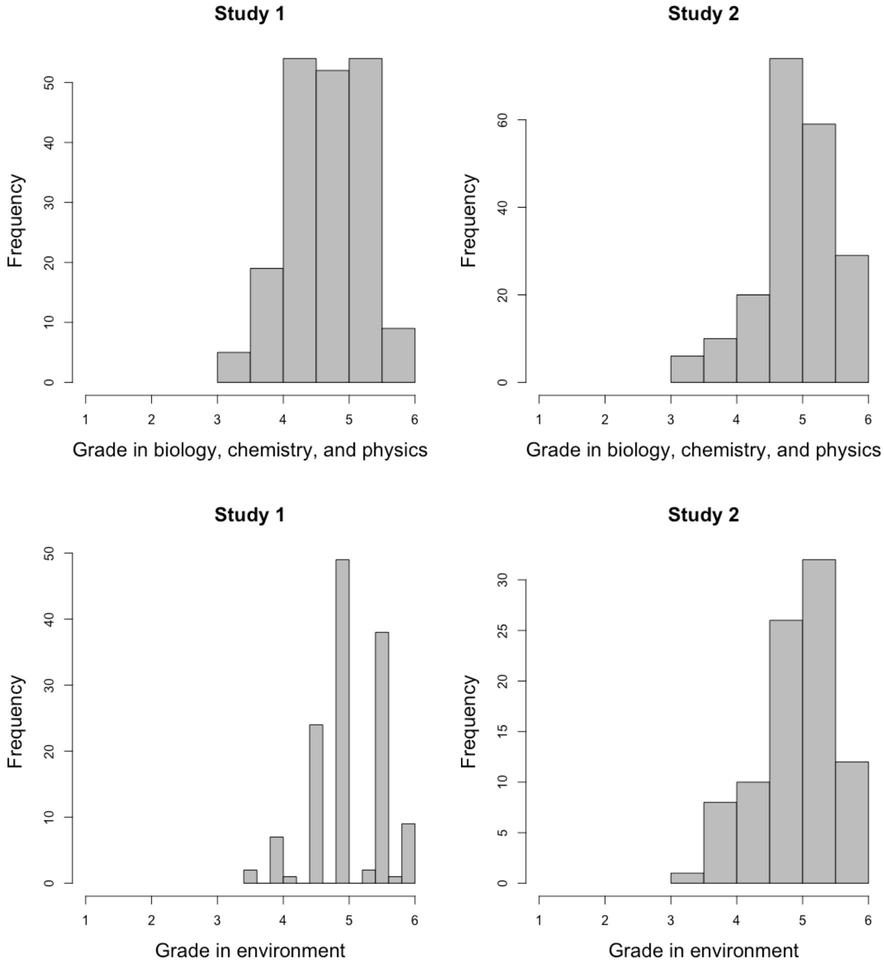
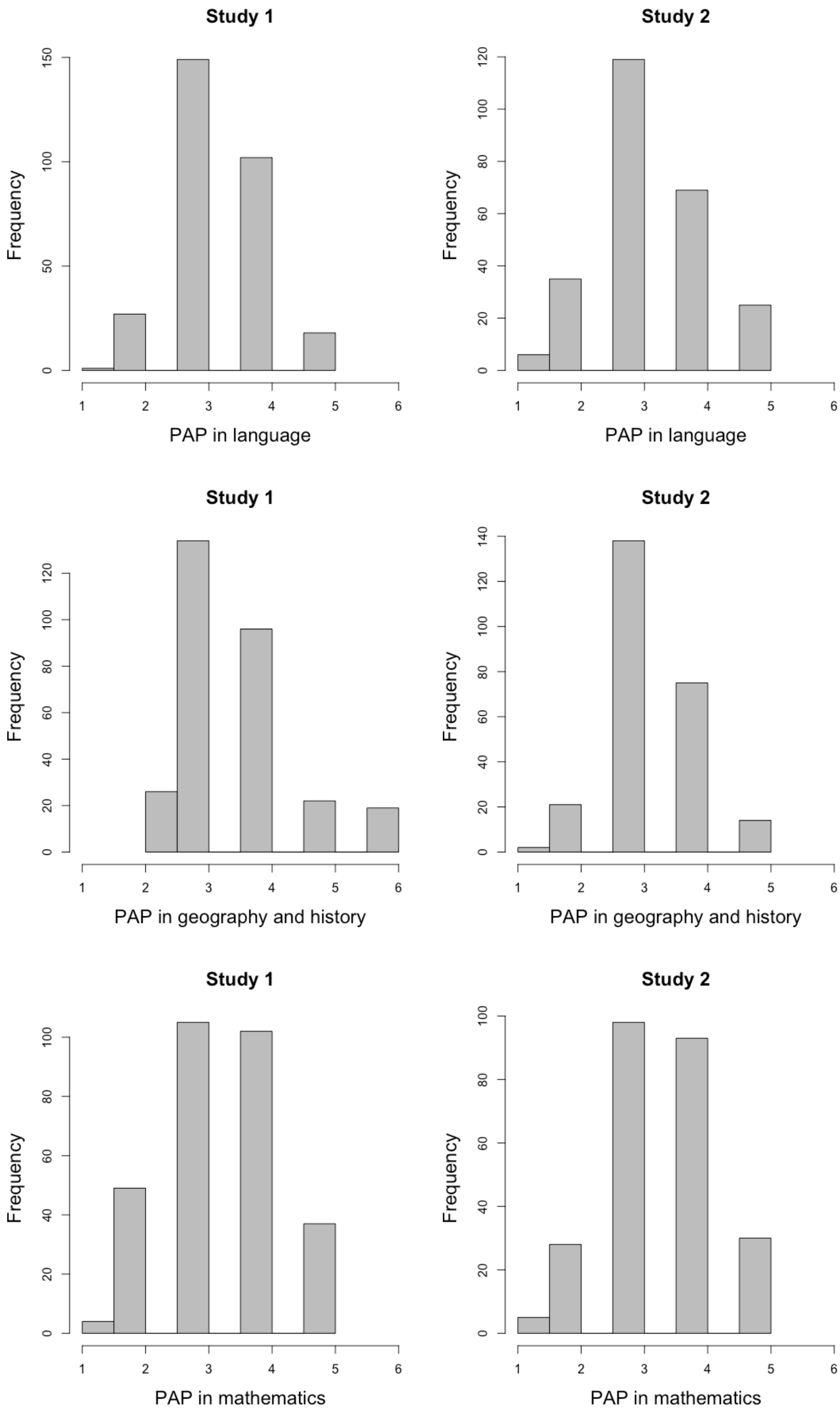

Study 1
Study 2
Frequency
Grade in biology, chemistry, and physics
Study 1
Study 2
Frequency
Grade in environment

Figure S2

Histogram of Parent-Reported Academic Performance (PAP) in Different Subjects of Study 1 (n = 301) and Study 2 (n = 465)



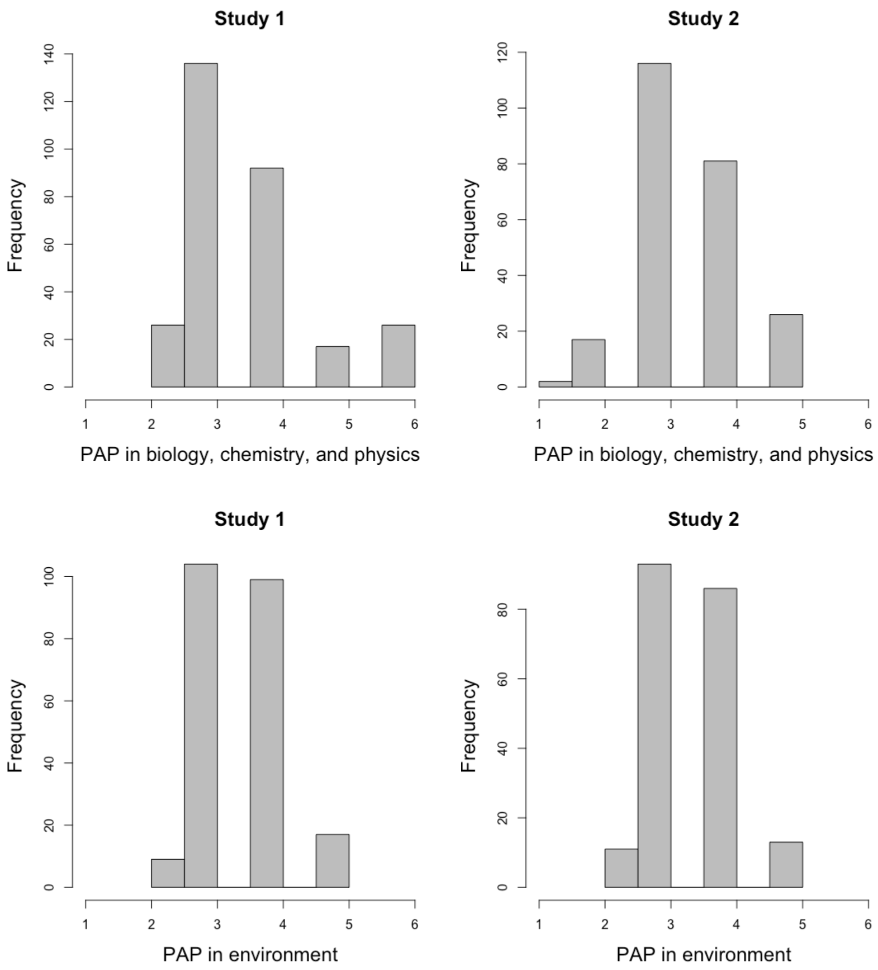
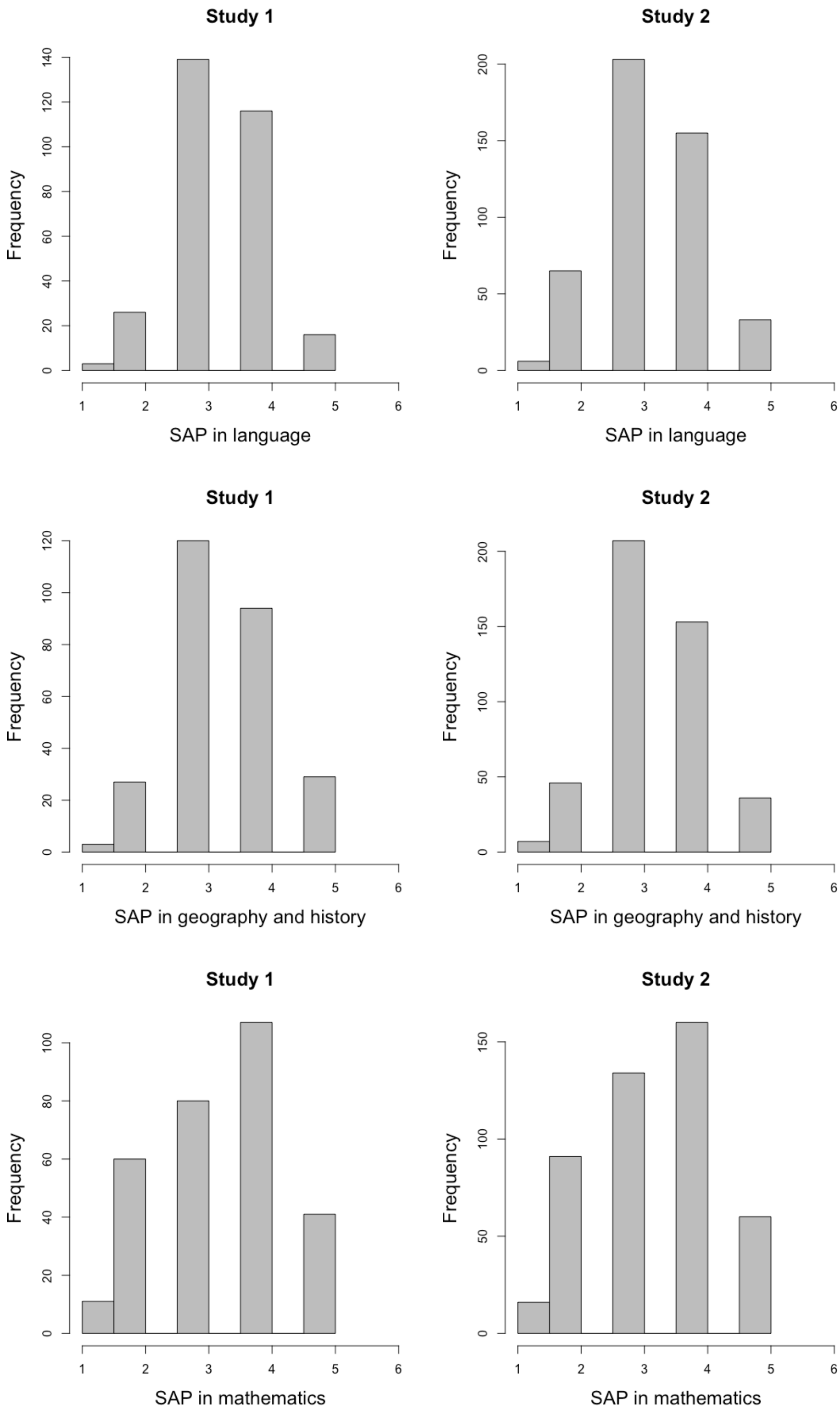

Study 1
Frequency
PAP in biology, chemistry, and physics
Study 2
Frequency
PAP in biology, chemistry, and physics
Study 1
Frequency
PAP in environment
Study 2
Frequency
PAP in environment

Figure S3

Histogram of Self-Reported Academic Performance (SAP) in Different Subjects of Study 1 (n = 301) and Study 2 (n = 465)



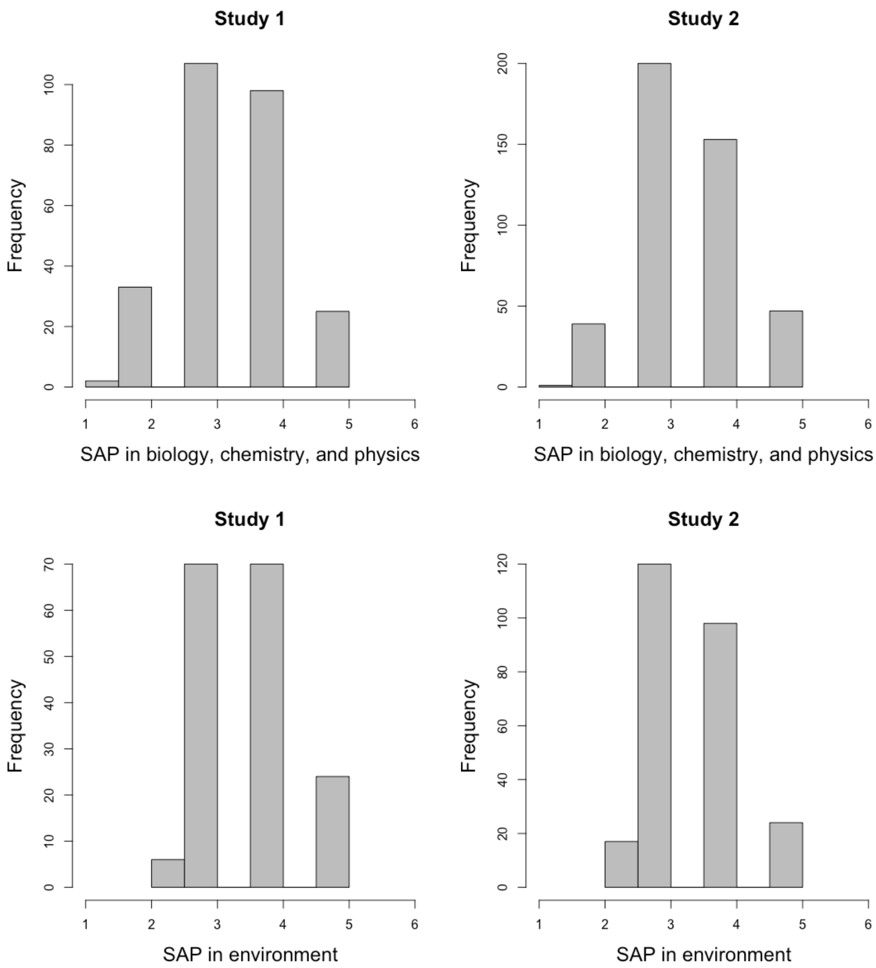
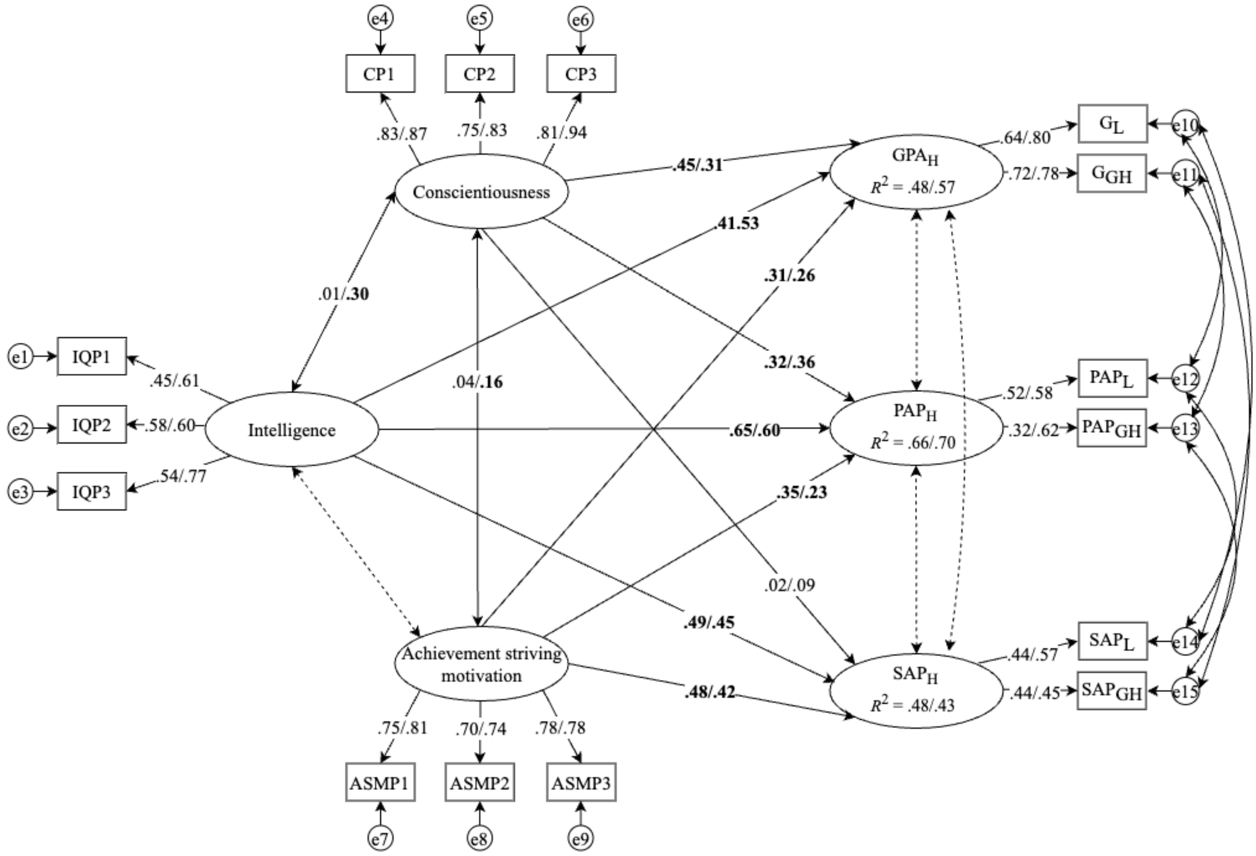

Study 1
Study 2
Frequency
SAP in biology, chemistry, and physics
Study 1
Study 2
Frequency
SAP in environment

Figure S4

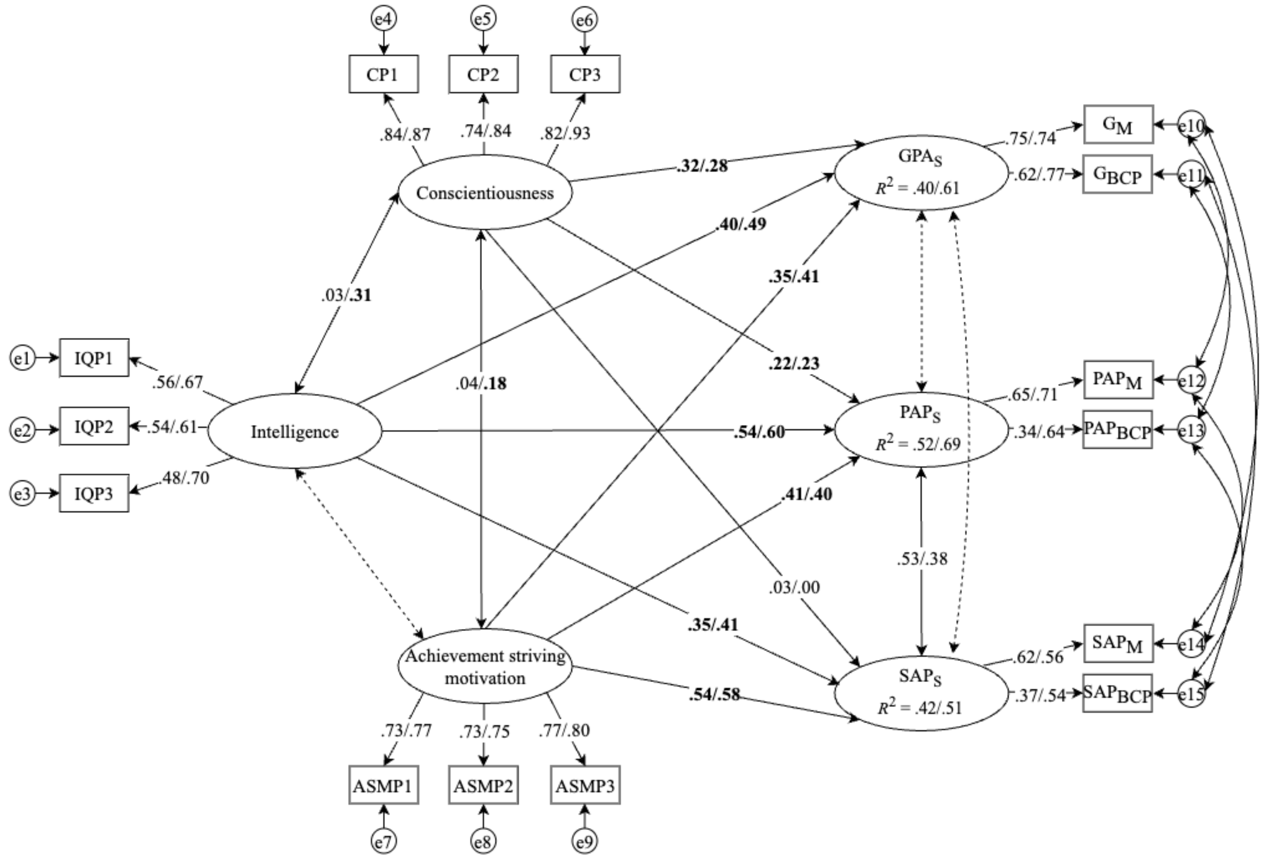
Structural Equation Model of Study 1 (n = 301) and Study 2 (n = 465) Focusing on the School-Subject Domain Humanities (Standardized Coefficients for Study 1 Are Shown Left of the Slash, Those for Study 2 Right of the Slash)



Note. Model fit was χ^2 = 127.038, p < .001, df = 73, comparative fit index (CFI) = .953, root-mean-square error of approximation (RMSEA) = .050, standardized root-mean-square residual (SRMR) = .049 for Study 1 and χ^2 = 103.121, p < .05, df = 73, CFI = .983, RMSEA = .030, SRMR = .045 for Study 2. Nonsignificant covariances between intelligence and achievement striving motivation and between GPA_H, PAP_H, and SAP_H were restricted to zero (see dashed lines). Significant coefficients with a p value below .05 are depicted in bold. In Study 1, intelligence was measured 7 years before the other constructs (Time 1); all other constructs were assessed at the same time (Time 2). In Study 2, all constructs were assessed at the same time (Time 1). IQP1 = intelligence parcel 1; IQP2 = intelligence parcel 2; IQP3 = intelligence parcel 3; CP1 = conscientiousness parcel 1; CP2 = conscientiousness parcel 2; CP3 = conscientiousness parcel 3; ASMP1 = achievement striving motivation parcel 1; ASMP2 = achievement striving motivation parcel 2; ASMP3 = achievement striving motivation parcel 3; GPA_H = grade point average in humanities; G_L = grade in language; G_{GH} = grade in geography and history; PAP_H = parent-reported academic performance in humanities; PAP_L = parent-reported academic performance in language; PAP_{GH} = parent-reported academic performance in geography and history; SAP_H = self-reported academic performance in humanities; SAP_L = self-reported academic performance in language; SAP_{GH} = self-reported academic performance in geography and history; e = residual.

Figure S5

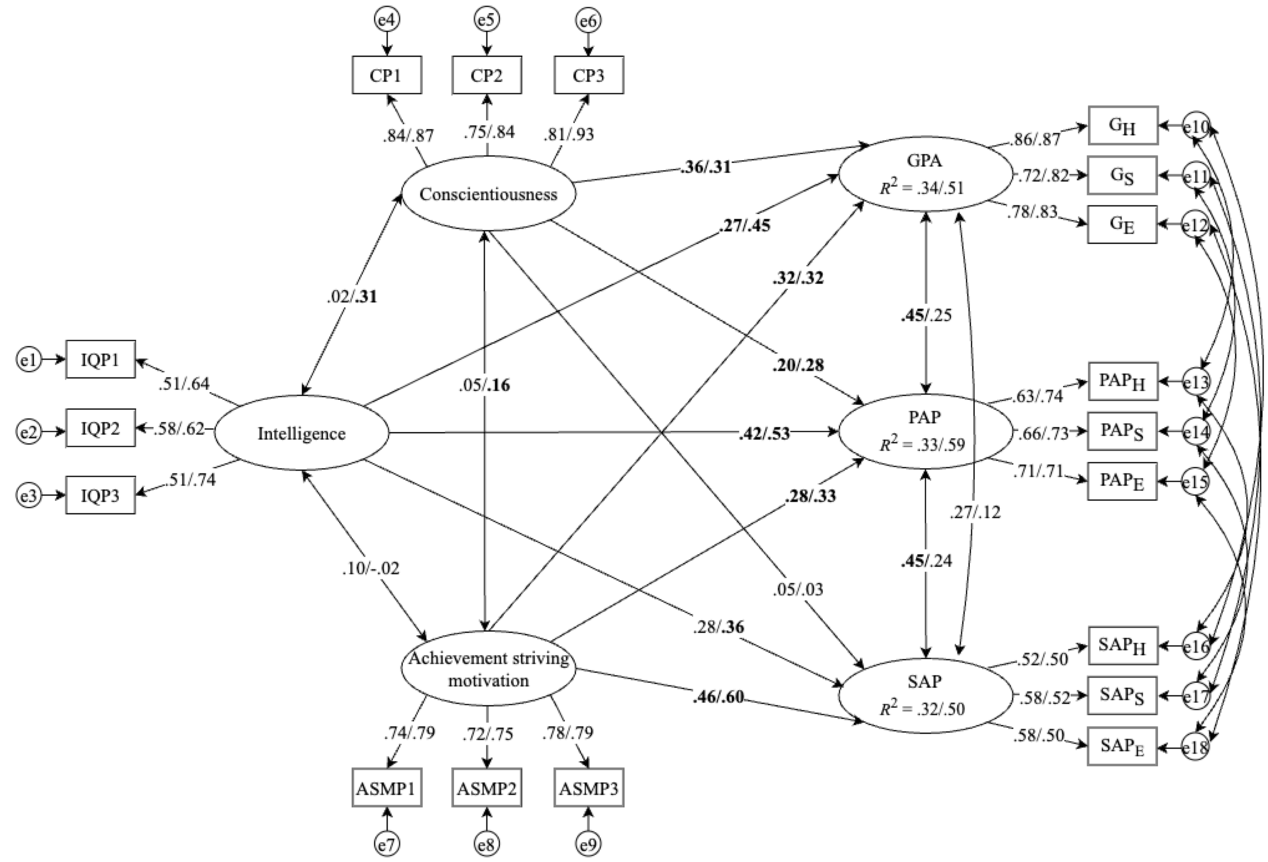
Structural Equation Model of Study 1 (n = 301) and Study 2 (n = 465) Focusing on the School-Subject Domain Science (Standardized Coefficients for Study 1 Are Shown Left of the Slash, Those for Study 2 Right of the Slash)



Note. Model fit was χ^2 = 127.262, p < .001, df = 72, comparative fit index (CFI) = .959, root-mean-square error of approximation (RMSEA) = .050, standardized root-mean-square residual (SRMR) = .051 for Study 1 and χ^2 = 135.845, p < .001, df = 72, CFI = .966, RMSEA = .045, SRMR = .047 for Study 2. Nonsignificant covariances between intelligence and achievement striving motivation, between GPA_S and PAP_S, and between GPA_S and SAP_S were restricted to zero (see dashed lines). Significant coefficients with a p value below .05 are depicted in bold. In Study 1, intelligence was measured 7 years before the other constructs (Time 1); all other constructs were assessed at the same time (Time 2). In Study 2, all constructs were assessed at the same time (Time 1). IQP1 = intelligence parcel 1; IQP2 = intelligence parcel 2; IQP3 = intelligence parcel 3; CP1 = conscientiousness parcel 1; CP2 = conscientiousness parcel 2; CP3 = conscientiousness parcel 3; ASMP1 = achievement striving motivation parcel 1; ASMP2 = achievement striving motivation parcel 2; ASMP3 = achievement striving motivation parcel 3; GPA_S = grade point average in science; G_M = grade in mathematics; G_{BCP} = grade in biology, chemistry, and physics; PAP_S = parent-reported academic performance in science; PAP_M = parent-reported academic performance in mathematics; PAP_{BCP} = parent-reported academic performance in biology, chemistry, and physics; SAP_S = self-reported academic performance in science; SAP_M = self-reported academic performance in mathematics; SAP_{BCP} = self-reported academic performance in biology, chemistry, and physics; e = residual.

Figure S6

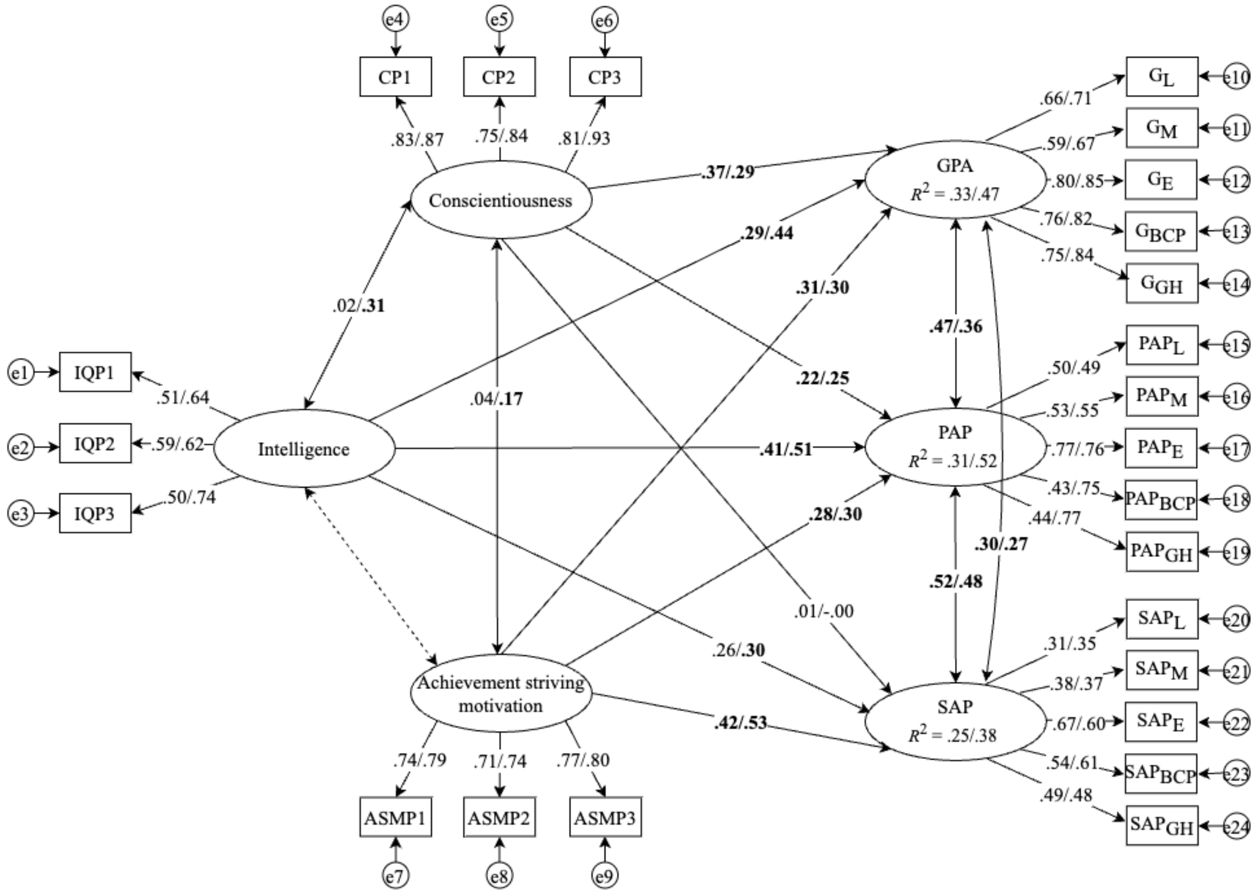
Structural Equation Model of Study 1 (n = 301) and Study 2 (n = 465) Focusing on Overall Estimates Without Restrictions of Nonsignificant Paths (Standardized Coefficients for Study 1 Are Shown Left of the Slash, Those for Study 2 Right of the Slash)



Note. Model fit was χ^2 = 197.716, *p* < .001, *df* = 111, comparative fit index (CFI) = .945, root-mean-square error of approximation (RMSEA) = .051, standardized root-mean-square residual (SRMR) = .057 for Study 1 and χ^2 = 181.936, *p* < .001, *df* = 111, CFI = .966, RMSEA = .038, SRMR = .053 for Study 2. None of the nonsignificant paths were restricted to zero. Significant coefficients with a *p* value below .05 are depicted in bold. In Study 1, intelligence was measured 7 years before the other constructs (Time 1); all other constructs were assessed at the same time (Time 2). In Study 2, all constructs were assessed at the same time (Time 1). IQP1 = intelligence parcel 1; IQP2 = intelligence parcel 2; IQP3 = intelligence parcel 3; CP1 = conscientiousness parcel 1; CP2 = conscientiousness parcel 2; CP3 = conscientiousness parcel 3; ASMP1 = achievement striving motivation parcel 1; ASMP2 = achievement striving motivation parcel 2; ASMP3 = achievement striving motivation parcel 3; GPA = grade point average; G_H = grade in humanities; G_S = grade in science; G_E = grade in environment; PAP = parent-reported academic performance; PAP_H = PAP in humanities; PAP_S = PAP in science; PAP_E = PAP in environment; SAP = self-reported academic performance; SAP_H = SAP in humanities; SAP_S = SAP in science; SAP_E = SAP in environment; e = residual.

Figure S7

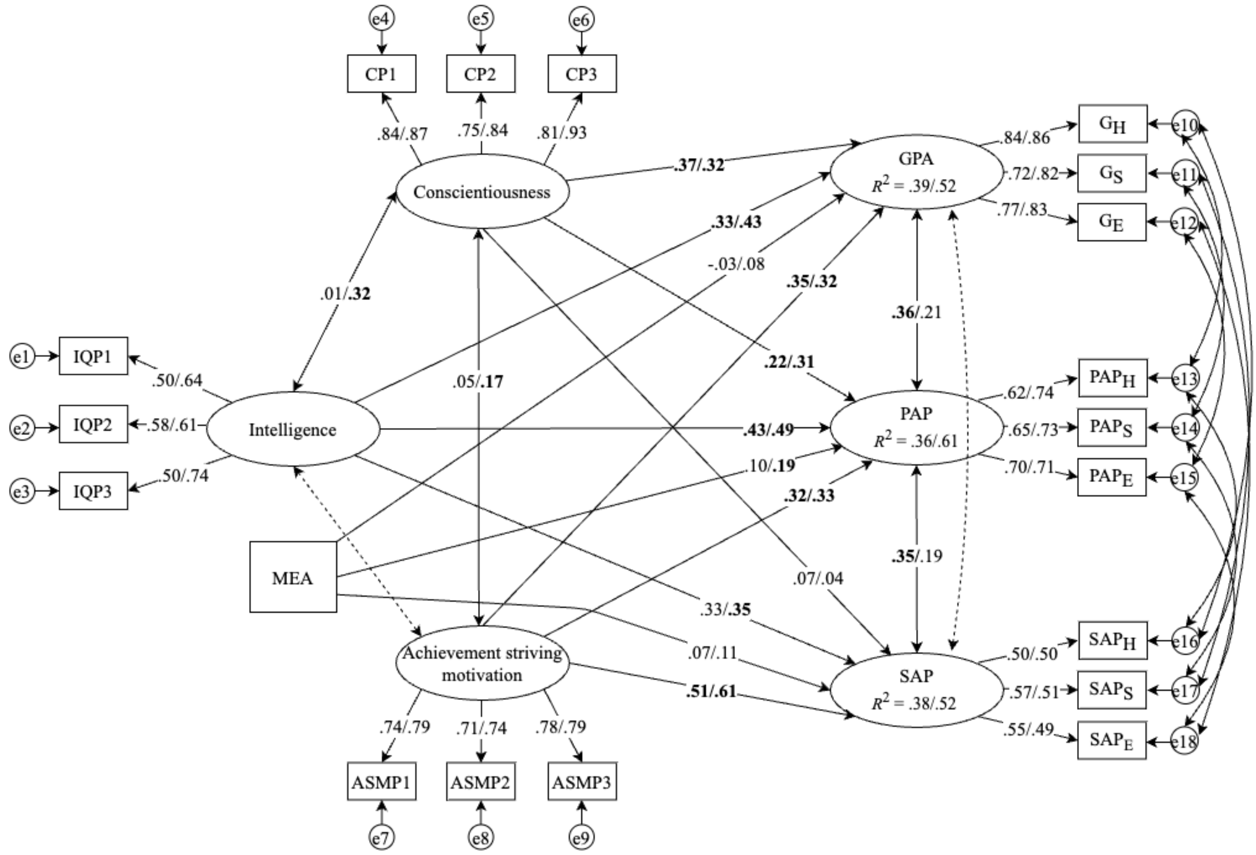
Structural Equation Model of Study 1 (n = 301) and Study 2 (n = 465) With Five Subjects as Indicators of Academic Achievement Measures (Standardized Coefficients for Study 1 Are Shown Left of the Slash, Those for Study 2 Right of the Slash)



Note. Model fit was χ^2 = 369.806, p < .001, df = 223, comparative fit index (CFI) = .930, root-mean-square error of approximation (RMSEA) = .048, standardized root-mean-square residual (SRMR) = .065 for Study 1 and χ^2 = 375.970, p < .001, df = 223, CFI = .946, RMSEA = .039, SRMR = .065 for Study 2. Nonsignificant covariances between intelligence and achievement striving motivation were restricted to zero (see dashed line). Correlations of residuals are not depicted in the figure for readability reasons. Significant coefficients with a p value below .05 are depicted in bold. In Study 1, intelligence was measured 7 years before the other constructs (Time 1); all other constructs were assessed at the same time (Time 2). In Study 2, all constructs were assessed at the same time (Time 1). IQP1 = intelligence parcel 1; IQP2 = intelligence parcel 2; IQP3 = intelligence parcel 3; CP1 = conscientiousness parcel 1; CP2 = conscientiousness parcel 2; CP3 = conscientiousness parcel 3; ASMP1 = achievement striving motivation parcel 1; ASMP2 = achievement striving motivation parcel 2; ASMP3 = achievement striving motivation parcel 3; GPA = grade point average; G_L = grade in language; G_M = grade in mathematics; G_E = grade in environment; G_{BCP} = grade in biology, chemistry, and physics; G_{GH} = grade in geography and history; PAP = parent-reported academic performance; PAP_L = PAP in language; PAP_M = PAP in mathematics; PAP_E = PAP in environment; PAP_{BCP} = PAP in biology, chemistry, and physics; PAP_{GH} = PAP in geography and history; SAP = self-reported academic performance; SAP_L = SAP in language; SAP_M = SAP in mathematics; SAP_E = SAP in environment; SAP_{BCP} = SAP in biology, chemistry, and physics; SAP_{GH} = SAP in geography and history; e = residual.

Figure S8

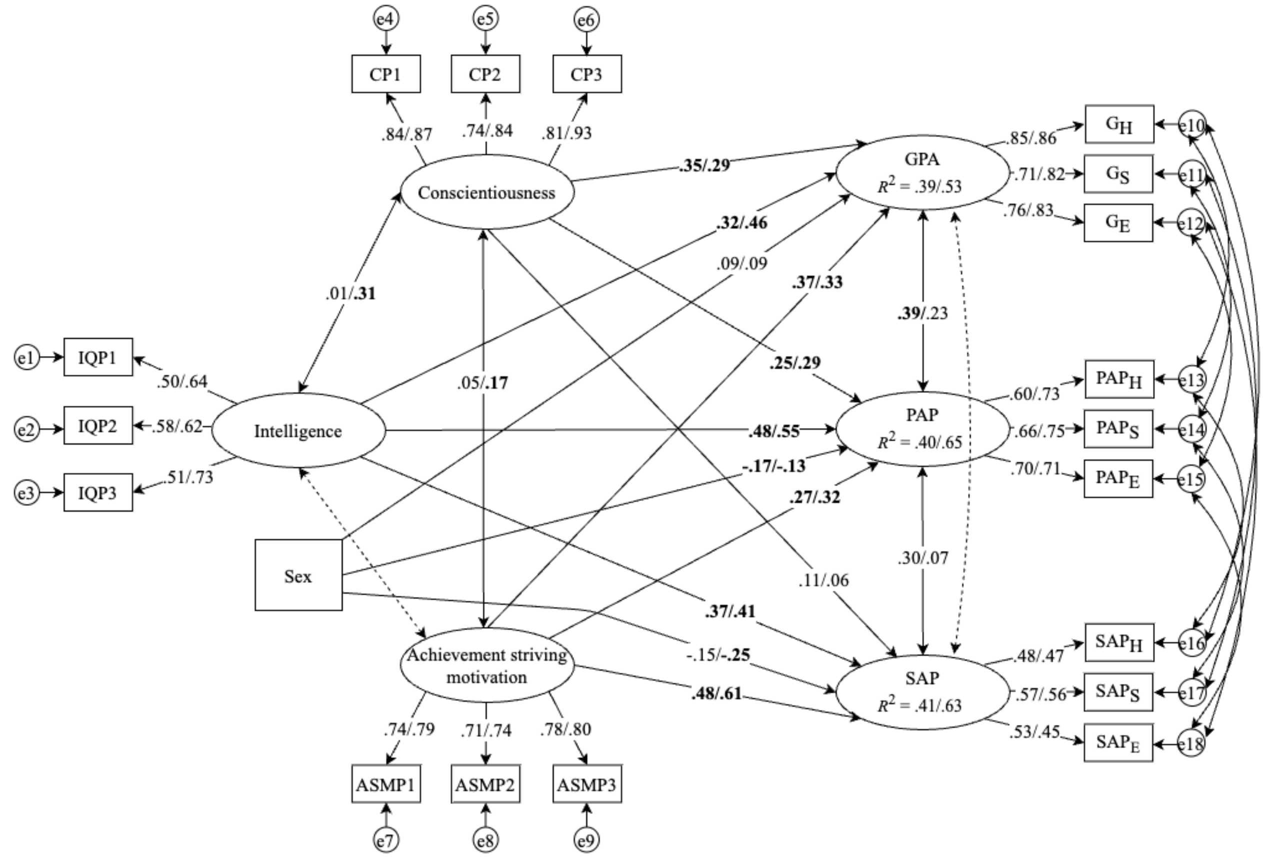
Structural Equation Model of the Post-Hoc Analyses Including Maternal Educational Attainment in Study 1 (n = 301) and Study 2 (n = 465; Standardized Coefficients for Study 1 Are Shown Left of the Slash, Those for Study 2 Right of the Slash)



Note. Maternal educational attainment was added as a dichotomous variable (0 = no postsecondary education, 1 = postsecondary education). Model fit was χ^2 = 246.734, p < .001, df = 128, comparative fit index (CFI) = .928, root-mean-square error of approximation (RMSEA) = .055, standardized root-mean-square residual (SRMR) = .065 for Study 1 and χ^2 = 208.818, p < .001, df = 128, CFI = .962, RMSEA = .038, SRMR = .057 for Study 2. Nonsignificant covariances between GPA and SAP and between intelligence and achievement striving motivation were restricted to zero (see dashed lines). Significant coefficients with a p value below .05 are depicted in bold. In Study 1, intelligence was measured 7 years before the other constructs (Time 1); all other constructs were assessed at the same time (Time 2). In Study 2, all constructs were assessed at the same time (Time 1). MEA = maternal educational attainment; IQP1 = intelligence parcel 1; IQP2 = intelligence parcel 2; IQP3 = intelligence parcel 3; CP1 = conscientiousness parcel 1; CP2 = conscientiousness parcel 2; CP3 = conscientiousness parcel 3; ASMP1 = achievement striving motivation parcel 1; ASMP2 = achievement striving motivation parcel 2; ASMP3 = achievement striving motivation parcel 3; GPA = grade point average; G_H = grade in humanities; G_S = grade in science; G_E = grade in environment; PAP = parent-reported academic performance; PAP_H = PAP in humanities; PAP_S = PAP in science; PAP_E = PAP in environment; SAP = self-reported academic performance; SAP_H = SAP in humanities; SAP_S = SAP in science; SAP_E = SAP in environment; e = residual.

Figure S9

Structural Equation Model of the Post-Hoc Analyses Including Sex in Study 1 (n = 301) and Study 2 (n = 465; Standardized Coefficients for Study 1 Are Shown Left of the Slash, Those for Study 2 Right of the Slash)



Note. Sex was added as a dichotomous variable (0 = male, 1 = female). Model fit was χ^2 = 261.542, p < .001, df = 128, comparative fit index (CFI) = .919, root-mean-square error of approximation (RMSEA) = .059, standardized root-mean-square residual (SRMR) = .070 for Study 1 and χ^2 = 242.596, p < .001, df = 128, CFI = .948, RMSEA = .044, SRMR = .064 for Study 2. Nonsignificant covariances between GPA and SAP and between intelligence and achievement striving motivation were restricted to zero (see dashed lines). Significant coefficients with a p value below .05 are depicted in bold. In Study 1, intelligence was measured 7 years before the other constructs (Time 1); all other constructs were assessed at the same time (Time 2). In Study 2, all constructs were assessed at the same time (Time 1). IQP1 = intelligence parcel 1; IQP2 = intelligence parcel 2; IQP3 = intelligence parcel 3; CP1 = conscientiousness parcel 1; CP2 = conscientiousness parcel 2; CP3 = conscientiousness parcel 3; ASMP1 = achievement striving motivation parcel 1; ASMP2 = achievement striving motivation parcel 2; ASMP3 = achievement striving motivation parcel 3; GPA = grade point average; G_H = grade in humanities; G_S = grade in science; G_E = grade in environment; PAP = parent-reported academic performance; PAP_H = PAP in humanities; PAP_S = PAP in science; PAP_E = PAP in environment; SAP = self-reported academic performance; SAP_H = SAP in humanities; SAP_S = SAP in science; SAP_E = SAP in environment; e = residual.

Supplemental References

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. Routledge.

Grob, A., & Hagmann-von Arx, P. (2018). *Intelligence and Development Scales–2 (IDS-2). Intelligenz- und Entwicklungsskalen für Kinder und Jugendliche [Intelligence and Development Scales for Children and Adolescents]*. Hogrefe.

Grob, A., Meyer, C. S., & Hagmann-von Arx, P. (2009). *Intelligence and Development Scales (IDS)*. Huber.

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), 1–55. https://doi.org/10.1080/10705519909540118

Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, *9*(4), 486–492. https://doi.org/10.1080/17405629.2012.686740

APPENDIX C: STUDY 3

Odermatt, S. D., Grieder, S., Schweizer, F., Bünger, A., & Grob, A. (2023). *The role of language proficiency in the assessment of cognitive and developmental functions in children: An analysis of the Intelligence and Development Scales–2.* Manuscript submitted for publication.

Manuscript submitted for publication.

The Role of Language Proficiency in the Assessment of Cognitive and Developmental Functions in Children: An Analysis of the Intelligence and Development Scales–2

Salome D. Odermatt, Silvia Grieder, Florine Schweizer, Anette Bünger, and Alexander Grob

Department of Psychology, University of Basel, Switzerland

Author Note

The authors declare the following conflict of interest: AG is a recipient of the royalties for the Intelligence and Development Scales–2 (IDS-2). This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. We would like to thank all participants and all test administrators for their support. Further, we thank Anita Todd for copyediting the manuscript. The data and study material are not publicly available owing to property rights. The data are available on request from the corresponding author. The analysis code can be accessed at https://osf.io/shrzk/?view_only=5a6f4444a2ff4c5d87885524c153ed86.

Correspondence concerning this article should be addressed to Salome D. Odermatt, Department of Psychology, University of Basel, Missionsstrasse 62, 4055 Basel, Switzerland. Email: salome.odermatt@unibas.ch

Abstract

Assessing cognitive and developmental functions in multilingual participants is challenging, as proficiency in the test language may interfere with test performance. We examined whether different language aspects (i.e., receptive and expressive language abilities, multilingualism) contribute independently to the variance in scores on cognitive and developmental functions of the Intelligence and Development Scales–2 (IDS-2). The sample comprised 826 children aged 5–10 years. Hierarchical regression analyses revealed that receptive language ability was significantly associated with almost all IDS-2 scores. Expressive language ability explained little additional variance, except for the intelligence composites, Verbal Reasoning, including corresponding subtests, and the basic skills subtests. Multilingualism explained variance above objectively measured language abilities only in subtests of Verbal Reasoning and verbal Long-Term Memory. Findings highlight the importance of considering language proficiency, particularly in tasks with high verbal demands, when assessing cognitive and developmental functions with the IDS-2 in participants at risk for linguistic disadvantages.

Keywords: language abilities, assessment, cognitive functions, developmental functions, Intelligence and Development Scales–2, multilingual children

1. Introduction

Assessments of individuals' cognitive abilities and developmental domains play a crucial role in a comprehensive evaluation in clinicians' daily practice. For example, performance on psychometric tests provides information about individuals' strengths and difficulties and forms the groundwork for the development of clinical hypotheses and intervention measures (Flanagan & Harrison, 2012). Yet, participants' test performance can be compromised by factors other than the examined construct, including aspects of measurement and participant characteristics, such as individuals' level of proficiency in the test language (American Educational Research Association et al., 2014). This is particularly relevant as worldwide migration rates have been increasing over the past decades (United Nations, 2022). For instance, in German-speaking countries, the national censuses have identified approximately 29% of the German (German Federal Statistical Office, 2023), 39% of the Swiss (Swiss Federal Statistical Office, 2022a), and 25% of the Austrian (Austrian Federal Statistical Office, 2022) population as currently having a migration background.

Because of this relatively high percentage of individuals with a migration background, the number of children in schools and in psychological services who need educational assessment but whose native language is not the test language of the region in which they are living has consequently increased. Furthermore, in some cases they may also have a different cultural background. Thus, the accurate assessment of these individuals' abilities may be hampered, as, for example, multilingual individuals often have insufficient knowledge of the test language, such as lower levels of language comprehension and grammar and smaller vocabularies in each of their languages (Bialystok et al., 2010; Hoff et al., 2012; Reich et al., 2002). Therefore, language issues may arise when conducting standardized tests (e.g., Guthke & Wiedl, 1996; Ortiz, 2019).[1] In particular, individuals with linguistically diverse backgrounds may have difficulties understanding verbal instructions or, if required, giving verbal responses during the test administration (Cormier et al., 2022; Weiss et al., 2006). These constraints depend on the verbal demands a test places on the participant with respect to complexity, length, and verbosity of instructions (Cormier et al., 2011), response options, as well as the proportion of verbal components within the tasks (Flanagan & Ortiz, 2001). Specifically, children who are bilingual or have a migration background tend to exhibit lower test scores especially in tasks that are more verbal dependent (Hagmann-von Arx et al., 2013; Schweizer et al., 2021). As a result, these children might not show their full potential, which bears the danger of biased test results and in consequence may lead to an underestimation of their true abilities (Hagmann-von Arx et al., 2013). In particular, the abilities of younger children who enter the school system may be underestimated, as attending formal educational institutions enhances the development of language proficiency (Grob et al., 2014). This is critical, as an underestimation of abilities, such as intelligence, can lead to possible negative consequences regarding the children's future school career and long-term development (Calero et al., 2013; Goldstein et al., 2015; Hessels, 1997; Klingner et al., 2007; Sullivan, 2011).

Thus, it is essential to evaluate the contribution of participants' proficiency in the test language to their performance on standardized assessments, to identify possible participant characteristics that could compromise the validity of the test score interpretation (see also fairness in testing; American Educational Research Association et al., 2014). Our aim in the present study was therefore to investigate the relative importance of different language aspects, namely, children's receptive language ability, expressive language ability, and multilingualism,[2] on test performance on various cognitive (i.e., intelligence, executive functions) and developmental (i.e., psychomotor skills, social-emotional skills, basic skills) functions of the Intelligence and Development Scales–2 (IDS-2; Grob & Hagmann-von Arx, 2018a). The IDS-2 is a paper-and-pencil test for children and adolescents between 5 and 20 years, based on the Intelligence and Development Scales for children between 5 and 10 years (IDS; Grob et al., 2009). The series also includes a version for children between 3 and 5 years, the Intelligence and Development Scales–Preschool (IDS-P; Grob, Reimann, et al., 2013). The IDS-2 is a multidimensional psychometric tool for obtaining a comprehensive assessment of individuals' cognitive and developmental functions using a single test battery (see Figure 1 and Table S1 in the Supplement for a summary of the IDS-2 domains included in our study; Grob & Hagmann-von Arx, 2018b). The IDS-2 was standardized between 2015 and 2017 in the German-speaking part of Switzerland, Germany, and Austria and was published in 2018. Subsequently, additional language adaptations have been released, such as in Dutch, English (UK), Italian, and Polish (Grob et al., 2018, 2019, 2021, 2022), or are currently in progress in several further countries (e.g., Brazil, Denmark, Finland, France, Norway, Sweden, Spain, and the United States).

The IDS-2 has rarely been used in previous research to investigate possible relations between individuals' proficiency in the test language and performance on the IDS-2. To date, Schweizer et al. (2021) conducted the only study so far, which examined mean-level differences between matched monolingual, simultaneously bilingual, and successively bilingual children and adolescents (each group: n = 132; M_{age} = 12.34 years) in the intelligence domain of the IDS-2. They found that successive bilinguals showed lower mean values than those of monolinguals—and to some extent also lower than those of simultaneous bilinguals—in the intelligence composites and in the group factors Verbal Reasoning and verbal Long-Term Memory. At the level of subtests, significant group differences were reported for Naming Categories, Naming Opposites, and Story Recall. No differences were detected between monolingual and simultaneously bilingual participants. However, evidence on the several other IDS-2 domains (e.g., executive functions, psychomotor skills, social-emotional skills, basic skills) is currently lacking.

With respect to the two other versions—the IDS and IDS-P—previous research examined mean-level differences between children with a migration background (defined as having a language other than German as their native language) and matched control samples (Grob, Meyer, & Hagmann-von Arx, 2013; Grob, Reimann, et al., 2013; Hagmann-von Arx et al., 2013). These studies reported no group differences in the intelligence composite (Grob, Meyer, & Hagmann-von Arx, 2013; Grob, Reimann, et al., 2013; Hagmann-von Arx et al., 2013), which assesses mainly fluid intelligence aspects (Baltes, 1987, 1990), and in most of the intelligence subtests, except for the subtest Auditory Memory. On this subtest, which measures verbal long-term memory, children with a migration background scored lower than the control group (Grob, Meyer, & Hagmann-von Arx, 2013; Grob, Reimann, et al., 2013), which has been explained by the fact that active language abilities may be necessary to solve the task (Grob, Reimann, et al., 2013). In addition, the majority of these studies found no mean-level differences in psychomotor skills, while mixed results were reported for social-emotional skills and mathematics (Grob, Meyer, & Hagmann-von Arx, 2013; Grob, Reimann, et al., 2013; Hagmann-von Arx et al., 2013). Regarding executive functions as well as the basic skills reading and spelling, which are also included in the IDS-2 but are not part of the IDS and IDS-P, we refer to other literature, even though comparability with the IDS-2 tasks may be limited to some degree.

Although previous research reported that individuals with linguistically diverse backgrounds scored lower than their monolingual peers in many domains, studies have indicated an advantage for bilingual individuals in executive functions, such as inhibitory control and cognitive flexibility (Barac et al., 2014; Grundy, 2020; Yurtsever et al., 2023). Specifically, bilingual children outperformed their monolingual peers mainly on nonverbal executive function tasks, while no differences or lower performance were found on verbal tasks (Bialystok & Craik, 2010; Foy & Mann, 2014). This advantage has been explained by the fact that both languages are constantly active in bilinguals (Kroll et al., 2012; Rodriguez-Fornells et al., 2002; Thierry & Wu, 2007) and managing these jointly activated languages results in the training of executive functions, as, for example, responding in the target language has to be controlled (Abutalebi & Green, 2008).

Concerning basic skills, language abilities are crucial for learning processes in school and hence for the successful acquisition of reading, spelling, and mathematics (e.g., Storch & Whitehurst, 2002). Previous research indicated that multilingual individuals and individuals with a migration background achieved lower test scores than their control peers in reading and spelling (Konsortium PISA.ch, 2019; Melby-Lervåg & Lervåg, 2014; Verhoeven, 2000). Moreover, there is evidence that vocabulary size relates to scholastic performance (e.g., Kastner et al., 2001) and the acquirement of literacy skills (Lee Swanson et al., 2008; Ouellette, 2006). Although previous studies on the IDS and IDS-P produced inconsistent results with respect to mathematics (Grob, Meyer, & Hagmann-von Arx, 2013; Grob, Reimann, et al., 2013; Hagmann-von Arx et al., 2013), several other studies reported group differences between individuals with and without a migration background in the natural sciences (Konsortium PISA.ch, 2018) and mathematics (e.g., Bos et al., 2007; Ehm et al., 2011; Paetsch et al., 2016). Moreover, basic skills are also described as "cultural skills," which seem to be related to previous language experiences in the school context (Grob & Hagmann-von Arx, 2018b; Köller & Baumert, 2008).

Nevertheless, a large part of the existing literature usually focused on group differences, such as between individuals with and without a migration background (Grob, Meyer, & Hagmann-von Arx, 2013; Grob, Reimann, et al., 2013; Hagmann-von Arx et al., 2013) or between monolinguals and bilinguals (Schweizer et al., 2021). This makes it difficult to disentangle the relative importance of language abilities in test performance, as other factors, such as cultural background, can also play a role. Moreover, individuals within these groups may also have distinct levels of proficiency in the test language and therefore be at different positions on the continuum of language abilities (Ortiz, 2019). This has not been taken into account in analyses comparing groups. However, to derive concrete recommendations for practice, a closer look at individuals' measurable language abilities has been suggested (Ortiz, 2019). To the best of our knowledge, only one recent study (Cormier et al., 2022) followed this approach and investigated the effects of participants' receptive and expressive language abilities in English on test performance in the Woodcock–Johnson Tests of Cognitive Abilities (4th ed.; Schrank et al., 2014). Results of this study indicated that both language abilities were significantly related to test performance in this cognitive test battery, with receptive language ability showing slightly higher associations than expressive language ability. However, the authors noted in their Limitations section that the sample did not include many English-language learners (Cormier et al., 2022), which might have restricted the generalizability of the findings. Moreover, these results also raise the question to what extent objective measures of receptive and expressive language abilities relate to performance in other important domains of the child development, beyond cognitive abilities.

1.1 The Present Study

Building on this literature, we aimed to examine how different language aspects (i.e., children's receptive language ability, expressive language ability, and multilingualism) contribute beyond each other to test performance on the cognitive and developmental functions assessed with the IDS-2 (Grob & Hagmann-von Arx, 2018a). We thereby sought to follow a holistic approach by integrating multiple language aspects on one side and various cognitive and developmental domains on the other. Specifically, we investigated the relations of distinct language abilities to test scores on the IDS-2 after controlling for relevant sociodemographic variables (i.e., sex and socioeconomic status [SES], represented by maternal educational background; Weiss & Saklofske, 2020).[3] We distinguished between children's receptive and expressive language abilities and relied on standardized assessments of these two language skills. Moreover, we added multilingualism in the final step to explore whether other components of having a linguistically diverse background contribute beyond objectively measured language abilities to the variance in test scores on the IDS-2. We employed the IDS-2 because this test battery offers a comprehensive assessment of important cognitive and developmental functions (see Figure 1 for an overview; Grob & Hagmann-von Arx, 2018b). Additionally, this test procedure is widely used by clinicians, such as psychologists and physicians, in countries that have a high number of immigrants (i.e., Switzerland, Germany, and Austria; Austrian Federal Statistical Office, 2022; German Federal Statistical Office, 2023; Swiss Federal Statistical Office, 2022a) and therefore many children with linguistically diverse backgrounds in psychological assessment settings. Hence, it is crucial that clinicians are aware of the composites, group factors, and subtests that relate to aspects of language in order to take these effects into account when administering the IDS-2 to participants at risk for linguistic disadvantages. However, there is currently no evidence concerning the relative importance of receptive and expressive language abilities on children's test performance in the IDS-2 intelligence domain, and relations between language aspects and the other IDS-2 domains have not yet been investigated. Thus, we also aimed to extend current knowledge regarding the validity of test-score interpretations of the IDS-2.

In line with previous research (e.g., Hagmann-von Arx et al., 2013), we assumed that as the verbal demands of a task increase—with respect to instructions, content, and response format—language abilities will be more strongly associated with the test score and explain more variance. Therefore, we hypothesized that performance on those subtests and intelligence group factors of the IDS-2 that we classified as "low linguistic loading" (see Table 1) would show the smallest positive associations with receptive and expressive language abilities. In contrast, we expected that performance on subtests and intelligence group factors of the IDS-2 that we categorized as "high linguistic loading" (see Table 1) would show the strongest positive associations with receptive and expressive language abilities. We based our classification on findings from previous studies presented above and on considerations of the verbal demands of the specific tasks. Concerning multilingualism, we formulated the following research question, as this represented an exploratory analysis: Does multilingualism explain variance beyond language abilities in test scores on the IDS-2?

2. Method

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study.

2.1 Participants

The sample consisted of participants from the IDS-2 standardization and validation study (N = 2,030, 52% female, age range 5–20 years). As the IDS-2 subtest Language Skills is administered only to participants aged 5–10 years, we omitted adolescents aged 11–20 years (n = 1,120) from our analyses. Additionally, we excluded participants with missing data concerning their receptive and expressive language abilities, native language, age, sex, or maternal educational background (n = 84). The final sample comprised 826 children (M_{age} = 8.06 years, SD = 1.66, age range: 5.02–10.99 years) with 51% being female. In terms of education, 59.6% of the participants' mothers had not gone beyond mandatory education or had vocational job training (i.e., no postsecondary education), while 40.4% of the participants' mothers had obtained a university degree (i.e., postsecondary education), which is higher than in the corresponding population (e.g., 29.2% of women currently hold a postsecondary degree in Switzerland;

Swiss Federal Statistical Office, 2022b). Regarding children's current education, they attended preschool (*n* = 5), kindergarten (*n* = 189), elementary school (*n* = 590), secondary school (*n* = 14), special education services (*n* = 23), other (*n* = 3), or unknown (*n* = 2). About 15% of the sample reported one or more medical or psychological diagnosis.[4]

Participants were from the German-speaking part of Switzerland (*n* = 529), Germany (*n* = 252), and Austria (*n* = 45). A total of 26.0% reported being multilingual (bilingual: 24.4%; trilingual or more than three languages: 1.6%), which we defined as speaking German (the test language) and having at least one other language as their native language. This percentage corresponds to recent studies on the proportion of multilingual individuals in the corresponding population (e.g., in Switzerland, 33% of children under the age of 15 years were exposed to two languages at home; in Germany, 21% of children under the age of 14 years spoke a language other than German in the family context; German Federal Statistical Office, 2022; Swiss Federal Statistical Office, 2021). Besides German, the children in our study spoke Russian (*n* = 32), English (*n* = 29), French (*n* = 26), Italian (*n* = 19), Turkish (*n* = 18), Spanish (*n* = 14), Albanian (*n* = 12), Serbian (*n* = 12), Tamil (*n* = 9), Portuguese (*n* = 7), Dutch (*n* = 5), Chinese (*n* = 4), Croatian (*n* = 4), and 23 other languages (each under *n* = 4). However, all participants had to speak the test language (i.e., German) sufficiently to take part in the study. Hence, all participants' educational language was German. The other 74.0% reported being monolingual German speakers and indicated German as their native language. Characteristics of monolingual and multilingual children are presented in Table S2 in the Supplement.

2.2 Procedure

Recruitment was conducted between 2015 and 2017 through day care centers, playgroups, kindergartens, schools, and school psychological services in Switzerland as well as through psychosocial institutions and universities in Germany and Austria. Participants were individually tested with the IDS-2 at their homes, at a laboratory of the university, or at the respective institutions' laboratories by one of several trained test administrators (i.e., undergraduate psychology students or psychologists). The administration of the IDS-2 took about 4 h, depending on participants' age and performance, as there are age-specific implementation rules and performance-based rules for ending testing. Testing could be divided into two sessions one week apart if the participant wished. Parents were asked to report demographic variables, such as their children's native language(s) and parental educational background, in a questionnaire administered by the test administrator prior to testing. The families received a monetary incentive for participation either as a gift card or in cash. The local ethics committee (BLINDED) approved the study protocol. Parents were asked to sign a consent form.

2.3 Instrument

The IDS-2 (Grob & Hagmann-von Arx, 2018a) assesses cognitive (i.e., intelligence and executive functions) and developmental (i.e., psychomotor skills, social-emotional skills, basic skills, and motivation and attitude) functions in children and adolescents aged 5–20 years. An overview and a detailed description of the IDS-2 domains included in our study are provided in Figure 1 and Table S1, respectively. The IDS-2 also assesses participants' language skills in terms of their level of language comprehension (i.e., receptive language ability) and language production (i.e., expressive language ability). Children's receptive language ability is measured by asking them to carry out recited instructions using multiple materials (e.g., a ball or a pencil); to assess their expressive language ability, children are asked to form meaningful sentences from spoken and pictorially depicted words. Several studies have documented the psychometric properties of the IDS-2 for the standardization sample (e.g., Grieder & Grob, 2020; Odermatt et al., 2022) , showing, for example, high correlations with other frequently used test procedures in German-speaking countries (Grob & Hagmann-von Arx, 2018b).

2.4 Statistical Analyses

Hierarchical regression analyses were conducted to examine whether different language variables (i.e., receptive language ability, expressive language ability, and multilingualism) predicted scores on composites, intelligence group factors, and subtests of the cognitive and developmental functions of the IDS-2, beyond sex and SES. We used age-standardized IDS-2 scores (M = 100, SD = 15, for Profile IQ, Full-Scale IQ, Screening IQ, and the seven intelligence group factors; M = 10, SD = 3, for other composites and subtests) and therefore omitted controlling for age. We formed the variable *multilingualism* using information regarding children's native language reported by the parents. In all regression analyses, we entered control variables (i.e., sex and SES) in Step 1, children's receptive language ability in Step 2, and their expressive language ability in Step 3. Multilingualism (i.e., being multilingual vs. being monolingual) was added in Step 4. We calculated separate analyses for each composite, intelligence group factor, and subtest of the IDS-2, resulting in 40 analyses.[5] To adjust for multiple testing, we corrected p values (hereafter denoted p_H) according to Hommel (1988). Moreover, reliabilities for the IDS-2 scores were indicated, including Cronbach's alpha for homogeneous subtests, reliabilities computed using a formula by Lienert and Raatz (1998) for composites, and retest reliabilities for single-item or heterogeneous subtests obtained from the IDS-2 manual (Grob & Hagmann-von Arx, 2018b). Analyses were performed using R (R Core Team, 2022).

3. Results

Table 2 presents the descriptive statistics of the predictors (i.e., sex, SES, receptive and expressive language abilities, multilingualism) and the IDS-2 scores. Reliability coefficients were high for composites and intelligence group factors and high to satisfactory for subtests of the IDS-2. Pearson correlations showed moderate associations between receptive language ability, expressive language ability, and multilingualism and therefore no multicollinearity between the language aspects was detected (see Table S3 in the Supplement). The proportion of variance explained (R^2) by each predictor in the IDS-2 scores is displayed in Figure 2 for the cognitive functions and in Figure 3 for the developmental functions. Results of the hierarchical regression analyses are reported in Tables S4 to S11 in the Supplement.[6]

3.1 Cognitive Functions

3.1.1 Intelligence

After controlling for sex and SES, children's receptive language ability was significantly associated with higher scores on Profile IQ (β = .57, p_H < .001), Full-Scale IQ (β = .56, p_H < .001), and Screening IQ (β = .49, p_H < .001), explaining between 23% and 32% of additional variance. Children's receptive language ability was also significantly related to each intelligence group factor score (highest for Verbal Reasoning: β = .55, p_H < .001; lowest for Processing Speed: β = .32, p_H < .001) and each subtest score (highest for Naming Opposites: β = .50, p_H < .001; lowest for Washer Design: β = .20, p_H < .001). Additional explained variance ranged between 4% and 28%. Expressive language ability, entered in Step 3, significantly correlated with higher scores on Profile IQ (β = .28, p_H < .001), Full-Scale IQ (β = .28, p_H < .001), and Screening IQ (β = .28, p_H < .001), accounting for 6% of additional variance in each composite. Expressive language ability was related to scores on each intelligence group factor, except for Processing Speed (β = .11, p_H = .665) and Visuospatial Short-Term Memory (β = .13, p_H = .111), additionally explaining between 2% and 8% of variance. At the level of subtests, children's expressive language ability was associated with scores on Shape Design (β = .14, p_H = .024), Digit and Letter Span (β = .20, p_H < .001), Matrices: Completion (β = .21, p_H < .001), Naming Categories (β = .26, p_H < .001), Story Recall (β = .25, p_H < .001), Mixed Digit and Letter Span (β = .21, p_H < .001), Rotated Shape Memory (β = .16, p_H = .006), and Naming Opposites (β = .31, p_H < .001). Between 2% and 7% of variance was explained beyond children's receptive language ability. Adding the variable multilingualism in Step 4 only explained variance in the subtests Story Recall (β = .13, p_H = .022) and Naming Opposites (β = .16, p_H < .001), accounting for 2% of additional variance in each of these subtests. No significant associations were found between multilingualism and scores on the intelligence composites or group factors (see Figure 2 and Tables S4 to S7 in the Supplement).

3.1.2 Executive Functions

Beyond the control variables sex and SES, children's receptive language ability was significantly related to higher scores on the executive functions composite (β = .42, p_H < .001) and on each subtest in the executive functions domain (highest for Listing Words: β = .38, p_H < .001; lowest for Drawing Routes: β = .18, p_H < .001), explaining between 3% and 17% of additional variance. In Step 3, children's expressive language ability was associated with scores on the executive functions composite (β = .20, p_H < .001) as well as with scores on the subtests Listing Words (β = .16, p_H = .007) and Divided Attention (β = .19, p_H < .001), explaining between 2% and 3% of additional variance. The variable multilingualism did not account for variance in any of the scores in the executive functions domain in Step 4 (see Figure 2 and Table S8 in the Supplement).

3.2 Developmental Functions

3.2.1 Psychomotor Skills

Concerning psychomotor skills, positive associations between children's receptive language ability and scores on the psychomotor skills composite (β = .33, p_H < .001) as well as all subtests (highest for Visuomotor Skills: β = .28, p_H < .001) were found, explaining between 5% and 11% of additional variance beyond the control variables. Expressive language ability, entered in Step 3, was related only to scores on the psychomotor skills composite (β = .15, p_H = .010; ΔR^2 = 2%) and not to scores on the subtests. The variable multilingualism did not explain any additional variance in the scores in the psychomotor skills domain in Step 4 (see Figure 3 and Table S9 in the Supplement).

3.2.2 Social-Emotional Skills

After controlling for sex and SES, children's receptive language ability was significantly associated with scores on the social-emotional skills composite (β = .21, p_H < .001) and the subtests Identifying Emotions (β = .19, p_H < .001) and Regulating Emotions (β = .17, p_H < .001), explaining between 3% and 4% of variance. No relationship was found with scores on the subtest Socially Competent Behavior (β = .10, p_H = .742). When children's expressive language ability was entered in Step 3, significant correlations were detected only with scores on the social-emotional skills composite (β = .18, p_H = .002) and the subtest Regulating Emotions (β = .17, p_H = .006), explaining between 2% and 3% of additional variance. The variable multilingualism did not explain additional variance in any of the scores in the social-emotional skills domain in Step 4 (see Figure 3 and Table S10 in the Supplement).

3.2.3 Basic Skills

Regarding basic skills, children's receptive language ability was significantly related to higher scores on each subtest (highest for Logical-Mathematical Reasoning: β = .46, p_H < .001), explaining between 12% and 21% of additional variance. In Step 3, expressive language ability was also significantly correlated with higher scores on each subtest (highest for Spelling: β = .31, p_H < .001) and accounted for between 6% and 8% of additional variance. No significant associations were found when the variable multilingualism was entered in Step 4 (see Figure 3 and Table S11 in the Supplement).

4. Discussion

In the present study, we sought to investigate the relative importance of different language aspects (i.e., receptive and expressive language abilities, multilingualism) for participants' test performance on various cognitive and developmental functions of the IDS-2 using a sample of children aged 5–10 years. Hierarchical regression analyses showed that children's receptive language ability was significantly related to all scores on the IDS-2 when accounting for sex and SES, except for the subtest Socially Competent Behavior. Children's expressive language ability explained overall little additional variance above receptive language ability in the IDS-2 scores. The highest amounts of additional explained variance were found for intelligence composites, the intelligence group factor Verbal Reasoning and its corresponding subtests Naming Categories and Naming Opposites, and the subtests in the basic skills domain (i.e., Logical-Mathematical Reasoning, Reading, and Spelling). In contrast, multilingualism, entered in the final step, explained variance beyond receptive and expressive language abilities solely in the two intelligence subtests Story Recall and Naming Opposites. Overall, language aspects explained the largest variance in scores in the intelligence and basic skills domains. The results emphasize the relevance of considering children's proficiency in the test language within the assessment of cognitive and developmental functions with the IDS-2, especially in tasks with high verbal demands, when testing participants at risk for linguistic disadvantages.

4.1 Receptive Language Ability and Test Performance on the IDS-2

In line with previous research (Cormier et al., 2022), children's receptive language ability explained variance in scores on almost all subtests, intelligence group factors, and composites of the IDS-2. Hence, the ability to understand verbal instructions is crucial to completing the IDS-2 tasks and therefore for children's test performance. This can be explained by the fact that all test directions are given verbally in the IDS-2. Moreover, about one third and approximately a quarter of the variance in scores in the intelligence and basic skills domains, respectively, was explained by receptive language ability. With respect to the intelligence domain, especially in the intelligence composites Profile IQ, Full-Scale IQ, and Screening IQ as well as in the group factor Verbal Reasoning and its corresponding subtests Naming Categories and Naming Opposites, high associations were found with children's receptive language ability. These findings are in line with our hypotheses (see Table 1) and previous research on group differences in the IDS-2 intelligence domain by Schweizer and colleagues (2021). As the intelligence group factor Verbal Reasoning is designed to measure the broad ability Comprehension-Knowledge according to the Cattell–Horn–Carroll (CHC) model (McGrew, 1997, 2009; Schneider & McGrew, 2018), which encompasses language-based knowledge (Schneider & McGrew, 2018), mainly crystallized and verbal aspects of intelligence are covered.

In terms of basic skills, children's receptive language ability explained variance in scores on all subtests, including Logical-Mathematical Reasoning, Reading, and Spelling. These findings are consistent with our hypotheses (see Table 1) and in line with previous research that has emphasized the importance of language abilities for the acquirement of basic skills (e.g., Storch & Whitehurst, 2002). In the case of Logical-Mathematical Reasoning, we even found one of the highest associations with children's receptive language ability across all analyses. One reason for our finding might be that the items in this subtest are presented according to a flexible interview approach (Ginsburg, 1997). This includes test instructions with extensive verbal explanations and the possibility of asking children specific questions about the solution paths they used. Therefore, a high level of language comprehension may be required to understand the tasks of the Logical-Mathematical Reasoning subtest in the IDS-2. However, as previous research on the IDS and IDS-P has provided mixed evidence (Grob, Meyer, & Hagmann-von Arx, 2013; Grob, Reimann, et al., 2013; Hagmann-von Arx et al., 2013), future studies are needed to clarify these inconsistencies.

In addition, children's receptive language ability was related to all scores in the executive functions domain, with the highest amounts of variance explained in the composite and in the subtests Listing Words and Divided Attention which was in line with our hypotheses (see Table 1). As both of these subtests include long and to some degree complex instructions, more language comprehension is demanded from the participant. Moreover, one explanation for the smaller associations between receptive language ability and scores on the other executive function subtests, Animal Colors and Drawing Routes, may be that these include pictorially depicted tasks, which might help participants understand test directions.

In accordance with our hypotheses (see Table 1), children's receptive language ability explained comparatively little variance in scores on subtests in the psychomotor skills domain. This result corresponds to studies that reported no mean-level differences between children with a migration background and control samples in the IDS and IDS-P psychomotor skills domains (Grob, Reimann, et al., 2013; Hagmann-von Arx et al., 2013). One explanation for this finding might be that in tasks in the IDS-2 psychomotor skills domain, test directions are accompanied by gestures from the test administrator, which serve to demonstrate the task during the instruction phase (e.g., test administrators show how to balance on a rope before the testing starts).

Concerning the social-emotional skills domain, children's receptive language ability explained the lowest amounts of variance in scores in this domain. Hence, children's language comprehension skills play a minor role in performance on these IDS-2 tasks. For the Socially Competent Behavior subtest, we did not find a significant association with children's receptive language ability, representing the only nonsignificant association with respect to this variable. One reason for this finding may be that the instructions in this subtest are presented with detailed pictorial illustrations of social situations, which might have helped participants understand the verbal directions. Moreover, the results for the social-emotional skills domain of the IDS reported in previous studies are inconsistent (Grob, Meyer, & Hagmann-von Arx, 2013; Hagmann-von Arx et al., 2013). For example, one study found a significant mean-level difference between participants with and without a migration background in the composite score in the IDS social emotional skills domain in older children aged 9–10 years, but not in younger children between 6 and 8 years (Hagmann-von Arx et al., 2013). However, they did not examine group differences at the subtest level. Since the few previous studies have provided mixed evidence, this finding should be treated with caution and ought to be replicated by further studies.

4.2 Expressive Language Ability and Test Performance on the IDS-2

Children's expressive language ability was related to 25 of the 40 IDS-2 scores that we have analyzed but explained only little additional variance beyond children's receptive language ability in most of these scores. This is in line with our classification of linguistic loading (see Table 1), as we assumed that most of the IDS-2 subtests and intelligence group factors make low to moderate linguistic demands.

Nevertheless, in accordance with our hypotheses and previous research (e.g., Melby-Lervåg & Lervåg, 2014; Paetsch et al., 2016; Schweizer et al., 2021; Verhoeven, 2000), we observed some considerable levels of additional explained variance in the intelligence group factor Verbal Reasoning and its corresponding subtests Naming Categories and Naming Opposites as well as in the basic skills subtests Logical-Mathematical Reasoning, Reading, and Spelling. Thus, these tasks require not only the ability to understand but also the ability to produce verbal information. For example, concerning basic skills, children have to verbally explain their solution processes in the subtest Logical-Mathematical Reasoning, or they need to answer questions about a previously read text in the subtest Reading. The strong associations we have found between receptive and expressive language abilities and basic skills might also be explained by the fact that language abilities are generally central to learning in the school context, not only for understanding the content of lessons but also for being able to communicate with others about it (Paetsch et al., 2016). Concerning the subtests Naming Categories and Naming Opposites, the participant has to answer the questions verbally, with no opportunity to respond non-verbally. In addition, the quality of the verbal response is also evaluated and embedded as part of the task content, as vocabulary and verbal logical reasoning are required (Grob & Hagmann-von Arx, 2018b). Hence, the linguistic demands represent an element of the measurement *and* an aspect of the construct itself that is intended to be assessed in these two subtests.

4.3 Multilingualism and Test Performance on the IDS-2

Finally, multilingualism explained variance beyond receptive and expressive language abilities in only two subtests of the intelligence domain, namely, Story Recall and Naming Opposites. These findings correspond to earlier literature that reported group differences in these IDS-2 subtests (Schweizer et al., 2021) and in the IDS and IDS-P subtest Auditory Memory (Grob, Meyer, & Hagmann-von Arx, 2013; Grob, Reimann, et al., 2013). We assume that—as we accounted for language abilities—other components of having a linguistically diverse background contributed to the variance in Story Recall and Naming Opposites. Nevertheless, we can provide only assumptions about possible explanations for our results since these were exploratory analyses. One reason for the significant association between multilingualism and Naming Opposites might be that this subtest of Verbal Reasoning measures crystallized components of intelligence according to the CHC model (McGrew, 1997, 2009; Schneider & McGrew, 2018). This includes, besides verbal aspects, “the ability to comprehend and communicate culturally-valued knowledge…developed through experience, learning and acculturation” (Schneider & McGrew, 2018, p. 114). Therefore, this subtest might be dependent on factors such as cultural background, socialization, and previous schooling and learning experiences.

The subtest Story Recall assesses verbal Long-Term Memory and encompasses a verbal-dependent presentation and recall format (Schweizer et al., 2021). In addition, the content of the story might also partly contain socialization or culture-specific aspects as it includes names (e.g., Judith, Daniel) that are commonly used in German-speaking countries and elements, such as “inflatable boat” or “oar”, which have to be remembered by the participants. Children from linguistically diverse backgrounds might be less familiar with such names or expressions and therefore may have more difficulties remembering them. However, our results should be considered preliminary, and future research is needed.

In contrast to previous research (e.g., Foy & Mann, 2014; Yurtsever et al., 2023), we did not find any advantage for multilingual children in the executive functions domain. This might be explained by the fact that previous studies found results in favor of bilinguals mainly when examining nonverbal executive functions, whereas the IDS-2 executive function tasks tend to be more dependent on verbal skills, as the associations with language abilities in our study have shown. However, it is important to note that there is a growing body of research that challenges the bilingual advantage hypothesis in general (Lowe et al., 2021; Paap et al., 2016).

4.4 Strengths, Limitations, and Future Directions

One strength of the present study is that we relied on a comprehensive test battery and could therefore provide a more complete picture and nuanced exploration of the relative importance of language aspects in test performance in various cognitive and developmental domains compared to previous studies. In line with this, we could also investigate relations at the level of subtests, intelligence group factors, and composites. Moreover, we used standardized assessments of children's receptive and expressive language ability and included therefore an objective measure of language abilities. We also consider it a main strength that we incorporated multiple language aspects to gain insights regarding their differential contribution to test performance, thus following an integrated approach. In addition, we included a rather large sample that was representative of the percentage of multilingual children in German-speaking countries (German Federal Statistical Office, 2022; Swiss Federal Statistical Office, 2021). We also controlled for sex and SES in the hierarchical regression analyses before entering the language aspects in the model, which allowed us to take into account possible confounding effects of relevant sociodemographic characteristics (Weiss & Saklofske, 2020).

Our study also has limitations that should be acknowledged in future research. Our sample included only children aged 5–10 years because the subtest Language Skills is not administered to adolescents in the IDS-2. As previous research found age effects in the IDS (Hagmann-von Arx et al., 2013) and in the intelligence domain of the IDS-2 (Schweizer et al., 2021), future research should examine the relations between objectively measured language abilities and test scores on the IDS-2 in adolescents. Although we controlled for SES, as represented by maternal educational background, a larger proportion of the children's mothers had completed postsecondary education in our study compared to census data of the corresponding population (e.g., Swiss Federal Statistical Office, 2022b), which reduces generalizability. In addition, we had only limited information regarding children's native languages and therefore could not consider other variables, such as when the children were first exposed to the test language or how often they spoke the test language in the family context. As the majority of the children in our study attended regular school settings and did not display psychological diagnoses, our findings may not be generalizable to children with developmental risk factors or disorders (e.g., children with autism spectrum disorder). Future studies should therefore investigate the role of language aspects in test performance in the IDS-2 specifically in children with special needs, as these children typically undergo assessments of cognitive and developmental functions in psychological practice. We further encourage future research to investigate possible effects of culture on test performance by focusing on samples with culturally diverse individuals (e.g., non-European origin) and including assessments of participants' level of acculturation. Finally, studies examining measurement invariance should be conducted to test the equivalence and reliability of the IDS-2 across groups with linguistically diverse samples. For test construction, we suggest developing test batteries with minimal verbal demands. In such batteries, nonverbal instructions and response options could be implemented through gestures, pictures, and novel digital approaches using sound and video elements.

4.5 Implications

When assessing cognitive and developmental functions in multilingual participants, clinicians encounter substantial challenges, as limited proficiency in the test language could mean an individual's true ability is underestimated (Hagmann-von Arx et al., 2013). According to current guidelines on standards for psychological testing, test administrators should therefore examine the validity of score interpretations for participants who have reduced proficiency in the test language (American Educational Research Association et al., 2014). In line with this, our study provides evidence to support the claim that information about participants' language proficiency should be gathered during the diagnostic process (Flanagan et al., 2007), if there are hints that the participant may have linguistic disadvantages in the test language (e.g., assessed as part of the anamnestic evaluation and clinical judgment). By doing this, it would be possible to determine where on the continuum of language proficiency the participants are located, which would lead to a more accurate consideration of possible language effects on individuals' test performance (Cormier et al., 2022).

Considering that the IDS-2 even includes a standardized measure of language skills, we therefore suggest that—when there is a suspicion that the participant may have difficulties in the test language—clinicians should assess children's receptive language abilities prior to the administration of all domains of the IDS-2. We further propose measuring children's expressive language abilities before administering the IDS-2 intelligence and basic skills domains to participants at risk for linguistic disadvantages. If proficiency in the test language cannot be guaranteed or multilingual participants show, despite sufficient language abilities, considerable difficulties in IDS-2 subtests with high verbal demands, a nonverbal or culture-fair test should be conducted. If there is no such test battery available for the specific domain, possible linguistic disadvantages have to be considered in the interpretation of the IDS-2 test scores. However, as basic skills represent "cultural skills" that are, for example, related to input from the school environment and therefore reflect previous language experiences (Grob & Hagmann-von Arx, 2018b; Köller & Baumert, 2008), associations between language abilities and performance on the basic skills subtests are to be expected. Hence, the validity of test score interpretations in this domain is not compromised by insufficient test language proficiency. However, participants' language abilities should still be taken into account in the interpretation of scores derived from the basic skills domain. Finally, although measuring participants' language abilities in addition to the cognitive or developmental domain that covers the main question of the psychological assessment will result in a somewhat longer duration of the test administration, it is essential to consider children's proficiency in the test language to draw accurate and valid conclusions for recommendations, treatment, and high-stakes diagnostic decisions.

Footnotes

1. Although there may also be cultural aspects that are associated with assessment in this context, in the present study we focused on the relative importance of language aspects for participants' test performance.
2. We formed the variable *multilingualism* using information regarding children's native language: Multilingual children spoke German (the test language) and had at least one other language as their native language, whereas monolingual children indicated German as their native language.
3. As we used age-standardized scores, we did not additionally control for age.
4. In this sample, the following medical and psychological diagnoses were reported: visual problems (n = 59), hearing problems (n = 6), motor problems (n = 16), speech problems (n = 16), giftedness (n = 21), intellectual disabilities (n = 19), scholastic problems (n = 15), behavioral problems (n = 20), autism spectrum disorder (n = 5), affective symptoms (n = 2), physical problems (n = 28), and other (n = 2).
5. We did not perform analyses for the basic skills composite because receptive and expressive language abilities are also contained in the calculation of this score. Moreover, we did not examine the IDS-2 domain motivation and attitude as its subtests are administered only to adolescents aged 11–20 years.
6. To control for effects of intelligence, we repeated the hierarchical regression analyses with intelligence (Profile IQ) included as a predictor of the IDS-2 domains executive functions, psychomotor skills, social-emotional skills, and basic skills. Although children's language abilities were no longer associated with many of the IDS-2 scores after accounting for intelligence, the overall pattern of results remained largely the same, as receptive and expressive language abilities were still particularly related to the subtests of the basic skills domain (see Tables S12 to S15 in the Supplement for full results). However, as the Profile IQ is highly correlated with language abilities (r = .50 to .60), these findings need to be treated with caution and further research should use a nonverbal intelligence assessment.

5. References

Abutalebi, J., & Green, D. W. (2008). Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Language and Cognitive Processes*, *23*(4), 557–582. https://doi.org/10.1080/01690960801920602

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*.

Austrian Federal Statistical Office. (2022). *Mehr als ein Viertel der österreichischen Gesamtbevölkerung hat Migrationshintergrund [More than a quarter of the total Austrian population has a migration background]*. https://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/bevoelkerung/bevoelkerungsstruktur/bevoelkerung_nach_migrationshintergrund/033240.html

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology*, *23*(5), 611–626.

Baltes, P. B. (1990). Entwicklungspsychologie der Lebensspanne: Theoretische Leitsätze [Lifespan developmental psychology: Guiding theoretical principles]. *Psychologische Rundschau*, *41*, 1–24.

Barac, R., Bialystok, E., Castro, D. C., & Sanchez, M. (2014). The cognitive development of young dual language learners: A critical review. *Early Childhood Research Quarterly*, *29*(4), 699–714. https://doi.org/10.1016/j.ecresq.2014.02.003

Bialystok, E., & Craik, F. I. M. (2010). Cognitive and linguistic processing in the bilingual mind. *Current Directions in Psychological Science*, *19*(1), 19–23. https://doi.org/10.1177/0963721409358571

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *13*(4), 525–531. https://doi.org/10.1017/S1366728909990423

Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Valtin, R., & Walther, G. (2007). Erste Ergebnisse aus IGLU: Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich [First results from IGLU: Student performance at the end of the fourth grade in international comparison]. *Sozialwissenschaftlicher Fachinformationsdienst soFid, Bildungsforschung*, *2007*(1), 9–46.

Calero, M. D., Fernández-Parra, A., López-Rubio, S., Carles, R., Mata, S., Vives, M. del C., Navarro, E., & Márquez, J. (2013). Variables involved in personal, social and school adjustment in a sample of preschool-aged children from different cultural backgrounds. *European Journal of Psychology of Education*, *28*(1), 133–155. https://doi.org/10.1007/s10212-012-0107-8

Cormier, D. C., Bulut, O., McGrew, K. S., & Kennedy, K. (2022). Linguistic influences on cognitive test performance: Examinee characteristics are more important than test characteristics. *Journal of Intelligence*, *10*(1), Article 1. https://doi.org/10.3390/jintelligence10010008

Cormier, D. C., McGrew, K. S., & Evans, J. J. (2011). Quantifying the "degree of linguistic demand" in spoken intelligence test directions. *Journal of Psychoeducational Assessment*, *29*(6), 515–533. https://doi.org/10.1177/0734282911405962

Ehm, J.-H., Duzy, D., & Hasselhorn, M. (2011). Das akademische Selbstkonzept bei Schulanfängern [The academic self-concept of pupils at the beginning of school]. *Frühe Bildung*, *0*(0), 37–45. https://doi.org/10.1026/2191-9186/a000008

Flanagan, D. P., & Harrison, P. L. (2012). *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed.). Guilford Press.

Flanagan, D. P., & Ortiz, S. O. (2001). *Essentials of cross-battery assessment*. Wiley.

Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2007). *Essentials of cross-battery assessment with C/D ROM* (2nd ed.). Wiley.

Foy, J. G., & Mann, V. A. (2014). Bilingual children show advantages in nonverbal auditory executive function task. *International Journal of Bilingualism*, *18*(6), 717–729. https://doi.org/10.1177/1367006912472263

German Federal Statistical Office. (2022). *Statistiken der Kinder- und Jugendhilfe [Statistics of child and youth services]*. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Soziales/Kindertagesbetreuung/Publikationen/Downloads-Kindertagesbetreuung/tageseinrichtungen-kindertagespflege-5225402227004.pdf

German Federal Statistical Office. (2023). *Bevölkerung nach Migrationshintergrund und Geschlecht [Population by migration background and gender]*. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/Tabellen/liste-migrationshintergrund-geschlecht.html

Ginsburg, H. P. (1997). *Entering the child's mind*. Cambridge University Press.

Goldstein, S., Princiotta, D., & Naglieri, J. A. (2015). *Handbook of intelligence. Evolutionary theory, historical perspective, and current concepts*. Springer.

Grieder, S., & Grob, A. (2020). Exploratory factor analyses of the Intelligence and Development Scales–2: Implications for theory and practice. *Assessment*, *27*(8), 1853–1869. https://doi.org/10.1177/1073191119845051

Grob, A., & Hagmann-von Arx, P. (2018a). *Intelligence and Development Scales–2 (IDS-2). Intelligenz- und Entwicklungsskalen für Kinder und Jugendliche [Intelligence and Development Scales for Children and Adolescents]*. Hogrefe.

Grob, A., & Hagmann-von Arx, P. (2018b). *Intelligence and Development Scales–2 (IDS-2). Intelligenz- und Entwicklungsskalen für Kinder und Jugendliche. Manual zu Theorie, Interpretation und Gütekriterien [Intelligence and Development Scales for Children and Adolescents. Manual on theory, interpretation and psychometric criteria]*. Hogrefe.

Grob, A., Hagmann-von Arx, P., Barnett, A., Stuart, N., & Vanzan, S. (2021). *Intelligence and Development Scales–2 (IDS-2). Intelligence and Development Scales for Children and Adolescents*. Hogrefe.

Grob, A., Hagmann-von Arx, P., Ferri, R., Rea, M., & Casagrande, M. (2022). *Intelligence and Development Scales–2. Scale di intelligenza e sviluppo per bambini e adolescenti [Intelligence and Development Scales for Children and Adolescents]*. Hogrefe.

Grob, A., Hagmann-von Arx, P., Jaworowska, A., Matczak, A., & Fecenec, D. (2019). *Intelligence and Development Scales–2. Inteligencji i Rozwoju dla Dzieci i Młodzieży [Intelligence and Development Scales for Children and Adolescents]*. Hogrefe.

Grob, A., Hagmann-von Arx, P., Ruiter, S. A. J., Timmerman, M. E., & Visser, L. (2018). *Intelligence and Development Scales–2 (IDS-2). Intelligentie- en ontwikkelingsschalen voor kinderen en jongeren. [Intelligence and Development Scales for Children and Adolescents]*. Hogrefe.

Grob, A., Keller, K., & Troesch, L. M. (2014). *ZWEITSPRACHE–Mit ausreichenden Deutschkenntnissen in den Kindergarten [Second language–with sufficient German skills in kindergarten]*. University of Basel.

Grob, A., Meyer, C. S., & Hagmann-von Arx, P. (2009). *Intelligence and Development Scales (IDS)*. Huber.

Grob, A., Meyer, C. S., & Hagmann-von Arx, P. (2013). *Intelligence and Development Scales (IDS): Intelligenz- und Entwicklungsskalen für Kinder von 5–10 Jahren [Intelligence and Development Scales for Children 5–10 years old]*. Hans Huber.

Grob, A., Reimann, G., Gut, J., & Frischknecht, M.-C. (2013). *Intelligence and Development Scales–Preschool (IDS-P): Intelligenz- und Entwicklungsskalen für das Vorschulalter [Intelligence and Development Scales for Preschool Children]*. Hans Huber.

Grundy, J. G. (2020). The effects of bilingualism on executive functions: An updated quantitative analysis. *Journal of Cultural Cognitive Science*, *4*(2), 177–199. https://doi.org/10.1007/s41809-020-00062-5

Guthke, J., & Wiedl, K. H. (1996). *Dynamisches Testen: Zur Psychodiagnostik der intraindividuellen Variabilität [Dynamic testing: On the psychodiagnostics of intraindividual variability]*. Hogrefe.

Hagmann-von Arx, P., Petermann, F., & Grob, A. (2013). Konvergente und diskriminante Validität der WISC-IV und der Intelligence and Development Scales (IDS) bei Kindern mit Migrationshintergrund [Convergent and discriminant validity of the WISC-IV and the Intelligence and Development Scales (IDS) in children with a migration background]. *Diagnostica*, *59*(4), 170–182. https://doi.org/10.1026/0012-1924/a000091

Hessels, M. (1997). Low IQ but high learning potential: Why Zeyneb and Moussa do not belong in special education. *Educational and Child Psychology*, *14*, 121–136.

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, *39*(1), 1–27. https://doi.org/10.1017/S0305000910000759

Hommel, G. (1988). A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika*, *75*, 383–386.

Kastner, J. W., May, W., & Hildman, L. (2001). Relationship between language skills and academic achievement in first grade. *Perceptual and Motor Skills*, *92*(2), 381–390. https://doi.org/10.2466/pms.2001.92.2.381

Klingner, J. K., Blanchett, W. J., & Harry, B. (2007). Race, culture, and developmental disabilities. In S. L. Odom, R. H. Horner, M. E. Snell, & J. Blacher (Eds.), *Handbook of developmental disabilities* (pp. 55–75). Guilford Press.

Köller, O., & Baumert, J. (2008). Entwicklung schulischer Leistungen [Development of scholastic performance]. In R. Oerter & L. Montada (Eds.), *Entwicklungspsychologie [Developmental psychology]* (Vol. 6, pp. 735–768). Beltz. https://pure.mpg.de/pubman/faces/ViewItemOverviewPage.jsp?itemId=item_2102046

Konsortium PISA.ch. (2018). *PISA 2015: Schülerinnen und Schüler der Schweiz im internationalen Vergleich [PISA 2015: Swiss students in international comparison]*. SBFI/EDK und Konsortium PISA.ch.

Konsortium PISA.ch. (2019). *PISA 2018: Schülerinnen und Schüler der Schweiz im internationalen Vergleich [PISA 2018: Swiss students in international comparison]*. SBFI/EDK and Konsortium PISA.ch.

Kroll, J. F., Dussias, P. E., Bogulski, C. A., & Kroff, J. R. V. (2012). Chapter seven - Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In B. H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 56, pp. 229–262). Academic Press. https://doi.org/10.1016/B978-0-12-394393-4.00007-8

Lee Swanson, H., Rosston, K., Gerber, M., & Solari, E. (2008). Influence of oral language and phonological awareness on children's bilingual reading. *Journal of School Psychology*, *46*(4), 413–429. https://doi.org/10.1016/j.jsp.2007.07.002

Lienert, G. A., & Raatz, U. (1998). *Testaufbau und Testanalyse [Test design and test analysis]*. Beltz.

Lowe, C. J., Cho, I., Goldsmith, S. F., & Morton, J. B. (2021). The bilingual advantage in children's executive functioning is not related to language status: A meta-analytic review. *Psychological Science*, *32*(7), 1115–1146. https://doi.org/10.1177/0956797621993108

McGrew, K. S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive Gf-Gc framework. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 151–119). Guilford Press.

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), 1–10. https://doi.org/10.1016/j.intell.2008.08.004

Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first- and second-language learners. *Psychological Bulletin*, *140*(2), 409–433. https://doi.org/10.1037/a0033890

Odermatt, S. D., Möhring, W., Grieder, S., & Grob, A. (2022). Cognitive and developmental functions in autistic and non-autistic children and adolescents: Evidence from the Intelligence and Development Scales–2. *Journal of Intelligence*, *10*(4), Article 4. https://doi.org/10.3390/jintelligence10040112

Ortiz, S. O. (2019). On the measurement of cognitive abilities in English learners. *Contemporary School Psychology*, *23*(1), 68–86. https://doi.org/10.1007/s40688-018-0208-8

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, *98*(3), 554–566. https://doi.org/10.1037/0022-0663.98.3.554

Paap, K. R., Johnson, H. A., & Sawi, O. (2016). Should the search for bilingual advantages in executive functioning continue? *Cortex*, *74*, 305–314. https://doi.org/10.1016/j.cortex.2015.09.010

Paetsch, J., Radmann, S., Felbrich, A., Lehmann, R., & Stanat, P. (2016). Sprachkompetenz als Prädiktor mathematischer Kompetenzentwicklung von Kindern deutscher und nicht-deutscher Familiensprache [Language competence as a predictor of mathematical competence development of children of German and non-German family language]. *Zeitschrift Für Entwicklungspsychologie und Pädagogische Psychologie*, *48*(1), 27–41. https://doi.org/10.1026/0049-8637/a000142

R Core Team. (2022). *R: A language and environment for statistical computing* (Version 2022.12.0+353) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Reich, H. H., Roth, H.-J., Dirim, I., Jørgensen, J. N., List, G., List, G., Neumann, U., Siebert-Ott, G., Steinmüller, U., Teunissen, F., Vallen, T., & Wurnig, V. (2002). *Spracherwerb zweisprachig aufwachsender Kinder und Jugendlicher: Ein Überblick über den Stand der nationalen und internationalen Forschung [Language acquisition of children and adolescents growing up bilingually: An overview of the state of national and international research]*. Behörde für Bildung und Sport [Authority for Education and Sports].

Rodriguez-Fornells, A., Rotte, M., Heinze, H.-J., Nösselt, T., & Münte, T. F. (2002). Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature*, *415*(6875), Article 6875. https://doi.org/10.1038/4151026a

Schneider, W. J., & McGrew, K. S. (2018). The Cattell–Horn–Carroll theory of cognitive abilities. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (4th ed., pp. 73–163). Guilford Press.

Schrank, F. A., McGrew, K. S., & Mather, N. (2014). *Woodcock—Johnson IV Tests of Cognitive Abilities*. Riverside Publishing.

Schweizer, F., Grieder, S., Bünger, A., & Grob, A. (2021). Vergleich von Intelligenztestleistungen bei monolingualen und bilingualen Kindern und Jugendlichen in den Intelligence and Development Scales–2 (IDS-2) [Comparison of intelligence test performance of monolingual and bilingual children and adolescents on the Intelligence and Development Scales–2 (IDS-2)]. *Diagnostica*, *67*(1), 36–46. https://doi.org/10.1026/0012-1924/a000260

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*, *38*(6), 934–947. https://doi.org/10.1037/0012-1649.38.6.934

Sullivan, A. L. (2011). Disproportionality in special education identification and placement of English language learners. *Exceptional Children*, *77*(3), 317–334. https://doi.org/10.1177/001440291107700304

Swiss Federal Statistical Office. (2021). *Erhebung zur Sprache, Religion und Kultur 2019 [Language, Religion, and Culture Survey 2019]*. https://www.bfs.admin.ch/bfs/de/home/statistiken/kataloge-datenbanken/medienmitteilungen.assetdetail.15384140.html

Swiss Federal Statistical Office. (2022a). *Bevölkerung nach Migrationsstatus [Population by migration status]*. https://www.bfs.admin.ch/bfs/de/home/statistiken/bevoelkerung/migration-integration/nach-migrationsstatuts.html

Swiss Federal Statistical Office. (2022b). *Bildungsstand der Wohnbevölkerung nach Alter und Geschlecht [Educational attainment of the resident population by age and gender]*. https://www.bfs.admin.ch/bfs/de/home/statistiken/kataloge-datenbanken/tabellen.assetdetail.22024469.html

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences*, *104*(30), 12530–12535. https://doi.org/10.1073/pnas.0609927104

United Nations. (2022). *World migration report 2022*. https://worldmigrationreport.iom.int/wmr-2022-interactive/

Verhoeven, L. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading*, *4*(4), 313–330. https://doi.org/10.1207/S1532799XSSR0404_4

Weiss, L. G., Harris, J. G., Prifitera, A., Courville, T., Rolfhus, E., Saklofske, D. H., & Holdnack, J. A. (2006). WISC-IV interpretation in societal context. The essentials and beyond. In L. G. Weiss, D. H. Saklofske, A. Prifitera, & J. A. Holdnack (Eds.), *WISC-IV: Advanced clinical interpretation* (pp. 1–49). Elsevier.

Weiss, L. G., & Saklofske, D. H. (2020). Mediators of IQ test score differences across racial and ethnic groups: The case for environmental and social justice. *Personality and Individual Differences*, *161*, Article 109962. https://doi.org/10.1016/j.paid.2020.109962

Yurtsever, A., Anderson, J. A. E., & Grundy, J. G. (2023). Bilingual children outperform monolingual children on executive function tasks far more often than chance: An updated quantitative analysis. *Developmental Review*, *69*, Article 101084. https://doi.org/10.1016/j.dr.2023.101084

Table 1

Cognitive and Developmental Functions From the Intelligence and Development Scales–2 Ranked According to Their Assumed Degree of Linguistic Loading

Domain	Low linguistic loading variable	Moderate linguistic loading variable	High linguistic loading variable
Intelligence	*Visual Processing*	*Auditory Short-Term Memory*	*Verbal Reasoning*
	Shape Design	Digit and Letter Span	Naming Categories
	Washer Design	Mixed Digit and Letter Span	Naming Opposites
	Processing Speed	*Long-Term Memory*	*Long-Term Memory*
	Parrots	Picture Recall	Story Recall
	Boxes		
	Abstract Reasoning		
	Matrices: Completion		
	Matrices: Odd One Out		
	Visuospatial Short-Term Memory		
	Shape Memory		
	Rotated Shape Memory		
Executive functions	Animal Colors	Listing Words	
	Drawing Routes	Divided Attention	
Psychomotor skills	Gross Motor Skills		
	Fine Motor Skills		
	Visuomotor Skills		
Social-emotional skills		Identifying Emotions	
		Regulating Emotions	
		Socially Competent Behavior	
Basic skills			Logical-Mathematical Reasoning
			Reading
			Spelling

Note. It is assumed that performances in subtests and in intelligence group factors of the IDS-2 that are classified as “low linguistic loading” would be least susceptible to language effects, whereas performances in subtests and in intelligence group factors of the IDS-2 that are classified as “high linguistic loading” would be most susceptible to language effects. Long-Term Memory is listed in two categories as it contains subtests with different linguistic loadings. Intelligence group factors are printed in italics.

Table 2

Descriptive Statistics of Predictors and Composites, Intelligence Group Factors, and Subtests From the Intelligence and Development Scales–2 (IDS-2)

Predictor or score	*M*	*SD*	*n*	%	Minimum	Maximum	Skewness	Kurtosis	Reliability
Predictor									
Sex (female)			424	51.3%					
SES (no postsecondary maternal education)			492	59.6%					
Receptive language ability[a]	10.33	2.99			1	19	-0.25	0.43	0.88
Expressive language ability[a]	10.28	3.08			1	19	-0.46	0.06	0.92
Multilingualism (yes)			215	26.0%					
IDS-2 scores									
Profile IQ[b]	100.42	15.28			55	145	-0.55	0.52	0.99
Full-Scale IQ[b]	100.41	15.3			55	142	-0.47	0.27	0.98
Screening IQ[b]	100.58	15.25			55	145	-0.19	0.12	0.97
Visual Processing[b]	99.75	15.67			55	141	-0.19	0.16	0.97
Processing Speed[b]	100.06	15.3			55	145	-0.1	0.53	0.97
Auditory Short-Term Memory[b]	100.53	15.25			55	145	-0.15	0.01	0.96
Visuospatial Short-Term Memory[b]	100.07	14.65			55	145	-0.14	0.57	0.95
Abstract Reasoning[b]	100.06	15.67			55	145	-0.04	0.19	0.97
Verbal Reasoning[b]	100.46	14.4			55	139	-0.54	0.46	0.97
Long-Term Memory[b]	100.03	14.74			55	142	-0.23	0.29	0.97
Shape Design[a]	10.18	3.15			1	19	-0.18	0.34	0.88
Washer Design[a]	10.29	3.08			1	19	-0.06	0.16	0.88
Parrots[a]	10.20	3.00			1	19	-0.34	0.57	0.87
Boxes[a]	10.09	3.07			1	19	-0.16	0.89	0.86
Digit and Letter Span[a]	10.50	2.84			1	19	-0.14	0.15	0.83
Mixed Digit and Letter Span[a]	10.66	2.94			1	19	-0.24	0.07	0.80
Shape Memory[a]	10.33	2.91			1	19	-0.01	0.69	0.80
Rotated Shape Memory[a]	10.22	2.41			1	19	-0.25	0.92	0.80
Matrices: Completion[a]	10.50	3.12			1	19	0.17	0.18	0.88
Matrices: Odd One Out[a]	10.26	3.01			1	19	0.09	0.24	0.88
Naming Categories[a]	10.51	2.89			1	19	-0.41	0.64	0.89
Naming Opposites[a]	10.56	2.91			1	17	-0.75	0.93	0.85
Story Recall[a]	10.23	3.04			1	19	-0.38	0.47	0.92
Picture Recall[a]	10.52	2.89			1	19	0.04	0.34	0.88

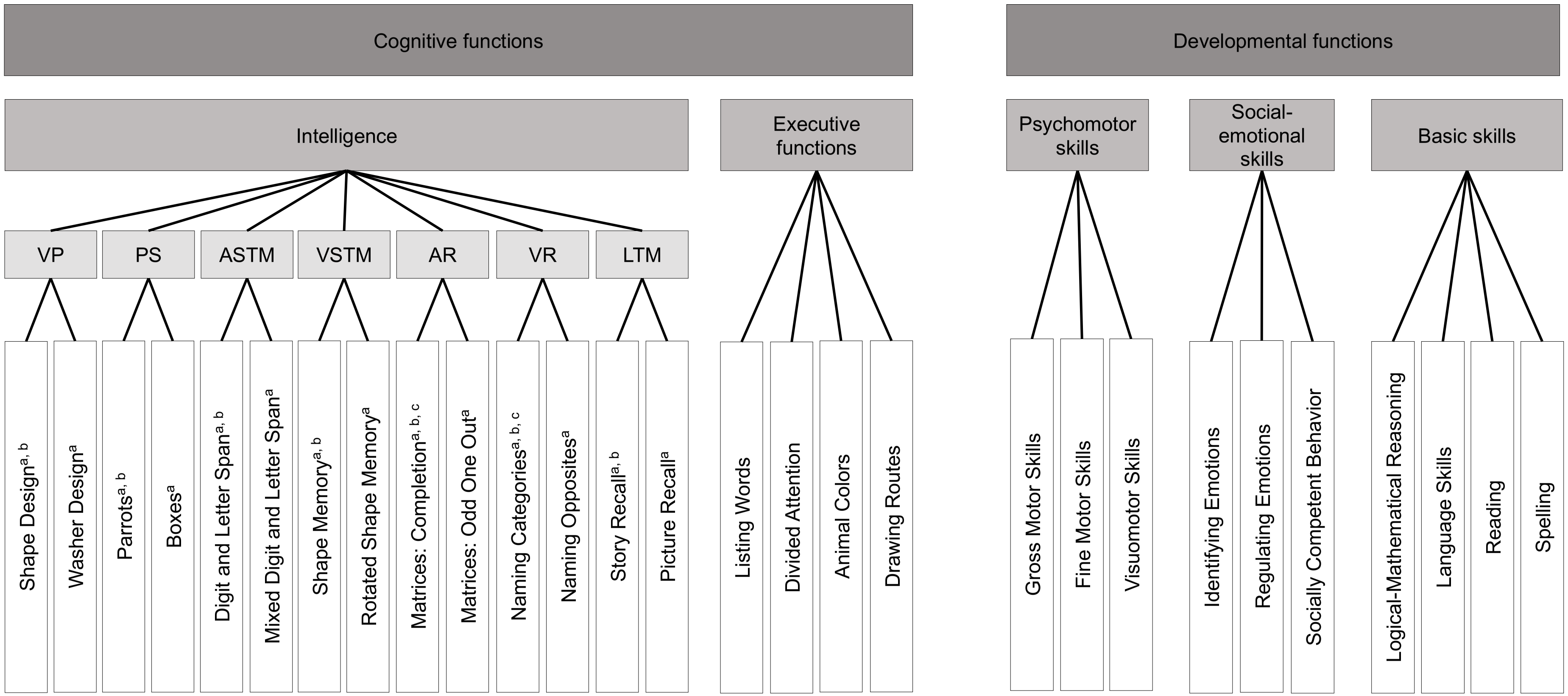
Predictor or score	*M*	*SD*	*n*	%	Minimum	Maximum	Skewness	Kurtosis	Reliability
Executive functions composite[b]	10.11	1.90			3.12	15.75	-0.28	0.35	0.97
Listing Words[c]	9.93	3.07			1	18	-0.09	-0.14	0.75
Divided Attention[b]	10.11	2.31			1.5	18	-0.45	0.96	0.91
Animal Colors[c]	10.08	3.03			1	19	0.01	0.04	0.72
Drawing Routes[b]	10.34	2.49			2.5	17	-0.33	-0.11	0.95
Psychomotor skills composite[b]	10.33	1.86			2.67	16.83	-0.32	0.6	0.97
Gross Motor Skills[a]	10.51	3.12			1	19	-0.1	0.03	0.71
Fine Motor Skills[b]	10.35	2.39			1.5	19	-0.1	0.4	0.95
Visuomotor Skills[b]	10.14	2.08			2.5	19	-0.11	0.69	0.94
Social-emotional skills composite[b]	10.50	1.92			1.33	14.67	-1.15	2.55	0.95
Identifying Emotions[c]	10.29	2.37			1	14	-1.43	2.19	0.85
Regulating Emotions[c]	10.49	2.66			1	15	-0.89	0.71	0.78
Socially Competent Behavior[c]	10.73	2.9			1	19	-0.58	1.06	0.71
Basic skills composite[b]	10.13	2.33			1	17	-0.42	0.88	0.98
Logical-Mathematical Reasoning[a]	10.22	3.03			1	19	-0.1	0.37	0.94
Reading[b]	9.72	2.73			1	17.5	-0.36	0.33	0.97
Spelling[a]	10.15	3.11			1	19	-0.38	0.49	0.72

Note. SES = socioeconomic status.

[a]Cronbach's alpha reported. [b]Reported reliability calculated according to a formula by Lienert and Raatz (1998). [c]Retest reliability reported.

Figure 1

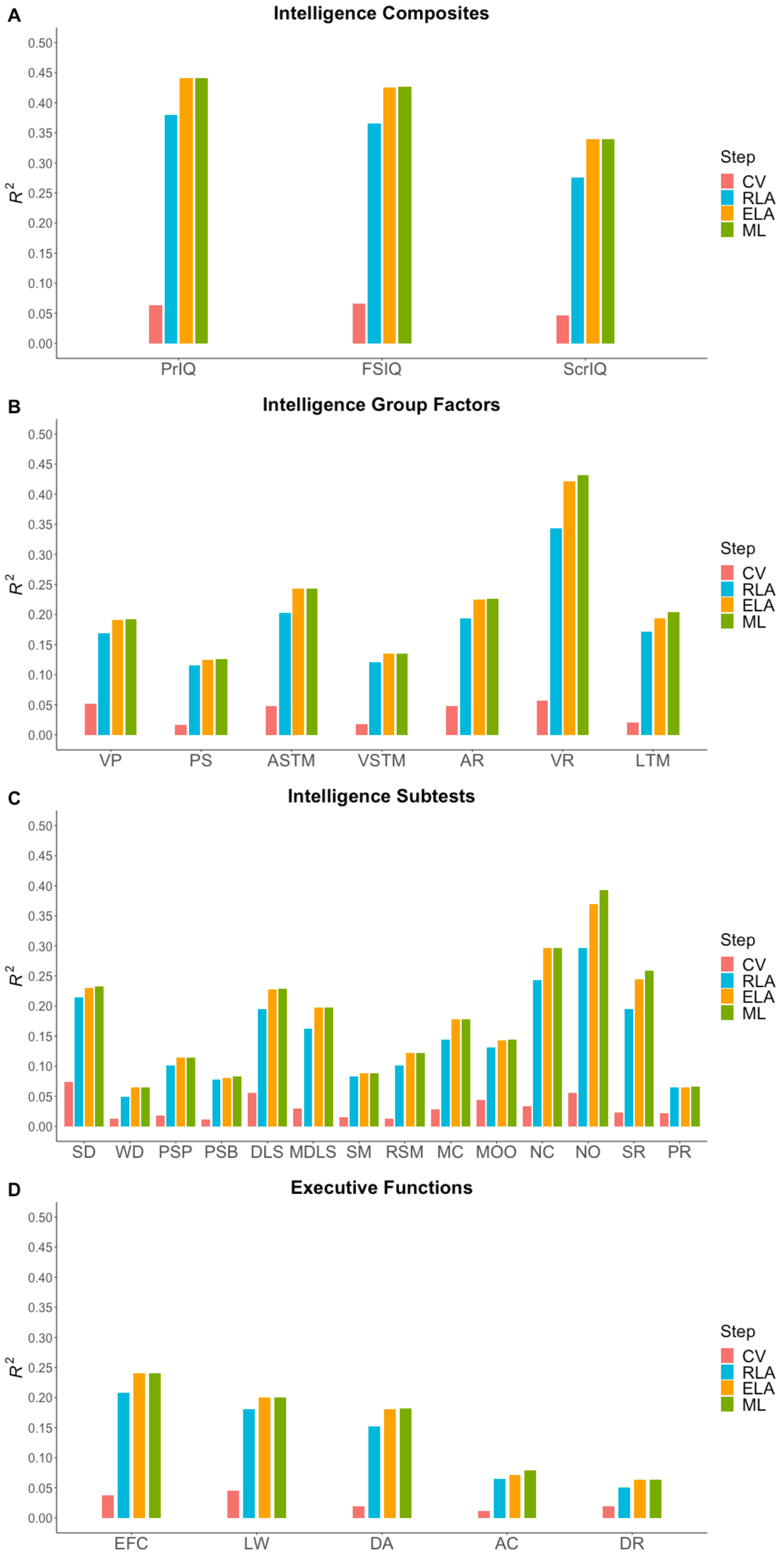
Overview of the Cognitive and Developmental Functions Assessed in the Intelligence and Development Scales–2 (IDS-2) and Included in Our Study



Note. Functions are depicted in dark gray, domains in medium gray, intelligence group factors in light gray, and subtests in white. Reading and Spelling only for ages 7 to 20 years. The domain motivation and attitude is not shown as it is only for ages 11 to 20 and was therefore omitted in our analyses. VP = Visual Processing; PS = Processing Speed; ASTM = Auditory Short-Term Memory; VSTM = Visuospatial Short-Term Memory; AR = Abstract Reasoning; VR = Verbal Reasoning; LTM = Long-Term Memory. [a] Subtests included in the Profile IQ. [b] Subtests included in the Full-Scale IQ. [c] Subtests included in the Screening IQ.

Figure 2

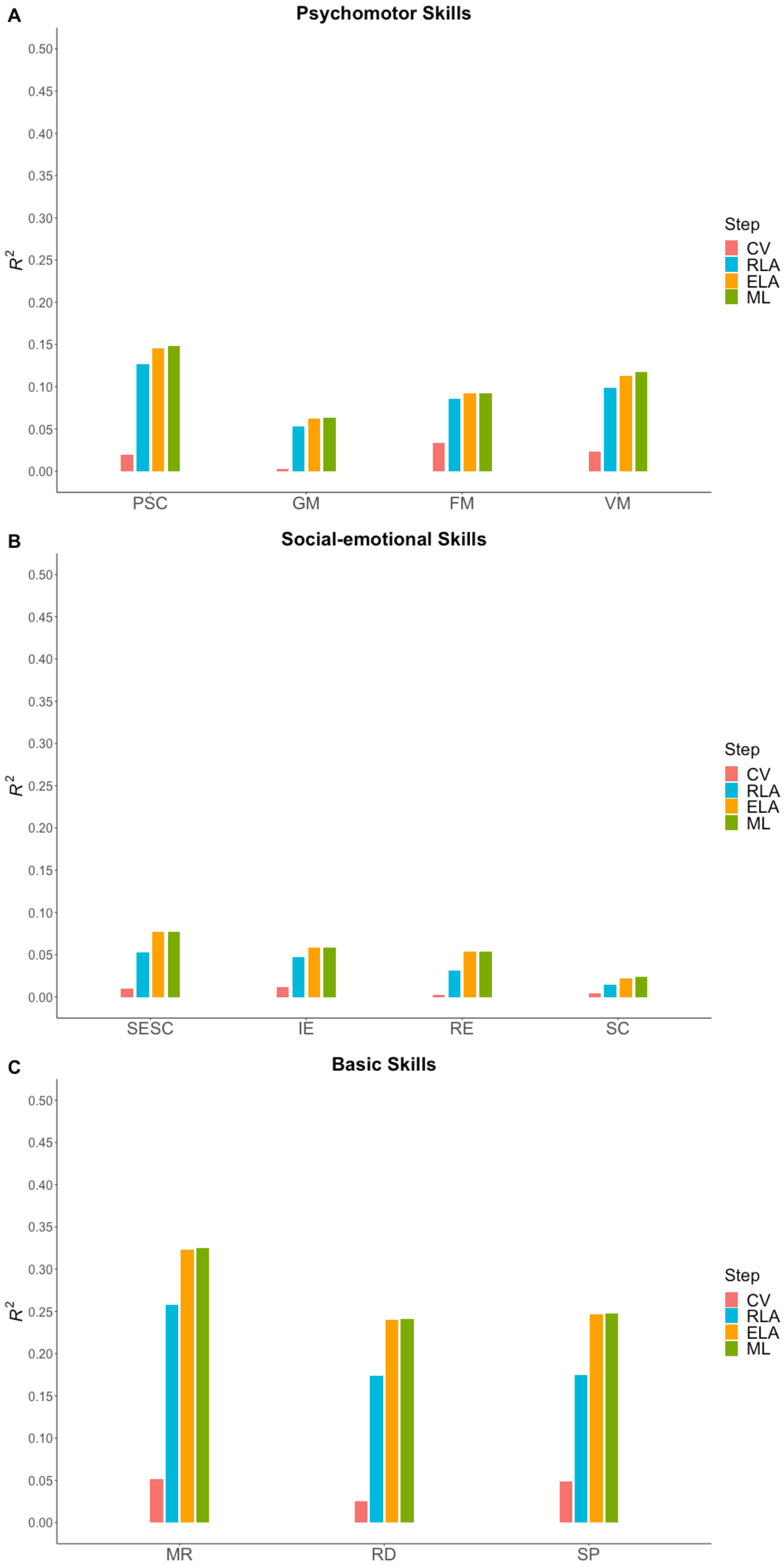
The Proportion of Variance Explained by Each Predictor in the Scores on the Cognitive Functions From the Intelligence and Development Scales–2



Note. The explained variance by each of the four predictors (i.e., Step 1: control variables, Step 2: receptive language ability, Step 3: expressive language ability, Step 4: multilingualism) is shown for (**A**) intelligence composites, (**B**) intelligence group factors, (**C**) intelligence subtests, and (**D**) executive functions composite and subtests of the Intelligence and Development Scales–2. CV = control variables (i.e., sex and socioeconomic status); RLA = receptive language ability; ELA = expressive language ability; ML = multilingualism; PrIQ = Profile IQ; FSIQ = Full-Scale IQ; ScrIQ = Screening IQ; VP = Visual Processing; PS = Processing Speed; ASTM = Auditory Short-Term Memory; VSTM = Visuospatial Short-Term Memory; AR = Abstract Reasoning; VR = Verbal Reasoning; LTM = Long-Term Memory; SD = Shape Design; WD = Washer Design; PSP = Parrots; PSB = Boxes; DLS = Digit and Letter Span; MDLS = Mixed Digit and Letter Span; SM = Shape Memory; RSM = Rotated Shape Memory; MC = Matrices: Completion; MOO = Matrices: Odd One Out; NC = Naming Categories; NO = Naming Opposites; SR = Story Recall; PR = Picture Recall; EFC = Executive functions composite; LW = Listing Words; DA = Divided Attention; AC = Animal Colors; DR = Drawing Routes.

Figure 3

The Proportion of Variance Explained by Each Predictor in the Scores on the Developmental Functions From the Intelligence and Development Scales–2



Note. The explained variance by each of the four predictors (i.e., Step 1: control variables, Step 2: receptive language ability, Step 3: expressive language ability, Step 4: multilingualism) is shown for (**A**) psychomotor skills composite and subtests, (**B**) social-emotional skills composite and subtests, and (**C**) basic skills subtests of the Intelligence and Development Scales–2. CV = control variables (i.e., sex and socioeconomic status); RLA = receptive language ability; ELA = expressive language ability; ML = multilingualism; PSC = Psychomotor skills composite; GM = Gross Motor Skills; FM = Fine Motor Skills; VM = Visuomotor Skills; SESC = Social-emotional skills composite; IE = Identifying Emotions; RE = Regulating Emotions; SC = Socially Competent Behavior; MR = Logical-Mathematical Reasoning; RD = Reading; SP = Spelling.

Supplementary Material

Table S1

Description of the Domains of the Intelligence and Development Scales–2 (IDS-2) Included in Our Study

Domain	Group factor	Subtest	No. of items	Description
Intelligence	Visual Processing (VP)	Shape Design (SD)	20	Reproduce geometric figures with the help of rectangles and triangles
		Washer Design (WD)	2–4[a]	Reproduce counter patterns according to a template
	Processing Speed (PS)	Parrots (PSP)	56–180[a]	Cross out parrots with two orange features that look to the left from rows of different parrots
		Boxes (PSB)	104–180[a]	Cross out groups of three or four boxes from rows of different groups of boxes
	Auditory Short-Term Memory (ASTM)	Digit and Letter Span (DLS)	40	Repeat number and letter sequences forward and backward
		Mixed Digit and Letter Span (MDLS)	36	Repeat mixed number and letter sequences forward and backward
	Visuospatial Short-Term Memory (VSTM)	Shape Memory (SM)	23	Remember figures and recognize them from a selection of figures and positions
		Rotated Shape Memory (RSM)	23	Remember figures and recognize them from a selection of rotated figures and positions
	Abstract Reasoning (AR)	Matrices: Completion (MC)	35	Understand how a figure changes and transfer these changes to a continuing figure
		Matrices: Odd One Out (MOO)	31	Select from presented pictures the one that does not fit with the others
	Verbal Reasoning (VR)	Naming Categories (NC)	34	Name categories for a group of pictures or words
		Naming Opposites (NO)	34	Name opposites of presented words
	Long-Term Memory (LTM)	Story Recall (SR)	19–32[a]	Listen to a semantically meaningful story and recall it after at least 20 min
		Picture Recall (PR)	11–21[a]	Look at a picture and recall key features and details after at least 20 min
Executive functions		Listing Words (LW)	2-4[a]	List words based on categories or starting letters
		Divided Attention (DA)	50-100[a]	Cross out parrots with two orange features that look to the left from different parrots and list animals
		Animal Colors (AC)	3	Say colors of animals as fast as possible
		Drawing Routes (DR)	14	Travel given routes as fast as possible once

Domain	Group factor	Subtest	No. of items	Description
Psychomotor skills		Gross Motor Skills (GM)	3	Balance on a rope, catch and throw a ball, and jump sideways over a rope
		Fine Motor Skills (FM)	6	Quickly screw nuts on and off bolts of different sizes and quickly thread beads of different sizes
		Visuomotor Skills (VM)	12	Move exactly between lines, draw figures, and reflect figures
Social-emotional skills		Identifying Emotions (IE)	10	Recognize and name emotions of children in photos
		Regulating Emotions (RE)	6-9[a]	Specify regulation strategies for the emotions of anger, fear, and grief
		Socially Competent Behavior (SC)	6-9[a]	Name socially competent behavior according to a presented social situation
Basic skills		Logical-Mathematical Reasoning (MR)	64	Solve logical-mathematical reasoning tasks
		Expressive Language Ability (ELA)	12	Form sentences from several words
		Receptive Language Ability (RLA)	15	Carry out instructions
		Reading (RD)	3	Including reading words, reading pseudo words, and text comprehension
		Spelling (SP)	40-60[a]	Word dictation

Note. The intelligence domain includes three intelligence composites: (1) Profile IQ (PrIQ) is based on all 14 intelligence subtests, (2) Full-Scale IQ (FSIQ) is based on one subtest from every intelligence group factor (i.e., Shape Design, Parrots, Digit and Letter Span, Shape Memory, Matrices: Completion, Naming Categories, Story Recall), and (3) Screening IQ (ScrIQ) is based on one subtest from the intelligence group factors Abstract Reasoning (i.e., Matrices: Completion) and Verbal Reasoning (i.e., Naming Categories). Reading and Spelling only for ages 7 to 20 years. The domain motivation and attitude is not shown as it is only for ages 11 to 20 years and was therefore omitted in our analyses. [a] Depending on age and skill.

Table S2
Characteristics of Monolingual and Multilingual Children

Characteristic	Monolingual *N* = 611				Multilingual *N* = 215				*t*	χ^2	*p*
	M	*SD*	*n*	%	*M*	*SD*	*n*	%			
Age (years)	8.09	1.67			7.98	1.64			-0.81		.418
Sex										117.89	1.000
Female			322	53%			102	47%			
Male			289	47%			113	53%			
SES (maternal education)										126.90	1.000
No postsecondary education			371	61%			121	56%			
Postsecondary education			240	39%			94	44%			
Receptive language ability	10.69	2.81			9.32	3.26			-5.47		**< .001**
Expressive language ability	10.70	2.85			9.10	3.40			-6.17		**< .001**

Note. SES = socioeconomic status. Independent-samples *t* tests for age, receptive language ability, and expressive language ability. χ^2 test for sex and SES. Significant results are presented in bold.

Table S3

Pearson Correlations of the Predictors Included in the Hierarchical Regression Analyses

	Variable	1	2	3	4
1	Sex	—			
2	SES	.03	—		
3	RLA	.12***	.16***	—	
4	ELA	.11**	.17***	.45***	—
5	ML	.05	-.04	.20***	.23***

Note. Sex: -1 = male, 1 = female; SES = socioeconomic status: -1 = no postsecondary maternal education, 1 = postsecondary maternal education; ML = multilingualism: -1 = yes, 1 = no. RLA = receptive language ability; ELA = expressive language ability.

*$p_H < .05$. **$p_H < .01$. ***$p_H < .001$.

Table S4

Hierarchical Regression Analyses for Sex, SES, Receptive Language Ability, Expressive Language Ability, and Multilingualism as Predictors of Scores on the Profile IQ, Full-Scale IQ, and Screening IQ From the Intelligence and Development Scales–2

Step	Predictor	Profile IQ		Full-Scale IQ		Screening IQ	
		β	p_H	β	p_H	β	p_H
1	Sex	.04	.999	-.01	.999	.03	.999
	SES	.25	**< .001**	.26	**< .001**	.21	**< .001**
	R^2	.06		.07		.05	
2	Sex	-.03	.999	-.08	.955	-.03	.999
	SES	.16	**< .001**	.17	**< .001**	.13	**.005**
	RLA	.57	**< .001**	.56	**< .001**	.49	**< .001**
	R^2	.38		.37		.28	
	ΔR^2	**.32*** **		**.30*** **		**.23*** **	
3	Sex	-.04	.999	-.09	.147	-.04	.999
	SES	.13	**< .001**	.14	**< .001**	.11	.086
	RLA	.45	**< .001**	.44	**< .001**	.37	**< .001**
	ELA	.28	**< .001**	.28	**< .001**	.28	**< .001**
	R^2	.44		.43		.34	
	ΔR^2	**.06*** **		**.06*** **		**.06*** **	
4	Sex	-.04	.999	-.10	.143	-.04	.999
	SES	.13	**< .001**	.14	**< .001**	.11	.088
	RLA	.45	**< .001**	.44	**< .001**	.37	**< .001**
	ELA	.28	**< .001**	.27	**< .001**	.28	**< .001**
	ML	.02	.999	.02	.999	.00	.999
	R^2	.44		.43		.34	
	ΔR^2	.00		.00		.00	

Note. Sex: -1 = male, 1 = female; SES = socioeconomic status: -1 = no postsecondary maternal education, 1 = postsecondary maternal education; ML = multilingualism: -1 = yes, 1 = no; RLA = receptive language ability; ELA = expressive language ability; p_H = *p* value adjusted with the correction by Hommel (1988). Significant results are presented in bold.

*p_H < .05. **p_H < .01. ***p_H < .001.

Table S5

Hierarchical Regression Analyses for Sex, SES, Receptive Language Ability, Expressive Language Ability, and Multilingualism as Predictors of Scores on the Seven Group Factors of the Intelligence Domain From the Intelligence and Development Scales–2

Step	Predictor	Visual Processing		Processing Speed		Auditory Short-Term Memory		Visuospatial Short-Term Memory		Abstract Reasoning		Verbal Reasoning		Long-Term Memory	
		β	p_H	β	p_H	β	p_H	β	p_H	β	p_H	β	p_H	β	p_H
1	Sex	-.13	.058	.05	.999	.05	.999	.03	.999	.08	.999	.04	.999	.07	.999
	SES	.19	**< .001**	.12	.210	.21	**< .001**	.13	.071	.20	**< .001**	.23	**< .001**	.12	.125
	R^2	.05		.02		.05		.02		.05		.06		.02	
2	Sex	-.17	**< .001**	.01	.999	-.00	.999	-.01	.999	.03	.999	-.02	.999	.02	.999
	SES	.14	**.011**	.07	.999	.15	**.002**	.08	.999	.14	**.006**	.15	**< .001**	.06	.999
	RLA	.35	**< .001**	.32	**< .001**	.40	**< .001**	.33	**< .001**	.39	**< .001**	.55	**< .001**	.40	**< .001**
	R^2	.17		.12		.20		.12		.19		.34		.17	
	ΔR^2	**.12*** **		**.10*** **		**.15*** **		**.10*** **		**.14*** **		**.28*** **		**.15*** **	
3	Sex	-.18	**< .001**	.00	.999	-.01	.999	-.01	.999	.02	.999	-.04	.999	.01	.999
	SES	.12	.062	.06	.999	.13	**.022**	.07	.999	.12	**.048**	.12	**.011**	.05	.999
	RLA	.27	**< .001**	.27	**< .001**	.30	**< .001**	.27	**< .001**	.30	**< .001**	.41	**< .001**	.32	**< .001**
	ELA	.17	**.001**	.11	.665	.23	**< .001**	.13	.111	.20	**< .001**	.32	**< .001**	.17	**.002**
	R^2	.19		.13		.24		.13		.22		.42		.19	
	ΔR^2	**.02** **		.01		**.04*** **		.01		**.03*** **		**.08*** **		**.02** **	
4	Sex	-.18	**< .001**	.00	.999	-.01	.999	-.01	.999	.02	.999	-.04	.999	.01	.999
	SES	.12	.091	.05	.999	.13	**.023**	.06	.999	.12	.074	.13	**.002**	.06	.999
	RLA	.28	**< .001**	.28	**< .001**	.30	**< .001**	.27	**< .001**	.31	**< .001**	.40	**< .001**	.31	**< .001**
	ELA	.18	**< .001**	.12	.462	.23	**< .001**	.14	.105	.21	**< .001**	.30	**< .001**	.15	**.017**
	ML	-.03	.999	-.04	.999	.00	.999	-.01	.999	-.03	.999	.10	.059	.11	.379
	R^2	.19		.13		.24		.13		.23		.43		.20	
	ΔR^2	.00		.00		.00		.00		.01		.01		.01	

Note. Sex: -1 = male, 1 = female; SES = socioeconomic status: -1 = no postsecondary maternal education, 1 = postsecondary maternal education; ML = multilingualism: -1 = yes, 1 = no; RLA = receptive language ability; ELA = expressive language ability. p_H = p value adjusted with the correction by Hommel (1988). Significant results are presented in bold.

*$p_H < .05$. **$p_H < .01$. ***$p_H < .001$.

Table S6

Hierarchical Regression Analyses for Sex, SES, Receptive Language Ability, Expressive Language Ability, and Multilingualism as Predictors of Scores on the Seven Subtests of the Intelligence Domain (Part 1) From the Intelligence and Development Scales–2

Step	Predictor	Shape Design		Parrots		Digit and Letter Span		Shape Memory		Matrices: Completion		Naming Categories		Story Recall	
		β	p_H	β	p_H	β	p_H	β	p_H	β	p_H	β	p_H	β	p_H
1	Sex	-.18	**< .001**	.05	.999	.05	.999	.02	.999	.02	.999	.03	.999	-.02	.999
	SES	.21	**< .001**	.12	.196	.23	**< .001**	.12	.227	.17	**< .001**	.18	**< .001**	.15	**.007**
	R^2	.07		.02		.06		.01		.03		.03		.02	
2	Sex	-.23	**< .001**	.02	.999	.00	.999	-.01	.999	-.02	.999	-.03	.999	-.07	.999
	SES	.15	**.002**	.07	.999	.17	**< .001**	.07	.999	.11	.228	.11	.223	.08	.999
	RLA	.38	**< .001**	.30	**< .001**	.38	**< .001**	.27	**< .001**	.35	**< .001**	.47	**< .001**	.42	**< .001**
	R^2	.21		.10		.20		.08		.14		.24		.20	
	ΔR^2	**.14*****		**.08*****		**.14*****		**.07*****		**.11*****		**.21*****		**.18*****	
3	Sex	-.24	**< .001**	.01	.999	-.01	.999	-.01	.999	-.03	.999	-.04	.999	-.09	.961
	SES	.13	**.011**	.06	.999	.15	**.002**	.07	.999	.09	.957	.08	.993	.06	.999
	RLA	.32	**< .001**	.24	**< .001**	.29	**< .001**	.23	**< .001**	.26	**< .001**	.36	**< .001**	.32	**< .001**
	ELA	.14	**.024**	.13	.141	.20	**< .001**	.09	.999	.21	**< .001**	.26	**< .001**	.25	**< .001**
	R^2	.23		.11		.23		.09		.18		.30		.24	
	ΔR^2	**.02***		.01		**.03*****		.01		**.04*****		**.06*****		**.04*****	
4	Sex	-.24	**< .001**	.01	.999	-.01	.999	-.01	.999	-.03	.999	-.04	.999	-.09	.787
	SES	.13	**.021**	.06	.999	.15	**.001**	.07	.999	.09	.990	.08	.990	.07	.999
	RLA	.33	**< .001**	.24	**< .001**	.29	**< .001**	.23	**< .001**	.26	**< .001**	.35	**< .001**	.30	**< .001**
	ELA	.15	**.010**	.13	.137	.20	**< .001**	.09	.999	.21	**< .001**	.25	**< .001**	.23	**< .001**
	ML	-.05	.999	-.01	.999	.01	.999	-.01	.999	-.02	.999	.03	.999	.13	**.022**
	R^2	.23		.11		.23		.09		.18		.30		.26	
	ΔR^2	.00		.00		.00		.00		.00		.00		**.02***	

Note. Sex: -1 = male, 1 = female; SES = socioeconomic status: -1 = no postsecondary maternal education, 1 = postsecondary maternal education; ML = multilingualism: -1 = yes, 1 = no; RLA = receptive language ability; ELA = expressive language ability; p_H = *p* value adjusted with the correction by Hommel (1988). Significant results are presented in bold.

*p_H < .05. **p_H < .01. ***p_H < .001.

Table S7

Hierarchical Regression Analyses for Sex, SES, Receptive Language Ability, Expressive Language Ability, and Multilingualism as Predictors of Scores on the Seven Subtests of the Intelligence Domain (Part 2) From the Intelligence and Development Scales–2

Step	Predictor	Washer Design		Boxes		Mixed Digit and Letter Span		Rotated Shape Memory		Matrices: Odd One Out		Naming Opposites		Picture Recall	
		β	p_H	β	p_H	β	p_H	β	p_H	β	p_H	β	p_H	β	p_H
1	Sex	-.03	.999	.03	.999	.04	.999	.04	.999	.12	.219	.05	.999	.14	**.034**
	SES	.11	.555	.10	.987	.17	**.001**	.10	.915	.17	**< .001**	.23	**< .001**	.05	.999
	R^2	.01		.01		.03		.01		.04		.06		.02	
2	Sex	-.05	.999	-.00	.999	-.00	.999	.00	.999	.08	.999	-.01	.999	.11	.376
	SES	.08	.999	.06	.999	.11	.336	.05	.999	.12	.092	.15	**< .001**	.02	.999
	RLA	.20	**< .001**	.26	**< .001**	.37	**< .001**	.30	**< .001**	.30	**< .001**	.50	**< .001**	.21	**< .001**
	R^2	.05		.08		.16		.10		.13		.30		.06	
	ΔR^2	**.04*** **		**.07*** **		**.13*** **		**.09*** **		**.09*** **		**.24*** **		**.04*** **	
3	Sex	-.06	.999	-.01	.999	-.01	.999	-.00	.999	.07	.999	-.03	.999	.11	.420
	SES	.06	.999	.05	.999	.09	.993	.04	.999	.11	.296	.12	**.013**	.01	.999
	RLA	.13	.165	.24	**< .001**	.28	**< .001**	.23	**< .001**	.25	**< .001**	.37	**< .001**	.20	**< .001**
	ELA	.14	.102	.06	.999	.21	**< .001**	.16	**.006**	.12	.268	.31	**< .001**	.02	.999
	R^2	.06		.08		.20		.12		.14		.37		.07	
	ΔR^2	.01		.00		**.04*** **		**.02** **		.01		**.07*** **		.01	
4	Sex	-.06	.999	-.01	.999	-.01	.999	-.00	.999	.08	.999	-.03	.999	.11	.444
	SES	.06	.999	.05	.999	.09	.996	.03	.999	.11	.428	.13	**.001**	.02	.999
	RLA	.14	.174	.24	**< .001**	.28	**< .001**	.24	**< .001**	.25	**< .001**	.35	**< .001**	.20	**< .001**
	ELA	.14	.118	.07	.999	.21	**< .001**	.17	**.005**	.13	.172	.28	**< .001**	.01	.999
	ML	-.00	.999	-.05	.999	-.02	.999	-.03	.999	-.04	.999	.16	**< .001**	.03	.999
	R^2	.06		.08		.20		.12		.14		.39		.07	
	ΔR^2	.00		.00		.00		.00		.00		**.02*** **		.01	

Note. Sex: -1 = male, 1 = female; SES = socioeconomic status: -1 = no postsecondary maternal education, 1 = postsecondary maternal education; ML = multilingualism: -1 = yes, 1 = no; RLA = receptive language ability; ELA = expressive language ability; p_H = *p* value adjusted with the correction by Hommel (1988). Significant results are presented in bold.

*$p_H < .05$. **$p_H < .01$. ***$p_H < .001$.

Table S8

Hierarchical Regression Analyses for Sex, SES, Receptive Language Ability, Expressive Language Ability, and Multilingualism as Predictors of Scores on the Composite and Subtests of the Executive Functions Domain From the Intelligence and Development Scales–2

Step	Predictor	Executive functions composite		Listing Words		Divided Attention		Animal Colors		Drawing Routes	
		β	p_H	β	p_H	β	p_H	β	p_H	β	p_H
1	Sex	.09	.998	.14	**.029**	.05	.999	.09	.999	-.05	.999
	SES	.17	**< .001**	.16	**.003**	.13	.106	.06	.999	.13	.103
	R^2	.04		.04		.02		.01		.02	
2	Sex	.04	.999	.09	.993	.01	.999	.06	.999	-.08	.999
	SES	.11	.302	.10	.549	.07	.999	.02	.999	.10	.959
	RLA	.42	**< .001**	.38	**< .001**	.37	**< .001**	.23	**< .001**	.18	**< .001**
	R^2	.21		.18		.15		.06		.05	
	ΔR^2	**.17*** **		**.14*** **		**.13*** **		**.05*** **		**.03*** **	
3	Sex	.03	.999	.08	.999	.00	.999	.05	.999	-.08	.999
	SES	.09	.990	.09	.995	.06	.999	.02	.999	.09	.999
	RLA	.33	**< .001**	.31	**< .001**	.29	**< .001**	.19	**< .001**	.13	.347
	ELA	.20	**< .001**	.16	**.007**	.19	**< .001**	.10	.999	.13	.389
	R^2	.24		.20		.18		.07		.06	
	ΔR^2	**.03*** **		**.02** **		**.03*** **		.01		.01	
4	Sex	.03	.999	.08	.999	.00	.999	.06	.999	-.08	.999
	SES	.09	.993	.09	.993	.06	.999	.01	.999	.09	.999
	RLA	.33	**< .001**	.31	**< .001**	.28	**< .001**	.21	**< .001**	.13	.354
	ELA	.21	**< .001**	.15	**.015**	.18	**< .001**	.11	.946	.13	.408
	ML	-.02	.999	.02	.999	.05	.999	-.09	.998	-.01	.999
	R^2	.24		.20		.18		.08		.06	
	ΔR^2	.00		.00		.00		.01		.00	

Note. Sex: -1 = male, 1 = female; SES = socioeconomic status: -1 = no postsecondary maternal education, 1 = postsecondary maternal education; ML = multilingualism: -1 = yes, 1 = no; RLA = receptive language ability; ELA = expressive language ability; p_H = *p* value adjusted with the correction by Hommel (1988). Significant results are presented in bold.

*p_H < .05. **p_H < .01. ***p_H < .001.

Table S9

Hierarchical Regression Analyses for Sex, SES, Receptive Language Ability, Expressive Language Ability, and Multilingualism as Predictors of Scores on the Composite and Subtests of the Psychomotor Skills Domain From the Intelligence and Development Scales–2

Step	Predictor	Psychomotor skills composite		Gross Motor Skills		Fine Motor Skills		Visuomotor Skills	
		β	p_H	β	p_H	β	p_H	β	p_H
1	Sex	.11	.268	.00	.999	.18	**< .001**	.09	.993
	SES	.08	.999	.05	.999	.02	.999	.12	.277
	R^2	.02		.00		.03		.02	
2	Sex	.08	.999	-.02	.999	.15	**.002**	.06	.999
	SES	.02	.999	.01	.999	-.02	.999	.07	.999
	RLA	.33	**< .001**	.23	**< .001**	.23	**< .001**	.28	**< .001**
	R^2	.13		.05		.09		.10	
	ΔR^2	**.11*** **		**.05*** **		**.06*** **		**.08*** **	
3	Sex	.07	.999	-.03	.999	.15	**.005**	.06	.999
	SES	.01	.999	-.00	.999	-.03	.999	.06	.999
	RLA	.27	**< .001**	.18	**.001**	.19	**< .001**	.22	**< .001**
	ELA	.15	**.010**	.11	.982	.09	.999	.13	.110
	R^2	.15		.06		.09		.11	
	ΔR^2	**.02** **		.01		.00		.01	
4	Sex	.07	.999	-.03	.999	.15	**.005**	.06	.999
	SES	.01	.999	-.00	.999	-.03	.999	.05	.999
	RLA	.27	**< .001**	.19	**< .001**	.19	**< .001**	.23	**< .001**
	ELA	.16	**.005**	.11	.801	.10	.999	.15	**.039**
	ML	-.05	.999	-.03	.999	-.01	.999	-.07	.999
	R^2	.15		.06		.09		.12	
	ΔR^2	.00		.00		.00		.01	

Note. Sex: -1 = male, 1 = female; SES = socioeconomic status: -1 = no postsecondary maternal education, 1 = postsecondary maternal education; ML = multilingualism: -1 = yes, 1 = no; RLA = receptive language ability; ELA = expressive language ability; p_H = *p* value adjusted with the correction by Hommel (1988). Significant results are presented in bold.

*p_H < .05. **p_H < .01. ***p_H < .001.

Table S10

Hierarchical Regression Analyses for Sex, SES, Receptive Language Ability, Expressive Language Ability, and Multilingualism as Predictors of Scores on the Composite and Subtests of the Social-Emotional Skills Domain From the Intelligence and Development Scales–2

Step	Predictor	Social-emotional skills composite		Identifying Emotions		Regulating Emotions		Socially Competent Behavior	
		β	p_H	β	p_H	β	p_H	β	p_H
1	Sex	.07	.999	.05	.999	.04	.999	.05	.999
	SES	.07	.999	.09	.999	.03	.999	.03	.999
	R^2	.01		.01		.00		.00	
2	Sex	.04	.999	.03	.999	.02	.999	.04	.999
	SES	.03	.999	.06	.999	-.00	.999	.02	.999
	RLA	.21	**< .001**	.19	**< .001**	.17	**< .001**	.10	.742
	R^2	.05		.05		.03		.01	
	ΔR^2	**.04*****		**.04*****		**.03*****		.01	
3	Sex	.03	.999	.02	.999	.01	.999	.03	.999
	SES	.02	.999	.05	.999	-.02	.999	.01	.999
	RLA	.14	.105	.14	.081	.10	.996	.06	.999
	ELA	.18	**.002**	.12	.471	.17	**.006**	.10	.999
	R^2	.08		.06		.05		.02	
	ΔR^2	**.03****		.01		**.02****		.01	
4	Sex	.03	.999	.02	.999	.01	.999	.03	.999
	SES	.01	.999	.05	.999	-.02	.999	.00	.999
	RLA	.14	.086	.14	.085	.10	.997	.07	.999
	ELA	.18	**.001**	.12	.502	.17	**.007**	.11	.993
	ML	-.02	.999	-.01	.999	-.00	.999	-.05	.999
	R^2	.08		.06		.05		.02	
	ΔR^2	.00		.00		.00		.00	

Note. Sex: -1 = male, 1 = female; SES = socioeconomic status: -1 = no postsecondary maternal education, 1 = postsecondary maternal education; ML = multilingualism: -1 = yes, 1 = no; RLA = receptive language ability; ELA = expressive language ability; p_H = *p* value adjusted with the correction by Hommel (1988). Significant results are presented in bold.

*$p_H < .05$. **$p_H < .01$. ***$p_H < .001$.

Table S11

Hierarchical Regression Analyses for Sex, SES, Receptive Language Ability, Expressive Language Ability, and Multilingualism as Predictors of Scores on the Subtests of the Basic Skills Domain From the Intelligence and Development Scales–2

Step	Predictor	Logical-Mathematical Reasoning		Reading		Spelling	
		β	p_H	β	p_H	β	p_H
1	Sex	-.07	.999	.05	.999	.21	**< .001**
	SES	.22	**< .001**	.15	.091	.06	.999
	R^2	.05		.03		.05	
2	Sex	-.12	**.021**	-.01	.999	.16	**.040**
	SES	.14	**.001**	.09	.999	.01	.999
	RLA	.46	**< .001**	.39	**< .001**	.36	**< .001**
	R^2	.26		.17		.17	
	ΔR^2	**.21***		**.14***		**.12***	
3	Sex	-.14	**< .001**	-.02	.999	.14	.094
	SES	.12	**.028**	.06	.999	-.02	.999
	RLA	.34	**< .001**	.26	**< .001**	.23	**< .001**
	ELA	.29	**< .001**	.29	**< .001**	.31	**< .001**
	R^2	.32		.24		.25	
	ΔR^2	**.06***		**.07***		**.08***	
4	Sex	-.14	**< .001**	-.02	.999	.14	.095
	SES	.11	.052	.05	.999	-.02	.999
	RLA	.35	**< .001**	.27	**< .001**	.23	**< .001**
	ELA	.30	**< .001**	.30	**< .001**	.30	**< .001**
	ML	-.05	.999	-.03	.999	.02	.999
	R^2	.33		.24		.25	
	ΔR^2	.01		.00		.00	

Note. Sex: -1 = male, 1 = female; SES = socioeconomic status: -1 = no postsecondary maternal education, 1 = postsecondary maternal education; ML = multilingualism: -1 = yes, 1 = no; RLA = receptive language ability; ELA = expressive language ability; p_H = *p* value adjusted with the correction by Hommel (1988). The basic skills composite was not investigated since the receptive and expressive language abilities are also included in the calculation of this score. Significant results are presented in bold.

*$p_H < .05$. **$p_H < .01$. ***$p_H < .001$.

Table S12

Hierarchical Regression Analyses for Sex, SES, Intelligence, Receptive Language Ability, and Expressive Language Ability as Predictors of Scores on the Composite and Subtests of the Executive Functions Domain From the Intelligence and Development Scales–2

Step	Predictor	Executive functions composite		Listing Words		Divided Attention		Animal Colors		Drawing Routes	
		β	p_H	β	p_H	β	p_H	β	p_H	β	p_H
1	Sex	.09	.938	.14	**.017**	.05	.974	.09	.956	-.06	.974
	SES	.16	**.001**	.15	**.004**	.12	.117	.05	.974	.12	.133
	R^2	.04		.04		.02		.01		.02	
2	Sex	.07	.972	.12	**.017**	.04	.974	.08	.974	-.07	.974
	SES	.01	.974	.03	.974	-.02	.974	-.04	.974	.04	.974
	IQ	.64	**< .001**	.52	**< .001**	.57	**< .001**	.36	**< .001**	.33	**< .001**
	R^2	.42		.29		.32		.14		.12	
	ΔR^2	**.38*** **		**.25*** **		**.30*** **		**.13*** **		**.10*** **	
3	Sex	.06	.974	.11	.093	.03	.974	.07	.974	-.07	.974
	SES	.01	.974	.03	.974	-.02	.974	-.04	.974	.04	.974
	IQ	.58	**< .001**	.44	**< .001**	.53	**< .001**	.34	**< .001**	.33	**< .001**
	RLA	.10	.559	.12	.159	.08	.974	.05	.974	.00	.974
	R^2	.42		.30		.33		.14		.12	
	ΔR^2	.00		.01		.01		.00		.00	
4	Sex	.05	.974	.10	.120	.02	.974	.07	.974	-.07	.974
	SES	.00	.974	.02	.974	-.02	.974	-.04	.974	.04	.974
	IQ	.56	**< .001**	.43	**< .001**	.51	**< .001**	.33	**< .001**	.32	**< .001**
	RLA	.09	.958	.12	.349	.07	.974	.05	.974	-.01	.974
	ELA	.05	.974	.03	.974	.05	.974	.01	.974	.03	.974
	R^2	.42		.31		.33		.14		.12	
	ΔR^2	.00		.01		.00		.00		.00	

Note. Sex: -1 = male, 1 = female; SES = socioeconomic status: -1 = no postsecondary maternal education, 1 = postsecondary maternal education; IQ = intelligence (Profile IQ); RLA = receptive language ability; ELA = expressive language ability; p_H = *p* value adjusted with the correction by Hommel (1988). As multilingualism did not explain additional variance in scores in the executive functions domain in the previous hierarchical regression analyses (see Table S8), this variable was omitted from the present analyses. Significant results are presented in bold.

*p_H < .05. **p_H < .01. ***p_H < .001.

Table S13

Hierarchical Regression Analyses for Sex, SES, Intelligence, Receptive Language Ability, and Expressive Language Ability as Predictors of Scores on the Composite and Subtests of the Psychomotor Skills Domain From the Intelligence and Development Scales–2

Step	Predictor	Psychomotor skills composite		Gross Motor Skills		Fine Motor Skills		Visuomotor Skills	
		β	p_H	β	p_H	β	p_H	β	p_H
1	Sex	.12	.095	.01	.974	.18	**< .001**	.10	.814
	SES	.07	.974	.04	.974	.01	.974	.11	.412
	R^2	.02		.00		.03		.02	
2	Sex	.10	.138	.00	.974	.17	**< .001**	.08	.850
	SES	-.06	.974	-.04	.974	-.08	.974	-.01	.974
	IQ	.52	**< .001**	.33	**< .001**	.36	**< .001**	.48	**< .001**
	R^2	.27		.10		.15		.24	
	ΔR^2	**.25*** **		**.10*** **		**.12*** **		**.22*** **	
3	Sex	.10	.285	-.00	.974	.17	**< .001**	.08	.946
	SES	-.06	.974	-.04	.974	-.08	.974	-.01	.974
	IQ	.48	**< .001**	.29	**< .001**	.33	**< .001**	.46	**< .001**
	RLA	.07	.974	.07	.974	.05	.974	.03	.974
	R^2	.27		.11		.16		.24	
	ΔR^2	.00		.01		.01		.00	
4	Sex	.09	.329	-.01	.974	.17	**< .001**	.08	.956
	SES	-.06	.974	-.04	.974	-.08	.974	-.01	.974
	IQ	.47	**< .001**	.28	**< .001**	.33	**< .001**	.46	**< .001**
	RLA	.06	.974	.06	.974	.05	.974	.02	.974
	ELA	.02	.974	.03	.974	.01	.974	.01	.974
	R^2	.27		.11		.16		.24	
	ΔR^2	.00		.00		.00		.00	

Note. Sex: -1 = male, 1 = female; SES = socioeconomic status: -1 = no postsecondary maternal education, 1 = postsecondary maternal education; IQ = intelligence (Profile IQ); RLA = receptive language ability; ELA = expressive language ability; p_H = *p* value adjusted with the correction by Hommel (1988). As multilingualism did not explain additional variance in scores of the psychomotor skills domain in the previous hierarchical regression analyses (see Table S9), this variable was omitted from the present analyses. Significant results are presented in bold.

*p_H < .05. **p_H < .01. ***p_H < .001.

Table S14

Hierarchical Regression Analyses for Sex, SES, Intelligence, Receptive Language Ability, and Expressive Language Ability as Predictors of Scores on the Composite and Subtests of the Social-Emotional Skills Domain From the Intelligence and Development Scales–2

Step	Predictor	Social-emotional skills composite		Identifying Emotions		Regulating Emotions		Socially Competent Behavior	
		β	p_H	β	p_H	β	p_H	β	p_H
1	Sex	.06	.974	.05	.974	.04	.974	.04	.974
	SES	.07	.974	.09	.963	.02	.974	.03	.974
	R^2	.01		.01		.00		.00	
2	Sex	.05	.974	.04	.974	.03	.974	.04	.974
	SES	-.00	.974	.04	.974	-.04	.974	-.00	.974
	IQ	.28	**< .001**	.19	**< .001**	.26	**< .001**	.15	**.005**
	R^2	.08		.04		.07		.03	
	ΔR^2	**.07*** **		**.03*** **		**.07*** **		**.03** **	
3	Sex	.04	.974	.03	.974	.02	.974	.03	.974
	SES	-.01	.974	.04	.974	-.04	.974	-.01	.974
	IQ	.22	**< .001**	.10	.974	.24	**< .001**	.13	.426
	RLA	.10	.974	.15	.132	.05	.974	.03	.974
	R^2	.09		.06		.07		.03	
	ΔR^2	.01		.02		.00		.00	
4	Sex	.03	.974	.02	.974	.01	.974	.03	.974
	SES	-.01	.974	.04	.974	-.05	.974	-.01	.974
	IQ	.17	**.023**	.06	.974	.20	**.004**	.11	.974
	RLA	.07	.974	.12	.715	.02	.974	.02	.974
	ELA	.13	.167	.10	.930	.12	.553	.07	.974
	R^2	.10		.06		.08		.03	
	ΔR^2	.01		.00		.01		.00	

Note. Sex: -1 = male, 1 = female; SES = socioeconomic status: -1 = no postsecondary maternal education, 1 = postsecondary maternal education; IQ = intelligence (Profile IQ); RLA = receptive language ability; ELA = expressive language ability; p_H = *p* value adjusted with the correction by Hommel (1988). As multilingualism did not explain additional variance in scores of the social-emotional skills domain in the previous hierarchical regression analyses (see Table S10), this variable was omitted from the present analyses. Significant results are presented in bold.

*$p_H < .05$. **$p_H < .01$. ***$p_H < .001$.

Table S15

Hierarchical Regression Analyses for Sex, SES, Intelligence, Receptive Language Ability, and Expressive Language Ability as Predictors of Scores on the Subtests of the Basic Skills Domain From the Intelligence and Development Scales–2

Step	Predictor	Logical-Mathematical Reasoning		Reading		Spelling	
		β	p_H	β	p_H	β	p_H
1	Sex	-.07	.974	.05	.974	.21	**< .001**
	SES	.22	**< .001**	.15	.078	.06	.974
	R^2	.05		.03		.05	
2	Sex	-.09	.124	.02	.974	.18	**< .001**
	SES	.06	.974	.02	.974	-.07	.974
	IQ	.62	**< .001**	.53	**< .001**	.54	**< .001**
	R^2	.41		.28		.32	
	ΔR^2	**.36*** **		**.25*** **		**.27*** **	
3	Sex	-.11	**.013**	.01	.974	.17	**< .001**
	SES	.06	.974	.02	.974	-.07	.974
	IQ	.53	**< .001**	.44	**< .001**	.49	**< .001**
	RLA	.16	**< .001**	.14	.308	.08	.974
	R^2	.43		.30		.32	
	ΔR^2	**.02*** **		.02		.00	
4	Sex	-.12	**.002**	-.00	.974	.16	**.002**
	SES	.05	.974	.00	.974	-.08	.974
	IQ	.47	**< .001**	.37	**< .001**	.42	**< .001**
	RLA	.13	**.034**	.10	.974	.04	.974
	ELA	.16	**< .001**	.19	**.002**	.20	**.001**
	R^2	.45		.32		.35	
	ΔR^2	**.02*** **		**.02** **		**.03** **	

Note. Sex: -1 = male, 1 = female; SES = socioeconomic status: -1 = no postsecondary maternal education, 1 = postsecondary maternal education; IQ = intelligence (Profile IQ); RLA = receptive language ability; ELA = expressive language ability; p_H = p value adjusted with the correction by Hommel (1988). The basic skills composite was not investigated since the receptive and expressive language abilities are also included in the calculation of this score. As multilingualism did not explain additional variance in scores of the basic skills domain in the previous hierarchical regression analyses (see Table S11), this variable was omitted from the present analyses. Significant results are presented in bold.

*p_H < .05. **p_H < .01. ***p_H < .001.

Supplemental References

Hommel, G. (1988). A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika*, *75*, 383–386.