Vocational Interest Development in Adolescence—Integrating Insights about Normative Change, Stability, and Influencing Factors

Dissertation

zur Erlangung des Doktorgrades

der Wirtschafts- und Sozialwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

vorgelegt von

M.Sc. Thomas Gfrörer

aus Hechingen

Tübingen

2021

1. Betreuer	Prof. Dr. Benjamin Nagengast
2. Betreuer	Prof. Dr. Ulrich Trautwein

Tag der mündlichen Prüfung:	08. 12. 2021

Dekan:	Prof. Dr. Josef Schmid
1. Gutachter:	Prof. Dr. Benjamin Nagengast
2. Gutachter	Prof. Dr. Richard Göllner

DANKSAGUNG

Zuallererst möchte ich mich ganz herzlich bei meinem Betreuungsteam bedanken: Dr. Gundula Stoll, Prof. Dr. Benjamin Nagengast, Prof. Dr. Ulrich Trautwein und Dr. Sven Rieger. Ich danke euch sehr für die zahlreichen Ideen, die großartige Unterstützung und die ermutigenden Worte während meiner Promotion.

Darüber hinaus möchte ich mich besonders bei Anna und Laura bedanken, die mich in unseren regelmäßigen Treffen aufgemuntert haben und mir bei Fragen jeglicher Art immer zur Seite standen.

Ein besonderer Dank geht an meine Familie und Freunde, die mich während den schwierigen Phasen meiner Promotion bedingungslos unterstützt haben. Ich habe unheimliches Glück, dass es euch gibt.

Der größte Dank geht an meine Freundin Janice, ohne die ich das alles nicht geschafft hätte. Ohne deine Unterstützung wäre diese Arbeit nicht möglich gewesen. Ich danke dir für alles was du für mich getan hast und freue mich auf unseren neuen gemeinsamen Lebensabschnitt.

ABSTRACT

Vocational interests are significant predictors for various life outcomes, educational decisions, and occupational choices. They are frequently assessed in practice through the application of interest inventories and used by vocational counselors to guide career-related decisions of students and jobseekers. In research, vocational interests are seen as relatively stable dispositions that develop over multiple years. Due to their stability and their impact on people's everyday life's, vocational interests are often included in models of individual differences. It is assumed that they describe patterns of persons general motives that are part of their personality. Theories about vocational interests suggest that they begin to develop over the course of adolescence—stability is assumed to increase and changes in interest intensity are expected. However, more empirical evidence is needed as current studies mainly focus on the description of vocational interest development in later life phases, such as the transition from late adolescence to young adulthood. Empirical studies that capture early life phases of development, such as the time period of late childhood and early adolescence (ages 11 to 14), are scarce. In addition, relatively little is known about possible factors that might influence the development of vocational interests. It is suggested that differences in personality characteristics and external factors could lead to differences in developmental trajectories.

The aim of the current dissertation was to do a comprehensive investigation of the development of vocational interests over the course of adolescence (ages 11 to 18). The development of interest stability, intensity, and gender differences was investigated. It was assumed that vocational interests increase in their stability over the course of adolescence (ages 11 to 18). Interest intensity was assumed to decrease from late childhood to early adolescence (ages 11 to 14) and to increase from middle to late adolescence (ages 15 to 18). Gender differences in vocational interests were assumed to increase from late childhood to early adolescence (ages 11 to 14) and being relatively stable afterwards. Besides these overall aims, the three studies included in the current dissertation focused on individual and contextual factors that could influence the development of vocational interests. It was assumed that personality characteristics are associated to vocational interest profile stability and that the engagement in leisure-related activities could influence the development of vocational interests. In all studies, vocational interests were measured based on Hollands (1997) RIASEC model.

The first study investigated the development of vocational interests over the course of adolescence. Besides examining mean-level change, gender differences in mean levels, and re-test correlations of vocational interests, the study focused on dispositional and situational

components of vocational interests. The investigation was inspired by recent theories, which assumed that vocational interests are dispositions that also consist of situation susceptible components. Data was used from a large-scale longitudinal sample (N = 3,876), where students from low and middle track schools in Germany were annually followed from fifth to eight grade (mean ages 11 to 14). The results suggest that vocational interests became more stable over the three-year period, as indicated by increases in re-test correlations for four of six interest dimensions. In addition, mean levels of vocational interests decreased descriptively from late childhood to early adolescence for all interest dimensions, except Social interests. Gender differences in mean levels were already large in terms of effect sizes around age 11 and further increased over time for all interest dimensions, except Artistic interests. Results from latent state-trait analysis suggest that vocational interests consist of both, stable and situation susceptible components. However, the proportion of the stable components increased over time.

The second study investigated the profile stability of vocational interests and its relation to personality traits, cognitive abilities, and gender. It was proposed that differences in individual characteristics could lead to differences in profile stability. The study investigated the research question in four different life phases: late childhood to early adolescence (ages 11 to 14), middle adolescence (ages 14 to 15), late adolescence to young adulthood (ages 17 to 23), and a longer time period over the course of young adulthood (ages 22 to 34). Data was used from four, previously conducted, large-scale longitudinal studies. Each life phase was consequently covered by a different sample from a different study. All the samples included students that lived in Germany. The results suggest that vocational interest profiles were moderately stable during the three-year time period from late childhood to early adolescence (ages 11 to 14) and highly stable during the one-year time period during middle adolescence (ages 14 to 15), the six-year time period from late adolescence to young adulthood (ages 17 to 23), and the twelve-year time period over the course of young adulthood (ages 22 to 34). Indicators of profile stability significantly varied between persons in each of the four life phases, suggesting that profile stability differed across participants. Gender was related to differences in profile stability in all life phases, with girls and women having significantly more stable profiles compared to boys and men. Associations of profile stability to personality traits and cognitive abilities were rather small. Consistent relationships were found for verbal cognitive abilities and the personality trait Extraversion, with higher scores being related to more stable vocational interest profiles.

The third study investigated the impact of engaging in unstructured out-of-school time science activities, such as reading a science book, watching a science TV show, or researching

on the internet about science, on the development of various constructs, including vocational interests. In line with theories about interest development, it was assumed that the engagement in unstructured out-of-school time science activities could foster the evolvement of vocational interests, as they possess advantageous properties for the initiation of situational interest. Data was used from a large-scale longitudinal sample (N = 2,655), where students from different school tracks in Germany were followed over three time points from ninth, eleventh to twelfth grade (mean ages 15, 17, and 18). Based on an outcome-wide longitudinal design for causal inference, the impact of unstructured out-of-school time science activities on the development of vocational interests was investigated. To account for self-selection effects, numerous confounder variables, such as pretests of vocational interests, other motivational variables, and ability-related constructs, were included in the analysis. The results suggest that the engagement in unstructured out-of-school time science activities had a robust influence on Investigative vocational interests, but not on the remaining interest dimensions.

The findings on interest stability, intensity, and gender differences over the course of adolescence (ages 11 to 18) were integrated from the three empirical studies. The findings suggest that stability of vocational interests increased over the course of adolescence. Interest intensity decreased from late childhood to early adolescence and increased from middle to late adolescence, as shown by changes in mean levels. Gender differences increased over the course of late childhood and early adolescence, as indicated by increasing mean levels between girls and boys. Deviations from the proposed general trends in interest stability, intensity, and gender differentiation are described in the general discussion. Findings of the three empirical studies are summarized and discussed regarding their implications for vocational interest development. It is proposed that experiencing activities can initiate the development of vocational interests over the course of adolescence. Practical implications, limitations, and an outlook for future research are provided at the end of the general discussion section.

ZUSAMMENFASSUNG

Berufliche Interessen sind wichtige Prädiktoren für unterschiedliche Lebensbereiche sowie für Bildungs- und Berufsentscheidungen. Berufliche Interessen werden häufig in der Praxis von Berufsberatern und Berufsberaterinnen anhand von Interessensinventaren erfasst. Die Informationen werden genutzt, um Karriereentscheidungen von Schülerinnen und Schülern sowie Arbeitssuchenden zu unterstützen. Die Forschung sieht berufliche Interessen als relativ stabile Dispositionen an, die sich über mehrere Jahre hinweg entwickeln. Aufgrund ihrer Stabilität und ihrem Einfluss auf das alltägliche Leben werden berufliche Interessen häufig in Modelle integriert, welche konsistente Unterschiede zwischen Personen charakterisieren. Es wird angenommen das berufliche Interessen generelle Motive von Menschen beschreiben, welche als Teil ihrer Persönlichkeit verstanden werden können. Theorien zu beruflichen Interessen postulieren den Beginn ihrer Entwicklung in der Adoleszenz. Es wird angenommen das die Stabilität beruflicher Interessen steigt und sich ihre Intensität verändert. Da sich ein Großteil der bisherigen Studien auf spätere Lebensphasen, wie den Übergang zwischen der späten Jugendphase und dem jungen Erwachsenenalter, konzentriert, ist mehr empirische Forschung nötig. Bisher gibt es wenig Studien die frühere Lebensphasen, wie beispielsweise die frühe Jugendphase (Alter 11 bis 14), untersuchten. Ebenso ist wenig über Faktoren bekannt, welche die Entwicklung beruflicher Interessen beeinflussen können. Es wird angenommen, dass Unterschiede in Persönlichkeitsmerkmalen und externalen Faktoren zu einer unterschiedlichen Entwicklung von beruflichen Interessen führen können.

Das Ziel der Dissertation ist eine umfassende Untersuchung der Entwicklung beruflicher Interessen über die Jugendphase hinweg (Alter 11 bis 18). Die Untersuchung fokussiert sich dabei auf den Verlauf der Stabilität, der Intensität und den Geschlechterunterschieden beruflicher Interessen. Es wird angenommen das die Stabilität beruflicher Interessen über den Verlauf der Jungendphase (Alter 11 bis 18) hinweg ansteigt. Die Intensität beruflicher Interessen soll in der frühen Jugendphase (Alter 11 bis 14) abnehmen und in der mittleren bis späten Jugendphase (Alter 15 bis 18) wieder ansteigen. Es wird vermutet das Geschlechterunterschiede sich über die frühe Jugendphase (Alter 11 bis 14) hinweg vergrößern, danach aber recht stabil bleiben. Neben diesen generellen Zielen untersuchen die drei empirischen Studien der Dissertation individuelle und kontextuelle Faktoren, welche die Entwicklung beruflicher Interessen beeinflussen können. Es wird angenommen das Persönlichkeitsmerkmale mit Unterschieden in der Profilstabilität beruflicher Interessen zusammenhängen und dass das Engagement in Freizeitaktivitäten einen Einfluss auf

die Entwicklung beruflicher Interessen haben kann. In allen Studien wurden berufliche Interessen basierend auf Hollands (1997) RIASEC Modell erfasst.

Die erste Studie untersuchte die Entwicklung beruflicher Interessen in der frühen Jugendphase. Neben der Beschreibung von Mittelwertunterschieden, Geschlechterunterschieden bei Mittelwerten und re-test Korrelationen, fokussierte sich die Studie auf dispositionale und situationale Komponenten beruflicher Interessen. Die Untersuchung wurde von Theorien angeregt, welche annehmen das berufliche Interessen Dispositionen sind, die auch aus situational beeinflussbaren Komponenten bestehen. Für die Analysen wurden Daten einer großen Längsschnittstudie (N = 3876) verwendet, bei der Schülerinnen und Schüler aus Haupt-, Real- und Mittelschulen über die Klassenstufen fünf bis acht hinweg (mittleres Alter 11 bis 14), jährlich befragt wurden. Die Ergebnisse deuten darauf hin, dass berufliche Interessen über die Zeitspanne von drei Jahren hinweg, stabiler werden. Vier der sechs Interessensdimensionen erlebten einen Anstieg in re-test Korrelationen. Laut den deskriptiven Befunden sanken die Mittelwerte beruflicher Interessen über die frühe Jugendphase hinweg in allen Interessensdimensionen außer Social. Die Effektstärken der Geschlechterunterschiede deuten bereits im Alter von 11 Jahren auf große Unterschiede zwischen Mädchen und Jungen hin. Über die Zeit hinweg vergrößern sich diese Unterschiede für alle Interessensdimensionen außer Artistic. Die Ergebnisse der latenten state-trait Analysen deuten darauf hin, dass berufliche Interessen sowohl aus stabilen als auch aus situational beeinflussbaren Komponenten bestehen. Es ist allerdings anzumerken das der Anteil der stabilen Komponenten über die Zeit hinweg zunimmt.

Die zweite Studie untersuchte die Profilstabilität beruflicher Interessen und ihren Zusammenhang mit Persönlichkeitseigenschaften, kognitiven Fähigkeiten und dem Geschlecht. Es wurde angenommen das Unterschiede in solchen individuellen Charakteristika zu Unterschieden in der Profilstabilität beruflicher Interessen führen können. Die Studie untersuchte die Forschungsfrage in vier unterschiedlichen Lebensphasen, der frühen Jugendphase (Alter 11 bis 14), der mittleren Jugendphase (Alter 14 bis 15), von der späten Jugendphase bis ins frühe Erwachsenenalter (Alter 17 bis 23) und über einen längeren Zeitraum im frühen Erwachsenenalter hinweg (Alter 22 bis 34). Die Daten der Studie stammen aus vier großen, bereits durchgeführten, Längsschnittstudien. Jede Lebensphase wurde somit anhand einer anderen Stichprobe von einer anderen Studie untersucht. Die Stichproben enthielten allesamt Schülerinnen und Schüler aus Deutschland. Die Ergebnisse deuten darauf hin, dass während der frühen Jugendphase (Alter 11 bis 14) die beruflichen Interessensprofile moderat stabil sind. Während der mittleren Jugendphase (Alter 14 bis 15), der späten Jugendphase und

dem frühen Erwachsenenalter (Alter 17 bis 23) sowie den zwölf Jahren über das frühe Erwachsenenalter hinweg (Alter 22 bis 34), sind berufliche Interessensprofile deutlich stabiler. Die Profilstabilitätsindikatoren variierten signifikant zwischen den Teilnehmenden in allen untersuchten Lebensphasen, was darauf schließen lässt, dass sich die Stabilität der Interessensprofile zwischen Personen unterscheidet. Das Geschlecht war in allen untersuchten Lebensphasen mit Unterschieden in der Profilstabilität verbunden. Mädchen und Frauen besaßen signifikant stabilere Interessensprofile im Vergleich zu Jungen und Männern. Zusammenhänge zwischen der Profilstabilität sowie Persönlichkeitseigenschaften und kognitiven Fähigkeiten waren eher klein. Konsistente Zusammenhänge wurden bei verbalen kognitiven Fähigkeiten und der Persönlichkeitseigenschaft Extraversion gefunden. Höhere Werte auf beiden Variablen hingen mit stabileren Interessensprofilen zusammen.

Die dritte Studie untersuchte wie sich das Engagement in unstrukturierten Aktivitäten außerhalb der Schule im Bereich Naturwissenschaft, wie beispielsweise das Lesen von Büchern mit naturwissenschaftlichen Inhalten, das Schauen einer Fernsehserie mit naturwissenschaftlichen Inhalten oder das Recherchieren von naturwissenschaftlichen Inhalten im Internet, auf die Entwicklung unterschiedlicher Konstrukte, einschließlich beruflicher Interessen, auswirkt. In Anlehnung an Theorien zur Interessensentwicklung wurde angenommen, dass das Engagement in solchen Aktivitäten berufliche Interessen beeinflussen kann. Die Daten zur dritten Studie stammen aus einer großen Längsschnittstudie (N = 2655) in der Schülerinnen und Schüler aller Schulformen aus Deutschland über drei Messzeitpunkte hinweg (Klasse neun, elf und zwölf; Alter 15, 17 und 18) befragt wurden. Basierend auf einem „outcome-wide longitudinal design“ für kausale Inferenz, wurde der Einfluss von unstrukturierten Aktivitäten außerhalb der Schule im Bereich Naturwissenschaft auf die Entwicklung beruflicher Interessen untersucht. Um für Selektionseffekte zu kontrollieren wurde eine Vielzahl an möglichen Konfundierungsvariablen, wie beispielsweise Pretests beruflicher Interessen, motivationale Variablen sowie fähigkeitsbezogene Konstrukte, berücksichtigt. Die Ergebnisse deuten darauf hin, dass das Engagement in unstrukturierten Aktivitäten außerhalb der Schule im Bereich Naturwissenschaft einen robusten Einfluss auf Investigative berufliche Interessen hat, allerdings nicht auf andere Interessensdimensionen.

Die Befunde der drei Studien zur Interessensstabilität, Intensität und den Geschlechterunterschieden wurden in der Diskussion integriert. Die Ergebnisse deuten darauf hin, dass die Stabilität von beruflichen Interessen über die Jungendphase (Alter 11 bis 18) hinweg ansteigt. Es wird angenommen das die Intensität von beruflichen Interessen über die frühe Jugendphase hinweg abnimmt und dass sie von der mittleren bis zur späten Jugendphase

wieder ansteigt. Daraufhin deuten die zeitlichen Mittelwertveränderungen aus den jeweiligen Studien. Basierend auf den Ergebnissen zu den ansteigenden Mittelwertunterschieden zwischen Mädchen und Jungen wird angenommen das Geschlechterunterschiede über die frühe Jugendphase hinweg zunehmen. Im Abschnitt zur generellen Diskussion, werden Abweichungen der Verläufe in den Bereichen Stabilität, Intensität und Geschlechterunterschiede diskutiert. Die Befunde der drei empirischen Studien werden zusammengefasst und ihre Implikationen für die Entwicklung beruflicher Interessen werden dargelegt. Es wird angenommen, dass das Erleben von Aktivitäten die Entwicklung beruflicher Interessen während der Jugendphase initiieren kann. Praktische Implikationen, Limitationen und einen Ausblick für zukünftige Forschungsthemen werden ebenfalls in der Diskussion beschrieben.

CONTENTS

1 INTRODUCTION AND THEORETICAL BACKGROUND

"What are you interested in?" is a common question in interpersonal encounters such as job interviews, study major applications and social gatherings. By elaborating your interests, you reveal more about yourself than your mere likes and dislikes towards certain activities. Findings from vocational interest research illustrate how influential the answer to that question may be (Stoll & Trautwein, 2017). Besides their influence on work-related outcomes like gross income (Stoll et al., 2017), job performance (Nye et al., 2012; Nye et al., 2017; Van Iddekinge et al., 2011), job persistence (Van Iddekinge et al., 2011), and job satisfaction (Hoff et al., 2020; Tsabari et al., 2005), vocational interests also predict work-unrelated outcomes like school performance (Rounds & Su, 2014; Su, 2012), college major choice (Päßler & Hell, 2012; Wille et al., 2020), and the persistence of this choice (Allen & Robbins, 2008), major life goals (Stoll, Einarsdóttir, et al., 2020), as well as relationship, marital, and perceived health status (Stoll et al., 2017). People's vocational interests consequently impact their lives far beyond work, especially through the influence of long-lasting life decisions (e.g., Usslepp et al., 2020). As vocational interests are a major predictor for occupational and study major choices (Holland, 1997; Päßler & Hell, 2012; Wille et al., 2020), they also possess a high practical relevance. Interest inventories are widely applied in vocational counselling (see American College Testing Program, 2009) and they are the basis of the Occupational Information Network (O*NET), which was established by the United States Department of Labor, to support the (online) career guidance process of jobseekers (Hansen, 2019; O*NET, 2021).

Despite their widespread practical application, relatively little is known about the development of vocational interests, especially in younger age groups (Rounds & Su, 2014; Stoll & Trautwein, 2017). Although current evidence about their development includes two meta-analyses (mean-level change: Hoff et al., 2018; stability: Low et al., 2005) that provide valuable insights about developmental patterns across broad life phases, detailed information for shorter developmental periods, such as yearly mean-level and stability changes, and on individual factors that may moderate interest stability and contextual influences that could alter interest development, is still missing. This is in line with former and current review articles (Hansen, 1984; Rounds & Su, 2014; Stoll & Trautwein, 2017), which called attention to the lack of multiwave studies—i.e., longitudinal studies with more than two time points—that focus on the development of vocational interests. Further studies based on multiwave data are needed, as they enable the investigation of developmental processes that were proposed by vocational interest theories (Gottfredson, 1981; Hidi & Renninger, 2006; Holland, 1997; Su et al., 2019).

Summarizing the main empirical findings about the development of vocational interests, the following three issues become apparent.

First, current meta-analyses include only a small number of studies during the period of late childhood and early adolescence (ages 11 to 14). In Hoff et al.'s (2018) meta-analysis about mean-level development over the life course, only 2 out of the 49 included studies examined the age range of 11 to 14. In addition, in Low et al.'s (2005) meta-analysis about stability, only 5 out of the 66 included studies examined the age range of 12 to 14. This small total number of studies illustrates a lack of research in vocational interests for age groups younger than 15. Over the course of late childhood and early adolescence, it is assumed that vocational interests are in the process of developing (Gottfredson, 1981; Tracey, 2001) and begin to transform into dispositional preferences (Holland, 1997; Low et al., 2005; Tracey, 2001). At the same time, young adolescents begin to explore the world of work in preparation for first career decisions, such as the choice of vocational educational training, which occurs around the age of 16 in many countries (Eurostat, 2021). This combination of internal transformations (e.g., inconsistencies in likes and dislikes) and increasing external requirements (e.g., dealing with the world of work) suggests that vocational interests could experience substantial changes in intensity and stability over the course of late childhood and early adolescence. However, to date, evidence for these crucial years is scarce (for exceptions see Päßler & Hell, 2020; Tracey, 2002). The first study of the current dissertation therefore addresses this issue by investigating the development of vocational interests during late childhood and early adolescence.

Second, theories about individual differences indicate that personality characteristics are associated with the development of vocational interests (Ackerman, 1996; Armstrong et al., 2008; Holland, 1997; Roberts & Wood, 2006; Rounds & Su, 2014; Schmidt, 2014; Su et al., 2019). Most studies that investigated this assumption focused on interest intensity (Ackerman, 1997; Barrick et al., 2003; Hoff, Song, Einarsdóttir, et al., 2020; Larson et al., 2002; Mount et al., 2005). However, little is known about the influence of personality characteristics on the stability of vocational interests. Studies about vocational interest stability indicate that the degree of stability significantly varies from person to person (e.g., Dunkleberger & Tyler, 1961; Etzel & Nagy, 2021; Hoyt et al., 1957; Stoll, Rieger, et al., 2020; Swanson & Hansen, 1988; Xu & Tracey, 2016; Zytowski, 1976). Although it is stated that these variations in stability might be associated with variations in personality characteristics (Hirschi, 2010; Schomburg & Tokar, 2003; Swanson, 1999), to date, no empirical study has investigated that relationship. Answering this question is also practically relevant, as most vocational counsellors rely on the stability of vocational interest profiles when they use them to suggest prospective occupations

for their counselees (Brown, 2002; Low et al., 2005; Savickas & Taber, 2006; Strong, 1931). The second study of the current dissertation therefore addresses this issue by investigating the relationship between the stability of vocational interests and personality characteristics.

Third, it is assumed that vocational interest development is also impacted by contextual factors (Holland, 1997). Although there is extensive evidence that vocational interests facilitate the selection into certain environments (e.g., school environments, Volodina et al., 2015; study majors, Päßler & Hell, 2012; Roloff Henoch et al., 2015; Wille et al., 2020; occupations, Wille et al., 2010, 2014), less is known on how these environments influence vocational interest development. Only a few studies investigated contextual influences on vocational interests, mostly focusing on the impact of educational (e.g., Etzel & Nagy, 2021) or occupational (e.g., (Meir & Navon, 1992; Schultz et al., 2017) environments in older age groups. The majority of them reported small effects, which suggests that environments might be more impactful during earlier life phases. Besides school, especially leisure-related environments seem influential, where children and adolescents spend approximately 30% to 50% of their waking time (Larson, 2000). In addition, leisure-related environments possess properties that are advantageous for triggering interest development, such as a high degree of autonomy and intrinsic value (Hidi & Renninger, 2006; Larson, 2000; Ryan & Deci, 2000; Su et al., 2019). Investigating the influence of leisure-related environments would provide more comprehensive insights about the role of experiences from different contexts on the development of vocational interests. The third study of the current dissertation addresses this issue by examining the impact of the engagement in leisure-related activities on the development of vocational interests.

The aim of the current dissertation is twofold. On the one hand, it addresses the three issues that were elaborated in the previous paragraphs, by investigating the influence of individual and contextual factors on the development of vocational interests—primarily in younger age groups. On the other hand, it provides a comprehensive investigation of the development of vocational interests, by integrating evidence on interest stability and intensity from different empirical studies under the umbrella of theoretically derived developmental principles. The main focus of the dissertation will be on the time period of adolescence (ages 11 to 18) because various processes of development are assumed to begin or occur during that specific life phase (Gottfredson, 1981; Holland, 1997; Su et al., 2019; Tracey, 2001). The combination of moderate interest stability (see Low et al., 2005) and the emerging need to deal with the world of work, makes adolescence an intriguing and promising time period to study vocational interest development.

The dissertation is structured in the following way. Chapter 1.1 introduces the role of situational consistency in personality and interest research. Chapter 1.2 gives an overview about the current conceptualization of vocational interests, their dispositional nature, and taxonomy. Chapter 1.3 provides an overview of indicators that can describe development in vocational interests, elaborates current theories about the development of vocational interests, and concludes with current empirical evidence. Chapter 1.4 focuses on possible factors that may impact the development of vocational interests, including individual and contextual influences. Chapter 1.5 summarizes methodological challenges that can be expected when investigating the development of vocational interests. Chapter 2 introduces the research questions to the three empirical studies, which are presented in Chapters 3 to 5. The first study addresses the missing descriptive information about the development of vocational interests during late childhood and early adolescence. The second study focuses on the relationship between the stability of vocational interests and personality characteristics. The third study investigates the influence of the engagement in leisure-related activities on the development of vocational interests. Chapter 6 summarizes, integrates, and discusses the results of the empirical studies, elaborates their relevance for practice and critically states their limitations.

1.1 Adolescence, Personality, and Situational Consistency

Adolescence is a phase that is accompanied by great changes in an individual's life (Choudhury et al., 2006; Crone & Dahl, 2012; Steinberg, 2005). Cognitive, affective, and biological development processes (Steinberg, 2005) are accompanied by shifts in social contexts (e.g., school environments) and roles (Crone & Dahl, 2012; Sawyer et al., 2018). Studies about brain development suggest that maturation processes and interconnectivity between parts of the brain are still evolving till late adolescence, resulting in different consequences for cognitive abilities, motivation, risk-taking, or self-regulation (Crone & Dahl, 2012; Steinberg, 2005). Cognitive abilities of adolescents are assumed to evolve, resulting in more enhanced thinking processes, including "abstract, multidimensional, planned and hypothetical thinking" (Steinberg, 2005, p. 70) and due to the onset of puberty, changes in arousal sensitivity and general motivation can be expected (Crone & Dahl, 2012; Steinberg, 2005). Besides these internal processes, adolescents are also influenced by changes in their environments and social normative roles (Sawyer et al., 2018). Educational environments such as schools change their structure when students transfer from primary to secondary education, resulting in increased performance requirements and individual responsibility of students to manage their school day (Harter et al., 1992). There are shifts in communication between parents and their children (Keijsers & Poulin, 2013), increases in parent-child conflicts (McGue et al., 2005), and changes in the structure and influence of peer groups (O'Brien & Bierman, 1988).

These biological, cognitive, and social changes over the course of adolescence suggest that individuals also sustainably transform their behavior. General behavioral patterns are often operationalized based on personality trait models. These models usually describe a broad set of traits that summarize individual characteristics, which capture relatively enduring patterns of behavior, such as being conscientious, extraverted, or open to new experiences (Ashton & Lee, 2007; DeYoung et al., 2013; McCrea & Costa, 1999). There is an ongoing debate in personality research about the consistency of these personality traits (Fleeson & Jayawickreme, 2015). On the one hand, there is the assumption that traits are consistent across situations (McCrea & Costa, 1999). Studies based on that approach identified a small set of stable personality traits, which are supposed to describe differences in general behavior between people (Ashton & Lee, 2007; DeYoung et al., 2013; McCrea & Costa, 1999). On the other hand, there is the assumption that traits are also inconsistent across situations (Fleeson & Jayawickreme, 2015). This approach assumes that personality traits unfold during a certain situation and their expressions are not always the same across situations (Mischel et al., 2002). For example, a person might

be more extraverted in some situations, such as at a party with close friends, but less extraverted in other situations, such as on the first day at a new job.

Personality characteristics that fit into this discussion about situational consistency include general interests. Situational consistency is also discussed in interest research (Hidi & Renninger, 2006; Silvia, 2001; Su et al., 2019). Models exist that capture both perspectives, situation-consistent interests, and situation-specific interest (Hidi & Renninger, 2006; Silvia, 2001; Su et al., 2019). There is research on interests that focuses on them being a stable aspect of personality that govern people's long-term choices and activities (Holland, 1997). Interests are assumed to be a good predictor of what people want out of life (Stoll, Einarsdóttir, et al., 2020). Other parts of interest research emphasize that interest also is similar to an emotion of curiosity (Silvia, 2008). It therefore has a motivating function for human behavior that initiates and attracts engagement in certain situations and activities (Silvia, 2001, 2008). Further models integrate both perspectives, suggesting that interests are consistent across situations, while interests have also unique functions and expressions during situations (Silvia, 2001; Su et al., 2019). In the current dissertation, a general perspective will be taken that highlights both concepts—situation-consistent and situation-specific. This view is complementary to approaches in personality research that integrate the conceptions of personality traits (i.e., a stable description of the personality of a person) and situation-specific personality states (i.e., expressions of a personality characteristic in a specific moment; e.g., Fleeson, 2004; Fleeson & Gallagher, 2009; Fleeson & Jayawickreme, 2015). Insights from the current dissertation can therefore be incorporated in that strain of personality research.

1.2 The Nature of Vocational Interests

To facilitate the understanding of vocational interest development, it is important to summarize the properties of vocational interests. A thorough description is necessary, as developmental mechanisms that are elaborated in later chapters are based on theoretical assumptions that are stated in the current chapter. The current chapter starts with an overall discussion about the dispositional and situational conceptualizations of interest. Empirical evidence is reported that provides support for the dispositional nature of vocational interests. Chapter 1.2.1 concludes by describing a theoretical approach that integrates dispositional and situational conceptualizations, the Trait-Situation Interest Dynamics (TSID) model (Su et al., 2019). In Chapter 1.2.2, the most widely used taxonomy in vocational interest research is elaborated, the RIASEC model by Holland (1997). Finally, Chapter 1.2.3 integrates Holland's (1997) RIASEC model and the TSID model (Su et al., 2019) to provide a more sophisticated understanding of the nature of vocational interests and to go beyond Holland's (1997) conceptualization for occupational preferences. The definition of vocational interests that is elaborated in the following chapter will be used throughout the remaining dissertation.

1.2.1 Dispositional and Situational Conceptualizations

The construct of interest has a long history in psychological research (Renninger & Hidi, 2011). About a century ago, interest began appearing in the works of Dewey (1913), Piaget (1940) and Thorndike (1935). Research endeavors that focused specifically on vocational interests date back to Strong (1943), Thurstone (1931) and Bordin (1943). Over time, a variety of interest conceptions, models and theories evolved, scattered across various disciplines of psychology (e.g., educational psychology, vocational psychology or biological psychology), postulating a diversity of developmental and structural assumptions (Renninger & Hidi, 2011). However, despite this variety, interests possess certain properties that are common across theoretical conceptualizations (Renninger & Hidi, 2011). The following characteristics are inherent to the construct of interest.

General Properties of Interest

Interest can be broadly defined as a preference—i.e., likes and dislikes—for certain activities, environments, or outcomes that determine the direction and the intensity of someone's behavior (Renninger & Hidi, 2011; Rounds & Su, 2014; Su et al., 2009, 2019). Interest is always oriented towards an object, implying that it has to be viewed in the respective context (Rounds & Su, 2014). For example, with regard to vocational interests, the object of

interest usually includes activities, environments, or outcomes that are located in the world of work (Rounds & Su, 2014). The contextualization also implies that interest is sustained and driven through the interplay between the person and properties of the respective environment (Holland, 1997). Certain environmental characteristics, such as the novelty of a task (Hidi & Renninger, 2006), can lead to the initiation of previously non-existent interest in a certain topic. In addition, interest can be comprised of both cognitive and affective components (Su et al., 2019). At the affect level, interest consists of positive emotions such as being fascinated, caught-up, or enlivened (Hidi & Renninger, 2006; Silvia, 2008; Su et al., 2019). At the cognitive level, interest is accompanied by thought processes that analyze whether the respective activities that initiated interest are intriguing, meaningful, or compatible with the perceptions of one's self (Hidi & Renninger, 2006; Su et al., 2019). Finally, there is a biological foundation for the construct of interest (see also Beltz et al., 2011; Betsworth et al., 1994; Moloney et al., 1991), illustrated by different patterns of brain activation when persons are highly interested in a specific topic or not (see Renninger & Hidi, 2011), by the influence of hormones on gender differences in interests (Beltz et al., 2011) and by findings of twin studies, which suggest that interests possesses hereditary components (e.g., Moloney et al., 1991).

Against the background of these general properties, the current dissertation specifically focuses on two general conceptualizations, dispositional interests, and situational interest.[1] Depending on the discipline and research tradition, interest is either described as a dispositional or a situational construct (Renninger & Hidi, 2011). As a disposition, interests are described as long-lasting and stable entities that are traitlike and not necessarily related to a current situation (Holland, 1997; Rounds & Su, 2014; Su et al., 2019). A person with dispositional interests in social topics, for example, will have preferences for social activities independent of certain situational structures. The person will autonomously seek out activities, situations, and environments that allow them to fulfill their social preferences. As a situational conception, interest is mainly bound to the situation, externally triggered, and usually accompanied by short-term occurrences of positive emotions, higher attention, and a deeper cognitive engagement (Hidi & Renninger, 2006). Situational interest is usually aroused by novel,

[1] There are further possibilities to categorize the conceptualizations of interest. Renninger and Hidi (2011), for example, divide conceptualizations of interest into the categories of development, emotion, task features, value, and vocational interests. However, Su et al. (2019) describe that most interest conceptualizations can be broadly categorized as situational or dispositional. In the current dissertation, situational and dispositional conceptualizations are important to understand the developmental mechanisms that will be derived in later chapters. Therefore, the two conceptualizations are explained in more detail.

exciting, and cognitively engaging features inherent to the situation (Hidi & Renninger, 2006). An example could be situational interest in science that is briefly triggered through the characteristics of a controversial school lesson, a novel experiment, or an exciting documentary (Hidi & Renninger, 2006).

Vocational Interests as Dispositional Preferences

In vocational psychology, interests are traditionally conceptualized as a dispositional construct (Holland, 1997; Rounds & Su, 2014; Su et al., 2019). This perspective evolved based on theoretical models and taxonomies, which postulated a traitlike nature of vocational interests (e.g., Holland, 1997; Rounds & Su, 2014). This assumption is supported by meta-analytic findings that show that vocational interests are relatively stable over the life course and that their stability is similar to other dispositional constructs (Low et al., 2005). For example, in early life phases, vocational interests are just as stable as personality traits; during young adulthood, they are even more stable (Low et al., 2005; Roberts & DelVecchio, 2000). Further evidence suggests that vocational interests are also, to some degree, hereditary (Kandler et al., 2014). According to a review by Kandler et al. (2014), vocational interests possess a significant degree of genetic variance (for primary studies see Betsworth et al., 1994; Harris et al., 2006; Kandler et al., 2011; Moloney et al., 1991), located within the range of 30% to 60%. As is the case for the degree of stability, the degree of genetic variance is similar between vocational interests and other dispositional constructs, such as personality traits (Kandler et al., 2014).

Vocational interests are assumed to be embedded in someone's personality (Ackerman & Heggestad, 1997; Holland, 1997; Roberts & Wood, 2006; Su et al., 2019). Some models of personality differentiated between core (i.e., personality traits represented in the Big Five taxonomy: Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism) and surface characteristics (e.g., motivation and values), implying that motivational constructs like vocational interests are subsequent to and an outcome of personality traits (see McCrea & Costa, 1999). However, recent empirical evidence suggests that this distinction may not be true (Kandler et al., 2011, 2014; Rieger et al., 2017). In addition, vocational interests are covered by the usual broad definition of personality[2]: "[…] personality is most commonly defined as the sum of all characteristics that reflect relatively enduring patterns of emotion, cognition, motivation and behavior in which one individual differs from others within a certain reference population (e.g. age group or culture)" (Kandler et al., 2014, p. 231). Therefore, recent

[2] In the current dissertation, when the term personality is mentioned, the definition of Kandler et al. (2014) applies.

theoretical approaches, such as the Neo-Socioanalytic Model of Personality (Roberts & Wood, 2006), integrated interests into their broader conceptualization of personality and assigned them similar properties to personality traits. This strengthens the view that vocational interests are dispositional preferences.

Including the Situation—The Trait-Situation Interest Dynamics Model

Although vocational interests are seen as dispositional constructs, meta-analytic evidence suggests that they are not perfectly stable (Low et al., 2005). Vocational interests are assumed to have the potential to substantially change, particularly in younger life stages (Rounds & Su, 2014). Rounds and Su (2014) therefore argued that the situational (or state) perspective must be included in vocational interest research, as it offers unique explanations for interest development, especially over the course of adolescence. The Trait-Situation Interest Dynamics (TSID) model proposed by Su et al. (2019) incorporated this idea, by integrating situational and dispositional conceptions of interests. It assumes that interests are relatively stable dispositions, which are reflected by an abstract and schematic mental representation. The mental representation stores previous experiences (e.g., general memories of conducting an activity), emotional responses (e.g., feeling happy, curious, or enlivened during the activity), and cognitive appraisals (e.g., questioning if the task is meaningful) of the object of interest. However, besides the dispositional component, interests also consist of a situational component that can be initiated by situational characteristics (Su et al., 2019). According to the TSID model, it is possible to initiate short-term situational interest, based on activities that are, for example, surprising, thought-provoking, or personally relevant (Hidi & Renninger, 2006). An accumulation of new experiences that arouse situational interest helps to strengthen and refine the abstract and schematic mental representation (Su et al., 2019). The TSID model therefore states that situational interest can initiate the development of dispositional interests.

1.2.2 The RIASEC Model

The most widely used taxonomy of vocational interests is Holland's (1997) model of interest classification (Rounds & Su, 2014). The model originates from his theory about vocational personalities and work environments (Holland, 1997). Besides its importance for research, the model is widely applied in the practice of vocational counselling (Brown, 2002; Hansen, 2019). Many interest inventories (Hansen, 2019), occupational exploration approaches (O*NET, 2021), and career guidance measures (Brown, 2002), are based on Holland's (1997) interest taxonomy. Due to its closeness to practice, some of the model's conceptions—such as the classification of interests and environments in six broad dimensions (Holland, 1997)—are

affected by an underlying idea of pragmatism. The broad aim of Holland's (1997) theory is to elaborate how people, with a certain set of personality characteristics, behave in occupational environments with certain properties (Holland, 1997). Based on the fit between the characteristics of a person and the characteristics of an environment, Holland (1997) derives several assumptions about career decisions, career satisfaction, occupational achievement, as well as the stability and change of occupational careers.

The RIASEC Dimensions

Holland (1997) describes six general dimensions: Realistic, Investigative, Artistic, Social, Enterprising, and Conventional (RIASEC).[3] These dimensions are defined as relatively stable dispositions that are an expression of someone's personality (Holland, 1997). They incorporate not only preferences for certain tasks, but also values, personality traits, and competencies. For example, a person with high scores on the dimension Investigative, possesses specific interests, values, personality traits, and competencies, which differ from persons characterized by other dimensions. Holland (1997) explains that the RIASEC dimensions are ideal-typical theoretical descriptions of certain personality types, which are used to measure the characteristics of a real person. The more a person resembles a respective RIASEC dimension, the more they will express the characteristics of the respective dimension in real life (Holland, 1997).

People who resemble the Realistic dimension have preferences for activities that entail the manipulation of objects, working with their hands, tools, and machines as well as physical tasks (Holland, 1997). There is an aversion to social activities that entail educational (e.g., teaching) or therapeutic (e.g., listening to someone's problems) tasks (Holland, 1997). Based on their preferences, Realistic people often develop abilities that encompass manual, mechanical, technical, agricultural, and electrical competencies. In line with their preferences, competencies, and values, Realistic people often tend to show personality traits such as being persistent, genuine, and practical as well as inflexible, materialistic, and hardheaded. According to the O*NET occupational data base, the following occupations are examples of a Realistic-

[3] Although Holland (1997) originally labelled the RIASEC dimensions as personality types, Stoll and Trautwein (2017) suggest that the term personality type is somewhat misleading. It suggests that people can be categorized into only one of the six personality types. However, a person is usually characterized by his or her expression on all of the six personality types. This circumstance is best illustrated by the profile perspective of vocational interests, which implies that people can be characterized by a set of scores on the respective RIASEC dimensions. Therefore, in the current dissertation the label dimension was chosen instead of the label personality type.

dominant environment: automotive engineer, forester, electrician, or animal trainer (O*NET, 2021).

The Investigative dimension is characterized by people having preferences for activities that encompass tasks like systematically observing and investigating physical, biological, and cultural phenomena (Holland, 1997). Investigative people have an aversion to persuasive (e.g., leading others) and repetitive activities (e.g., assembly line tasks) and often develop competencies that comprise mathematical and scientific skills. They value scholarly achievements, independence, being ambitious, or being logical (Holland, 1997). In line with their preferences, competencies, and values, Investigative people often tend to show personality traits such as being analytical, introspective, and curious as well as pessimistic, radical, and reserved. According to the O*NET occupational data base, the following occupations possess an Investigative-dominant environment: chemist, biostatistician, physicist, or sociologist (O*NET, 2021).

People who resemble the dimension Artistic prefer to do creative tasks like drawing and acting as well as taking part in activities that are unsystematic, ambiguous, and free (Holland, 1997). Artistic people have aversions towards explicit, systematic, and ordered activities and usually evolve advanced skills in language, art, drama, music, and writing (Holland, 1997). They value self-expression, equality, being imaginative, and being courageous (Holland, 1997). In line with their preferences, competencies, and values, Artistic people often tend to show personality traits such as being emotional, open, and imaginative as well as impractical, impulsive, and disorderly. According to the O*NET occupational data base, the following occupations possess an Artistic-dominant environment: actor, craft artist, photographer, or graphic designer (O*NET, 2021).

People who resemble the Social dimension have preferences for activities that entail informing, teaching, training, and developing others (Holland, 1997). They usually have aversions towards activities that include tools or machines and often develop a wide range of interpersonal and educational competencies (Holland, 1997). Social people value social and ethical activities as well as being helpful (Holland, 1997). In line with their preferences, competencies, and values, a Social person often tends to show personality traits such as being agreeable, empathic, and friendly (Holland, 1997). According to the O*NET occupational data base, the following occupations possess a Social-dominant environment: recreation worker, nurse, physical therapist, or career counsellor (O*NET, 2021).

The Enterprising dimension is characterized by people having preferences for tasks such as manipulating and leading other people (Holland, 1997). They have an aversion to observational and systematic tasks (Holland, 1997) and usually evolve competencies that contain leadership and business administrative skills (Holland, 1997). Enterprising people possess traditional values, such as valuing economic or political achievement (Holland, 1997). In line with their preferences, competencies, and values, Enterprising people often tend to show personality characteristics such as being enthusiastic, extroverted, and ambitious as well as forceful, domineering, and resourceful. According to the O*NET occupational data base, the following occupations possess an Enterprising-dominant environment: general manager, chief executive, retail salesperson, or financial manager (O*NET, 2021).

A person who scores high on the dimension Conventional likes to work in a structured way and according to given rules (Holland, 1997). Conventional activities are tasks such as keeping records, filing materials, and processing data (Holland, 1997). Conventional people have an aversion to free and unsystematic tasks (Holland, 1997) and evolve advanced clerical, computational, and business system skills (Holland, 1997). Due to their preferences, competencies, and values, Conventional people often tend to show personality characteristics such as being orderly, efficient, and methodical as well as unimaginative, inflexible, and careful. According to the O*NET occupational data base, the following occupations possess a Conventional-dominant environment: office clerk, loan officer, judicial law clerk, or archivist (O*NET, 2021).

Gender Differences on the RIASEC Dimensions

Gender differences in vocational interests have been known for almost a century (see Einarsdóttir & Rounds, 2009; Fouad, 1999). Although Holland (1997) is less specific about describing gender differences in his theory, empirical evidence suggests that there are large and robust differences between females and males on the majority of his interest dimensions (Su et al., 2009). Based on a meta-analysis that included about half a million respondents across 81 samples, Su et al. (2009) investigated gender differences in vocational interests across several taxonomies. According to the results, females were more interested in topics that were related to people (e.g., leading a group of people, teaching, or caring for the elderly), whereas males were more interested in topics that were related to things (e.g., working with tools, building things, or creating something out of wood). These gender differences also corresponded to mean-level differences on scales that captured the RIASEC dimensions (for a depiction see Figure 1). Females possessed significantly stronger Artistic ($\Delta d = 0.35$), Social ($\Delta d = 0.68$) and Conventional ($\Delta d = 0.33$) interests, whereas males possessed significantly stronger Realistic

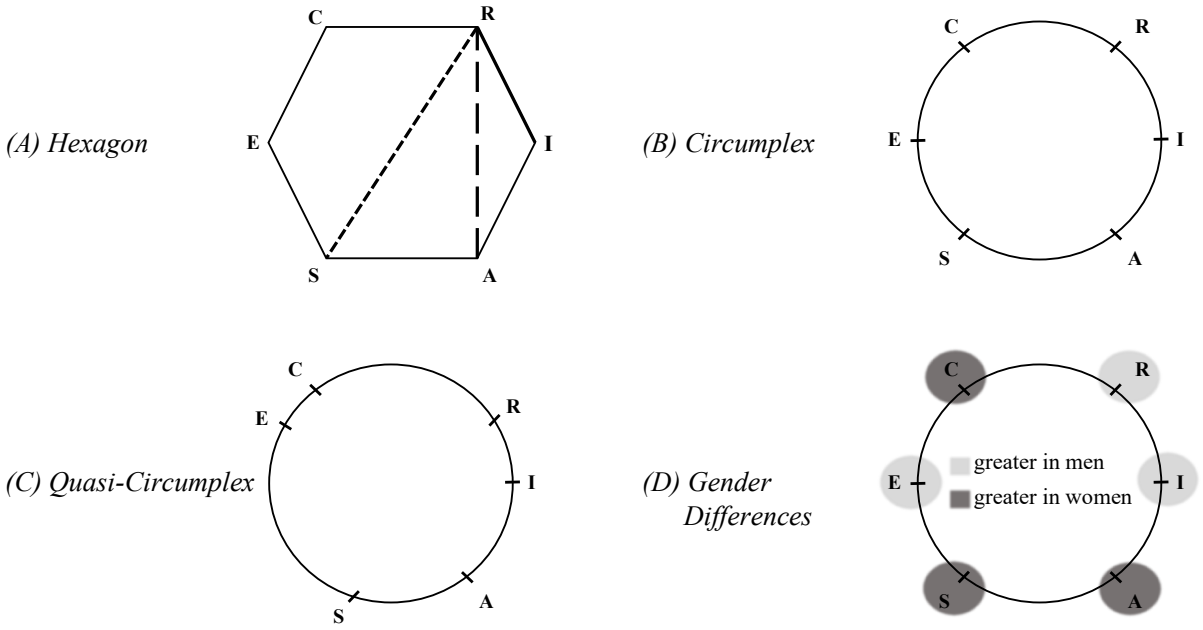
($\Delta d = 0.84$) and Investigative ($\Delta d = 0.26$) interests. The biggest differences in interests between males and females were found on the dimensions Realistic and Social. As effect sizes ranged between $\Delta d = 0.26$ and $\Delta d = 0.84$, gender differences in vocational interests are considered robust and substantial (Su et al., 2009), with some researchers arguing that they are the largest among gender differences in personality characteristics (Lubinski, 2000). Because gender differences play an important role in vocational interest research, assumptions about their origin and development will be thoroughly elaborated in Chapter 1.3.

Interrelations of the RIASEC Dimensions—Variable and Profile Perspectives

Holland (1997) extensively describes the interrelations of the RIASEC dimensions. According to his calculus hypothesis, the RIASEC dimensions can be arranged on the angles of a hexagon (see Figure 1, picture A). The proximity between the dimensions on the hexagon represents their theoretical and empirical closeness (Holland, 1997). This implies that interest dimensions that are close to each other on the hexagon, such as the dimensions Realistic and Investigative, are assumed to have a much stronger relationship than interest dimensions that are opposite to each other on the hexagon, such as the dimensions Realistic and Social (Holland, 1997).

Figure 1

Depictions of Holland's (1997) RIASEC Dimensions



Note. The figure depicts various interrelations of Holland's (1997) interest dimensions; gender differences in picture (D) are based on the meta-analysis of Su et al. (2009), who investigated gender differences on vocational interests.

Despite evidence that Holland's (1997) proposed ordering of the RIASEC dimensions is appropriate (Anderson et al., 1997; Darcy & Tracey, 2007; Rounds & Day, 1999), current studies suggest that the distances between the dimensions are not equal (e.g., Gupta et al., 2008). Instead of a hexagon, it is suggested that the RIASEC dimensions should be arranged on a circumplex (see Figure 1, picture B). The circumplex structure makes it possible to visualize the unequal distances between the RIASEC dimensions by locating them on a circle (i.e., quasi-circumplex; see Figure 1, picture C). Besides the circumplex structure, alternative models for vocational interests were proposed by Gati (1979), Prediger (1982) and Rounds and Tracey (1996). However, the circumplex has emerged as the favored structural model for Holland's (1997) dimensions as it was replicated across a wide range of samples (Tracey & Rounds, 1993; for a comparison between the models see Nagy et al., 2010).

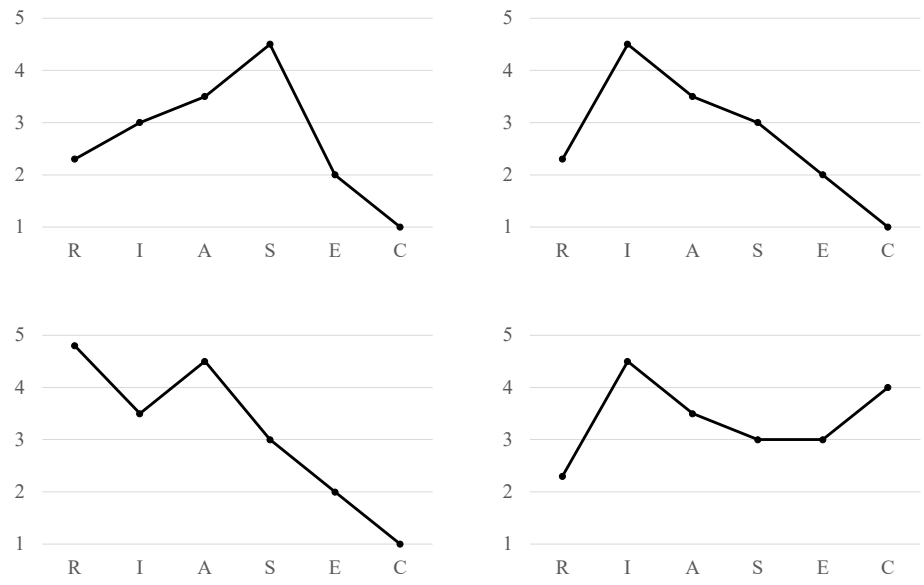
Besides the variable perspective, Holland (1997) emphasizes a profile perspective on vocational interests. A vocational interest profile is the set of a person's scores on all RIASEC dimensions (see Figure 2). The profile perspective is very prominent in Holland's (1997) theory, as he describes specific profile properties, such as consistency and differentiation (i.e., the so-called secondary constructs). Profile consistency implies that profiles with dominant interest dimensions (i.e., the dimensions that possess the highest scores) that are theoretically close to each other (e.g., Realistic and Investigative) are more consistent compared to profiles with dominant interest dimensions that are not theoretically close to each other (e.g., Realistic and Social; Holland, 1997). Profile differentiation focuses on the dispersion of all interest scores within a profile. Differentiated interest profiles have clear peaks and lows, whereas undifferentiated profiles usually possess similar expressions on all interest dimensions. Advances in methodological research have led to an increase in studies that investigated profile characteristics of vocational interests (Etzel et al., 2019; Etzel & Nagy, 2021; Gurtman, 1993; Nagy et al., 2009; Vock et al., 2013) and their influence on various outcomes (Tracey et al., 2014). Holland (1997) stated that consistent and differentiated profiles possess a better predictive validity in comparison to inconsistent and undifferentiated profiles. However, empirical evidence for that assumption is rather weak (Tracey et al., 2014), suggesting that consistency and differentiation are less practically relevant than proposed.

Current longitudinal studies that examined the development of vocational interests mainly focused on the variable perspective. They primarily investigated the development of interest dimensions over the course of certain life phases (Hoff et al., 2018; Hoff, Song, Einarsdóttir, et al., 2020; Päßler & Hell, 2020; Stoll, Rieger, et al., 2020). This approach was

adapted from other research areas that focused on the development of dispositional constructs. For example, studies about the life course development of personality traits often examined mean-level changes of single traits over longer periods of time (Roberts et al., 2006; Soto et al., 2011; Soto & Tackett, 2015; Van den Akker et al., 2014). In the current dissertation the development of vocational interests is investigated from both—variable and profile—perspectives. In contrast to profile consistency and differentiation (Tracey et al., 2014), the construct of profile stability is assumed to be theoretically and practically relevant (Low et al., 2005; Xu & Li, 2020), as vocational interest profiles are applied in the practice of career guidance (Hansen, 2019). Profile stability can be broadly defined as the temporal stability of the set of scores within a profile (Livingston et al., 2003). In Chapter 1.3, indicators are described that capture the development of vocational interests on the variable and the profile levels.

Figure 2

Depictions of Four Individual RIASEC Profiles



Note. The picture depicts four hypothetical RIASEC profiles of four individuals, with different dominant interest dimensions and shapes.

1.2.3 Beyond RIASEC—Integrating the TSID Model

The RIASEC model (Holland, 1997) and the TSID model (Su et al., 2019) are central to the current dissertation. Both models show similarities and differences in how they depict the construct of interest. The TSID model elaborates a general conceptualization of interests without having a content focus (e.g., occupational or educational), whereas the RIASEC model

specifically describes interests from an occupational perspective. However, the TSID model is strongly influenced by theories, assumptions, and empirical findings from vocational interest research (Su et al., 2019), including interest conceptions of Holland (1997). For a more sophisticated understanding of the nature of vocational interests, the differences and similarities of both models must be carved out. Furthermore, to comprehensively understand the development of vocational interests, especially over the course of adolescence, insights from both models are necessary. Holland's (1997) description of vocational interests helps to understand their development over broader life phases. Changes during smaller time periods and explanations for different developmental trajectories of vocational interests can be better understood by interest properties that are based on the TSID model.

Commonalities between the RIASEC and the TSID models

The RIASEC and the TSID models state that interests are traitlike dispositions, which are part of someone's personality (Holland, 1997; Su et al., 2019). Both models see interests as relatively stable entities that are, to some degree, hereditary and influenced by underlying biological factors. In this context, both models also emphasize the motivating function of interests, their contextualization, and their prolonged influence on human behavior. For example, they assume that interests are important predictors of the decisions we make, including minor choices, such as choosing leisure time activities (Kerby & Ragan, 2002), and long-lasting life decisions, such as choosing an educational over a vocational track (Usslepp et al., 2020). Besides the assumption that interests influence the selection into environments, both models emphasize that interests are also impacted by the environment. For example, Holland (1997) postulates that interests may adapt through socialization after entering an environment, while Su et al. (2019) underlines that interests can be shaped through an accumulation of positive or negative experiences in certain environments. This implies that both models see interests as traitlike dispositions that can be altered through the interaction with the environment.

Furthermore, both models see interests as part of someone's personality and they emphasize that interests are embedded in the development of personality. This implies that interests do not exist in a vacuum, but are influenced by and co-develop with a variety of different personality characteristics (Hoff, Song, Einarsdóttir, et al., 2020; Holland, 1997; Su et al., 2019). The TSID model is very specific in that matter. Su et al. (2019) explicitly state in one of their six propositions that interests evolve in relation to other individual characteristics. In addition, Su et al. (2019) locate interests in a nomological net of other variables, including personality traits (i.e., operationalized within the Big Five framework; McCrea & Costa, 1999)

and cognitive abilities. Holland (1997) makes similar assumptions about the embeddedness of interests. He suggests that over time, individuals will try out various activities and their interests will develop from single preferences to a well-defined set of dispositional interests. During that process, individuals will also evolve competencies, personality traits, and values that are in line with their set of dispositional interests (Holland, 1997).

Differences between the RIASEC and the TSID models

A subtle, but important difference between the RIASEC and the TSID models is how they locate interests within the framework of personality. According to Holland's (1997) descriptions, interests can be considered as a byproduct of the RIASEC dimensions (Su et al., 2019). This implies that depending on the dominant RIASEC dimension, people possess certain expressions of interests, personality traits, competencies, and values. From this perspective, interests are strictly speaking a downstream construct of the RIASEC personality types and located at a lower level of abstraction (see Su et al., 2019). Su et al. (2019), on the other hand, describe interests as part of someone's personality and not as a byproduct, which implies no lower level of abstraction.

The RIASEC and the TSID models also differ in their foci concerning the definition and description of interests. Holland (1997) for example provides a detailed taxonomy of vocational interests by describing the content of the six RIASEC dimensions. The TSID model, on the other hand, offers no such classification. This makes it difficult to derive any operationalization of interests based on the TSID model. However, the TSID model thoroughly describes in which way dispositional interests are stored—in the form of a mental representation of the object of interest—and which processes—affective and cognitive—are involved when someone experiences interest during a certain activity. This is achieved by integrating dispositional and situational perspectives, which offers a more in-depth description about what interests are. Holland (1997) on the other hand, provides no such explanations. Especially the integration of dispositional and situational components of interests is one of the major differences between the RIASEC and the TSID models.

Integrating the RIASEC and the TSID models

In the following, theoretical assumptions from both models will be integrated to gain a more sophisticated conceptualization of vocational interests. First, in line with Su et al. (2019) and Holland (1997) it is assumed that vocational interests are part of someone's personality and relatively enduring dispositions. This assumption is in line with empirical evidence on the stability of vocational interests (e.g., Low et al., 2005; Stoll, Rieger, et al., 2020). However, in

contrast to Holland (1997) and in line with Su et al. (2019), it is assumed that vocational interests are not necessarily a byproduct of overarching personality types, but rather are on an equal level of abstraction as personality characteristics. Evidence against the consideration that vocational interests are downstream constructs of personality is provided by Kandler et al. (2014). Second, in line with Holland (1997), the RIASEC taxonomy will be used as a content-specific classification for vocational interests. There are several studies that support the structural validity of the six interest dimensions (e.g., Nagy et al., 2010; Tracey & Rounds, 1993). Third, in line with Su et al. (2019), it is assumed that vocational interests are stored as mental representations that entail affective and cognitive components of the objects of interest. These mental representations are assumed to become refined based on the accumulation of new experiences. Because new experiences usually entail situational interest, it is also assumed that vocational interests comprise both dispositional and situational components. This assumption is supported by studies about school subject interests that suggest that interests may consist of dispositional and situational components (Rieger et al., 2017).

1.3 The Development of Vocational Interests in Adolescence

The purpose of Chapter 1.2 was to establish a sophisticated understanding of the nature of vocational interests. Chapter 1.3 elaborates theoretical approaches and empirical evidence about the development of vocational interests. The empirical indicators that are usually used to investigate continuity and change are described in Chapter 1.3.1. The chapter comprises indicators that focus on the variable perspective, such as mean-level change and re-test correlations, as well as indicators that focus on the profile perspective, such as profile stability. In Chapter 1.3.2 developmental theories are described and principles for the development of vocational interests during adolescence are derived. Finally, in Chapter 1.3.3 current empirical evidence on the development of vocational interests is reported, based on the indicators that were introduced earlier in Chapter 1.3.1.

1.3.1 Indicators of Vocational Interest Development

There are various ways to investigate the development of a dispositional construct. In the current dissertation, the focus will be on indicators that capture development on the variable (i.e., between-person) and profile (i.e., within-person) levels (Low & Rounds, 2007). An indicator that measures changes in interest intensity on the variable level is mean-level change (Low & Rounds, 2007). Mean-level development is investigated for a variety of dispositional constructs (e.g., personality traits; Roberts et al., 2006) and indicates changes in the expression of the intensity of the construct across all individuals within a sample (Low & Rounds, 2007). Mean-level changes over time are often labelled as normative change because they illustrate general changes in a construct that are found for a majority of people, which might correspond to underlying maturation processes (Roberts et al., 2006). Besides mean-level changes over time, mean-level differences between groups, for example gender, can also indicate differences in interest intensity between groups. This implies that the investigation of mean-level differences over time could be used to investigate the development of gender differences.

An indicator that measures stability on the variable level is re-test correlation. Instead of interest intensity, re-test correlations describe the stability of the rank orders of persons within a sample (Low & Rounds, 2007). Normative change and stability are distinct from one another (Low & Rounds, 2007; Roberts et al., 2006). It is possible that rank orders between persons do not change over time, however, mean levels do, and vice versa. Furthermore, even when mean levels show no shifts and re-test correlations are close to $r = 1$, change is still possible. For example, mean-level changes of subgroups in a sample could cancel each other

out, implying no overall mean-level shifts, or changes occur only on the individual level. This illustrates that continuity and change of vocational interests must be investigated from various perspectives.

According to the TSID model, interests may consist of a dispositional and a situational component (Su et al., 2019). Latent state-trait models (Steyer et al., 1999, 2015), which originate from the realm of confirmatory factor analysis, allow researchers to model such properties. Based on the calculation of state and trait variance proportions, vocational interests can be partitioned into long-term stable and short-term changing components. Although the TSID model stated that vocational interests could be composed of dispositional and situational components (Su et al., 2019), currently there is no empirical evidence on how large these components are. The application of latent state-trait models would improve the understanding of the nature of vocational interests. In Chapter 3, vocational interests are investigated with latent state-trait models over the course of late childhood and early adolescence. As the study examines the degree of dispositional and situational components of vocational interests, state and trait variance proportions are relevant indicators for the respective research question.

Mean-level change and re-test correlations are indicators on the between-person level (i.e., they express continuity and change between persons). An indicator on the within-person level (i.e., indicators that express continuity and change for a respective person) that is traditionally investigated in vocational interest research is profile stability (Low & Rounds, 2007). Profile stability is usually investigated by computing a profile stability coefficient for each individual in the sample (Low & Rounds, 2007). By aggregating the profile stability coefficients across all individuals in the sample, it is possible to investigate the average amount of within-person stability. An advantage of the within-person approach, is that stability can vary from person to person (Asendorpf, 1992). This implies that variations in stability can be associated to external criteria, such as individual or contextual factors. This approach allows researchers to connect within-person constructs (i.e., profile stability) to between-person differences. There are various ways to compute a profile stability indicator (Livingston et al., 2003; Low & Rounds, 2007; Xu & Li, 2020), of which each focuses on different profile characteristics (i.e., profile elevation, profile spread, or rank order of the profile elements). In Chapter 4, the stability of vocational interests and their predictors during various periods of adolescence and later life phases are investigated. Therefore, specific indicators that capture profile stability are thoroughly elaborated in Chapter 4.

There are additional indicators that capture the development of vocational interests (e.g. structural stability or congruence stability; Low & Rounds, 2007). However, mean levels, re-test correlations, state-trait variances and profile stability indicators are central to the current dissertation. They capture important information about the development of interest intensity and stability. In addition, focusing on the respective indicators enables the investigation of continuity and change from the between-person as well as the within-person perspective. Each of the studies included in the current dissertation therefore focuses on one of these indicators. As theories about vocational interests make no explicit assumptions about the development of these indicators, theoretical and empirical work will be used to derive assumptions about how they evolve over the course of adolescence.

1.3.2 Principles of Vocational Interest Development

In the current chapter, three principles are stated that summarize the development of vocational interests over the course of adolescence (ages 11 to 18). The focus will be on the derivation of universal principles that describe general developmental trends in interest stability, intensity, and gender differentiation.[4] In line with the interest conception from Chapter 1.2, both Holland's (1997) theory of vocational personalities and work environments, which includes the most widely used taxonomy for vocational interests (Hoff et al., 2018), and the TSID model (Su et al., 2019), which integrates conceptions of situational and dispositional interests, were used. In addition, an emphasis was laid on Gottfredson's (1981) theory of circumscription and compromise, as it focuses specifically on the development of occupational preferences from childhood to late adolescence. Due to the integration of vocational interests in the broad framework of personality, insights from personality trait development (e.g., (Roberts & Wood, 2006; Soto et al., 2011; Soto & Tackett, 2015; Van den Akker et al., 2014) will be used to deepen the explanation of the developmental principles and to integrate evidence beyond vocational interest research.

It is important to note that it is beyond the scope of the current dissertation to test all assumptions and hypotheses of the presented theories—it is also not expedient. The aim is

[4] Normative change and stability are intuitive constructs that are widely applied in vocational interest research (Low & Rounds, 2007). Principles that focus on these two aspects of development are informative and provide a basis for empirical investigations. In addition, gender differences in vocational interests are frequently investigated (Su et al., 2009), but relatively little is known about their development. Therefore, a specific focus was laid on the development of gender differences.

however, to describe key aspects of the respective theories and to integrate them under the umbrella of three developmental principles. The purpose of the principles is to guide assumptions about the development of indicators that capture stability and change in vocational interests. The empirical studies of the current dissertation were therefore not able to directly test these principles but can describe the development of indicators that capture stability and change in vocational interests. As the respective principles were specifically derived for the current dissertation to generate a framework where findings from different empirical studies could be integrated, future studies should examine the proposed developmental processes.

Cumulative Learning Principle

Holland's (1997) theory focuses on the work context and emphasizes that socialization processes in work environments are a major factor for vocational interest development. After selecting a work environment, people usually go through one of four processes: contextualization, reactive adjustment, active adjustment, or quitting (for an overview of these processes, see Stoll & Trautwein, 2017). These processes are driven by the congruence between the person's interests and the characteristics of the environment (Holland, 1997). When congruence is high (i.e., interests match to the requirements of the environment), interests that were responsible for the selection of the environment are reinforced and stabilize over time (contextualization; Stoll & Trautwein, 2017). When congruence is low (i.e., interests do not match to the requirements of the environment), people either change their interests (reactive adjustment), change the environment they are in (active adjustment), or quit participation in the environment (quitting), to establish a state of high congruence (Holland, 1997; Stoll et al., 2017).

However, the theoretical mechanisms that were proposed by Holland (1997) primarily focus on adulthood. It could be argued that during adolescence these processes are less pronounced, as adolescents are less independent and autonomous in their environmental choices in comparison to adults (see Spear & Kulbok, 2004). This may be especially true for influential contexts such as school, neighborhood, or family environments, which are usually allocated and not selected. The four processes that were described by Holland (1997) could therefore provide more explanatory power when focusing on explicit person-environment relationships, but less in understanding the general development of vocational interests over the course of adolescence.

Holland (1997) provides not much information about vocational interest development in younger age groups. According to his theory, children and adolescents possess unspecific

likes and dislikes for certain activities. Over time, when children and adolescents experience more activities, their likes and dislikes start to stabilize. This implies that vocational interests become more stable based on cumulative learning and increase in stability over the course of adolescence (Holland, 1997). The solidification of vocational interests is based on internal (e.g., personal satisfaction or enjoyment) and external (e.g., rewards from the environment or positive feedback) reinforcement processes, which transform vocational interests from single likes and dislikes into relatively stable dispositions. Holland (1997) also emphasizes the role of the environment. As experiences are usually situated in certain contexts, solidification is always a reciprocal process between the current expression of vocational interests and the environment. Holland (1997) also describes the role of biological factors in vocational interest development. For example, people with excellent physical characteristics in height, strength, or hand-eye coordination are more likely to choose certain activities that require these characteristics (e.g., in sports, construction, or landscaping).

Holland's (1997) proposition that interests develop based on the accumulation of experiences was integrated and refined by the TSID model (Su et al., 2019). According to the TSID model, stable interests are reflected by a mental representation of the object of interest. This mental representation can be described as a latent construct that corresponds to dispositional interests. New experiences that trigger situational interest help to refine and strengthen this mental representation. The solidification of interests is therefore driven by an amassment of positive experiences that are initiated through curious, thought-provoking, or surprising activities (Hidi & Renninger, 2006; Renninger & Hidi, 2011; Su et al., 2019). As different environments offer different experiences, the TSID model also illustrates the importance of the environment for interest solidification, which is in line with the person-environment interactions proposed by Holland (1997). Person-environment interactions can be understood as a reciprocal process that gradually strengthens (or weakens) interest in certain areas. Experiences within environments shape vocational interests, while vocational interests are also responsible for the choice of these environments (Su et al., 2019).

According to the TSID model, the mental representation that incorporates experiences from various situations can also be enriched by simply observing certain activities (Su et al., 2019). This assumption is in line with social learning theory (Bandura & Walters, 1963) and implies that not only direct but also vicarious experiences are assumed to foster interest development (Su et al., 2019). Vicarious experiences could be initiated by activities such as watching a documentary, having a conversation, or simply by thinking about a person doing an activity that is in line with someone's own interest. As soon as the vicarious experiences trigger

situational interest and the respective affective and cognitive responses (Su et al., 2019), it is possible that they initiate interest development.

Against the background of these theoretical assumptions, one characteristic of the development of vocational interests can be described as the *cumulative learning principle*. In line with Holland (1997) and the TSID model (Su et al., 2019), the principle states that vocational interests solidify over the course of adolescence based on an accumulation of experiences that are in line with someone's developing interests. Due to that process, adolescents constantly refine and enrich their mental representations of their objects of interest, including a broad range of affective (e.g., feeling enjoyment) and cognitive (e.g., positive ability-beliefs) appraisals (Su et al., 2019). Over time, as adolescents accumulate more experiences, they get a better sense about the activities they do and do not like. Vocational interests therefore evolve from incoherent preferences for single activities to sophisticated dispositions (Holland, 1997). The reengagement in activities as an initiator for interest solidification is underscored by many interest theories (Hidi & Renninger, 2006; Holland, 1997; Renninger & Hidi, 2011; Silvia, 2001). Based on the proposition of the cumulative learning principle, it is assumed that stability indicators of vocational interests (e.g., re-test correlations and profile stability) should increase over the course of adolescence.

Restriction Initiates Growth Principle

From childhood to adolescence, Gottfredson (1981) describes two overarching processes that are responsible for the development of occupational preferences—the phase of circumscription, which is followed by the phase of compromise. According to Gottfredson (1981), occupational preferences arise when there is a match between someone's self-concept (i.e., someone's view of oneself) and the perceived properties of an occupation. Self-concept consists of a variety of orientations that are gradually integrated over the course of adolescence (Gottfredson, 1981). During circumscription, the orientations toward size and power (ages 3 to 5), gender (ages 6 to 8), prestige (ages 9 to 13), and views about one's unique self (age 14 and older), such as abilities, values, and personality traits (Gottfredson, 1981), are integrated in the self-concept of children and adolescents. Simultaneously to the integration of these orientations, children and adolescents gradually rule out occupations that do not match to their developing self-concept. Based on that process, children and adolescents narrow down their choices for suitable occupations. Although circumscription seems crude at first, it is a lasting process that provides the basis for the phase of compromise (Gottfredson, 1981). After circumscription excluded unsuitable options (i.e., around the age of 14), adolescents recognize that many occupations that match to their self-concept are not reachable. This could be because some

professions might have too high skill requirements or are limited in their access due to the labor market. During the phase of compromise, adolescents are therefore urged to sacrifice occupational roles that seem more compatible to their self-concept for those that are observed as more accessible (Gottfredson, 1981).

Based on Gottfredson's (1981) theory, it is possible to derive assumptions about the development of interest intensity over the course of adolescence. During circumscription, adolescents gradually neglect occupations that seem unsuitable for their developing self-concept. Circumscription could be consequently characterized by decreasing interest intensity as adolescents are constantly confronted with the decision of which occupations they do not like. When the phase of compromise begins, adolescents have a restricted pool of suitable occupations, a better sense of the occupations that are suitable to them and a better sense about their self-concept (e.g., who they are, what they believe in and what their skills are). Adolescents are now able to pursuit and reengage in activities that correspond to suitable occupations, which could lead to an increase in interest intensity.

Tracey (2002) is more specific about the influence of ability self-concept on the development of vocational interests. He argues that from late childhood to early adolescence (ages 11 to 14), children and adolescents get a better sense about the activities they are good at and simultaneously improve in their knowledge about the world of work. In addition, they start to connect achievements in certain school subjects to the requirements of certain occupations (Tracey, 2002). Over time, when a more realistic assessment of their ability self-concept takes place, adolescents start to realize that not all occupations seem appropriate for them. This more realistic view of someone's own abilities could lead to decreases in interest intensity in some areas. Tracey (2002) states that during childhood, interests are assumed to be undifferentiated and uniformly high in a wide range of areas, whereas during adolescence, interests become more differentiated and consequently decrease in some areas (Tracey, 2002).

Besides the influence of self-concept, recent studies suggest (see Hoff et al., 2018; Päßler & Hell, 2020) that the disruption hypothesis—a developmental principle originating from personality research (Soto & Tackett, 2015)—should be applied to vocational interests. The pattern of personality trait development during adolescence is usually characterized by the maturation principle (Roberts & Wood, 2006). The maturation principle states that people "become more socially dominant, agreeable, conscientious, and emotionally stable with age", implying that mean levels in these personality traits do increase (Roberts & Wood, 2006), p. 19). However, recent studies suggest that early adolescence is characterized by a different

developmental pattern, a period of disruption (Soto et al., 2011; Soto & Tackett, 2015; Van den Akker et al., 2014). The disruption hypothesis proposes that the otherwise increasing personality traits conscientiousness, agreeableness, and openness experience temporary dips during the time period of early adolescence (Soto et al., 2011; Soto & Tackett, 2015; Van den Akker et al., 2014).

Based on meta-analytic findings, Hoff et al. (2018) suggested that a similar principle should be applied to the development of vocational interests. According to the disruption hypothesis, vocational interest intensity is assumed to decrease during early adolescence (ages 11 to 14) and increase afterwards from middle to late adolescence (ages 15 to 18). This proposition is in line with the assumptions that were derived from the processes of circumscription and compromise (Gottfredson, 1981) and the differentiation hypothesis of Tracey (2002). Hoff et al. (2018) argue very similar to Tracey (2002). During early adolescence, achievements in school subjects are connected to the ability requirements of certain career paths (Hoff et al., 2018). Simultaneously, school environments become more competitive and performance-oriented (Hoff et al., 2018). These processes lead to negative ability-related experiences in school environments that ultimately initiate decreases in vocational interest intensity (Hoff et al., 2018). However, Hoff et al. (2018) argue that disruption is only temporary and that adolescents are motivated by the challenges they face to adapt and reorganize. Vocational interest intensity therefore starts to increase again over the course of middle and late adolescence. Hoff et al. (2018) labelled that process as disruption breeds growth.

Against the background of these theoretical assumptions, another characteristic of the development of vocational interests can be described as the *restriction initiates growth principle*. In line with Gottfredson (1981), the term restriction was used, because adolescents are assumed to circumscribe their occupational choices, when they do not match to their gender, social status, abilities, values, and personality traits. The principle implies that vocational interests decrease in intensity from late childhood to early adolescence (ages 11 to 14) and increase in intensity from middle to late adolescence (ages 15 to 18). With increasing age, children and adolescents start to get a more nuanced self-concept, a better sense about what they are good at, and a more realistic view of the world of work (Gottfredson, 1981; Hoff et al., 2018; Tracey, 2002). When children and adolescents start to explore various activities and occupations, they realize that not all occupations match to their self-concept. Over time, they gradually neglect the activities and occupations that do not match to their self-concept. This process leads to a differentiation of vocational interests, which is characterized by overall decreases in interest intensity. Tracey (2002) and Hoff et al. (2018) imply that this process is

necessary for adolescents to ascertain which activities and occupations they do not like. After the period of restriction, the onset of growth begins. As adolescents have restricted their occupational choices (Gottfredson, 1981), they are now able to reengage in the activities they view as suitable and consequently further deepen their vocational interests. According to the propositions of the principle, mean levels of vocational interests should decrease from late childhood to early adolescence (ages 11 to 14) and increase from middle to late adolescence (ages 15 to 18).

Gender Differentiation Principle

Meta-analytic evidence indicates that there are robust gender-specific vocational interests (Su et al., 2009), with males being more interested in things (i.e., dimensions Realistic, Investigative and Conventional) and females being more interested in people (i.e., Artistic, Social and Enterprising; Su et al., 2009). Theories indicate that socialization effects could play an important role in the emergence of gender differences (Gottfredson, 1981). It is assumed that children and adolescents incorporate gender stereotypes through experiences in certain environments, which in turn may lead to a solidification of gendered vocational interests (Einarsdóttir & Rounds, 2020). Besides the influence of the environment, there is also evidence that prenatal sex hormone levels can influence gender differences in vocational interests (Beltz et al., 2011). Therefore, not only environmental but also biological factors may impact gender differences in vocational interests.

According to Gottfredson (1981), gender differences in vocational interests begin to manifest relatively early in life, usually during the age span of 6 to 8. Although children are already subtly aware of gender differences much earlier, during that time period they begin to consciously recognize that specific (stereotypical) occupations, behaviors and characteristics are attributed to each gender (Gottfredson, 1981). The understanding of the concept of gender roles becomes more nuanced over time. According to Gottfredson (1981), vocational interests depend on the match between one's self-concept and one's view of occupations. Gender differences in vocational interests should therefore occur during time periods where children become aware of their own sex roles and the ones that are attributed to certain occupations. According to Gottfredson (1981), this process starts already during childhood (ages 6 to 8), becomes more intensive during early adolescence (ages 11 to 14) and begins to stabilize during middle adolescence (around age 15). As gender differences are assumed to be stable for the remainder of adolescence, the period of childhood and early adolescence sets the foundation for gender differences in later life stages (Gottfredson, 1981).

The theory of Gottfredson (1981) implies that boys and girls gradually incorporate gender stereotypical experiences into their developing self-concept. These experiences are obtained early in life and have a long-lasting effect on vocational interests (Gottfredson, 1981). The TSID model describes how affective and cognitive components of such experiences are incorporated into vocational interests (Su et al., 2019). According to the model, an accumulation of new experiences could lead to increases (or decreases) in gender differences of vocational interests. For example, if a girl possesses equal trait levels in Investigative and Artistic interests and she constantly stays in environments that only reinforce and offer gender stereotypical activities, according to the TSID model, she will develop a relatively stable Artistic interest disposition. In addition, when the environments display Artistic activities as appropriate for females, according to Gottfredson (1981), she will incorporate that perception into her self-concept and further strengthen her preference for Artistic occupations. Over time, the girl would consequently develop a well-pronounced Artistic interest disposition, despite also having a dormant interest for Investigative activities.

Against the background of these theoretical assumptions, another characteristic of the development of vocational interests can be described as the *gender differentiation principle*. It states that gender differences in vocational interests are already present during late childhood (age 11), become more pronounced from late childhood to early adolescence (ages 11 to 14) and are relatively stable from middle to late adolescence (ages 15 to 18). Gender differences are assumed to develop based on two mechanisms. First, children and adolescents incorporate the concept of gender into their self-concept and start to neglect occupations that do not match to their gender (Gottfredson, 1981). This leads to girls expressing stereotypically female vocational interests such as Social and Artistic interests and boys expressing stereotypically male vocational interests such as Realistic and Investigative interests. Second, due to environmental influences, children and adolescents gradually incorporate direct and vicarious experiences about gender stereotypical activities. These gender stereotypical experiences are then stored into their developing mental representation of the object of interest (Su et al., 2019), leading to robust differences in dispositional interests. In line with the gender differentiation principle, gender differences on mean levels of vocational interests should be present during late childhood, increase from late childhood to middle adolescence, and be relatively stable from middle to late adolescence.

Table 1

A Summary of the Three Developmental Principles and their Assumptions

Principle	Assumption	Indicators
Cumulative Learning	Vocational interests solidify over the course of adolescence (ages 11 to 18).	Increases in stability indicators, such as re-test and profile correlations are expected.
Restriction Initiates Growth	Vocational interests decrease in intensity from late childhood to early adolescence (ages 11 to 14) and increase in intensity from middle to late adolescence (ages 15 to 18).	Decreases in mean levels that are followed by increases in mean levels are expected.
Gender Differentiation	Gender differences in vocational interests are already present during late childhood (age 11), become more pronounced from late childhood to early adolescence (ages 11 to 14), and are relatively stable from middle to late adolescence (ages 15 to 18).	Increases in gender differences on mean levels are expected.

Note. The elaboration of the theoretical processes behind the assumptions of the developmental principles can be found in Chapter 1.3.2.

1.3.3 Empirical Evidence on Vocational Interest Development

The current chapter summarizes empirical evidence on the development of vocational interests over the course of adolescence. The findings are structured based on the three developmental principles (i.e., cumulative learning, restriction initiates growth, and gender differentiation) that were proposed in the previous chapter. The aim of the chapter is to describe studies that investigated the development of stability, intensity, and gender differentiation. In addition, it is examined if stability, intensity, and gender differences develop in the ways proposed by the three developmental principles. Missing information on the development of vocational interests will be carved out as well. All the reviewed studies used indicators that

were explained at the beginning of Chapter 1.3, either focusing on the variable or profile perspective.

Cumulative Learning Principle

According to the cumulative learning principle, it is assumed that stability of vocational interests gradually increases over the course of adolescence. Low et al. (2005) conducted a meta-analysis about the stability of vocational interests over the life course, comprising retest correlations from 66 studies. Their results indicated that over the course of adolescence, stability slightly increased from early adolescence (ages 12 to 14; r = .51) and middle adolescence (ages 14 to 16; r = .53) to late adolescence (ages 16 to 18; r = .55). Although the evidence is in line with the cumulative learning principle, in comparison to other life phases, increases in stability over the course of adolescence are small (Low et al., 2005). According to Low et al. (2005), the biggest increase in stability can be found from late adolescence (ages 16 to 18; r = .55) to young adulthood (ages 18 to 22; r = .72), with stability peaking during young adulthood (ages 25 to 30; r = .77).

The meta-analysis from Low et al. (2005) provides important insights about the development of stability over the course of broad life phases. However, information about yearly stability changes during shorter developmental periods is still scarce, especially for younger age groups. Findings from Tracey (2002) suggest that increases in interest stability already begin over the course of late childhood and early adolescence (ages 11 to 14). He investigated the development of vocational interests based on two cohorts of students, which were followed from fifth to sixth (N = 126) and from seventh to eighth grade (N = 221), respectively. Re-test correlations of vocational interests were higher in the older cohort in comparison to the younger cohort, indicating that vocational interests became more stable with age (Tracey, 2002). However, so far, there is mixed evidence for the assumption that vocational interests already increase in stability from late childhood to early adolescence, as Päßler and Hell (2020) reported decreasing stabilities for some of the interest dimensions with age. Tracey (2002) and Päßler and Hell (2020) are so far the only studies that investigated the development of vocational interests during late childhood and early adolescence.

In contrast to findings from earlier life phases, there is robust evidence on yearly stability changes of vocational interests over the course of middle and late adolescence (e.g., Hoff et al., 2019; Hoff, Song, Einarsdóttir, et al., 2020; Xu & Tracey, 2016). Although increases are small, they support the general assumption that stability increases with age. Empirical studies using a profile perspective yield similar findings (Etzel & Nagy, 2021; Stoll, Rieger, et

al., 2020; Xu & Tracey, 2016; Zytowski, 1976). They are in line with the cumulative learning principle, as profile stability increases with age. So far, there is no empirical evidence about profile stability of vocational interests during the time period of late childhood and early adolescence.

Current empirical evidence provides support for the cumulative learning principle, as increases in vocational interest stability are frequently reported over the course of adolescence. Although these increases are smaller in comparison to later life phases (e.g., the transition from adolescence to adulthood; Low et al., 2005), they are consistent across studies. However, despite meta-analytic evidence (see Low et al., 2005), it is still unclear if vocational interests already begin to solidify over the course of late childhood and early adolescence—current evidence is so far inconclusive. In addition, previous studies had limitations in sample properties, as they either compared different cohorts (see Tracey, 2002) or experienced changes in the measurement context (see Päßler & Hell, 2020) over time. Furthermore, the findings of Päßler and Hell (2020) cannot be interpreted as age-related changes in stability because the sample in each wave comprised students from different grade levels. This illustrates that longitudinal studies are missing, where the same participants from homogenous grade levels are followed over multiple time points. Therefore, more empirical evidence is needed that investigates the stability of vocational interests over the course of late childhood and early adolescence. The first study of the current dissertation addresses this missing evidence, by investigating stability of vocational interests over the course of late childhood and early adolescence, based on a sample where participants were annually followed from 5th to 8th grade. This longitudinal design provides evidence on yearly stability changes. In addition, the second study of the current dissertation provides evidence on profile stability of vocational interests over the course of late childhood and early adolescence, focusing on the profile instead of the variable perspective.

Restriction Initiates Growth Principle

The restriction initiates growth principle states that interest intensity in vocational interests decreases over the course of late childhood and early adolescence and increases over the course of middle and late adolescence. The meta-analysis of Hoff et al. (2018) investigated the development of mean levels of vocational interests over the life course, comprising effect sizes from 49 studies. Vocational interests were operationalized in terms of the RIASEC framework. Over the course of adolescence (ages 11 to 18), the development of mean levels can be characterized by two different patterns (Hoff et al., 2018). During late childhood and early adolescence (ages 11 to 14), mean levels of vocational interests decreased in general.

Although decreases were significant for only two of the six RIASEC dimensions, five of them indicated decreases in terms of standardized effect sizes (-.02 < *d* < -.30). During middle and late adolescence, mean levels of vocational interests increased in general (ages 14 to 18; Hoff et al., 2018). Although the increases were significant for only two of the six dimensions, five of them indicated increases in terms of standardized effect sizes (.06 < *d* < .18). The findings of Hoff et al. (2018) suggest that the development of interest intensity is in line with the propositions that were derived from the restriction initiates growth principle.

Evidence on the development of vocational interests from middle to late adolescence (ages 14 to 18) seems quite robust, as Hoff et al. (2018) comprised data from 26 samples from these developmental periods. However, evidence during late childhood and early adolescence (ages 11 to 14) seems less robust, as Hoff et al. (2018) comprised data from only two studies from these developmental periods. The study of Päßler and Hell (2020) was not included in the meta-analysis of Hoff et al. (2018) and provides further empirical evidence for the development of vocational interests from late childhood to early adolescence. Päßler and Hell (2020) reported significant decreases in mean levels for the dimensions Realistic, Investigative, Artistic, and Social, over time. The results are in line with the assumptions from the restriction initiates growth principle and similar to the findings of Hoff et al. (2018). However, although the findings of Päßler and Hell (2020) provide important insights, their study possesses certain limitations that makes it difficult to interpret age-related mean-level changes. The study design of Päßler and Hell (2020) captured the transition from elementary to secondary school (i.e., after the first measurement wave, students were allocated to different school tracks) and was accompanied by a change in the context of measurement over time. At the first measurement wave, students filled out the questionnaire within the classroom with a teacher present, whereas in later waves, questionnaires were filled out at home. In addition, due to its multi-cohort design where each wave comprised students from different grade levels (grades T1 = 4th to 6th, T2 = 5th to 7th, T3 = 6th to 8th), some students experienced the transition from elementary to secondary school at different ages. The investigation of vocational interest development is consequently confounded by these properties.

In general, current empirical evidence supports the restriction initiates growth principle which states that interest intensity decreases from late childhood to early adolescence and increases from middle to late adolescence. This is illustrated by the meta-analysis from Hoff et al. (2018) that reported general decreases in mean levels from late childhood to early adolescence and general increases in mean levels from middle to late adolescence. However, despite the robust evidence on the development from middle to late adolescence, studies that

provide insights about the development from late childhood to early adolescence are scarce and possess a certain range of limitations in their study designs. Therefore, further longitudinal studies are needed that extend the works of Hoff et al. (2018) and Päßler and Hell (2020). This need is addressed by the first study of the current dissertation, which examines yearly mean-level changes of vocational interests from late childhood to early adolescence. As the measurement context stayed the same over time in this study, it is supposed to provide more robust evidence than Päßler and Hell (2020).

Gender Differentiation Principle

The gender differentiation principle states that gender differences are assumed to increase from late childhood to early adolescence and remain constant over the course of middle and late adolescence. Hoff et al. (2018) reported in their meta-analysis that gender differences in vocational interests are already present over the course of late childhood and early adolescence (ages 11 to 14). Gender differences were quite substantial, with boys being more interested in Realistic activities ($d = 0.15$) and girls in Social activities ($d = 0.52$). Unfortunately, Hoff et al. (2018) merely reported gender differences for Realistic and Social interests, but not the other interest dimensions. However, Tracey (2002) as well as Päßler and Hell (2020) investigated gender differences for all interest dimensions over the course of late childhood and early adolescence. According to Tracey (2002), boys scored higher on Realistic and Investigative interests, whereas girls scored higher on Artistic, Social, Enterprising and Conventional interests. These findings were generally in line with meta-analytic evidence on gender differences (Su et al., 2009). The results reported by Päßler and Hell (2020) were similar, except for Investigative interests, were girls scored higher values than boys.

Besides robust evidence on gender differences on mean levels, there is mixed evidence on the assumption that they increase over the course of late childhood and early adolescence. One the one hand, Hoff et al. (2018) reported that gender differences increased in Social and Realistic interest over the course of late childhood and early adolescence (ages 11 to 14). On the other hand, Päßler and Hell (2020) reported no significant increases in gender differences on vocational interests over the course of late childhood and early adolescence. Tracey (2002) did not investigate the development of gender differences over time. According to Hoff et al. (2018), gender difference remained relatively stable in later life periods, such as middle and late adolescence. The findings of the meta-analysis of Su et al. (2009) provide similar evidence. In their meta-analysis they also investigated age-related effects, by examining if the average age of participants (age 13 to 43) in a sample was related to gender differences in vocational interests. Although they reported slight changes in gender differences with increasing average

age, these changes were rather small and mostly limited to the interest dimension Social. In addition, the general moderator effects of age on gender differences were not significant. This suggests that that gender differences on vocational interests remain relatively stable from early adolescence to middle adulthood.

In line with the gender differentiation principle, there is robust evidence (see Hoff et al., 2018; Päßler & Hell, 2020; Tracey, 2002) that gender differences are already established during late childhood and early adolescence. In addition, evidence is quite robust that gender differences remain relatively stable from middle to late adolescence (see Hoff et al., 2018; Su et al., 2009). However, there is mixed evidence on the assumption that gender differences on vocational interests increase over the course of early adolescence. Some studies provide evidence for (see Hoff et al., 2018) and some against this assumption (see Päßler & Hell, 2020). Furthermore, some studies did investigate gender differences on a limited number of interest dimensions (see Hoff et al., 2018). Therefore, to get a clearer picture about the development of gender differences over the course of late childhood and early adolescence, further studies are needed that investigate the assumptions of the gender differentiation principle. This research gap is addressed by the first study of the current dissertation because it investigates gender differences and their development over the course of late childhood and early adolescence on all of Holland's (1997) interest dimensions.

1.4 Factors that Influence the Development of Vocational Interests

Chapter 1.3 introduced three overarching principles (i.e., cumulative learning principle, restriction initiates growth principle, and gender differentiation principle) that describe the development of interest stability, intensity, and gender differentiation over the course of adolescence. The principles proposed developmental trajectories of vocational interests that should occur for most adolescents. Chapter 1.4 describes mechanisms that could influence the development of vocational interests. They may cause individuals to partially deviate from these trajectories. The current chapter will thereby focus on the role of experiences. Many interest theories emphasize the importance of experiences as an initiator for interest development (Gottfredson, 1981; Hidi & Renninger, 2006; Holland, 1997; Renninger & Hidi, 2011; Su et al., 2019). For example, according to the TISD model, positive experiences lead to an accumulation of situational interest that could cause the formation of dispositional interests in a certain area (Su et al., 2019). In addition, more experiences lead to adolescents being more confident about activities they like and dislike, which could ultimately lead to more stable interests (Holland, 1997; Su et al., 2019). Therefore, the following chapter focuses on factors that could initiate differences in experiences. On the one hand, it focuses on individual factors (i.e., characteristics, such as personality traits or cognitive abilities) because they are usually responsible for the choice of activities that someone engages in and the way someone processes and incorporates experiences from these activities. On the other hand, it focuses on contextual factors (i.e., the environment) because they are primarily responsible for the range of activities that are provided to someone.

1.4.1 Individual Factors

Theories about vocational interests state that they are embedded in the broad framework of personality (Holland, 1997; Su et al., 2019). For example, Su et al. (2019) explain that interests are located within a nomological net that also includes relations with personality traits and cognitive abilities. In line with that assumption, many personality models integrate three relatively stable dispositional characteristics: cognitive abilities, personality traits and interests (Ackerman, 1996; Ackerman & Heggestad, 1997; Armstrong et al., 2008; Roberts & Wood, 2006; Schmidt, 2014). Apart from these integrative theories, there is also meta-analytic evidence that empirically links vocational interests with personality traits and cognitive abilities (Ackerman & Heggestad, 1997; Barrick et al., 2003; Larson et al., 2002; Mount et al., 2005). Furthermore, there is empirical evidence that vocational interests co-develop with personality

traits over the course of late adolescence (Hoff, Song, Einarsdóttir, et al., 2020), illustrating that both constructs might also be related longitudinally not only cross-sectionally. Therefore, individual factors that could influence the development of vocational interests are personality traits and cognitive abilities.

Empirical Evidence Based on the Variable Perspective

Most studies that investigated the relationship between vocational interests, personality traits and cognitive abilities took a variable perspective. Mount et al. (2005) conducted a meta-analysis about the interrelations between vocational interests and personality traits, illustrating that there are clear pairings between them. Scales of personality traits—measured based on the Big Five framework (McCrea & Costa, 1999)—were correlated to scales of vocational interests—operationalized based on Holland's (1997) interest model. Correlations were the highest between Investigative and Openness (r = .25), Artistic and Openness (r = .41), Social and Extraversion (r = .29), Enterprising and Extraversion (r = .40), as well as Conventional and Conscientiousness (r = .19). Correlations were smaller between Social and Openness (r = .13) as well as Social and Agreeableness (r = .17). The relationships between the remaining personality traits and interest dimensions were negligible. Similar findings were reported by previous meta-analyses (Barrick et al., 2003; Larson et al., 2002).

Building on meta-analytic cross-sectional findings (Barrick et al., 2003; L. M. Larson et al., 2002; Mount et al., 2005), Hoff, Song, Einarsdóttir, et al. (2020) investigated the co-development of personality traits and vocational interests. Based on assumptions of the TESSERA (Triggering situations, Expectancy, States/State expressions, and Reactions; Wrzus & Roberts, 2017) model of personality development, they argued that changes in personality traits are associated with changes in vocational interests. The TESSERA model proposes—similar to the TSID model (Su et al., 2019)—that having experiences are responsible for the development of personality. Shared situational content between personality traits and vocational interests could explain the co-development of both constructs (Hoff, Song, Einarsdóttir, et al., 2020). For example, people with strong Enterprising interests usually prefer to engage in Enterprising activities, such as leading other people, selling things to others, or persuading others of their ideas (Hoff, Song, Einarsdóttir, et al., 2020). These activities require behaviors, such as being talkative, energetic, and assertive, which are characteristics that are assigned to the personality trait Extraversion. The engagement in Enterprising activities consequently requires not only Enterprising interests, but also that people behave in an extraverted manner (Hoff, Song, Einarsdóttir, et al., 2020). This implies that when people continually seek out situations that entail Enterprising activities, they develop not only

Enterprising interests but also behaviors related to the personality trait Extraversion (Hoff, Song, Einarsdóttir, et al., 2020). Hoff, Song, Einarsdóttir, et al. (2020) reported that changes in Openness were highly correlated (r = .88) with changes in Artistic interests, changes in Extraversion were moderately correlated (r = .60) with changes in Enterprising interests and changes in Conscientiousness were moderately correlated (r = .35) with changes in Conventional interests.

Like research about personality traits, most studies that investigated the relationship between vocational interests and cognitive abilities took a variable perspective. Ackerman and Heggestad (1997) integrated scales of vocational interests, personality traits, and cognitive abilities, based on meta-analytic evidence and an extensive review of empirical studies, into four trait complexes: a clerical/conventional trait complex, a science/math trait complex, an intellectual/cultural trait complex and a social trait complex. The clerical/conventional trait complex includes Conventional interests that are assumed to be related to perceptual speed ability (i.e., the ability to correctly compare pictures, numbers, or letters). It is assumed that the engagement in Conventional activities usually entails structured administrative tasks, which require and foster perceptual speed ability (Ackerman & Heggestad, 1997). The science/math trait complex encompasses Realistic and Investigative interests and includes math reasoning abilities as well as spatial abilities. The science/math trait complex partly overlaps with the intellectual/cultural trait complex, which entails Investigative and Artistic interests, as well as a high degree of crystallized intelligence. In comparison to the other trait complexes, the social trait complex is not related to cognitive abilities and includes the interest dimensions Enterprising and Social.

Empirical Evidence Based on the Profile Perspective

Most integrative models (Ackerman & Heggestad, 1997; Armstrong et al., 2008; Roberts & Wood, 2006; Schmidt, 2014) describe the relationships between vocational interests, personality traits and cognitive abilities on the variable level. However, there is the assumption that personality characteristics can also influence vocational interests on the profile level (Swanson, 1999). Vock et al. (2013) compared vocational interest profiles of gifted and highly achieving students to those students with lower scores on an intelligence measure. Their results suggest that profiles of gifted and highly achieving students possessed more pronounced interests on the dimensions Realistic and Investigative, compared to those students with lower scores on an intelligence measure. Similarly, Perera and McIlveen (2018) investigated the relationship between profile constellations of vocational interests and the expression of personality traits. They reported that specific profile constellations were correlated with certain

personality traits. For example, participants with higher values on the personality trait openness to experience, had vocational interest profiles that were dominated by strong Investigative interests (Perera & McIlveen, 2018). These findings illustrate that personality traits and cognitive abilities can influence vocational interests on the profile level.

Most studies that focus on vocational interest profiles investigate the impact of personality traits and cognitive abilities on the intraindividual expression of vocational likes and dislikes (Etzel & Nagy, 2021; Perera & McIlveen, 2018; Rounds et al., 1987; Xu & Li, 2020). However, another crucial characteristic of vocational interest profiles is their stability (Savickas & Taber, 2006). Profile stability implies that the expression of the respective profile (i.e., the mean-level scores of the interest dimensions within the profile) is relatively stable over longer periods of time. Swanson (1999) as well as Schomburg and Tokar (2003) assume that that differences in personality are responsible for differences in vocational interest profile stability. Personality characteristics influence the experiences that we have. Besides predicting achievements in different life situations (Cheng & Furnham, 2012; Roberts et al., 2007; Schmitt, 2014; Spengler et al., 2018; Stoll et al., 2017), cognitive abilities and personality traits characterize a wide range of people's behaviors (e.g., Ackerman & Heggestad, 1997; Armstrong et al., 2008; Kandler et al., 2014; Roberts & Wood, 2006), as well as influence their (e.g., educational or occupational) life choices (Päßler & Hell, 2012; Usslepp et al., 2020) and the selection of different environments (Jackson et al., 2020; Päßler & Hell, 2012). As the solidification of vocational interests is driven by the accumulation of experiences (Holland, 1997; Su et al., 2019), differences in personality characteristics could explain variations in the degree of vocational interest stability.

Although there is currently no empirical evidence on the association between personality traits and cognitive abilities and vocational interest profile stability, Schomburg and Tokar (2003) and Hirschi (2010) associated stability of vocational interests to other constructs. Schomburg and Tokar (2003) examined the influence of private self-consciousness on vocational interest stability over a 12-week time interval in a small sample of university students (N = 108). Private self-consciousness is described as "a regular tendency to be attentive to or aware of private aspects of the self, such as inner feelings, thoughts, and motives" (Schomburg & Tokar, 2003, p. 369). They assumed that people with higher private self-consciousness possess higher profile stability and rank-order consistency. They computed a stability indicator for each of the RIASEC dimensions and an overall profile stability indicator. Private self-consciousness had only small effects on the stability of Enterprising interests and did not have an effect on the remaining interest dimensions. In addition, general profile stability was not

associated to private self-consciousness. In their discussion section, Schomburg and Tokar (2003) highlighted that personality traits should be investigated as predictors for vocational interest profile stability.

Hirschi (2010) investigated the association between profile stability and profile characteristics, such as profile differentiation and profile congruence, as well as the association between profile stability and career maturation. The sample consisted of $N = 292$ students from eighth grade, who attended school in the German-speaking part of Switzerland. Hirschi (2010) assumed that more differentiated and more congruent profiles were more stable, referring to Holland (1997) who stated that people with the respective profile characteristics possess more consistent vocational behaviors and hence a more stable career development. In addition, Hirschi (2010) proposed that higher scores of career maturation were related to higher profile stability, as students were more sure about their occupational choices and hence more consistent in their likes and dislikes (Holland, 1997). According to the results of Hirschi (2010), the profile characteristics of differentiation and congruence had an influence on profile stability. More differentiated and more congruent profiles predicted less change in profiles and consequently higher profile stability. However, career maturation had merely small to negligible effects on profile stability.

Besides existing empirical evidence and the considerations about personality traits and cognitive abilities, gender could also influence profile stability of vocational interests. For example, there is robust evidence for differences in mean levels between males and females (Su et al., 2009). This implies that gender differences could occur not only in interest intensity, but also in interest stability. Although there is meta-analytic evidence that suggests that stability is not influenced by gender (Low et al., 2005), recent studies provide evidence that suggests otherwise (e.g., Stoll, Rieger, et al., 2020; Xu & Tracey, 2016). Xu and Tracey (2016) as well as Stoll, Rieger, et al. (2020) report gender differences in profile stability coefficients, indicating more stable profiles for females than males. Although theses gender differences are small, they are constant across studies. In addition, these effects are found not only in adolescence (e.g., Xu & Tracey, 2016) but also in adulthood (e.g., Stoll, Rieger, et al., 2020).

1.4.2 Environmental Factors

Theories about vocational interests state that experiences and repetitive engagement in activities can initiate interest development (Holland, 1997; Su et al., 2019). As different environments (e.g., school, leisure, or study-major environments) provide different possibilities

for children and adolescents to engage in a certain range of activities (Su et al., 2019), changes in environments could lead to changes in the development of vocational interests. These person-environment interactions are emphasized by many interest theories (Gottfredson, 1981; Holland, 1997; Su et al., 2019), including Holland (1997), who proposed an adjustment of vocational interests through socialization effects after entering a new environment.

Studies that investigated socialization effects indicate that vocational interests can change in interaction with the environment (e.g., Etzel & Nagy, 2021; Meir & Navon, 1992; Schultz et al., 2017). Schultz et al. (2017) examined the influence of work environments on the development of vocational interests across a 20-year time span during adulthood. They reported that women who were employed in Realistic (d = 0.42), Artistic (d = 0.29) and Conventional (d = 0.31) occupations, increased in their interest intensity in the respective areas, compared to women who were not employed in these occupations. Meir and Navon (1992) provide similar evidence. Participants of their study experienced increases in Conventional interests after becoming a bank teller, an occupation that is dominated by Conventional activities (O*NET, 2021). The findings of Schultz et al. (2017) and Meir and Navon (1992) illustrate that interest intensity in a specific area can adjust after entering a content-related environment. However, the findings of Schultz et al. (2017) suggest that such effects need to evolve over a relatively long time (i.e., 20 years). Besides the impact on single interest dimensions, Etzel and Nagy (2021) show that vocational environments also influence vocational interest profiles. After entering vocational training, participants increased in person-environment congruence, indicating that vocational interest profiles adjusted to the characteristics of the vocational environment (Etzel & Nagy, 2021).

Current studies about socialization effects mainly focused on the time period of adulthood (e.g., Etzel & Nagy, 2021; Meir & Navon, 1992; Schultz et al., 2017). However, it could be assumed that environmental influences are more impactful during life phases where interests are less stable (Low et al., 2005)—for example, over the course of adolescence. Golle et al. (2019) compared the development of vocational interests between students that chose an academic track and students that chose vocational training. The students were followed over a six-year period, from late adolescence to young adulthood. Golle et al. (2019) found moderate differences on the dimensions Investigative (d = -0.22), Social (d = -0.19), and Enterprising (d = -0.41), indicating that students who attended vocational training developed less interest in these areas. Although students were followed over a period of only six years, effect sizes were similar in magnitude to those reported by Schultz et al. (2017), who followed participants over

a period of 20 years. This indicates that influences of the environment could be more impactful in younger age groups.

According to the TSID model (Su et al., 2019), it could be assumed that differences in vocational interests arise because environments differ in the activities they provide. For example, Golle et al. (2019) reported that students who attended the academic track had more pronounced Investigative interests in comparison to students who attended the vocational training. It could be argued that Investigative interests evolve, because the academic track provided a more enriching environment in terms of Investigative activities. Academic school tracks usually prepare students for higher education (i.e., university, specialized college, university of applied science, university of cooperative education) and include more investigative elements compared to vocational training. A similar argumentation could also apply to the results reported by Meir and Navon (1992) and Schultz et al. (2017).

It is important to note that current studies merely focused on the influence of school (Golle et al., 2019) and work environments (Etzel & Nagy, 2021; Meir & Navon, 1992; Schultz et al., 2017). However, especially during younger life phases, such as adolescence, other environments seem quite impactful too. Larson (2000) investigated the duration of time adolescents spent in certain environments, based on an experience sampling study. According to his results, besides school, adolescents spent between 30 to 50% of their waking time in structured (e.g., extracurricular courses) or unstructured (e.g., hobbies) leisure settings. Leisure activities could have beneficial properties for interest development, as they are associated with positive emotions, are voluntary in nature, and possess intrinsic value (Hong, Milgram, et al., 1993). These properties can initiate short-term situational interest in a specific topic (Eccles et al., 1998; Hidi & Renninger, 2006; Ryan & Deci, 2000). Long-term engagement in leisure activities could therefore foster dispositional interests. However, so far, most of the evidence on the relationship between vocational interests and leisure activities is based on cross-sectional studies (e.g., Leuty et al., 2016; Miller, 1991). More evidence is needed that longitudinally investigates the influence of leisure activities on the development of vocational interests.

1.5 Methodological Challenges in Vocational Interest Research

There are methodological challenges that are inherent to the investigation of the development of dispositional constructs, such as vocational interests. Therefore, with respect to the current dissertation, certain methodological particularities must be considered. All the following methodological particularities are in some part an issue in the three empirical studies of the current dissertation. They are addressed in detail in the analysis sections of Chapters 3 to 5.

As vocational interests are assumed to only change slightly over longer periods of time (Hoff et al., 2018; Holland, 1997; Low et al., 2005; Rounds & Su, 2014), studies that investigate their development need multiple measurement waves that span multiple years. Only then may noteworthy changes in vocational interests be detected. Although studies that investigate a one-year stability of vocational interests might be insightful during transitional periods, such as after finishing high school and during the beginning of college, they provide less information about developmental mechanisms that need to unfold over multiple years. In addition, as only minor changes can be expected of vocational interests over multiple years (see effect sizes in Hoff et al., 2018), studies not only need to have multiple measurement waves that capture multiple years, but they also need an adequate statistical power to provide adequate significance tests for the small effect sizes (Cohen, 1988). To achieve an adequate power for small effect sizes, usually large sample sizes are needed. This implies that longitudinal studies that investigate the development of vocational interests need to follow many participants over multiple years to provide accurate significance tests that detect even changes with small effect sizes. The current dissertation focuses on data sets from education research, which usually include large samples of students that were drawn from classes or schools. The handling of further methodological features, such as the multi-level structure of that type of data (McNeish et al., 2017), is thoroughly described in the respective studies.

According to the TSID model (Su et al., 2019), vocational interests could consist of components that are susceptible to situations. To investigate this, latent state-trait models are used (Bishop et al., 2015; Geiser et al., 2015; Steyer et al., 2015). Latent state-trait models partition variance components of items and determine the amount of variance that is due to a stable component and due to a situation-specific component (Bishop et al., 2015; Geiser et al., 2015; Steyer et al., 2015). Based on such an analysis, it can be investigated if the underlying construct is better characterized by a stable trait component or by a short-term changing component (Bishop et al., 2015; Geiser et al., 2015; Steyer et al., 2015). However, there are

different forms of latent state-trait models. For the current dissertation, the generalized second order growth model is utilized (see Bishop et al., 2015). The model specifies a long-term developing trait component, in addition to the situational susceptible components (Bishop et al., 2015). This specification is a good fit for vocational interest research, as theories such as the TSID model assume that they are dispositions that develop over time (Su et al., 2019). Geiser et al. (2015) emphasized that measurement invariance is an important prerequisite for the specification of generalized second order growth models. Models that fail to achieve measurement invariance may result in biased estimations of variance components (Geiser et al., 2015). As vocational interests possess robust and large gender differences, measurement invariance across gender might be an issue. In addition, the existing vocational interest inventories for children and adolescents (e.g., Tracey & Caulum, 2015; Tracey & Ward, 1998) were not investigated for measurement invariance over time as well as gender. This is surprising, as especially during younger age periods, the understanding of the construct might change over time. Therefore, to adequately estimate dispositional and situational susceptible components of vocational interests, interest inventories in younger age groups must be tested for measurement invariance.

Inherent to the theory of Holland (1997) is the profile perspective on vocational interests. This requires that comprehensive investigations about the development of vocational interests focus not only on the variable but also on the profile level. There are approaches that emphasize vocational interest profiles of individuals can be estimated with model-based profiles (Gurtman & Balakrishnan, 1998; Nagy et al., 2009). A popular way to estimate model-based profiles is the structural summary method (Gurtman & Balakrishnan, 1998), which estimates profiles of individuals with a cosine fit function. Based on that statistical analysis, profile characteristics such as the elevation (i.e., average interest intensity of the profile elements), differentiation (i.e., the spread of the profile elements) or orientation (i.e., the dominant interest dimensions) of a profile can be summarized (Etzel et al., 2019). A fit statistic that indicates the degree of variance that was explained by the cosine fit function is used to judge the fit of the model profile to the individual profile (Etzel et al., 2019; Gurtman & Balakrishnan, 1998). As the structural summary method enables researchers to summarize key profile characteristics with three statistics, it is an elegant approach that is becoming increasingly popular in vocational interest research. However, despite its advantages, the approach requires that the model-based profiles fit to the data of the individual profiles. If model-based profiles do not fit to individual profiles, statements about the estimated profile characteristics could be biased. In addition, the structural summary method requires that

vocational interests adhere to the circumplex structure (Etzel et al., 2019). As this structure is assumed to evolve during middle adolescence (Tracey, 2002), the estimation of model-based interest profiles might not be suitable in younger age groups. Although current research emphasizes the advantages of this approach (see Etzel et al., 2019), in the current dissertation descriptive individual profiles are used instead of model-based profiles.

It is assumed that vocational interest can be altered through the repetitive engagement in activities (Su et al., 2019). However, the investigation of that assumption comes with inherent methodological challenges. Interests are both an important predictor for the selection of activities (Holland, 1997; Päßler & Hell, 2012; Wille et al., 2020) and an outcome of engagement in activities (Su et al., 2019). This circumstance is comparable to the distinction of selection and socialization effects in other research areas (e.g., Monahan et al., 2011). Specific study designs and analysis methods are needed that can distinguish the effects of activity engagement from possible selection effects, especially when experimental data are not available (Monahan et al., 2011). To investigate the causal effect of activity engagement on vocational interest development certain properties need to apply. Study designs should include longitudinal samples that measure interests before and after the engagement in certain activities. This temporal order is important, as analysis methods must control for pretest of vocational interests before activity engagement (Foster, 2010; VanderWeele et al., 2020). In addition, analysis methods need to be applied that account for possible confounders. Confounders are variables that influence both activity engagement and vocational interests (VanderWeele et al., 2020). Only when confounder control and the appropriate longitudinal design are applied, causal effects of activity engagement on vocational interests can be estimated (VanderWeele et al., 2020). Therefore, the estimation of causal effects of activity engagement on vocational interest development requires large longitudinal samples that collect a variety of confounder variables.

2 AIMS AND RESEARCH QUESTIONS

The aim of the current dissertation is to do a comprehensive investigation of the development of vocational interests over the course of adolescence. Based on the three developmental principles, missing empirical evidence on interest stability (i.e., re-test correlations and profile correlations), intensity (i.e., mean-level changes), and gender differentiation (i.e., gender differences in mean-level changes) is collected and integrated. Besides these overall aims, the current dissertation specifically focuses on three issues in vocational interest research. These issues comprise a lack of descriptive information about the development of vocational interests during late childhood and early adolescence, open questions on the association between individual characteristics and the stability of vocational interest profiles and a lack of studies about the influence of leisure environments on the development of vocational interests. The following studies correspond to the three issues.

The first study focuses on the development of vocational interests over the course of late childhood and early adolescence. According to the three developmental principles it is assumed that interest intensity decreases, stability increases, and gender differences increase over the course of late childhood and early adolescence. These assumptions were derived from theoretical works of Gottfredson (1981), Holland (1997), and Su et al. (2019) as well as empirical studies from Tracey (2002), Päßler and Hell (2020), Hoff et al. (2018), and Low et al. (2005). Because multiwave longitudinal data over the course of late childhood and early adolescence—where the same participants from the same grade levels are followed over time—are not available so far, the first study (Chapter 3) is based on a large-scale multiwave longitudinal sample where participants were annually followed from fifth to eighth grade (mean ages 11 to 14). Empirical evidence is collected on mean-level development, stability, and gender differences, as these indicators are in correspondence to the assumptions of the three developmental principles. The second part of the study targets the examination of the assumed dispositional nature of vocational interests. According to the TSID model (Su et al., 2019), vocational interests are dispositional constructs that also consist of situational components. This is in line with assumptions from Braun et al. (2020) and Hertzog and Nesselroade (1987) that most psychological constructs consist of both dispositional and situational components. For that purpose, vocational interests were investigated based on models from latent state-trait theory (Bishop et al., 2015; Geiser et al., 2015; Geiser & Lockhart, 2012; Steyer et al., 1999, 2015).

The second study focuses on the association between individual characteristics and the stability of vocational interest profiles. Personality traits and cognitive abilities were selected

as relevant constructs. This selection was justified based both on the assumption that vocational interests are embedded in personality (Roberts & Wood, 2006; Su et al., 2019) and on (meta-analytic) empirical evidence, which suggests that vocational interests are related to personality traits (Barrick et al., 2003; Hoff, Song, Einarsdóttir, et al., 2020; Larson et al., 2002; Mount et al., 2005) and cognitive abilities (Ackerman, 1996, 1997; Ackerman & Beier, 2003; Ackerman & Heggestad, 1997). Because recent studies suggest that there are differences in profile stability between males and females (e.g., Stoll, Rieger, et al., 2020), gender was additionally examined in the study. The second study (Chapter 4) consequently investigated the association of personality traits, cognitive abilities, and gender with vocational interest profile stability—making it the first study that investigates the relationship between individual factors and profile stability. Because profile stability is important during phases where career decisions occur, the second study used different samples that comprised transitional periods. Profile stability was investigated from late childhood to early adolescence (ages 11 to 14), when career choices are not imminent; during middle adolescence (ages 14 to 15), when adolescents can opt out of school for vocational training; from late adolescence to young adulthood (ages 17 to 23), when adolescents usually have to choose their study major or plan for further educational arrangements; and over the course of young adulthood (ages 22 to 34), when interests are assumed to be relatively stable and students finish university to start their professional careers.

The third study focuses on the influence of experiences on the development of vocational interests. Based on Su et al. (2019), it is assumed that different environments provide different possibilities for children and adolescents to experience and reengage in a certain range of activities. Therefore, differences in the activities that are experienced are assumed to lead to changes in the development of vocational interests. However, so far, relatively little is known about environmental influences on vocational interests, especially during the time period of adolescence. Consequently, the third study (Chapter 5) investigates the influence of leisure-related activities on the development of vocational interests from middle to late adolescence. Over the course of adolescence, leisure-related activities are assumed to be overarching and influential because adolescents spend a great amount of their waking time within structured or unstructured leisure-related settings (Larson, 2002). In addition, leisure environments could provide different experiences to adolescents and foster a differentiated development of vocational interests, as adolescents are more autonomous in choosing these environments compared to other contexts such as school. The third study therefore investigated in which way the engagement in unstructured leisure science activities influenced a broad range of outcomes,

such as competencies, self-concepts, grades, vocational aspirations, and vocational interests. For the current dissertation, only the results that affect vocational interests are integrated.

3 STUDY 1

THE DEVELOPMENT OF VOCATIONAL INTERESTS IN EARLY ADOLESCENCE: STABILITY, CHANGE, AND STATE-TRAIT COMPONENTS

Gfrörer, T., Stoll, G., Rieger, S., Trautwein, U., & Nagengast, B. (2021). The Development of Vocational Interests in Early Adolescence: Stability, Change, and State-Trait Components. *European Journal of Personality*, 089020702110356. https://doi.org/10.1177/08902070211035630

Abstract

Vocational interests predict major life outcomes such as job performance, college major choice, and life goals. It is therefore important to gain a better understanding of their development during the crucial years of late childhood and early adolescence, when trait-like interests are starting to develop. The present study investigated the development of vocational interests in a longitudinal sample, comprising $N = 3{,}876$ participants—assessed at four time points from ages 11 to 14. Stability, state-trait variance components, mean-level development, and gender differences in mean-levels of Realistic, Investigative, Artistic, Social, Enterprising, and Conventional (RIASEC) dimensions were examined. Stabilities were moderate for all dimensions, but Realistic, Investigative, Social, and Conventional interests became more stable over time. For Realistic, Artistic, Social, and Conventional interests, the trait variance increased over time. At age 14, all dimensions had substantial trait variance components. The mean-levels of Realistic, Investigative, Artistic, and Conventional interests decreased over the 3 years ($-0.44 < d < -0.24$). Initial gender differences—with girls having higher Artistic and Social interests and boys having higher Realistic and Investigative interests—increased over time. By investigating the development of vocational interests in late childhood and early adolescence, we complement previous findings and provide first insights about state-trait proportions in early adolescence.

Keywords: vocational interests; development; multiwave longitudinal study; late childhood & early adolescence

The Development of Vocational Interests in Early Adolescence: Stability, Change, and State-Trait Components

Vocational interests shape people's lives in many different ways. They are important predictors of work- and achievement-related outcomes, such as job performance (Nye et al., 2012, 2017; Van Iddekinge et al., 2011), gross income (Stoll et al., 2017), and academic performance (Rounds & Su, 2014; Su, 2012). In addition, they predict life decisions with long-lasting consequences such as the choice to attend a higher educational school track (Usslepp et al., 2020), choice of college major (Päßler & Hell, 2012; Wille et al., 2020), and the persistence of these choices (Allen & Robbins, 2008), major life goals (Stoll, Einarsdóttir, et al., 2020), and even life outcomes such as getting married and having children (Stoll et al., 2017).

Early adolescence is supposed to be particularly relevant for vocational interest development because, on the one hand, it is assumed that vocational interests are not yet fully developed at that time (Gottfredson, 1981; Holland, 1997; Low et al., 2005), and on the other hand, adolescents are encouraged by their parents (Kracke, 1997; Whiston & Keller, 2004) as well as their schools (Gysbers, 2005; Noack et al., 2010) to explore occupational opportunities and various career paths (Gati et al., 2019). In addition, late childhood and early adolescence is generally regarded as a crucial period for an individual's development and maturation (Petersen, 1987)—indicating that this life phase constitutes a transitional phase in young people's lives. However, to date, only a few studies (see Päßler & Hell, 2020; Tracey, 2002) have investigated stability and change in vocational interests during these crucial years.

The present study addresses this gap in research by providing detailed descriptive information about the longitudinal development of vocational interests over the course of late childhood and early adolescence. We used data from a large longitudinal study (N = 3,876, 136 classes) in which students in Germany were followed across four time points, at intervals of 1 year each, from fifth to eighth grade (ages 11 to 14). To provide a detailed picture of various aspects of development, we investigated three complementary indicators. Besides approaches that are commonly used, such as mean-level change and stability, as well as gender differences in mean-level change, we also investigated proportions of state and trait variance components, to gain further insights about the nature of vocational interests.

Indicators of Development in Vocational Interests

The development of vocational interests is usually investigated with various indicators that capture different perspectives on continuity and change—with mean-level changes and

retest correlations constituting the two most common methods (Low & Rounds, 2007). Overall shifts in mean levels represent changes in interest intensity averaged across an entire sample. Mean-level changes are often described as normative change, indicating "generalizable patterns of personality development that apply to most people" (Roberts et al., 2006, p. 1). Retest correlations are frequently used as indicators of the stability of constructs, capturing the relative rank order between individuals over time (Low & Rounds, 2007).

Additional information about the stability of and change in a construct can be obtained by disentangling its stable, long-term changing and short-term fluctuating components through the use of state and trait variance components (Geiser et al., 2015). This approach allows researchers to investigate whether a construct is better characterized as exhibiting long-lasting irreversible trait development or short-term reversible fluctuations (Bishop et al., 2015; Geiser et al., 2015). As recent theories described that vocational interests may consist of both, state and trait components (Su et al., 2019), this approach could facilitate the understanding of how large the respective state and trait components are and how they change over time. Disentangling state and trait variance components is especially promising during late childhood and early adolescence, when major biological (e.g., puberty) and social transitions (e.g., changes in requirements in school life and changes in peer constellations) are assumed to initiate long-lasting, irreversible trait development (Soto & Tackett, 2015). Separating proportions of state and trait variance has already been used in other research areas (e.g., Braun et al., 2020; Schmukle & Egloff, 2005), however, we are not aware of any study that has to date investigated state and trait variance components in vocational interests.

Theoretical Assumptions and Empirical Findings on the Development of Vocational Interests in Late Childhood and Early Adolescence

Longitudinal research on the development of vocational interests in childhood and adolescence is still scarce to date. Even though the existing meta-analyses on stability (Low et al., 2005) and mean-level change (Hoff et al., 2018) both comprise the time period of late childhood and early adolescence as the youngest age group, their aggregated coefficients are based on only five (age group 12 to 14 for Low et al.; 2005) and two (age group 11 to 14 for Hoff et al.; 2018) primary studies. Of these studies (Knapp & Knapp, 1984; Kuder, 1964, 1975; Lubinski et al., 1995; Tracey, 2002; Zytowski, 1976), only one (Tracey, 2002) investigated vocational interest development during early adolescence, whereas the others either validated interest inventories or measured interests only once during the respective time frame.

In addition, all studies included in the meta-analyses comprised only two time points. Therefore, the aggregated coefficients provided information about stability and change in different age groups but did not necessarily resemble longitudinal pattern of development across multiple time points. Recent multiwave studies (Hoff et al., 2020; Stoll, Rieger, et al., 2020) with older samples have demonstrated that the pattern of development observed in one sample across time can differ from the pattern suggested by the existing meta-analyses.

Due to the lack of empirical findings, we drew on assumptions from three relevant theoretical frameworks to develop expectations about how vocational interests should develop during late childhood and early adolescence. First, we drew on Holland's (1997) theory of vocational personalities and work environments, which constitutes the most widely used framework for interest classification (Hoff et al., 2018). Second, we used the Trait Situation Interest Dynamic (TSID; Su et al., 2019) model, which integrates conceptions of state and trait interests. Third, we used Gottfredson's (1981) theory of circumscription and compromise, which focuses specifically on the development of occupational preferences from childhood to adolescence. To complement these theoretical assumptions, we tie in the existing empirical evidence and describe how our study builds on these findings.

Late Childhood and Early Adolescence as a Phase of Increasing Interest Stability?

According to Holland (1997), people can be characterized by six general interest dimensions: Realistic, Investigative, Artistic, Social, Enterprising, and Conventional (RIASEC; for more information, see Table 1). He emphasized that vocational interests are personal dispositions, and he described the RIASEC dimensions as "personality types" (p. 2). Vocational interests are therefore defined as preferences that are "trait-like" and consequently seen as relatively stable over longer periods of time (Rounds & Su, 2014, p. 1).

Holland (1997) made only a few predictions about how vocational interests should develop over the course of late childhood and early adolescence. He stated that stability is fostered through cumulative learning on the basis of experiences, where preferences for single activities are gradually rewarded and reinforced—for example, through personal satisfaction or external rewards. Preferences for certain activities that are reinforced eventually transform into a disposition, whereas preferences that are not reinforced are neglected. Although changes in vocational interests are possible across the entire life span, such changes become less likely with age (Holland, 1997). This also implies that changes in vocational interests are more likely during late childhood and early adolescence. Hence, investigating the development of vocational interests during that respective time period seems particularly vital.

Table 1
Reliability Coefficients and Item Examples

Interests	Items	Sample items (wording)	Preferred activities	ω (T1, T2, T3, T4)
		"Are you interested in the following things?"		
R	6	"building something"	Practical activities	.85, .86, .87, .89
I	6	"experimenting in a lab"	Problem solving activities and analytical thinking	.80, .81, .83, .84
A	6	"drawing pictures"	Creative activities	.77, .80, .81, .79
S	6	"helping others"	Teaching, caring, and informing	.83, .86, .86, .85
E	6	"leading a group"	Manipulating and leading others	.81, .80, .78, .79
C	6	"organizing things"	Structured tasks such as organizing and sorting	.81, .83, .83, .81

Note. ω = Omega Total (coefficient was computed in the framework of confirmatory factor analysis, and as required, the factor variances of the measurement models were constrained to one); R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional.

Recently, Su et al. (2019) proposed the TSID model by suggesting that, despite their generally high stabilities, vocational interests also comprise parts that are influenced by situational characteristics. Stable interests reflect a mental representation of an object, including the related affective and cognitive responses. If interests are less stable and are highly prone to the influence of situational characteristics, this mental representation cannot be assumed to be fully developed yet. Su et al. (2019) assumed a fluent transition between the stable and the situational component of interests, which resembles a state-trait continuum (Braun et al., 2020). According to the TSID model, interests solidify through an accumulation of positive experiences—such as novel activities that arouse curiosity, provide surprising information or are cognitively engaging (Hidi & Renninger, 2006; Renninger & Hidi, 2011)—or are altered through an accumulation of negative experiences (Su et al., 2019). By describing reengagement in specific activities as a key mechanism of interest development, the TSID model builds on Holland's assumption of cumulative learning as well as processes described by other interest theories (Hidi & Renninger, 2006; Holland, 1997; Renninger & Hidi, 2011; Silvia, 2001). In addition, people have experiences that are embedded in a multiplicity of different environments. People usually choose environments that offer activities that are in line with their interests. Their experiences strengthen (or weaken) their interests in these activities and increase (or

decrease) the probability that they will choose similar environments in the future. The TSID model therefore emphasizes that interests also develop on the basis of a dynamic interplay between the person and the environment (Su et al., 2019).

Empirical results on vocational interest stability support Holland's (1997) assumptions. The meta-analysis by Low et al. (2005) showed that vocational interests reflect relatively high retest correlations, especially during adulthood (peaking around $r = .65$ and $r = .77$), and already demonstrate moderate retest correlations during early adolescence ($r = .51$). In addition, their results suggest that retest correlations are higher in older age groups like late adolescence and young adulthood. Taken together, these results indicate that vocational interests possess moderate stability in late childhood and early adolescence, but their stability increases with age.

However, the empirical basis for the stability of vocational interests is still sparse. The meta-analysis by Low et al. (2005) comprised only 5 studies in the youngest age group (12-14 years) and of these five studies, only Tracey (2002) investigated the development of vocational interests across two time points during adolescence. Tracey's (2002) sample consisted of two cohorts of U.S. students that were followed from fifth to sixth ($N = 126$) and from seventh to eighth ($N = 221$) grade. He reported an increase in stability, with moderate retest correlations ($.35 < r < .62$) for the interval between fifth and sixth and high retest correlations ($.72 < r < .81$) for the interval between seventh and eighth grade.

Päßler and Hell (2020) investigated the development of vocational interests in a longitudinal sample of 541 Swiss children between the ages of 10 and 12. The sample comprised three waves, each including students from different grade levels (T1 = Grades 4 to 6, T2 = Grades 5 to 7, T3 = Grades 6 to 8). Päßler and Hell (2020) reported moderate stabilities between the first two waves ($.37 < r < .44$) and small to moderate stabilities between the last two ($.17 < r < .47$). The current study builds on findings from Tracey (2002) and Päßler and Hell (2020) by investigating whether the differences in stability they reported can also be found when a sample of participants enrolled in the same grade levels are followed over time.

Normative Change in Late Childhood and Early Adolescence—a Period of Decreasing Interests?

Gottfredson (1981) described the development of occupational preferences from childhood to adolescence. According to Gottfredson, the overarching processes that are central for shaping adolescents' vocational interests are *circumscription* and *compromise*. During circumscription, children and adolescents gradually integrate the concepts of size and power (ages 3 to 5), gender (ages 6 to 8), prestige (ages 9 to 13), and views of their own unique self

(age 14 and older), such as abilities, values, and personality, into their self-concept (Gottfredson, 1981). This integration takes place in the respective environment or social context, and self-concept is therefore very likely to be influenced by and to interact with the environment as it develops. Different environments may consequently lead to different self-concepts. Through their experiences and interactions with other people in different environments or social contexts, children and adolescents gradually rule out occupations that do not match their developing self-concept. According to Gottfredson (1981), vocational interests therefore depend on the internal comparison between one's views of oneself and one's views of certain occupations. After the phase of circumscription, Gottfredson postulated a phase of compromise in which adolescents begin to realize that some of the occupations they regard as suitable for their self-concept might not be available to them. Accordingly, adolescents shift their interests to occupations that might be less compatible with their self-concepts but more accessible. The process of compromise again highlights that interests interact with the respective environment as these interests develop

Against this background, it could be assumed that through the process of circumscription, vocational interests become more differentiated with age because adolescents gradually figure out which activities they do not like. This differentiation on the individual level could manifest in mean-level decreases on the group level (e.g., in RIASEC dimensions). In addition, there is evidence of a similar pattern of mean-level decreases in other motivational constructs, for example, the intrinsic value of school subjects (Gaspard et al., 2017) and academic self-concept (Cole et al., 2001) as well as an increase in the specialization of school-subject domains over time (Denissen et al., 2007). Furthermore, the process of circumscription could also initiate a general disinterest in the world of work at first, as the gradual identification of unsuitable occupations may be accompanied by an accumulation of negative experiences. This circumstance would also manifest in mean-level decreases of interests during these specific years (e.g., in RIASEC dimensions; see also Tracey, 2002).

Evidence of decreasing mean levels was identified in the meta-analysis by Hoff et al. (2018), who reported mean-level decreases for five of the six RIASEC dimensions during late childhood and early adolescence (ages 11 to 14). Two of these five interest dimensions showed decreases that were statistically significant, with small effect sizes between d = -0.30 and d = -0.17. These results were based on two studies, of which Tracey's (2002) two-cohort study was once again the only one that was published. He reported mean-level decreases in four of the six RIASEC dimensions for each cohort, from fifth to sixth (i.e., decreases in Realistic, Investigative, Artistic, and Enterprising interests) and from seventh to eighth grade (i.e.,

decreases in Realistic, Investigative, Enterprising, and Conventional interests). Päßler and Hell (2020) obtained similar results, reporting significant decreases in Realistic, Investigative, Artistic, and Social interests from ages 10 to 12.

Late Childhood and Early Adolescence as a Phase of Gender Differentiation?

Gottfredson's (1981) theory suggests that the manifestation of gender differences in vocational interests has already begun by the age of 6. At this age, children begin to understand the concept of gender roles and start to recognize cues, such as stereotypical gender-specific clothing and activities (Gottfredson, 1981). With increasing age, children understand more abstract cues, such as differences in stereotypical gender-specific personality characteristics (Gottfredson, 1981). According to Gottfredson (1981), the behavior of children during this time can be characterized as "concern[ed] with doing what is appropriate for one's sex" (p. 569). Gender differences in vocational interests should therefore begin to manifest during childhood and should be present over the course of adolescence and adulthood.

Indeed, there is empirical evidence for robust and large gender differences in the mean levels of vocational interests in adolescence and adulthood. Su et al. (2009) conducted a meta-analysis on gender differences in vocational interests in which they included cross-sectional studies with participants who ranged in age from 13 to 43. The results showed that across all age groups, males had higher mean-level scores in Realistic and Investigative interests, whereas females had higher scores in Artistic, Social, and Conventional interests. The effect sizes for these gender differences were quite large, with $|d| = 0.68$ for Social interest and $|d| = 0.84$ for Realistic interest. Similar gender differences for all age groups for Realistic and Social interests were also reported in the meta-analysis by Hoff et al. (2018).

A similar pattern in gender differences can already be found during late childhood. Tracey (2002) reported gender differences for fifth-grade students, with boys scoring higher on Realistic and Investigative interests and girls higher on Artistic, Social, and Conventional interests. Similar results were reported by Päßler and Hell (2020). Although meta-analytic results (Hoff et al, 2018) have indicated that the gender gap between Social and Realistic interests widens over the course of late childhood and early adolescence, this pattern was not confirmed in the longitudinal study by Päßler and Hell (2020).

The Present Study

We investigated the development of vocational interests over the course of late childhood and early adolescence. We used large-scale longitudinal data (N = 3,786) assessed at four time points in grades 5, 6, 7, and 8. We used prior empirical and theoretical work to derive our three main research questions and to guide our expectations. We did not preregister the assumptions of the following study.

First, we investigated the stability of vocational interests. We computed retest correlations for each scale between the scales scores from all adjacent time points. On the basis of prior results, we expected that vocational interests would be moderately stable in terms of retest correlations (r around .50; see Low et al., 2005). In addition, on the basis of assumptions derived from Holland (1997), we expected that retest correlations would increase over the course of late childhood and early adolescence.

Second, we investigated state and trait variance components and how they change over time. We computed state and trait variance proportions for each scale at each time point. According to the assumption of Su et al. (2019), we assumed that vocational interests consist of both, state and trait variance proportions. In addition, according to Holland (1997) as well as Su et al. (2019) we assumed that the trait variance proportions are more substantial than the state variance proportions, and that the trait variance proportion would increase over time.

Third, we investigated how mean levels in vocational interests develop from ages 11 to 14. Because current evidence based on multiwave data is scarce, we decided to provide comprehensive descriptive information about the yearly development of vocational interests on the basis of manifest mean levels. To account for measurement error and test for different types of growth, we also inspected latent mean-level trajectories. On the basis of the results from Hoff et al. (2018), Tracey (2002), and Päßler and Hell (2020) and the theoretical considerations derived from Gottfredson (1981), we expected a general decrease in mean levels over the course of late childhood and early adolescence. However, in line with the results from Hoff et al. (2018) and the general stability of vocational interests, we assumed that, over the 3 years, these decreases would be rather small (d < 0.30). No assumptions were made concerning the types of mean-level trajectories (i.e., linear or quadratic) because no prior evidence was available. We exploratorily tested which growth type fitted the data better.

Fourth, we investigated gender differences in vocational interests. In line with Su et al. (2009), we expected that boys would have higher mean levels in Realistic and Investigative interests and girls in Artistic, Social, and Conventional interests. In addition, according to the theoretical assumptions derived from Gottfredson (1981), an increase in gender differences over

time could be assumed. However, because the empirical findings from Hoff et al. (2018) and Päßler and Hell (2020) were inconclusive, we made no assumptions about increases in gender differences over time, but investigated them exploratorily.

Method

Sample

We used data from the Tradition and Innovation (TRAIN; Jonkmann et al., 2013) study.[5] TRAIN is a large-scale longitudinal study in which students from lower, intermediate, and multitrack schools (a combination of the lower and intermediate tracks) were followed annually from fifth to eighth grade (T1, T2, T3, and T4). The study protocol and data protection measures of the TRAIN Study were reviewed and approved by the responsible departments and the independent data protection official of the Ministry of Culture, Youth, and Sports at Baden-Württemberg and the Saxon Ministry of State for Education and Culture. Participation was voluntary, and students could participate only with active parental consent.

The schools were chosen from the states of Baden-Württemberg and Saxony. From every school (99 overall), one to two classes (136 overall) participated in the study. The sample sizes were n = 2,894 (46% female) students at T1 (Grade 5), n = 2,936 (45% female) students at T2 (Grade 6), n = 2,993 (46% female) students at T3 (Grade 7), and n = 3,060 (46% female) students at T4 (Grade 8). The sample size increased slightly over time because additional students were included at later time points. The reasons were repetition, changes in a class, or the relocation of a participant's family. The overall number of students who participated at a minimum of one time point was N = 3,876, which was also the sample used in the current analysis (for details about the sample composition and dropout, see Supplement A). It should be noted that although TRAIN comprises a large longitudinal sample, it is not representative of all students in the German states of Baden-Württemberg and Saxony because it does not include students from the highest school track. TRAIN was designed to investigate the influence of learning environments from lower, intermediate, and multitrack schools on educational outcomes, and higher track students were therefore not part of the sample (Jonkmann et al., 2013).

We conducted two types of attrition analyses at each time point. First, we compared participants who dropped out of the study at a respective time point with the participants who remained in the study. Second, we compared participants who joined the study at a respective time point with the participants who were already participating before that time point. Differences in interest measures, sociodemographic variables, achievement measures, and personality characteristics (i.e. Big Five) were investigated. In general, differences in interest

[5] This data set has already been used in several other publications. An overview is given here: https://uni-tuebingen.de/en/43704. None of these studies investigated the development of vocational interests, nor included vocational interests in the study.

measures, sociodemographic variables, and personality characteristics were small and, in the majority of cases, not significant. Comparing participants who dropped out of the study with the remaining participants, the standardized mean differences ranged from $|d|$ = 0.00 to $|d|$ = 0.39 ($|Mdn|$ = 0.10), such that the participants who dropped out generally had lower values—a finding that also applied to participants who joined the study late.

For achievement measures, the largest differences were found for school grades, which ranged from $|d|$ = 0.11 to $|d|$ = 0.84, indicating worse grades for participants who dropped out or joined the study late in comparison with the other participants in the sample. An explanation for these differences could be that participants who dropped out or joined the study late included students who repeated a grade, and poor performance is usually the major criterion for repeating a grade. Although differences in report-card grades were quite high, differences in standardized achievement were much lower. For participants who joined the study late, differences in math and German abilities were usually not significant and small in magnitude with standardized effect sizes ranging from $|d|$ = 0.01 to $|d|$ = 0.22 ($|Mdn|$ = 0.10). Differences between participants who dropped out and the remaining participants were highest at the fourth time point with standardized effect sizes ranging from $|d|$ = 0.40 to $|d|$ = 0.56—differences at the other time points were usually small (for detailed results, see Supplement A).

Instruments

Vocational interests were operationalized in accordance with the RIASEC framework. As there is currently no interest inventory that was designed for the transition from childhood to adolescence, two frequently-used and validated interest inventories were modified and combined. For adolescents older than 14, the Revised General Interest Structure-Test (AIST-R; Allgemeiner Interessen Strukturtest; Bergmann & Eder, 2005) is used most often in German-speaking countries (e.g., Germany, Austria, and Switzerland). For children, the German version of the Inventory of Children's Activities (ICA; Tracey & Ward, 1998; German version (ICA-D): von Maurice, 2006) is used. To ensure and enhance the comprehensibility of the interest inventory for the young age group of the TRAIN study, experts from the research field of vocational interests, selected items from the German version of the ICA, items from the AIST-R and constructed a few new items, focussing on activities that were familiar to children at age 11, but still suitable (i.e. not too childish) for adolescents at the age of 14. They followed the recommendations made by Tracey (2002) about measuring children's vocational interests by asking about "familiar activities, rather than unfamiliar occupations" (Tracey, 2002, p. 149).

The final questionnaire consisted of 11 items from the AIST-R, 15 items from the ICA and 11 new items (see Supplement B).

Every RIASEC dimension was assessed with six items per time point. Students answered the question "How much do you like this activity?" on a 5-point Likert scale ranging from 1 (*not at all*) to 5 (*very*). Sample items and reliability coefficients for all scales and every time point can be found in Table 1. Omega total (McNeish, 2018) ranged from $.77 < \omega < .89$. An overview of the item-total correlations and the factor loadings of the items, can be found in Supplement B. For every RIASEC dimension, a one-dimensional factor model represented the data adequately (for an overview of the model fit indices, see Supplement C). Evidence for the convergent and discriminant validity of the interest measure can be found in Supplement B.

Statistical Analyses

We provide all analysis scripts of the current study in an OSF repository (see: https://osf.io/tuys8/?view_only=cb07ca448fdf4dce998e250ae635a419).

Measurement invariance

To adequately attribute latent mean-level differences to changes in the latent construct, strong measurement invariance is required. Therefore, we investigated measurement models with different degrees of invariance across time as well as across the Time × Gender interaction. Invariance across time implies that the meaning and interpretation of the construct does not change over time. Invariance across gender and time implies that boys and girls attribute the same meaning and interpretation to the construct and that its meaning and interpretation do not change over time. We investigated configural measurement invariance (i.e., structural invariance over time and the Time × Gender interaction), weak measurement invariance (i.e., factor loading invariance over time and the Time × Gender interaction) and strong measurement invariance (i.e., factor loading and intercept invariance over time and the Time × Gender interaction).

The following fit indices and values were used to judge whether the overall fit of the models was adequate: Incremental fit indices such as the Comparative Fit Index (CFI) and the Tucker-Lewis Index (TLI) with values of .90 or higher, the Standardized Root Mean Square Residual (SRMR) with a value of .08 or lower, as well as the Root Mean Square Error of Approximation (RMSEA) with a value of .06 or lower (Hu & Bentler, 1999).

For evaluating the changes in incremental fit indices (i.e., CFI, SRMR and RMSEA) between the respective models (i.e., configural, weak and strong invariance), we applied the

recommendations made by Chen (2007). We used the recommended values to test for factor loading invariance with an adequate sample size, with values ≥ .010 for ΔCFI, additionally accompanied by values of ≥ .015 for ΔRMSEA or ≥ .030 for ΔSRMR, indicating substantive changes in testing for factor loading invariance (Chen, 2007). To test for intercept invariance, values of ≥ .010 for ΔCFI, additionally accompanied by values of ≥ .015 for ΔRMSEA or ≥ .010 for ΔSRMR were assumed to be substantial.

According to the fit indices, all the longitudinal models with strong invariance provided an adequate fit (for fit indices, see Supplement C). However, for the models that indicated invariance across the Time × Gender interaction, only partial strong invariance could be applied. Partial strong invariance suggests that the majority of the respective item intercepts were invariant between boys and girls, but a few of them were not. Aiming for partial invariance is a suitable approach when invariance cannot be achieved on all items (Putnick & Bornstein, 2016).

To achieve partial strong invariance, intercepts of single items with large gender differences are freed in accordance with the weak measurement invariance model. To achieve partial strong invariance in our study, we freed one item intercept for Social, two item intercepts for Realistic, Investigative, Enterprising, and Conventional, and three item intercepts for Artistic. All other item intercepts were invariant across time and gender. The final partial strong invariance models indicated an adequate overall model fit (according to Chen, 2007). For a detailed description of the measurement invariance results see Supplement C.

Mean levels and retest correlations

To examine the manifest mean-level change and correlation coefficients, we specified a saturated path model that included all the manifest scale scores for the six RIASEC dimensions (see a depiction of the model in Supplement D). The six manifest scale scores were allowed to be freely correlated across all time points. Descriptive statistics, mean-level differences, and retest correlations for adjacent time points were derived from this model. To examine the mean-level differences with a standardized effect size, we computed Cohen's *d* for repeated measures (Lakens, 2013; for the formula, see Supplement C). With regard to the meta-analytic findings from Hoff et al. (2018) and the general stability of interests, we considered the changes over the 3-year time period ranging from $d = 0.20$ to $d = 0.30$ to be substantial.

To estimate the model parameters, we used a maximum likelihood estimator that is robust against non-normality and nonindependence of data by choosing the analysis option TYPE = COMPLEX in Mplus 8 (Muthén & Muthén, 1998–2017). The estimator is labeled

MLR (Maximum Likelihood Robust) in Mplus 8, and its χ^2 test statistic is asymptotically equivalent to the Yuan-Bentler T_2 test statistic (Muthén & Muthén, 1998–2017; Yuan & Bentler, 2000). For significance testing, we provided 95% confidence intervals for all parameters in the saturated path model and for all parameters in the subsequent models. Additionally, all subsequent models were also specified with the MLR estimator in Mplus 8 (Muthén & Muthén, 1998–2017).

Computing proportions of state and trait variance

To compute the proportions of state and trait variance, we specified a generalized second-order growth model (GSGM; Bishop, Geiser, & Cole, 2015) with time-invariant factor loadings and time-invariant intercepts for each interest dimension separately (for a depiction of the model, see Supplement D). The GSGM is a hybrid between a latent state-trait model (LST; Steyer et al., 1999) and a growth curve model (GCM; McArdle & Epstein, 1987). The GSGM consisted of three components that were represented by three latent variables: a component that is stable over time, representing initial trait levels; a component that captures growth over time, representing trait growth; and time-point-specific components, representing time-point-specific fluctuations (or state residuals). An advantage of the GSGM is the possibility of separating the model-implied variance components of the indicators that are due to the trait, trait growth, state residual, and measurement error (Geiser et al., 2015). We included the growth component in the model because, according to theory, vocational interests are dispositions that develop over time. Therefore, long-term changes such as trait growth should be accounted for and included as part of the proportion of reliable trait variance. A model without a growth component would merely represent state variability that implies no trait development over time (Geiser et al., 2015). The intercept factor of the growth curve model part was identified by fixing the factor loading of the first item of each time point to 1 and the intercept to 0 (see Supplement D for a model depiction).

The model-implied variance components were used to calculate the consistency (CO), occasion specificity (OS), and reliability (REL; Geiser et al., 2015) of the interest constructs. CO represents the proportion of the variance of the indicators that is due to the trait and trait-growth components, whereas OS represents the proportion of the variance that is due to the state residual component (see Supplement D for the formulas). The state residual component OS represents the influences of the situation and the Person × Situation interactions that were not accounted for by the trait (Geiser et al., 2015). REL represents the reliable part of the

variance that was not due to measurement error, consisting of the sum of CO and OS (Geiser et al., 2015).

The total model-implied variance of every item in the GSGM can therefore be segmented into the components OS, CO, and measurement error (ε = 1 – REL; Geiser et al., 2015). For a better understanding of how much of the model-implied variance was due to trait, trait growth, state residual, and measurement error, we averaged the coefficients across all items separately for every time point. CO, OS, and the proportion of measurement error were therefore depicted for every time point separately, but they were averaged across the items within a time point.

Specifying growth type

For every RIASEC dimension, we specified two GSGMs, one with linear and one with quadratic trait growth. To decide whether the model with the quadratic growth type fits the data substantively better, we determined whether the changes in the descriptive fit indices between the model that contained only the linear growth and the model that additionally contained the quadratic growth were substantial. If this was the case, we decided to choose the model with a quadratic growth factor. To assess the changes in model fit, we used the values proposed by Chen (2007). Because these recommendations were explicitly given for measurement invariance testing, we used them with caution and only as approximate guides. As no assumptions about the growth type could be derived from current studies, our approach here was to investigate which growth type best described the data.

Analyzing gender differences

To examine gender differences in the mean levels, we used multigroup modeling. We specified a multigroup saturated path model to obtain manifest mean-level scores for boys and girls. For the GSGM multigroup models, we first specified an overall model, where both groups, boys and girls, had linear as well as quadratic growth components. We then determined whether the means of the respective quadratic growth factors were statistically significantly different from zero. If this was not the case, we specified another model in which we constrained the quadratic growth type for the respective group. If the change in the descriptive fit indices according to the values proposed by Chen (2007) was negligible, we decided to keep the constrained model. Due to missing prior evidence, the approach of choosing the appropriate growth type was again data driven.

Missing data

To deal with the occurrence of missing data, we used full information maximum likelihood estimation (FIML; see, e.g., Enders, 2001) in all models. To make the Missing at Random (MAR) assumption more plausible and to reduce possible selection bias due to attrition, we included multiple auxiliary variables that could explain the causes of the missing data, thus improving parameter estimation (Collins et al., 2001; Enders, 2008; Graham, 2003). The auxiliary variables were included in the estimation of the model in line with the saturated correlates model approach (Graham, 2003) through the Mplus 8 AUXILIARY command (Asparouhov & Muthen, 2008). In this approach, auxiliary variables are included in the analysis through correlations with variables that are part of the actual analysis, but these auxiliary variables are not part of the structural model. All variables that indicated significant differences during the attrition analysis were included as auxiliary variables in all analyses (for an overview of these variables, see Supplement A).

Nested data structure

In the present data set, students were nested in classes, which led to a hierarchical data structure that could cause an underestimation of the standard errors. Because classes were not the focus of our analysis, we treated them as a nuisance (McNeish et al., 2017) and therefore relied on cluster-robust standard errors to correct the underestimation. We used the implemented analysis option in Mplus 8, TYPE = COMPLEX, which computes the standard errors on the basis of a sandwich estimator method that accounts for the violation of the independence of observations assumption (McNeish et al., 2017). The intraclass correlations of the RIASEC scales with class as the cluster variable ranged from .02 to .11. We decided not to include school as another cluster variable because only the intraclass correlations for Investigative, Enterprising, and Conventional at the first time point were noteworthy with an ICC of .09. The intraclass correlations of the majority of the variables were below .04 (see Supplement A for an overview of the intraclass correlations).

Results

The exact *p*-Values of all parameters depicted in the results section and a correlation matrix containing all scales can be found in the respective output files on the OSF repository (see: https://osf.io/tuys8/?view_only=cb07ca448fdf4dce998e250ae635a419).

Stability of Vocational Interests

Retest correlations

To investigate the stability of vocational interests, we computed retest correlations for the vocational interest scales. In line with our assumptions, the 1-year retest correlations for all interest dimensions were moderate, ranging from $r = .41$ to $r = .64$ (see Table 2). The retest correlations for the entire 3-year interval were lower, ranging from $r = .32$ to $r = .49$ for all interest dimensions.

Realistic, Investigative, Social, and Conventional showed significant increases in retest correlations from the first to the last interval, with increases ranging from $\Delta r = .06$ to $\Delta r = .08$ (see Table 2). From the first to the second interval, the retest correlations for these four interest dimensions showed significant increases that ranged from $\Delta r = .07$ to $\Delta r = .11$. From the second to the last interval, the retest correlations for Realistic ($\Delta r = .00$) and Social ($\Delta r = .00$) did not change and reached a plateau, whereas the retest correlations for Investigative ($\Delta r = -.03$) and Conventional ($\Delta r = -.05$) decreased, but not statistically significant.

For Artistic and Enterprising, the retest correlations also increased from the first to the last interval (see Table 2), by $\Delta r = .03$ and $\Delta r = .04$, respectively, but not significantly (see Table 2). Whereas the retest correlations for Artistic hardly changed at all, the retest correlations for Enterprising showed a significant increase ($\Delta r = .12$) between the first two intervals and a significant decrease ($\Delta r = -.08$) between the last two intervals, resulting in the nonsignificant change ($\Delta r = .04$) from the first to the last interval.

Table 2
Manifest Retest Correlations of RIASEC Scales and their Change over Time

Int.	Time Interval	*r* [95% CI]	Time Interval Change	Δ*r* [95% CI]
R	1 to 2	.55 [.51, .59]		
	2 to 3	.64 [.60, .67]	1 to 2 and 2 to 3	.08 [.04, .13]
	3 to 4	.64 [.60, .67]	2 to 3 and 3 to 4	.00 [-.05, .05]
	1 to 4	.49 [.45, .53]	1 to 2 and 3 to 4	.08 [.03, .13]
I	1 to 2	.43 [.38, .47]		
	2 to 3	.54 [.50, .58]	1 to 2 and 2 to 3	.11 [.06, .17]
	3 to 4	.51 [.47, .55]	2 to 3 and 3 to 4	-.03 [-.08, .02]
	1 to 4	.33 [.28, .38]	1 to 2 and 3 to 4	.08 [.02, .14]
A	1 to 2	.53 [.49, .57]		
	2 to 3	.58 [.54, .62]	1 to 2 and 2 to 3	.05 [.00, .11]
	3 to 4	.56 [.52, .60]	2 to 3 and 3 to 4	-.02 [-.07, .03]
	1 to 4	.42 [.38, .46]	1 to 2 and 3 to 4	.03 [-.02, .09]
S	1 to 2	.49 [.44, .53]		
	2 to 3	.56 [.51, .60]	1 to 2 and 2 to 3	.07 [.01, .12]
	3 to 4	.55 [.51, .59]	2 to 3 and 3 to 4	-.00 [-.05, .05]
	1 to 4	.41 [.36, .46]	1 to 2 and 3 to 4	.07 [.01, .12]
E	1 to 2	.43 [.38, .47]		
	2 to 3	.55 [.51, .58]	1 to 2 and 2 to 3	.12 [.07, .17]
	3 to 4	.47 [.43, .51]	2 to 3 and 3 to 4	-.08 [-.13, -.03]
	1 to 4	.34 [.29, .39]	1 to 2 and 3 to 4	.04 [-.02, .10]
C	1 to 2	.41 [.36, .46]		
	2 to 3	.52 [.47, .56]	1 to 2 and 2 to 3	.11 [.06, .17]
	3 to 4	.47 [.43, .51]	2 to 3 and 3 to 4	-.05 [-.11, .01]
	1 to 4	.32 [.27, .38]	1 to 2 and 3 to 4	.06 [.00, .12]

Note. $N = 3{,}875$; Int. = RIASEC interest dimensions; *r* = retest correlation between two time intervals; Δ*r* = difference between adjacent retest correlations; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional.

Proportions of state and trait variance of vocational interests

To investigate vocational interest stability while accounting for trait development, we computed components of state and trait variance (see Table 3). In line with our assumptions, the CO coefficients (i.e., the proportion of the variance in the indicators due to the trait and trait-growth components) for Realistic, Artistic, Social, and Conventional significantly increased by 5% to 12% from T1 to T4 (see Figure 1).

From T1 to T4, the OS coefficients (i.e., the proportion of variance due to the state residual component, representing the influence of the situation and Person × Situation interactions) did not show significant change for Realistic, Artistic, or Social but showed a

significant decrease for Conventional. In line with our assumption, at T4, the CO coefficients for Realistic, Artistic, Social, and Conventional were significantly higher than the OS coefficients, indicated by nonoverlapping confidence intervals. A similar pattern was already present at T1, except for Conventional, for which differences between OS and CO were not significant (Δ = -.05, 95% CI [-.11, .01]).

In contrast to our expectations, from T1 to T4, the CO coefficients for Investigative and Enterprising did not change significantly. In addition, from T1 to T4, the OS coefficient for Investigative showed a significant increase, whereas the OS coefficient for Enterprising did not change significantly. At T4, CO coefficients for Enterprising (Δ = .06, 95% CI [-.01, .13]) and Investigative (Δ = .02, 95% CI [-.04, .07]) did not differ significantly from the respective OS coefficients, a result that was not in line with our assumptions.

Table 3

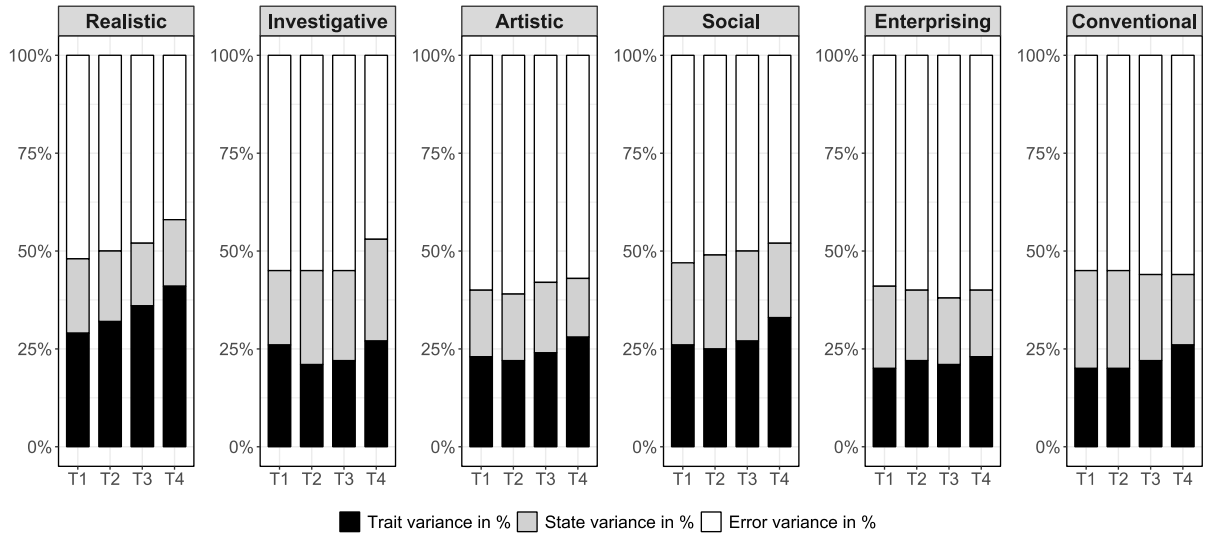
Consistency and Occasion Specificity of RIASEC Dimensions per Time Point

Int.	TL	CO [95% CI]	Δ [95% CI]	OS [95% CI]	Δ [95% CI]
	1	.29 [.26, .32]		.19 [.16, .22]	
	2	.32 [.30, .34]	.03 [.01, .06]	.18 [.16, .21]	-.01 [-.04, .03]
R_L	3	.36 [.33, .38]	.04 [.02, .05]	.16 [.14, .19]	-.02 [-.05, .01]
	4	.41 [.37, .44]	.05 [.02, .08]	.17 [.14, .21]	.01 [-.03, .05]
	14		.12 [.08, .16]		-.02 [-.06, .03]
	1	.26 [.23, .28]		.19 [.16, .21]	
	2	.21 [.19, .23]	-.05 [-.07, -.03]	.24 [.22, .27]	.05 [.03, .08]
I_L	3	.22 [.19, .24]	.01 [-.01, .02]	.23 [.20, .25]	-.02 [-.04, .01]
	4	.27 [.24, .30]	.06 [.04, .07]	.26 [.22, .29]	.03 [.00, .06]
	14		.02 [-.02, .05]		.07 [.03, .11]
	1	.23 [.20, .25]		.17 [.14, .19]	
	2	.22 [.21, .24]	.00 [-.02, .01]	.17 [.15, .20]	.01 [-.02, .04]
A_L	3	.24 [.22, .26]	.01 [.00, .03]	.18 [.16, .21]	.01 [-.01, .04]
	4	.28 [.26, .30]	.04 [.03, .06]	.15 [.13, .18]	-.03 [-.06, .00]
	14		.06 [.03, .08]		-.01 [-.05, .02]
	1	.26 [.23, .29]		.21 [.18, .24]	
	2	.25 [.22, .28]	-.01 [-.03, .01]	.24 [.22, .27]	.04 [.01, .07]
S_L	3	.27 [.24, .30]	.02 [.00, .03]	.23 [.20, .26]	-.01 [-.04, .01]
	4	.33 [.31, .36]	.07 [.04, .09]	.19 [.15, .23]	-.04 [-.08, -.01]
	14		.07 [.04, .11]		-.02 [-.06, .03]
	1	.20 [.18, .23]		.21 [.18, .24]	
	2	.22 [.20, .23]	.01 [-.01, .03]	.18 [.16, .20]	-.03 [-.07, .00]
E_L	3	.22 [.20, .23]	.00 [-.02, .02]	.17 [.15, .19]	-.01 [-.04, .02]
	4	.23 [.20, .27]	.02 [-.01, .05]	.17 [.14, .21]	.00 [-.04, .05]
	14		.03 [-.01, .07]		-.04 [-.09, .01]
	1	.20 [.17, .23]		.25 [.21, .28]	
	2	.20 [.18, .23]	.00 [-.02, .02]	.25 [.23, .28]	.01 [-.02, .04]
C_L	3	.22 [.20, .24]	.02 [.00, .03]	.22 [.20, .25]	-.03[-.06, .00]
	4	.26 [.23, .29]	.04 [.02, .07]	.18 [.15, .21]	-.04 [-.08, -.01]
	14		.06 [.02, .10]		-.07 [-.11, -.02]

Note. Int. = RIASEC interest dimensions; TL = time lag; CO = consistency coefficient; Δ = differences in the consistency and occasion specificity coefficient over time; OS = occasion specificity coefficient; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional; $_L$ = linear growth factor; 1–4 = time points 1 to 4; 14 = lag between T1 to T4.

Figure 1

Time-point-specific proportions of state and trait variance in the overall sample.



Note. Consistency (i.e., trait variance), occasion specificity (i.e., state variance), and measurement error (i.e., error variance) are depicted.

Mean-Level Development of Vocational Interests

To investigate the mean-level development of vocational interests, we examined the coefficients in the saturated path model (see Table 4) and the trajectories of the latent mean levels. In line with our assumptions, the mean levels of the majority of the interest dimensions decreased from T1 to T4 (Figure 2). The mean levels for Realistic, Investigative, Artistic, and Conventional decreased significantly over the 3-year time period, with effect sizes ranging from d = -0.44 to d = -0.24. The latent mean levels of these interest dimensions showed a significant linear decrease over time, and including a quadratic growth factor did not substantially improve the model fit according to the indices in Table 5 (for an overview of all model fit indices, see Supplement C).

By contrast, Social was the only interest dimension for which the mean levels increased slightly (d = 0.07), whereas the mean levels for Enterprising did not change (d = -0.01) from T1 to T4. The latent mean levels for Social and Enterprising showed a linear trend over time, and including a quadratic growth factor did not substantially improve the model fit. For Social, the mean levels of the latent growth factors indicated a significant increase, and for Enterprising, a nonsignificant decrease (see Table 6 and Supplement C).

Table 4

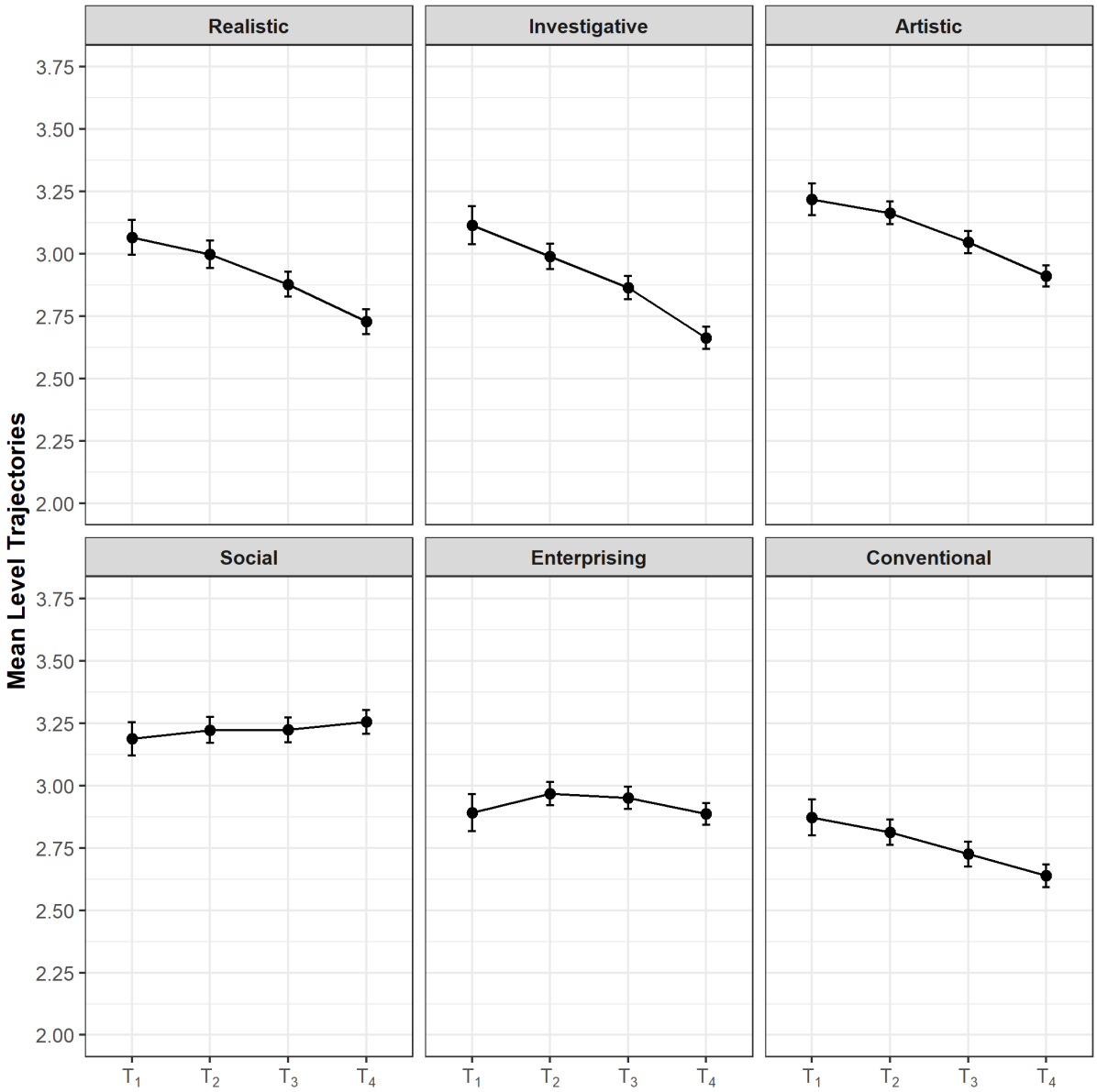
Descriptive Statistics and Mean-Level Change in the RIASEC Scales in the Overall Sample

Int.	TL	*M* [95% CI]	*SD* [95% CI]	Δ [95% CI]	*d*
R	1	3.07 [3.00, 3.14]	1.16 [1.14, 1.19]		
	2	3.00 [2.94, 3.05]	1.13 [1.10, 1.15]	-0.07 [-0.13, -0.01]	-0.06
	3	2.88 [2.83, 2.93]	1.12 [1.10, 1.15]	-0.12 [-0.17, -0.07]	-0.11
	4	2.73 [2.68, 2.78]	1.13 [1.10, 1.15]	-0.15 [-0.19, -0.11]	-0.13
	14			-0.34 [-0.41, -0.27]	-0.30
I	1	3.11 [3.04, 3.19]	1.08 [1.05, 1.10]		
	2	2.99 [2.94, 3.04]	1.03 [1.00, 1.06]	-0.13 [-0.19, -0.06]	-0.12
	3	2.87 [2.82, 2.91]	0.99 [0.96, 1.02]	-0.12 [-0.17, -0.08]	-0.12
	4	2.66 [2.62, 2.71]	0.99 [0.96, 1.02]	-0.20 [-0.25, -0.16]	-0.20
	14			-0.45 [-0.53, -0.37]	-0.44
A	1	3.22 [3.16, 3.28]	1.02 [0.99, 1.05]		
	2	3.16 [3.12, 3.21]	0.98 [0.96, 1.01]	-0.05 [-0.11, 0.00]	-0.05
	3	3.05 [3.00, 3.09]	0.98 [0.95, 1.00]	-0.12 [-0.16, -0.07]	-0.12
	4	2.91 [2.87, 2.95]	0.96 [0.93, 0.98]	-0.14 [-0.18, -0.09]	-0.14
	14			-0.31 [-0.37, -0.25]	-0.31
S	1	3.19 [3.12, 3.25]	1.04 [1.01, 1.07]		
	2	3.22 [3.17, 3.28]	1.02 [0.99, 1.05]	0.04 [-0.02, 0.09]	0.03
	3	3.22 [3.17, 3.27]	0.98 [0.96, 1.01]	0.00 [-0.04, 0.04]	0.00
	4	3.26 [3.21, 3.30]	0.96 [0.94, 0.99]	0.03 [-0.01, 0.08]	0.03
	14			0.07 [0.01, 0.13]	0.07
E	1	2.89 [2.82, 2.97]	1.07 [1.04, 1.10]		
	2	2.97 [2.92, 3.01]	0.97 [0.95, 1.00]	0.08 [0.01, 0.14]	0.07
	3	2.95 [2.91, 3.00]	0.92 [0.89, 0.94]	-0.02 [-0.06, 0.03]	-0.02
	4	2.89 [2.84, 2.93]	0.89 [0.87, 0.91]	-0.06 [-0.10, -0.02]	-0.07
	14			-0.01 [-0.08, 0.07]	-0.01
C	1	2.87 [2.80, 2.94]	1.04 [1.01, 1.07]		
	2	2.81 [2.76, 2.86]	1.00 [0.98, 1.03]	-0.06 [-0.12, 0.00]	-0.06
	3	2.73 [2.68, 2.77]	0.96 [0.93, 0.99]	-0.09 [-0.15, -0.04]	-0.09
	4	2.64 [2.59, 2.68]	0.92 [0.89, 0.95]	-0.09 [-0.13, -0.05]	-0.09
	14			-0.24 [-0.30, -0.17]	-0.24

Note. N = 3,875; Int. = RIASEC interest dimensions; TL = time lag; M = mean level, SD = standard deviation; Δ = mean-level differences between the time points; d = Cohen's d for repeated measurement; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional; 1–4 = time points 1 to 4; 14 = lag between T1 to T4.

Figure 2

Manifest Mean Levels of the Scale scores for the RIASEC Dimensions



Note. 95% confidence intervals are depicted; T_{1-4} = time points 1–4; *n* = 3,875.

Table 5
Fit Indices for Single and Multigroup Generalized Second-Order Growth Model

Int.	Model	Growth	CFI	TLI	RMSEA	SRMR	χ^2	*df*
R	Overall	L	0.963	0.963	0.030	0.044	1228.52*	270
		Q	0.965	0.964	0.030	0.046	1183.61*	266
	Multigroup	QQ	0.942	0.941	0.033	0.051	1541.80*	545
		QL	0.939	0.939	0.033	0.052	1591.82*	549
I	Overall	L	0.911	0.909	0.041	0.086	1998.54*	270
		Q	0.926	0.923	0.037	0.080	1709.17*	266
	Multigroup	QQ	0.907	0.906	0.041	0.074	2098.53*	545
		QQ	0.907	0.906	0.041	0.074	2098.53*	545
A	Overall	L	0.919	0.917	0.038	0.067	1798.12*	270
		Q	0.926	0.923	0.037	0.065	1659.15*	266
	Multigroup	QQ	0.900	0.899	0.039	0.064	1946.27*	544
		QL	0.896	0.895	0.040	0.065	2010.69*	548
S	Overall	L	0.934	0.933	0.037	0.068	1728.05*	270
		Q	0.945	0.943	0.034	0.080	1490.88*	266
	Multigroup	QQ	0.927	0.926	0.036	0.073	1773.64*	546
		LL	0.913	0.914	0.039	0.069	2004.72*	554
E	Overall	L	0.955	0.954	0.026	0.047	980.29*	270
		Q	0.955	0.953	0.026	0.047	975.92*	266
	Multigroup	QQ	0.941	0.941	0.030	0.048	1358.64*	545
		QQ	0.941	0.941	0.030	0.048	1358.64*	545
C	Overall	L	0.966	0.965	0.024	0.061	876.90*	270
		Q	0.968	0.966	0.024	0.057	838.85*	266
	Multigroup	QQ	0.947	0.946	0.030	0.052	1362.34*	545
		LQ	0.946	0.946	0.030	0.053	1381.48*	549

Note. Int. = RIASEC interest dimensions; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; χ^2 = Chi-Square Statistic; *df* = Degrees of Freedom; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional; L = linear growth; Q = quadratic growth; QL = growth type of the respective group, the first letter refers to girls, the second to boys.
* $p < .01$.

Table 6

Intercept and Slope of Single and Multigroup Generalized Second-Order Growth Models

		Intercept		Slope		
Group	Int.	*M* [95% CI]	σ^2 [95% CI]	*M* [95% CI]	σ^2 [95% CI]	*r*(I,S) [95% CI]
O	R_L	3.37 [3.30, 3.43]	0.46 [0.39, 0.52]	-0.10 [-0.12, -0.08]	0.02 [0.01, 0.04]	-.20 [-.34, -.07]
	I_L	3.64 [3.56, 3.72]	0.89 [0.79, 0.99]	-0.16 [-0.19, -0.13]	0.11 [0.09, 0.13]	-.60 [-.66, -.55]
	A_L	3.63 [3.57, 3.70]	0.64 [0.56, 0.72]	-0.11 [-0.14, -0.09]	0.06 [0.04, 0.07]	-.38 [-.46, -.31]
	S_L	3.19 [3.12, 3.26]	0.68 [0.57, 0.79]	0.06 [0.03, 0.08]	0.07 [0.04, 0.10]	-.46 [-.56, -.36]
	E_L	2.87 [2.81, 2.92]	0.18 [0.13, 0.22]	-0.01 [-0.03, 0.00]	0.01 [0.00, 0.01]	-.47 [-.58, -.36]
	C_L	3.10 [3.02, 3.19]	0.61 [0.51, 0.72]	-0.14 [-0.17, -0.10]	0.05 [0.02, 0.07]	-.40 [-.53, -.28]
B	R_L	3.66 [3.59, 3.72]	0.31 [0.23, 0.39]	-0.07 [-0.09, -0.05]	0.02 [0.01, 0.04]	-.38 [-.53, -.22]
	I_Q	3.68 [3.58, 3.78]	1.01 [0.88, 1.15]	-0.04 [-0.07, -0.01]	0.10 [0.08, 0.12]	.34 [.23, .45]
	A_L	3.47 [3.39, 3.56]	0.82 [0.68, 0.96]	-0.17 [-0.20, -0.13]	0.08 [0.05, 0.10]	-.54 [-.63, -.45]
	S_L	2.99 [2.91, 3.07]	0.69 [0.56, 0.82]	0.02 [-0.01, 0.05]	0.07 [0.04, 0.11]	-.53 [-.63, -.42]
	E_Q	3.00 [2.91, 3.10]	0.28 [0.18, 0.39]	-0.02 [-0.04, -0.01]	0.01 [-0.01, 0.04]	-.11 [-.93, .72]
	C_Q	2.99 [2.89, 3.09]	0.51 [0.29, 0.73]	0.03 [0.01, 0.06]	0.04 [0.01, 0.07]	-.21 [-.76, .33]
G	R_Q	2.96 [2.88, 3.03]	0.41 [0.31, 0.51]	-0.03 [-0.05, -0.01]	0.03 [0.02, 0.05]	.28 [.10, .47]
	I_Q	3.48 [3.37, 3.59]	1.09 [0.95, 1.24]	-0.04 [-0.07, -0.01]	0.10 [0.08, 0.13]	.38 [.29, .47]
	A_Q	3.94 [3.85, 4.04]	0.57 [0.43, 0.70]	-0.04 [-0.06, -0.02]	0.04 [0.02, 0.06]	.32 [.15, .48]
	S_L	3.41 [3.33, 3.50]	0.58 [0.46, 0.70]	0.11 [0.08, 0.15]	0.07 [0.04, 0.09]	-.53 [-.65, -.41]
	E_Q	2.93 [2.85, 3.02]	0.37 [0.18, 0.55]	-0.03 [-0.05, -0.01]	0.02 [0.00, 0.04]	.11 [-.31, .52]
	C_L	2.95 [2.87, 3.04]	0.59 [0.46, 0.72]	-0.18 [-0.22, -0.15]	0.06 [0.03, 0.08]	-.44 [-.60, -.28]

Note. Int. = RIASEC interest dimensions; *M* = mean of intercept or slope factor; σ^2 = variance of intercept or slope factor; *r*(I,S) = correlation between intercept and slope; O = overall sample; B = boys; G = girls; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional; $_L$ = linear growth; $_Q$ = quadratic growth.

Gender Differences in the Mean Levels of Vocational Interests

To investigate the mean-level differences between boys and girls, we examined the coefficients in the multigroup saturated path model (see Table 7) and the multigroup GSGM. In line with our assumptions, we found gender differences in the manifest mean levels for Realistic and Investigative, with boys having higher mean levels than girls. For boys and girls, manifest mean levels for Realistic and Investigative decreased significantly from T1 to T4 (see Tables 8 and 9), with standardized effect sizes ranging from d = -0.44 to d = -0.54 for girls and d = -0.24 to d = -0.35 for boys. For Realistic and Investigative, the gender differences in the manifest mean levels increased from T1 to T4 (see Figure 3). The latent mean levels for Realistic and the corresponding model fits (see Table 5) indicated a linear trajectory for boys and a quadratic trajectory for girls. For Investigative, the latent mean levels indicated a quadratic trajectory for girls and boys.

In line with our assumptions, we found gender differences in manifest mean levels for Artistic and Social (see Table 7), with girls having higher mean levels than boys. Both boys and girls showed significant decreases in their mean levels for Artistic, with effect sizes of d = -0.30 and d = -0.36, respectively. Their developmental pattern differed, however, for Social (see Figure 3), where boys did not show significant changes in their mean levels from T1 to T4, and girls showed a significant increase (d = 0.22). Differences between boys and girls in their manifest mean levels in Social interest increased over time. For Artistic, gender differences in the manifest mean levels did not increase over time (see Table 7). The latent mean levels for Artistic indicated a linear trajectory for boys and a quadratic trajectory for girls (see Table 5). For Social, the latent mean levels indicated a linear trajectory for girls and boys.

In contrast to our assumptions, for Conventional, boys demonstrated higher manifest mean levels than girls at T3 and T4. In addition, at T3 and T4, boys had higher mean levels than girls for Enterprising. From T1 to T4, boys and girls did not show significant changes in manifest mean levels for Enterprising. However, both groups showed significant mean-level decreases in Conventional, with d = -0.32 for girls and d = -0.17 for boys. Despite having similar manifest mean-level change patterns, the manifest mean-level differences between boys and girls on both interest dimensions increased from T1 to T4 (see Figure 3). According to the fit indices (see Table 5), the latent mean levels for Conventional indicated a quadratic trajectory for boys and a linear trajectory for girls. For Enterprising, the latent mean levels indicated a quadratic trajectory for girls and boys.

Table 7

Differences in Means, Standard Deviations, and Mean-Level Changes between Boys and Girls

Int.	TL	Δ*M* [95% CI]	*d*	Δ*SD* [95% CI]	δ [95% CI]	*d*
	1	0.92 [0.82, 1.01]	0.86	0.01 [-0.06, 0.07]		
	2	0.86 [0.78, 0.94]	0.83	0.04 [-0.01, 0.10]	-0.05 [-0.16, 0.06]	-0.05
R	3	1.01 [0.92, 1.10]	1.01	0.00 [-0.06, 0.06]	0.15 [0.05, 0.25]	0.15
	4	1.11 [1.04, 1.19]	1.14	0.06 [0.01, 0.12]	0.10 [0.00, 0.20]	0.11
	14				0.20 [0.08, 0.32]	0.20
	1	0.21 [0.11, 0.31]	0.20	0.03 [-0.03, 0.08]		
	2	0.23 [0.16, 0.31]	0.23	0.04 [-0.01, 0.09]	0.02 [-0.09, 0.13]	0.02
I	3	0.35 [0.27, 0.44]	0.36	0.01 [-0.03, 0.06]	0.12 [0.03, 0.21]	0.12
	4	0.39 [0.31, 0.47]	0.40	0.05 [0.00, 0.09]	0.04 [-0.05, 0.13]	0.05
	14				0.18 [0.06, 0.30]	0.19
	1	-0.65 [-0.75, -0.55]	-0.68	0.10 [0.04, 0.16]		
	2	-0.72 [-0.79, -0.64]	-0.79	0.14 [0.10, 0.19]	-0.07 [-0.17, 0.03]	-0.07
A	3	-0.68 [-0.77, -0.60]	-0.75	0.08 [0.03, 0.13]	0.04 [-0.05, 0.12]	0.06
	4	-0.62 [-0.71, -0.53]	-0.69	0.04 [-0.01, 0.09]	0.07 [-0.01, 0.14]	0.08
	14				0.03 [-0.09, 0.16]	0.06
	1	-0.48 [-0.59, -0.37]	-0.47	0.05 [-0.01, 0.11]		
	2	-0.58 [-0.66, -0.50]	-0.59	0.11 [0.06, 0.17]	-0.10 [-0.22, 0.01]	-0.11
S	3	-0.59 [-0.68, -0.50]	-0.63	0.06 [0.00, 0.12]	-0.01 [-0.09, 0.07]	-0.01
	4	-0.73 [-0.81, -0.64]	-0.82	0.06 [0.00, 0.11]	-0.14 [-0.23, -0.04]	-0.15
	14				-0.25 [-0.37, -0.13]	-0.27
	1	0.05 [-0.06, 0.15]	0.04	0.03 [-0.03, 0.08]		
	2	0.04 [-0.04, 0.11]	0.04	0.04 [-0.01, 0.09]	-0.01 [-0.12, 0.10]	-0.01
E	3	0.10 [0.02, 0.19]	0.11	0.03 [-0.02, 0.08]	0.07 [-0.02, 0.16]	0.07
	4	0.17 [0.09, 0.25]	0.19	0.01 [-0.04, 0.06]	0.07 [-0.03, 0.16]	0.07
	14				0.12 [0.00, 0.24]	0.12
	1	0.06 [-0.03, 0.16]	0.06	0.06 [0.00, 0.11]		
	2	-0.02 [-0.11, 0.06]	-0.02	0.06 [0.00, 0.11]	-0.09 [-0.20, 0.03]	-0.08
C	3	0.12 [0.04, 0.20]	0.12	0.03 [-0.02, 0.08]	0.14 [0.05, 0.23]	0.15
	4	0.20 [0.13, 0.28]	0.22	0.02 [-0.03, 0.07]	0.09 [0.00, 0.17]	0.10
	14				0.14 [0.03, 0.26]	0.15

Note. Positive values imply higher values for boys. Int. = RIASEC interest dimensions; TL = time lag; Δ*M* = mean-level differences between boys and girls; *d* = Cohen's *d* for the mean-level differences between boys and girls; Δ*SD* = differences in standard deviations between boys and girls; δ = differences in mean-level differences between boys and girls; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional; 1–4 = time points 1 to 4; 14 = lag between T1 to T4.

Table 8

Descriptive Statistics and Mean-Level Change in RIASEC Scales in the Sample of Girls

Int.	TL	*M* [95% CI]	*SD* [95% CI]	Δ [95% CI]	*d*
	1	2.57 [2.48, 2.65]	1.07 [1.02, 1.11]		
	2	2.52 [2.46, 2.59]	1.01 [0.97, 1.05]	-0.04 [-0.12, 0.04]	-0.04
R	3	2.33 [2.26, 2.39]	1.01 [0.96, 1.05]	-0.20 [-0.27, -0.13]	-0.20
	4	2.12 [2.06, 2.18]	0.95 [0.90, 0.99]	-0.21 [-0.27, -0.14]	-0.21
	14			-0.45 [-0.54, -0.36]	-0.44
	1	2.99 [2.90, 3.09]	1.06 [1.02, 1.10]		
	2	2.86 [2.81, 2.92]	1.00 [0.96, 1.04]	-0.13 [-0.23, -0.04]	-0.13
I	3	2.68 [2.61, 2.74]	0.97 [0.93, 1.01]	-0.19 [-0.25, -0.12]	-0.19
	4	2.45 [2.39, 2.51]	0.95 [0.91, 0.98]	-0.23 [-0.29, -0.16]	-0.24
	14			-0.54 [-0.65, -0.43]	-0.54
	1	3.57 [3.49, 3.65]	0.91 [0.86, 0.96]		
	2	3.56 [3.51, 3.61]	0.84 [0.81, 0.88]	-0.01 [-0.08, 0.06]	-0.01
A	3	3.42 [3.36, 3.48]	0.87 [0.83, 0.91]	-0.14 [-0.20, -0.08]	-0.16
	4	3.25 [3.19, 3.31]	0.88 [0.85, 0.91]	-0.17 [-0.23, -0.12]	-0.20
	14			-0.32 [-0.41, -0.23]	-0.36
	1	3.44 [3.36, 3.53]	0.98 [0.94, 1.03]		
	2	3.54 [3.48, 3.60]	0.92 [0.88, 0.95]	0.10 [0.01, 0.19]	0.10
S	3	3.55 [3.48, 3.62]	0.90 [0.86, 0.95]	0.01 [-0.05, 0.07]	0.01
	4	3.65 [3.58, 3.72]	0.86 [0.82, 0.91]	0.10 [0.04, 0.16]	0.11
	14			0.21 [0.11, 0.30]	0.22
	1	2.87 [2.78, 2.96]	1.06 [1.01, 1.10]		
	2	2.95 [2.89, 3.01]	0.95 [0.92, 0.99]	0.08 [-0.01, 0.17]	0.08
E	3	2.89 [2.83, 2.95]	0.90 [0.86, 0.94]	-0.06 [-0.12, 0.01]	-0.06
	4	2.79 [2.73, 2.85]	0.88 [0.85, 0.91]	-0.10 [-0.17, -0.03]	-0.11
	14			-0.08 [-0.18, 0.02]	-0.08
	1	2.83 [2.75, 2.92]	1.01 [0.97, 1.05]		
	2	2.83 [2.76, 2.89]	0.97 [0.94, 1.01]	-0.01 [-0.10, 0.08]	-0.01
C	3	2.66 [2.60, 2.73]	0.94 [0.90, 0.98]	-0.17 [-0.23, -0.10]	-0.17
	4	2.52 [2.46, 2.59]	0.90 [0.87, 0.94]	-0.14 [-0.20, -0.07]	-0.15
	14			-0.31 [-0.40, -0.22]	-0.32

Note. n = 1,726; Int. = RIASEC interest dimensions; TL = time lag; M = mean level, SD = standard deviation; Δ = mean-level differences between the time points; d = Cohen's d for repeated measurement; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional; 1–4 = time points 1 to 4; 14 = lag between T1 to T4.

Table 9

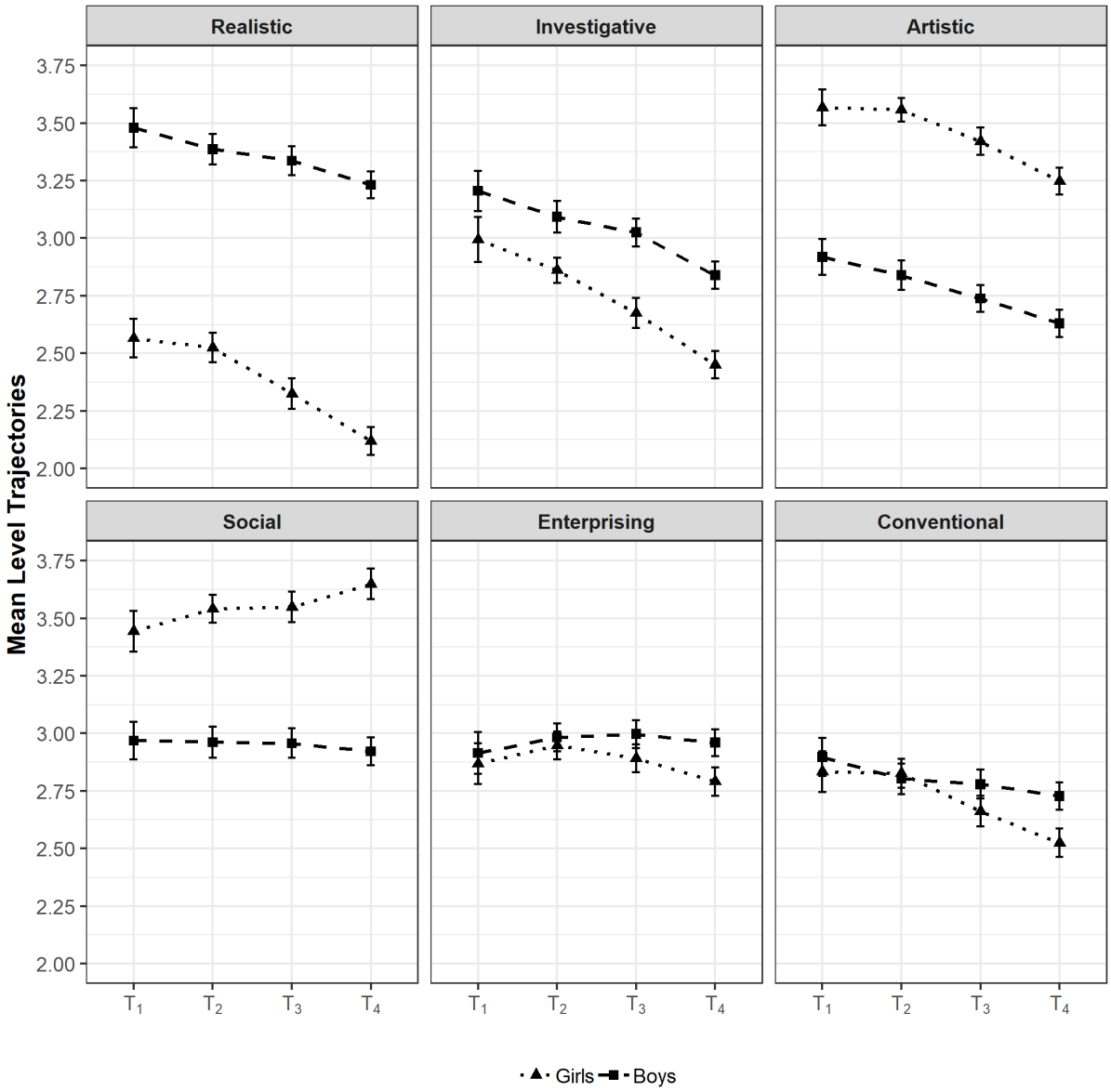
Descriptive Statistics and Mean-Level Change of RIASEC Scales in the Sample of Boys

Int.	TL	*M* [95% CI]	*SD* [95% CI]	Δ [95% CI]	*d*
	1	3.48 [3.40, 3.57]	1.07 [1.03, 1.11]		
	2	3.39 [3.32, 3.45]	1.06 [1.02, 1.10]	-0.09 [-0.17, -0.01]	-0.09
R	3	3.34 [3.27, 3.40]	1.01 [0.97, 1.04]	-0.05 [-0.12, 0.02]	-0.05
	4	3.23 [3.17, 3.29]	1.01 [0.97, 1.04]	-0.10 [-0.16, -0.05]	-0.10
	14			-0.25 [-0.34, -0.16]	-0.24
	1	3.21 [3.12, 3.29]	1.09 [1.05, 1.12]		
	2	3.09 [3.02, 3.16]	1.04 [1.01, 1.08]	-0.11 [-0.20, -0.03]	-0.11
I	3	3.03 [2.96, 3.09]	0.98 [0.95, 1.02]	-0.07 [-0.13, 0.00]	-0.07
	4	2.84 [2.78, 2.90]	0.99 [0.95, 1.03]	-0.19 [-0.25, -0.12]	-0.19
	14			-0.37 [-0.46, -0.28]	-0.35
	1	2.92 [2.84, 3.00]	1.01 [0.97, 1.05]		
	2	2.84 [2.78, 2.90]	0.98 [0.95, 1.02]	-0.08 [-0.16, -0.01]	-0.08
A	3	2.74 [2.68, 2.80]	0.95 [0.91, 0.99]	-0.10 [-0.16, -0.04]	-0.10
	4	2.63 [2.57, 2.69]	0.92 [0.88, 0.95]	-0.11 [-0.17, -0.05]	-0.12
	14			-0.29 [-0.37, -0.21]	-0.30
	1	2.97 [2.89, 3.05]	1.04 [0.99, 1.08]		
	2	2.96 [2.89, 3.03]	1.03 [0.99, 1.07]	-0.01 [-0.09, 0.07]	-0.01
S	3	2.96 [2.89, 3.02]	0.97 [0.93, 1.00]	-0.01 [-0.06, 0.05]	-0.01
	4	2.92 [2.86, 2.98]	0.92 [0.88, 0.96]	-0.04 [-0.10, 0.03]	-0.04
	14			-0.05 [-0.12, 0.03]	-0.05
	1	2.92 [2.82, 3.01]	1.08 [1.04, 1.12]		
	2	2.98 [2.92, 3.04]	0.99 [0.96, 1.03]	0.07 [-0.02, 0.15]	0.07
E	3	3.00 [2.94, 3.06]	0.93 [0.90, 0.96]	0.01 [-0.05, 0.08]	0.02
	4	2.96 [2.90, 3.02]	0.89 [0.86, 0.93]	-0.04 [-0.09, 0.02]	-0.04
	14			0.05 [-0.04, 0.13]	0.05
	1	2.90 [2.81, 2.98]	1.07 [1.03, 1.11]		
	2	2.80 [2.74, 2.87]	1.03 [1.00, 1.07]	-0.09 [-0.17, -0.02]	-0.09
C	3	2.78 [2.72, 2.84]	0.97 [0.93, 1.00]	-0.02 [-0.09, 0.04]	-0.02
	4	2.73 [2.67, 2.79]	0.93 [0.89, 0.97]	-0.05 [-0.11, 0.00]	-0.05
	14			-0.17 [-0.25, -0.09]	-0.17

Note. n = 2,098; Int. = RIASEC interest dimensions; TL = time lag; M = mean level, SD = standard deviation; Δ = mean-level differences between the time points; d = Cohen's d for repeated measurement; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional; 1–4 = time points 1 to 4; 14 = lag between T1 to T4.

Figure 3

Manifest Mean Levels of the Scale Scores for the RIASEC Dimensions for Girls and Boys



Note. 95% confidence intervals are depicted; T_{1-4} = time points 1–4; n_{girls} = 1,726; n_{boys} = 2,098.

Discussion

The current study used multiple indicators of continuity and change to examine the development of vocational interests over the course of late childhood and early adolescence (i.e., ages 11 to 14), in a sample in which the same participants were followed over four time points. The multiwave data made it possible to investigate mean-level trajectories and changes in stability coefficients as well as gender differences in vocational interests over time—information that is currently not available for the crucial period of late childhood and early adolescence. In addition, by investigating state and trait variance proportions new empirical evidence was generated that informs about the relative proportions of state and trait components in vocational interests.

Our study had three major findings: First, vocational interests showed moderate retest correlations (around $r = .50$) over the course of late childhood and early adolescence, and the retest correlations for the interest dimensions Realistic, Investigative, Social, and Conventional showed significant increases over time. In addition, the findings on state and trait variance proportions indicated that vocational interests generally possessed a substantial proportion of trait variance that further increased over time. Second, over the course of late childhood and early adolescence, Realistic, Investigative, Artistic, and Conventional, showed decreases in their mean levels. The effect sizes for these changes were small to moderate, and the latent mean-level trajectories indicated linear trends over time. Third, we found gender differences in the mean levels of all interest dimensions that increased over the course of late childhood and early adolescence for all interest dimensions—except Artistic.

Increasing Stability in Interests during Late Childhood and Early Adolescence

The moderate ($.41 < r < .64$) retest correlations in our study are in line with the meta-analytic findings from Low et al. (2005), who reported an average retest correlation of $r = .51$. In addition, the majority of the interest dimensions possessed retest correlations that increased with age.

The proportions of state and trait variance in the current study indicate that the majority of vocational interests possess a substantial trait component that increases with age. Because no other studies have investigated state and trait variance proportions in vocational interests, we will compare them to constructs that have a similar dispositional nature. Rieger et al. (2017) reported proportions of trait variance for personality traits using the same data set, indicating that the proportions of trait variance in personality traits are similar in magnitude to the ones reported for vocational interests. Rieger et al. (2017) also reported proportions of trait variance

for interests in school subjects that were averaged across all time points, with estimates ranging from .27 to .30. The proportions of trait variance found at T4 in the current study were similar in magnitude, ranging from .23 to .33, except for Realistic, which had a proportion of trait variance of .41.

Our results are in line with Holland's (1997) and the TSID model's (Su et al., 2019) theoretical assumptions. Holland (1997) stated that stability is fostered through cumulative learning on the basis of experiences, where preferences for single activities are gradually rewarded and reinforced. Similar assumptions are made by the TSID model, which implies that interests solidify through an accumulation of positive experiences or are altered by an accumulation of negative experiences.

This assumption would also be in line with Gottfredson (1981) because having experiences could help adolescents gain a better sense of their self-concept (i.e., their views of who they are, what is suitable for them, and what their skills are). An improved understanding of one's self-concept could also lead to a further solidification of interests. If adolescents have a better sense of who they are, what they like, and what they are good at, they may also have a better sense of which activities seem suitable to them.

In line with these thoughts, we assume that the increase in stability and the increase in the proportions of trait variance could be explained by an increase in opportunities to have autonomous experiences that are in line with one's interests. During early adolescence, school (e.g., increased focus on vocational orientation) and peer group environments (Larson & Richards, 1991) as well as relationships with parents (Collins & Russell, 1991; Keijsers & Poulin, 2013; Larson & Richards, 1991) undergo noticeable changes. Accompanied by increased autonomy from their parents (Keijsers & Poulin, 2013), these changes can provide adolescents with a larger range of opportunities to experience new activities (e.g., extracurricular activities). In addition, adolescents become increasingly able to self-select suitable environments and to reengage in the activities they are interested in. In line with Holland (1997) and the TSID model (Su et al., 2019), such an increase in self-determination during the transition from childhood to adolescence should lead to a solidification of vocational interests (Ryan & Deci, 2000).

In addition to this general trend of interest solidification, we also found differences between the six interest dimensions. The Realistic, Investigative, Social, and Conventional dimensions increased in stability, but this was not the case for Artistic or Enterprising. These differences illustrate that not all interest dimensions solidify in the same way over the course of late childhood and early adolescence. The six RIASEC dimensions capture different

categories of activities of which some might be more accessible to adolescents than others. If solidification is fostered by having experiences, it can be assumed that activities that are experienced less often or that are less accessible develop differently. For example, we found increases in stability for Realistic and Social but not for Enterprising. Realistic activities often comprise practical tasks (i.e., building things out of wood, playing outside, or creating things with one's hands) that are already familiar to children and adolescents. Similarly, activities that are related to Social interests mainly describe interactions between people, indicating that Realistic and Social activities are more prominent in the lives of children and young adolescents. By contrast, Enterprising activities, such as leading a group, selling things, and organizing events, are less prominent in the lives of children and young adolescents. Consequently, adolescents might have less opportunities to experience this type of activities, the mental representation of Enterprising activities might be less developed in younger ages, and the respective interest dimension might therefore be less stable (Su et al., 2019).

Decreasing Mean Levels during Late Childhood and Early Adolescence

In line with our expectations, the majority of Holland's (1997) RIASEC dimensions decreased over the course of late childhood and early adolescence, a finding also reported by Hoff et al. (2018) and Päßler and Hell (2020). However, in comparison with Hoff et al. (2018), we found significant decreases in four instead of two RIASEC dimensions and larger effect sizes (Current study: $-0.44 < d < -0.01$; Hoff et al., 2018: $-0.30 < d < -0.02$). Päßler and Hell (2020) found results that were more similar to our study as they also reported significant mean-level decreases for four of the six RIASEC dimensions. In the current study, we found mean-level decreases in Realistic, Investigative, Artistic, and Conventional interests, whereas Päßler and Hell (2020) reported mean-level decreases in Realistic, Investigative, Artistic, and Social interests.

We argue that these overall mean-level decreases could indicate a differentiation process by which adolescents gradually figure out which activities and occupations they do not like—similar explanations were already provided by Tracey (2002) and Krapp (2002). The differentiation could be initiated on the basis of the process of circumscription (Gottfredson, 1981). Adolescents gradually figure out which activities they do not like by circumscribing possible occupational choices on the basis of how well the activities fit with their developing self-concept. This implies that over time, fewer occupations seem suitable to them. This misfit between one's view of oneself and the view of possible occupations could lead to a decrease in interest intensity and consequently to a decrease in the mean levels of vocational interests.

A different explanation can be derived from personality research, where decreases in mean levels of conscientiousness, agreeableness, and openness have been reported over the course of late childhood and early adolescence (Denissen et al., 2013; Soto et al., 2011; Soto & Tackett, 2015; Van den Akker et al., 2014). The temporary declines in the otherwise increasing trajectories of personality traits during adolescence are described as a disruption period that is assumed to originate from the physiological, social, and normative changes that occur during that time (Soto & Tackett, 2015). Hoff et al. (2018) and Päßler and Hell (2020) argued that the disruption hypothesis should be extended to vocational interests. They stated that due to the challenges that occur during early adolescence, at least one of the basic needs of competence, autonomy, and relatedness—which are necessary for developing interest (Ryan & Deci, 2000)—cannot be regularly fulfilled, which results in a general decrease in interest.

In contrast to our assumptions, we found mean-level increases in Social interest for girls, a finding that was in line with Hoff et al. (2018) and that differed from Päßler and Hell (2020), where the mean levels of Social interest decreased over time. Although we had expected all interest dimensions to decrease, there might be plausible reasons for why Social interests, by contrast, increased in our sample for girls. People with higher scores on Social interests usually prefer activities such as listening to someone's problems, helping other people, or engaging in social situations (Holland, 1997). They also tend to be characterized as empathic, friendly, and sociable (Holland, 1997). There is evidence that similar behaviors and characteristics seem to be relevant in social interactions and especially for popularity in peer groups (Parkhurst & Hopmeyer, 1998). Late childhood and early adolescence are characterized by an increase in the orientation toward (Berndt, 1979; Fuligni et al., 2001) and the time spent in peer groups (Larson & Richards, 1991). Therefore, it could be assumed that increases in peer-group importance are accompanied by increases in Social interests—which might be beneficial for integration into and socialization in peer groups. These increases in Social interests may also resemble general increases in people-oriented interests (i.e., Social, Artistic, and Enterprising interests) reported by Hoff et al. (2018) for late adolescence and older age groups. In returning to the multidimensionality of vocational interests, this suggests that differences in developmental patterns might also be influenced by the functions these interests can fulfill in adolescents or the outcomes that are associated with these interests.

Increasing Gender Differences during Late Childhood and Early Adolescence

We found gender differences for Realistic, Investigative, Social, and Artistic as early as fifth grade. Similar findings were reported by Tracey (2002) and Päßler and Hell (2020). These

findings are in line with Gottfredson (1981), who proposed that gender differences begin to manifest at the age of 6. In addition, we also found increases in gender differences in all interest dimensions except Artistic. These findings were in line with Hoff et al. (2018), who reported increases in gender differences for Realistic and Social interests. However, the current study is the first to report increases in gender differences for a majority of interest dimensions. At T4, we found gender differences on all interest dimensions, with Realistic, Investigative, Enterprising, and Conventional demonstrating higher values for boys, and Artistic and Social demonstrating higher values for girls. These results were largely in line with the findings of the meta-analysis by Su et al. (2009), who reported higher mean levels for girls on Artistic, Social, and Conventional and higher mean levels for boys on Realistic, Investigative, and Enterprising.

Increases in gender differences over the course of late childhood and early adolescence are in line with the increased gender orientation in this life phase described by Gottfredson (1981). Children and adolescents incorporate gender as an important aspect of their self and are consequently more drawn to activities and occupations that represent the matching gender. In addition, gender differences in vocational interests could be strengthened by an increase in peer-group orientation during late childhood and early adolescence (Berndt, 1979; Fuligni et al., 2001; Larson & Richards, 1991). There is evidence that peer groups during late childhood are either more female-dominant or male-dominant (Rubin et al., 2006; Wilson et al., 2011). Therefore, children and adolescents might not only incorporate gender into their self-concept, but they might also socialize within same-sex peer groups, which might in turn contribute to the development of more gendered interests.

Integrating Current Findings into the Life Course Development of Vocational Interests

The current study complements a recent trend of emerging multiwave studies in vocational interest research (Hoff et al., 2020; Päßler & Hell, 2020; Stoll, Rieger, et al., 2020). Taken together, the findings from these multiwave longitudinal studies provide information about the pattern of vocational interest development that spans an age range from late childhood and early adolescence (ages 10 to 12; Päßler & Hell, 2020), late adolescence to young adulthood (ages 16 to 24; Hoff et al., 2020), and a 10-year time period across young adulthood (ages 19 to 29; Stoll, Rieger, et al., 2020). The present study investigated the age range of 11 to 14 and therefore extends the existing longitudinal findings.

Regarding stabilities in vocational interests the current study is the first to demonstrate that the stability of vocational interests has already begun to increase at younger ages. This indicates that increases in stability are already beginning as early as late childhood, continue

afterwards from late adolescence to young adulthood (Hoff et al., 2020; Stoll, Rieger, et al., 2020), and halt during young adulthood (Stoll, Rieger, et al., 2020). However, increases in stability during late childhood and early adolescence are relatively small in comparison with the increases that occur in later life stages (Hoff et al., 2020).

Regarding state and trait variance components of vocational interests, the current study provides first insights that trait variance proportions increase across late childhood and early adolescence. However, as proposed by Su et al. (2019), also state variance proportions were substantial, especially during late childhood, suggesting that vocational interests can consist of both, state and trait components. Although empirical evidence is missing so far, according to Holland (1997), we would assume that trait variance proportions further increase over the life course, while state variance proportions further decrease. More research with older samples is needed to test these assumptions.

Regarding normative change in vocational interests, the current study supports previous findings for late childhood and early adolescence. Taken together, the longitudinal mean-level trajectories indicate that the mean levels of vocational interests decrease over the course of late childhood and early adolescence (Päßler & Hell, 2020), increase from late adolescence to young adulthood (Hoff et al., 2020; Stoll, Rieger, et al., 2020), and remain relatively stable during young adulthood (Stoll, Rieger, et al., 2020). This suggests that the disruption hypothesis, which was originally proposed in personality trait research, can also be applied to vocational interests (for a similar argumentation, see Hoff et al., 2018; Päßler & Hell, 2020).

Previous findings have indicated that gender differences are already present during late childhood and early adolescence (Päßler & Hell, 2020) and remain relatively stable across young adulthood (Stoll, Rieger, et al., 2020). However, the results of the current study suggest that gender differences in vocational interests increase during late childhood and early adolescence. This finding was very prominent as all interest dimensions, except Artistic, exhibited increases in gender differences over time.

In regards to the life span development of vocational interests, our study provides important information on the developmental pattern in late childhood and early adolescence. However, the picture is not yet complete and more longitudinal studies are needed, that focus on other relevant life phases—including for example the entrance in school or the transition from work life to retirement.

Limitations and Future Outlook

This study has several strengths such as a large multiwave longitudinal sample that captures an important—and so far, understudied—life phase, as well as the investigation of three different indicators of stability and change. However, there are some limitations that should be mentioned. First, the students included in our sample were from the low and intermediate school tracks in Germany, whereas academic track students were not part of the sample. This could limit the generalizability of our results. However, the general lower, intermediate, and higher tracks in Germany are not tracked by content type (e.g., Technical, Economic, or Nursing School), which indicates that although school tracks differ in their performance requirements, they do not differ in the content-related experiences they offer. In addition, the theory of vocational interests does not propose different developmental trajectories for students from different academic or school tracks. Nevertheless, future studies on interest development in younger age groups should also include students from the highest school track in order to capture the entire student population.

Second, only partial measurement invariance could be achieved across the Time × Gender interaction, indicating that some of the items were interpreted differentially by boys and girls (Putnick & Bornstein, 2016). For example, we found intercept differences between boys and girls on the item "studying the behavior of animals and plants". This means that boys and girls do not attribute the same meaning and interpretation to this item, for example, girls could interpret this as an activity that focuses on reading a biology book, which would entail Artistic elements because reading is involved. Boys, on the other hand, could view the activity as physical because studying the behavior of animals and plants is usually done in nature. However, valid inferences about differences in latent means can still be drawn as long as two factor loadings and the intercepts of the respective factor model are invariant across groups (Byrne et al., 1989; van de Schoot et al., 2012). In addition, Schmitt et al. (2011) demonstrated only small differences in latent mean levels between the full and partial measurement invariance models across groups—when partial invariance was applied. This indicates that partial measurement invariance should not have largely affected our results because the majority of the item intercepts were invariant across groups, indicating that, overall, boys and girls had the same understanding of the six dimensions of vocational interests.

Third, the goal of the current study was to investigate mean-level change, retest correlations and state-trait variance proportions of vocational interests in a descriptive way. We did not probe for the effects of specific factors that might have driven continuity and change in vocational interests. Accordingly, we can only speculate about the reasons and processes that

led to the changes in mean levels, retest correlations, and proportions of state and trait variance. Future studies should focus on the factors that could influence vocational interest development. In the current study, we discussed several mechanisms that might influence vocational interest development. However, future research will need to investigate these mechanisms to provide further insights into the processes that drive vocational interest development.

Conclusion

The present study illustrates that vocational interest development over the course of late childhood and early adolescence can be characterized by an increase in stability, an increase in trait variance proportions, a decrease in mean levels, and an increase in gender differences. Our results demonstrate that longitudinal studies based on multiwave data can provide crucial descriptive information and improve the understanding of vocational interest development. These descriptive insights could lay the groundwork for future research that focuses on underlying processes and predictors of mean-level changes and stability in vocational interests.

Data Accessibility Statement

Because of reasons of data protection, the dataset used in the present investigation is not openly accessible, but the variance-covariance matrices as well as mean vectors of the variables can be found in the analysis scripts that are accessible through the following link: https://osf.io/tuys8/?view_only=cb07ca448fdf4dce998e250ae635a419.

References

Ackerman, P. L. (1997). Personality, Self-Concept, Interests, and Intelligence: Which Construct Doesn't Fit? *Journal of Personality*, *65*(2), 171–205. https://doi.org/10.1111/j.1467-6494.1997.tb00952.x

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, Personality, and Interests: Evidence for Overlapping Traits. *Psychological Bulletin*, *121*(2), 219–245. https://doi.org/10.1037/0033-2909.121.2.219

Allen, J., & Robbins, S. B. (2008). Prediction of college major persistence based on vocational interests, academic preparation, and first-year academic performance. *Research in Higher Education*, *49*, 62–79. https://doi.org/10.1007/s11162-007-9064-5

Asparouhov, T., & Muthen, B. (2008). *Auxiliary variables predicting missing data*. https://www.statmodel.com/download/AuxM2.pdf

Bergmann, C., & Eder, F. (2005). *Allgemeiner Interessen-Struktur-Test mit Umwelt-Struktur-Test (UST-R)—Revision [General Interest Structure Test and Environmental Structure Test—Revision]*. Beltz.

Berndt, T. J. (1979). Developmental changes in conformity to peers and parents. *Developmental Psychology*, *15*(6), 608–616.

Bishop, J., Geiser, C., & Cole, D. A. (2015). Modeling latent growth with multiple indicators: A comparison of three approaches. *Psychological Methods*, *20*(1), 43–62. https://doi.org/10.1037/met0000018

Braun, L., Göllner, R., Rieger, S., Trautwein, U., & Spengler, M. (2020). How state and trait versions of self-esteem and depressive symptoms affect their interplay: A longitudinal experimental investigation. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspp0000295

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Methods*, *105*(3), 456–466. https://doi.org/10.1037/0033-2909.105.3.456

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Cole, D. A., Maxwell, S. E., Martin, J. M., Peeke, L. G., Seroczynski, A. D., Tram, J. M., Hoffman, K. B., Ruiz, M. D., Jacquez, F., & Maschman, T. (2001). The development of multiple domains of child and adolescent self-concept: A cohort sequential longitudinal design. *Child Development*, *72*(6), 1723–1746. https://doi.org/10.1111/1467-8624.00375

Collins, L. M., Schafer, J. L., & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, *6*(4), 330–351. https://doi.org/10.1037/1082-989X.6.4.330

Collins, W. A., & Russell, G. (1991). Mother-child and father-child relationships in middle childhood and adolescence: A developmental analysis. *Developmental Review*, *11*(2), 99–136. https://doi.org/10.1016/0273-2297(91)90004-8

Denissen, J. J. A., van Aken, M. A. G., Penke, L., & Wood, D. (2013). Self-regulation underlies temperament and personality: An integrative developmental framework. *Child Development Perspectives*, *7*(4), 255–260. https://doi.org/10.1111/cdep.12050

Denissen, J. J. A., Zarrett, N. R., & Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: Longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Development*, *78*(2), 430–447. https://doi.org/10.1111/j.1467-8624.2007.01007.x

Enders, C. K. (2001). The impact of nonnormality on full information maximum-likelihood estimation for structural equation models with missing data. *Psychological Methods*, *6*(4), 352–370. https://doi.org/10.1037//I082-989X.6.4.352

Enders, C. K. (2008). A note on the use of missing auxiliary variables in full information maximum likelihood-based structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *15*(3), 434–448. https://doi.org/10.1080/10705510802154307

Fuligni, A. J., Eccles, J. S., Barber, B. L., & Clements, P. (2001). Early adolescent peer orientation and adjustment during high school. *Developmental Psychology*, *37*(1), 28–36. https://doi.org/10.1037/0012-1649.37.1.28

Gaspard, H., Häfner, I., Parrisius, C., Trautwein, U., & Nagengast, B. (2017). Assessing task values in five subjects during secondary school: Measurement structure and mean level

differences across grade level, gender, and academic subject. *Contemporary Educational Psychology*, *48*, 67–84. https://doi.org/10.1016/j.cedpsych.2016.09.003

Gati, I., Levin, N., & Landman-Tal, S. (2019). Decision-Making Models and Career Guidance. In J. A. Athanasou & H. N. Perera (Eds.), *International Handbook of Career Guidance* (2nd ed.). Springer International Publishing.

Geiser, C., Keller, B. T., Lockhart, G., Eid, M., Cole, D. A., & Koch, T. (2015). Distinguishing state variability from trait change in longitudinal data: The role of measurement (non)invariance in latent state-trait analyses. *Behavior Research Methods*, *47*(1), 172–203. https://doi.org/10.3758/s13428-014-0457-z

Gottfredson, L. S. (1981). Circumscription and compromise: A developmental theory of occupational aspirations. *Journal of Counseling Psychology*, *28*(6), 545–579. https://doi.org/10.1037/0022-0167.28.6.545

Graham, J. W. (2003). Adding missing-data-relevant variables to FIML-based structural equation models. *Structural Equation Modeling*, *10*(1), 80–100. https://doi.org/10.1207/S15328007SEM1001_4

Gysbers, N. C. (2005). Comprehensive school guidance programs in the United States: A career profile. *International Journal for Educational and Vocational Guidance*, *5*(2), 203–215. https://doi.org/10.1007/s10775-005-8800-7

Hidi, S., & Renninger, K. A. (2006). The Four-Phase Model of Interest Development. *Educational Psychologist*, *41*, 111–127. https://doi.org/10.1207/s15326985ep4102_4

Hoff, K. A., Briley, D. A., Wee, C. J. M., & Rounds, J. (2018). Normative changes in interests from adolescence to adulthood: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *144*(4), 426–451. https://doi.org/10.1037/bul0000140

Hoff, K. A., Song, Q. C., Einarsdóttir, S., Briley, D. A., & Rounds, J. (2020). Developmental structure of personality and interests: A four-wave, 8-year longitudinal study. *Journal of Personality and Social Psychology*, *118*(5), 1044–1064. https://doi.org/10.1037/pspp0000228

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Psychological Assessment Resources.

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

Jonkmann, K., Rose, N., & Trautwein, U. (2013). *Tradition und Innovation: Entwicklungsverläufe an Haupt- und Realschulen in Baden-Württemberg und Mittelschulen in Sachsen – Abschlussbericht für die Länder Baden-Württemberg und Sachsen [unpublished project report].* Universität Tübingen.

Keijsers, L., & Poulin, F. (2013). Developmental changes in parent–child communication throughout adolescence. *Developmental Psychology*, *49*(12), 2301–2308. https://doi.org/10.1037/a0032217

Knapp, R. R., & Knapp, L. (1984). *COPS Interest Inventory manual.* EdITS.

Kracke, B. (1997). Parental behaviors and adolescents' career exploration. *The Career Development Quarterly*, *45*, 341–350. https://doi.org/10.1002/j.2161-0045.1997.tb00538.x

Krapp, A. (2002). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction*, *12*(4), 383–409. https://doi.org/10.1016/S0959-4752(01)00011-1

Kuder, G. F. (1964). *Kuder General Interest Survey—Form E.* Science Research Associates.

Kuder, G. F. (1975). *Kuder General Interest Survey manual.* Science Research Associates.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 1–12. https://doi.org/10.3389/fpsyg.2013.00863

Larson, R., & Richards, M. H. (1991). Daily companionship in late childhood and early adolescence: Changing developmental contexts. *Child Development*, *62*(2), 284–300. https://doi.org/10.1111/j.1467-8624.1991.tb01531.x

Low, K. S. D., & Rounds, J. (2007). Interest change and continuity from early adolescence to middle adulthood. *International Journal for Educational and Vocational Guidance*, *7*(1), 23–36. https://doi.org/10.1007/s10775-006-9110-4

Low, K. S. D., Yoon, M., Roberts, B. W., & Rounds, J. (2005). The stability of vocational interests from early adolescence to middle adulthood: A quantitative review of

longitudinal studies. *Psychological Bulletin*, *131*(5), 713–737. https://doi.org/10.1037/0033-2909.131.5.713

Lubinski, D., Benbow, C. P., & Ryan, J. (1995). Stability of vocational interests among the intellectually gifted from adolescence to adulthood: A 15-year longitudinal study. *Journal of Applied Psychology*, *80*(1), 196–200. https://doi.org/10.1037/0021-9010.80.1.196

McArdle, J. J., & Epstein, D. (1987). Latent growth curves within developmental structural equation models. *Child Development*, *58*(1), 110–133. https://doi.org/10.2307/1130295

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, *23*(3), 412–433. https://doi.org/10.1037/met0000144

McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods*, *22*(1), 114–140. https://doi.org/10.1037/met0000078

Mount, M. K., Barrick, M. R., Scullen, S. M., & Rounds, J. (2005a). Higher Order Dimensions of the Big Five Personality Traits and the Big Six Vocational Interest Types. *Personnel Psychology*, *58*(2), 447–478. https://doi.org/10.1111/j.1744-6570.2005.00468.x

Muthén, B., & Muthén, L. K. (1998-2017). *Mplus User's Guide* (8th ed.). Muthén & Muthén.

Noack, P., Kracke, B., Gniewosz, B., & Dietrich, J. (2010). Parental and school effects on students' occupational exploration: A longitudinal and multilevel analysis. *Journal of Vocational Behavior*, *77*(1), 50–57. https://doi.org/10.1016/j.jvb.2010.02.006

Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2017). Interest congruence and performance: Revisiting recent meta-analytic findings. *Journal of Vocational Behavior*, *98*, 138–151. https://doi.org/10.1016/j.jvb.2016.11.002

Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2012). Vocational interests and performance: A quantitative summary of over 60 years of research. *Perspectives on Psychological Science*, *7*(4), 384–403. https://doi.org/10.1177/1745691612449021

Parkhurst, J. T., & Hopmeyer, A. (1998). Sociometric popularity and peer-perceived popularity: Two distinct dimensions of peer status. *The Journal of Early Adolescence*, *18*(2), 125–144. https://doi.org/10.1177/0272431698018002001

Päßler, K., & Hell, B. (2012). Do interests and cognitive abilities help explain college major choice equally well for women and men? *Journal of Career Assessment*, *20*(4), 479–496. https://doi.org/10.1177/1069072712450009

Päßler, K., & Hell, B. (2020). Stability and change in vocational interests from late childhood to early adolescence. *Journal of Vocational Behavior*, *121*, 103462. https://doi.org/10.1016/j.jvb.2020.103462

Petersen, A. C. (1987). The nature of biological-psychosocial interactions: The sample case of early adolescence. In R. M. Lerner & T. T. Foch (Eds.), *Child psychology. Biological-psychosocial interactions in early adolescence* (pp. 35–61). Lawrence Erlbaum Associates, Inc.

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Renninger, K. A., & Hidi, S. (2011). Revisiting the Conceptualization, Measurement, and Generation of Interest. *Educational Psychologist*, *46*, 168–184. https://doi.org/10.1080/00461520.2011.587723

Rieger, S., Göllner, R., Spengler, M., Trautwein, U., Nagengast, B., & Roberts, B. W. (2017). Social Cognitive Constructs Are Just as Stable as the Big Five Between Grades 5 and 8. *AERA Open*, *3*(3), 1–9. https://doi.org/10.1177/2332858417717691

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

Rounds, J., & Su, R. (2014). The nature and power of interests. *Current Directions in Psychological Science*, *23*(2), 98–103. https://doi.org/10.1177/0963721414522812

Rubin, K. H., Bukowski, W. M., & Parker, J. G. (2006). Peer interactions, relationships, and groups. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of Child Psychology* (pp. 571–645). John Wiley & Sons, Inc. https://doi.org/10.1002/9780470147658.chpsy0310

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Schmitt, N., Golubovich, J., & Leong, F. T. L. (2011). Impact of measurement invariance on construct correlations, mean differences, and relations with external correlates: An illustrative example using Big Five and RIASEC measures. *Assessment*, *18*(4), 412–427. https://doi.org/10.1177/1073191110373223

Schmukle, S. C., & Egloff, B. (2005). A Latent State-Trait Analysis of Implicit and Explicit Personality Measures. *European Journal of Psychological Assessment*, *21*(2), 100–107. https://doi.org/10.1027/1015-5759.21.2.100

Silvia, P. J. (2001). Interest and Interests: The Psychology of Constructive Capriciousness. *Review of General Psychology*, *5*(3), 270–290. https://doi.org/10.1037/1089-2680.5.3.270

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology*, *100*(2), 330–348. https://doi.org/10.1037/a0021717

Soto, C. J., & Tackett, J. L. (2015). Personality Traits in Childhood and Adolescence: Structure, Development, and Outcomes. *Current Directions in Psychological Science*, *24*(5), 358–362. https://doi.org/10.1177/0963721415589345

Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state-trait theory and research in personality and individual differences. *European Journal of Personality*, *13*(5), 389–408. https://doi.org/10.1002/(SICI)1099-0984(199909/10)13:5<389::AID-PER361>3.0.CO;2-A

Stoll, G., Einarsdóttir, S., Chelsea Song, Q., Ondish, P., Sun, J.-T., & Rounds, J. (2020). The Roles of Personality Traits and Vocational Interests in Explaining What People Want out of Life. *Journal of Research in Personality*, *86*, 103939. https://doi.org/10.1016/j.jrp.2020.103939

Stoll, G., Rieger, S., Lüdtke, O., Nagengast, B., Trautwein, U., & Roberts, B. W. (2017). Vocational interests assessed at the end of high school predict life outcomes assessed 10 years later over and above IQ and Big Five personality traits. *Journal of Personality and Social Psychology*, *113*(1), 167–184. https://doi.org/10.1037/pspp0000117

Stoll, G., Rieger, S., Nagengast, B., Trautwein, U., & Rounds, J. (2020). Stability and change in vocational interests after graduation from high school: A six-wave longitudinal study. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspp0000359

Su, R. (2012). *Power of vocational interests and interest congruence in predicting career success [unpublished doctoral dissertation]*. https://www.ideals.illinois.edu/bitstream/handle/2142/34329/Su_Rong.pdf?sequence=1&isAllowed=y

Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological Bulletin*, *135*(6), 859–884. https://doi.org/10.1037/a0017364

Su, R., Stoll, G., & Rounds, J. (2019). The nature of interests: Toward a unifying theory of trait-situation interest dynamics. In C. D. Nye & J. Rounds (Eds.), *Vocational interests: Rethinking their role in understanding workplace behavior and practice* (pp. 11–38). Taylor & Francis/Routledge. https://doi.org/10.4324/9781315678924-2

Tracey, T. J. G. (2002). Development of interests and competency beliefs: A 1-year longitudinal study of fifth- to eighth-grade students using the ICA-R and structural equation modeling. *Journal of Counseling Psychology*, *49*(2), 148–163. https://doi.org/10.1037//0022-0167.49.2.148

Tracey, T. J. G., & Ward, C. C. (1998). The structure of children's interests and competence perceptions. *Journal of Counseling Psychology*, *45*(3), 290–303. https://doi.org/10.1037/0022-0167.45.3.290

Usslepp, N., Hübner, N., Stoll, G., Spengler, M., Trautwein, U., & Nagengast, B. (2020). RIASEC interests and the Big Five personality traits matter for life success—But do they already matter for educational track choices? *Journal of Personality*, *88*(5), 1007–1024. https://doi.org/10.1111/jopy.12547

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

Van den Akker, A. L., Deković, M., Asscher, J., & Prinzie, P. (2014). Mean-level personality development across childhood and adolescence: A temporary defiance of the maturity

principle and bidirectional associations with parenting. *Journal of Personality and Social Psychology*, *107*(4), 736–750. https://doi.org/10.1037/a0037248

Van Iddekinge, C. H., Roth, P. L., Putka, D. J., & Lanivich, S. E. (2011). Are you interested? A meta-analysis of relations between vocational interests and employee performance and turnover. *Journal of Applied Psychology*, *96*(6), 1167–1194. https://doi.org/10.1037/a0024343

von Maurice, J. (2006). *ICA-D. Deutschsprachige Version des Inventory of Children's Activities – Revised (ICA-R, Tracey & Ward, 1998)[unpublished instrument]*. Otto-Friedrich-Universität Bamberg.

Whiston, S. C., & Keller, B. K. (2004). The influences of the family of origin on career development: A review and analysis. *The Counseling Psychologist*, *32*(4), 493–568. https://doi.org/10.1177/0011000004265660

Wille, E., Stoll, G., Gfrörer, T., Cambria, J., Nagengast, B., & Trautwein, U. (2020). It takes two: Expectancy-value constructs and vocational interests jointly predict STEM major choices. *Contemporary Educational Psychology*, *61*, 101858. https://doi.org/10.1016/j.cedpsych.2020.101858

Wilson, T., Karimpour, R., & Rodkin, P. C. (2011). African American and European American students' peer groups during early adolescence: Structure, status, and academic achievement. *The Journal of Early Adolescence*, *31*(1), 74–98. https://doi.org/10.1177/0272431610387143

Yuan, K.-H., & Bentler, P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology*, *30*(1), 165–200. https://doi.org/10.1111/0081-1750.00078

Zytowski, D. G. (1976). Long-term prrofile stability of the Kuder Occupational Interest Survey. *Educational and Psychological Measurement*, *36*(3), 689–692. https://doi.org/10.1177/001316447603600315

Supplement Material

Supplement A: Additional Sample Information

Sample Composition

The number of participants who attended lower track schools was n = 1,678 (43.2% of participants who participated in the study at a minimum of one time point) and the number of participants who attended intermediate track schools was n = 881 (22.7% of participants who participated in the study at a minimum of one time point). Both school tracks—lower and intermediate—came from the state of Baden-Württemberg. The number of participants who attended multitrack schools (a combination of the lower and intermediate tracks) was n = 1,321 (34.0% of participants who participated in the study at a minimum of one time point). All of the participants from multitrack schools were from the state of Saxony because the format of multitrack schools did not exist in the state of Baden-Württemberg during that time. Participants who attended the highest school track in Germany were not included in the study.

For 87.6% of the participants who participated in the study at a minimum of one time point, background information was available on where they were born. A total of 93.4% of these participants were born in Germany. For 86.0% of the participants who participated in the study at a minimum of one time point, information about their migration status was available. Of these participants, 66.7% had no migration background, 11.8% reported that at least one parent was not born in Germany, 16.0% reported that both parents were not born in Germany but the participant was, and 5.5% reported that both parents and the participant were not born in Germany.

To investigate differences between participants who joined the study late and those who dropped out of the study, we compared joiners with the participants who were already in the study and dropouts with the participants who continued to participate in the study. This was done for every time point separately. Information about participants who joined and dropped out of the study can be found in Table A1.

Table A1

Information about Participants who Joined the Study Late and Dropped out of the Study

	T1 (Grade 5)	T2 (Grade 6)	T3 (Grade 7)	T4 (Grade 8)	Pooled data
Complete sample all variables	2,894	2,936	2,993	3,060	3,876
Dropped out of the study		-226	-359	-235	
Changed classes		8 (3.5%)	28 (7.8%)	10 (4.3%)	
Moved		63 (27.9%)	72 (20.1%)	66 (28.1%)	
Repeated a grade		39 (17.3%)	50 (13.9%)	46 (19.6%)	
Other / no information		116 (51.3%)	209 (58.2%)	113 (48.1%)	
Joined the study later		+268	+416	+302	
Changed classes		18 (6.7%)	98 (23.6%)	8 (2.6%)	
Moved		107 (39.9%)	65 (15.6%)	93 (30.8%)	
Repeated a grade		50 (18.7%)	28 (6.7%)	74 (24.5%)	
Other / no information		93 (34.7%)	225 (54.0%)	127 (42.1%)	

Note. Rieger et al. (2017) provide similar information.

Attrition Analysis

In Table A2, mean-level differences in terms of effect sizes (Cohen's *d*) between participants who dropped out of the study and participants who continued to participate in the study at the subsequent time point are reported. In terms of effect sizes, achievement-related variables showed high to moderate differences, school motivation constructs showed small to moderate differences, and personality characteristics showed small differences. A similar picture could be found for differences between participants who joined the study late and participants who were already part of the study. All variables that indicated significant differences in terms of effect sizes were included as auxiliary variables in all analyses.

Table A2
Differences (Cohen's d) between Students who Dropped Out and Continued

	T1 » T2	T2 » T3	T3 » T4
Dropouts			
Gender (male = 1)	OR = 1.015	OR = 1.087	OR = **1.134**
Socioeconomic status (HISEI)	0.02	**-0.18**	0.03
migration	OR = **3.023**	OR = **1.965**	OR = **2.563**
Math grade	**0.20**	**0.48**	**0.69**
German grade	0.11	**0.58**	**0.70**
English grade	0.16	**0.31**	**0.61**
Math achievement	**-0.22**	**-0.34**	**-0.56**
German achievement	**-0.21**	**-0.23**	**-0.40**
Vocational interests			
Realistic	0.01	0.00	-0.04
Investigative	0.11	-0.10	-0.12
Artistic	**0.36**	-0.05	**-0.19**
Social	**0.20**	-0.02	-0.12
Enterprising	0.14	0.09	0.02
Conventional	**0.30**	0.00	-0.05
Big Five			
Conscientiousness	0.17	-0.11	**-0.32**
Neuroticism	**0.17**	0.09	**0.27**
Openness	**0.21**	-0.07	**-0.27**
Agreeableness	**0.18**	0.02	**-0.21**
Extraversion	0.12	0.01	-0.17
Other constructs			
Self-concept in math	0.00	**-0.15**	**-0.39**
Self-concept in German	0.13	**-0.14**	-0.01
Self-concept in English	0.10	0.00	**-0.28**
Interest in math	**0.20**	-0.09	**-0.21**
Interest in German	**0.20**	0.07	0.05
Interest in English	0.12	0.05	-0.15
Effort in math	0.08	**-0.23**	**-0.35**
Effort in German	0.06	-0.08	-0.14
Effort in English	-0.03	-0.05	**-0.29**

Note. Bold parameters are significantly different from zero at $p < .05$, two-tailed. Positive values imply higher values for the participants who dropped out of the study, significant variables were included as auxiliaries in the main analysis; OR = Odds Ratio.

Table A3

Differences (Cohen's d) between Students who Joined Late and those who were already in the Study

	T1 » T2	T2 » T3	T3 » T4
Joined the study late			
Gender (male = 1)	OR = **1.362**	OR = 0.995	OR = 1.093
Socioeconomic status (HISEI)	-0.14	**0.14**	0.11
migration	OR = 1.191	OR = 1.122	OR = 1.312
Math grade	**0.68**	**0.29**	**0.84**
German grade	**0.60**	**0.37**	**0.65**
English grade	**0.64**	**0.38**	**0.64**
Math achievement	**-0.22**	0.10	-0.01
German achievement	**-0.18**	0.10	0.02
Vocational interests			
Realistic	0.00	-0.07	-0.11
Investigative	0.00	**-0.16**	-0.14
Artistic	-0.05	-0.05	**-0.19**
Social	0.00	0.03	0.01
Enterprising	0.00	0.09	0.04
Conventional	0.00	0.00	**-0.19**
Big Five			
Conscientiousness	-0.10	-0.01	**-0.31**
Neuroticism	0.14	**0.15**	0.10
Openness	-0.08	0.03	-0.13
Agreeableness	-0.12	0.05	**-0.30**
Extraversion	-0.07	0.01	-0.06
Other constructs			
Self-concept in math	-0.12	0.00	**-0.25**
Self-concept in German	**-0.15**	**-0.15**	-0.09
Self-concept in English	-0.06	**-0.18**	0.03
Interest in math	0.03	-0.01	**-0.23**
Interest in German	0.04	-0.03	**-0.18**
Interest in English	0.07	-0.11	-0.13
Effort in math	-0.10	-0.09	**-0.37**
Effort in German	**-0.24**	-0.10	**-0.34**
Effort in English	-0.09	**-0.18**	**-0.18**

Note. Bold parameters are significantly different from zero at $p < .05$, two-tailed. Positive values imply higher values for the participants who joined the study late, significant variables were included as auxiliaries in the main analysis; OR = Odds Ratio.

Clustering Information: School and Class Level

Table A4
Intraclass Correlations of RIASEC Scale Scores

Cluster Variable	Interests	T1	T2	T3	T4
Class	Realistic	.08	.04	.03	.03
	Investigative	.10	.03	.03	.02
	Artistic	.07	.03	.02	.02
	Social	.08	.04	.04	.04
	Enterprising	.10	.03	.03	.03
	Conventional	.11	.04	.04	.04
School	Realistic	.07	.03	.02	.02
	Investigative	.09	.03	.03	.01
	Artistic	.05	.02	.01	.02
	Social	.07	.04	.04	.03
	Enterprising	.09	.03	.03	.03
	Conventional	.09	.04	.03	.04

Note. N = 3,351; T1–T4 = time points 1 to 4; The ICCs were computed in a saturated multilevel path model (TYPE = TWOLEVEL) in Mplus 8; Missing data were treated with FIML.

Supplement B: Interest Inventory

Item Wording and Item Scale Correlations

Note: The information of item wording is not available because of copyright restrictions.

Table B1

Items, their Origins, and their Item-Scale Correlations

Scale	Item	"How much do you like this activity?"	Origin	r_{it1}	r_{it2}	r_{it3}	r_{it4}
R	01		ICA-R	.52	.56	.59	.58
	07		ICA-R	.63	.66	.67	.73
	13		AISTR (altered)	.59	.61	.57	.63
	19		AISTR	.70	.72	.73	.80
	25		New	.63	.70	.71	.71
	31		AISTR	.67	.69	.73	.77
I	02		AISTR	.55	.52	.56	.58
	08		AISTR	.53	.54	.49	.54
	14		ICA-R	.53	.56	.58	.64
	20		ICA-R	.67	.64	.68	.70
	26		ICA-R	.59	.60	.61	.61
	32		New	.49	.58	.58	.62
A	03		ICA-R	.54	.54	.55	.56
	09		AISTR	.60	.62	.63	.61
	15		ICA-R	.54	.50	.54	.50
	21		New	.49	.51	.54	.53
	27		New	.51	.49	.49	.45
	33		AISTR	.52	.59	.61	.59
S	04		AISTR	.53	.61	.63	.63
	10		AISTR	.60	.66	.64	.65
	16		ICA-R	.65	.69	.69	.68
	22		ICA-R	.65	.68	.65	.66
	28		New	.58	.59	.60	.58
	34		AISTR	.63	.65	.66	.65
E	05		ICA-R	.43	.38	.36	.39
	11		ICA-R	.66	.64	.60	.65
	17		ICA-R	.52	.50	.51	.55
	23		New	.61	.53	.54	.51
	29		New	.64	.68	.66	.66
	35		New	.53	.56	.51	.47
C	06		ICA-R	.44	.48	.50	.52
	12		AISTR (altered)	.65	.66	.62	.60
	18		ICA-R	.69	.71	.69	.69
	24		ICA-R	.64	.63	.61	.57

30	New	.54	.56	.57	.55
36	New	.52	.57	.56	.51

Note. r_{it} = Correlation of the item with the scale when the respective item is dropped; the items were translated by the authors of the study.

Factor Loadings of Factor Models

Table B2

Factor Loadings for the Configural Invariance Measurement RIASEC Models

Constructs	Item	Loading T1 (se)	Loading T2 (se)	Loading T3 (se)	Loading T4 (se)
R	01	0.76(0.03)	0.76(0.02)	0.79(0.02)	0.79(0.02)
	07	1.03(0.02)	1.00(0.02)	1.00(0.02)	1.03(0.02)
	13	0.90(0.02)	0.91(0.02)	0.81(0.02)	0.80(0.02)
	19	1.14(0.03)	1.13(0.02)	1.14(0.02)	1.20(0.01)
	25	1.02(0.03)	1.07(0.02)	1.07(0.02)	1.01(0.02)
	31	1.15(0.02)	1.14(0.02)	1.19(0.02)	1.17(0.02)
I	02	0.98(0.03)	0.88(0.03)	0.92(0.03)	0.95(0.03)
	08	0.95(0.03)	0.94(0.03)	0.88(0.03)	0.90(0.02)
	14	0.92(0.03)	1.00(0.03)	0.98(0.03)	1.00(0.02)
	20	1.25(0.03)	1.14(0.03)	1.18(0.02)	1.17(0.02)
	26	1.11(0.03)	1.09(0.03)	1.11(0.02)	1.07(0.03)
	32	0.78(0.04)	0.95(0.03)	0.93(0.03)	0.91(0.03)
A	03	0.97(0.03)	0.95(0.03)	0.95(0.03)	1.03(0.03)
	09	1.11(0.03)	1.12(0.03)	1.14(0.03)	1.22(0.03)
	15	1.01(0.03)	0.94(0.03)	0.94(0.03)	0.88(0.03)
	21	0.98(0.03)	1.03(0.03)	1.04(0.03)	1.05(0.03)
	27	0.96(0.03)	0.90(0.03)	0.85(0.03)	0.75(0.03)
	33	0.97(0.03)	1.06(0.03)	1.09(0.02)	1.09(0.02)
S	04	0.86(0.03)	0.96(0.02)	0.97(0.02)	1.01(0.03)
	10	0.98(0.03)	1.02(0.02)	1.01(0.02)	1.03(0.02)
	16	1.04(0.03)	1.03(0.02)	1.04(0.02)	1.04(0.02)
	22	1.09(0.02)	1.02(0.02)	1.00(0.02)	1.00(0.02)
	28	0.99(0.03)	0.96(0.02)	0.96(0.02)	0.92(0.03)
	34	1.05(0.03)	1.09(0.02)	1.02(0.02)	1.01(0.03)
E	05	0.74(0.04)	0.66(0.03)	0.64(0.04)	0.67(0.03)
	11	1.20(0.03)	1.19(0.02)	1.19(0.03)	1.23(0.02)
	17	0.90(0.03)	0.87(0.03)	0.91(0.03)	1.01(0.03)
	23	1.01(0.03)	0.90(0.03)	0.95(0.03)	0.89(0.03)
	29	1.22(0.03)	1.33(0.02)	1.34(0.03)	1.31(0.03)
	35	0.93(0.03)	1.04(0.03)	0.97(0.03)	0.90(0.03)
C	06	0.78(0.03)	0.83(0.03)	0.88(0.02)	0.91(0.03)
	12	1.11(0.03)	1.09(0.02)	1.04(0.02)	1.06(0.02)
	18	1.21(0.03)	1.18(0.02)	1.14(0.02)	1.22(0.02)
	24	1.11(0.03)	1.06(0.02)	1.06(0.02)	1.00(0.03)
	30	0.93(0.03)	0.92(0.02)	0.95(0.02)	0.95(0.03)
	36	0.86(0.03)	0.92(0.03)	0.93(0.03)	0.85(0.03)

Note. se = Standard Error.

Convergent and Discriminant Validity

In addition to structural and reliability information, we provide evidence on the convergent and discriminant validity of our interest measure. We correlated vocational interests with personality traits, self-concepts and competencies. We measured the Big Five personality traits of openness, conscientiousness, extraversion, agreeableness, and neuroticism. In line with the meta-analysis by (Mount et al., 2005), the highest correlations are assumed to be between Investigative and Openness (r = .25), Artistic and Openness (r = .41), Social and Extraversion (r = .29), Enterprising and Extraversion (r = .40), as well as Conventional and Conscientiousness (r = .19). Smaller correlations are assumed between Social and Openness (r = .13) as well as Social and Agreeableness (r = .17).

Intercorrelations are depicted in Tables B3 and B4 such that the dark grey areas indicate the strong relationships and the light grey areas the weak relationships that Mount et al. (2005) proposed. At the first time point, the relationships that Mount et al. (2005) proposed were present. In addition, the proposed relationships showed the strongest correlations between personality traits and vocational interests (except Enterprising). However, at the first time point, all the personality traits were moderately correlated. At the fourth time point, the pattern of intercorrelations was close to the one reported by Mount et al. (2005).

Self-concept was measured for two school subjects: math and German. In line with the study by Ackerman et al. (1995), vocational interests and self-concept should be correlated in the following way: math self-concept and Realistic (r = .38), math self-concept and Investigative (r = .22), math self-concept and Conventional (r = .61), verbal self-concept and Artistic (r = .40), as well as verbal self-concept and Social (r = .16). At both time points, we found the proposed intercorrelations for math self-concept, although they were a little bit lower than they were in Ackerman et al. (1995). Similar results were found for German self-concept.

Expectations of the relationships between vocational interests and competencies were derived from Ackerman and Heggestad (1997). On the basis of evidence from various studies, Investigative and Realistic interests' relationships with math abilities were incorporated into the math/science trait complex. The relationships between verbal abilities and Artistic interests were incorporated into the intellectual/cultural trait complex. For math competencies, we found the respective pattern at both time points such that there were positive correlations with Investigative interests and negative correlations with Artistic and Social interests. However, we did not find positive correlations between Artistic interests and verbal competencies at either time point.

In summary, these results support the convergent and discriminant validity of our interest measure. The majority of vocational interests' proposed relationships with personality traits, self-concepts, and competencies were reflected in our data. The pattern improved over time, such that the correlations were more in line with the expected correlational pattern at the fourth time point.

Table B3

Manifest Correlations between RIASEC Scales and Associated Constructs at T1

	R	I	A	S	E	C
O	.28 [.24, .33]	.35 [.30, .39]	.40 [.36, .44]	.37 [.33, .41]	.36 [.31, .41]	.35 [.31, .40]
C	.23 [.18, .28]	.27 [.23, .31]	.32 [.27, .36]	.31 [.26, .35]	.29 [.24, .34]	.35 [.30, .40]
E	.21 [.16, .25]	.24 [.19, .28]	.29 [.24, .33]	.31 [.26, .35]	.33 [.28, .38]	.26 [.21, .30]
A	.13 [.07, .18]	.23 [.19, .28]	.32 [.28, .36]	.38 [.33, .43]	.25 [.19, .30]	.27 [.22, .32]
N	.16 [.11, .21]	.11 [.06, .16]	.12 [.07, .18]	.17 [.12, .22]	.20 [.15, .25]	.20 [.15, .26]
SK. M	.20 [.16, .25]	.13 [.09, .16]				.17 [.14, .21]
SK. G		.16 [.13, .19]	.25 [.21, .28]	.18 [.15, .21]	.15 [.11, .18]	.15 [.12, .18]
Co. M		.10 [.04, .17]	-.12 [-.17, -.06]	-.12 [-.18, -.07]		-.11 [-.17, -.04]
Co. G		.10 [.05, .16]			-.15 [-.21, -.09]	-.19 [-.25, -.14]

Note. O = Openness; C = Conscientiousness; E = Extraversion; A = Agreeableness; N = Neuroticism; SK. M = Self-concept Math; SK. G = Self-concept German; SK. E = Self-concept English; Co. M = Competencies Math; Co. G = Competencies German; Co. E = Competencies English; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional; Dark grey implies strong correlations, light grey implies weak correlations. Assumptions of the intercorrelations between the Big Five and RIASEC scales were derived from the meta-analysis of Mount et al. (2005); Assumptions of the intercorrelations between the Competence and RIASEC scales were derived from the meta-analytic evidence about trait-complexes from Ackerman and Heggestad (1997); Assumptions of the intercorrelations between Self-Concept and RIASEC scales were derived from Ackerman et al. (1995).

Table B4

Manifest Correlations between RIASEC Scales and Associated Constructs at T4

	R	I	A	S	E	C
O	.11 [.07, .15]	.21 [.18, .25]	.30 [.27, .33]	.20 [.16, .23]	.20 [.17, .24]	.16 [.13, .20]
C	.15 [.10, .19]	.18 [.14, .21]	.17 [.14, .20]	.20 [.16, .24]	.17 [.14, .20]	.21 [.18, .24]
E		.10 [.06, .14]	.19 [.15, .22]	.22 [.18, .26]	.24 [.21, .28]	.12 [.08, .15]
A			.17 [.13, .20]	.26 [.23, .30]	.11 [.08, .14]	
N						
SK. M	.28 [.24, .33]	.14 [.11, .17]				.20 [.16, .23]
SK. G			.15 [.12, .17]	.14 [.12, .17]		
Co. M	.14 [.07, .21]	.10 [.04, .15]		-.13 [-.18, -.07]		
Co. G	-.17 [-.26, -.09]					-.23 [-.29, -.16]

Note. O = Openness; C = Conscientiousness; E = Extraversion; A = Agreeableness; N = Neuroticism; SK. M = Self-concept Math; SK. G = Self-concept German; SK. E = Self-concept English; Co. M = Competencies Math; Co. G = Competencies German; Co. E = Competencies English; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional; Dark grey implies strong correlations, light grey implies weak correlations. Assumptions of the intercorrelations between the Big Five and RIASEC scales were derived from the meta-analysis of Mount et al. (2005); Assumptions of the intercorrelations between the Competence and RIASEC scales were derived from the meta-analytic evidence about trait-complexes from Ackerman and Heggestad (1997); Assumptions of the intercorrelations between Self-Concept and RIASEC scales were derived from Ackerman et al. (1995).

Supplement C: Information about Model Fit and Model Parameters

Measurement Invariance

In the present study, the following fit indices and values were used to judge whether the overall fit was adequate: incremental fit indices such as the Comparative Fit Index (CFI) and the Tucker-Lewis Index (TLI) with values of .90 or higher, the Standardized Root Mean Square Residual (SRMR) with a value of .08 or lower, as well as the Root Mean Square Error of Approximation with a value of .06 or lower (Hu & Bentler, 1999). To judge changes in the descriptive fit indices, we used the following values proposed by Chen (2007): values ≥ .010 for ΔCFI, additionally accompanied by a value of ≥ .015 for ΔRMSEA or ≥ 0.030 for ΔSRMR, indicated substantive changes in tests for factor loading invariance. To test for intercept invariance, a value of ≥ .010 for ΔCFI, additionally accompanied by a value of ≥ .015 for ΔRMSEA or ≥ 0.010 for ΔSRMR, seemed substantial.

For each RIASEC dimension, we specified the following models and compared them to each other. First, we specified configural measurement invariance models, were only the structure of the factor models was invariant over time and the Time × Gender interaction (i.e., invariant between boys, girls and over time). Second, we specified weak measurement invariance models, were the factor loadings of the factor models were invariant over time and the Time × Gender interaction (i.e., invariant between boys, girls and over time). Third, we specified strong measurement invariance models, were the factor loadings and the intercepts of the factor models were invariant over time and the Time × Gender interaction (i.e., invariant between boys, girls and over time). We compared the weak and the strong measurement invariance models to the configural measurement invariance models to investigate if substantive changes in fit indices occurred after adding more constraints to the models (the results can be found in Table C1 and C2).

For Realistic interest, strong measurement invariance applied over time, with adequate overall fit for the strong measurement invariance model and negligible changes in descriptive fit indices when more constraints were applied. However, for invariance across time and group, two intercepts had to be freed (Item 1: “Build something out of metal or wood”; Item 2: “Work with machines and tools”) that would otherwise be invariant between boys and girls. Although ΔCFI = 0.012 between the configural and partial strong measurement invariance model was slightly above the proposed value, the accompanying ΔRMSEA and ΔSRMR values were below the proposed values. This indicates that for the Time × Gender interaction, only partially strong measurement invariance applies.

For Investigative interest, strong measurement invariance applied over time, with adequate overall fit for the strong measurement invariance model and negligible changes (although ΔCFI was equal to .01, the ΔRMSEA and ΔSRMR values were below their proposed values) in descriptive fit indices when more constraints were applied. However, for invariance across time and group, two intercepts had to be freed (Item 1: "Study the behavior of animals and plants"; Item 2: "Watch a science show") that would otherwise be invariant between boys and girls. Although ΔCFI = 0.021 between the configural and partial strong measurement invariance model was slightly above the proposed value, the accompanying ΔRMSEA and ΔSRMR values were below the proposed values. This indicates that for the Time × Gender interaction, only partially strong measurement invariance applies.

For Artistic interest, strong measurement invariance applied over time, with adequate overall fit for the strong measurement invariance model and negligible changes (although ΔCFI was above .01, ΔRMSEA and ΔSRMR values were below their proposed values) in descriptive fit indices when more constraints were applied. However, for invariance across time and group, three intercepts had to be freed (Item 1: "Make things pretty (decorating and embellishing)"; Item 2: "Make up a story"; Item 3: "Design clothes") that would otherwise be invariant between boys and girls. Although ΔCFI = 0.039 between the configural and partial strong measurement invariance model was above the proposed value, the accompanying ΔRMSEA and ΔSRMR values were below the proposed values. In addition, the overall model fit was adequate with the CFI and TLI near .90 and the RMSEA and SRMR under the proposed values. This indicates that for the Time × Gender interaction, only partially strong measurement invariance applies.

For Social interest, strong measurement invariance applied over time, with an adequate overall fit for the strong measurement invariance model and negligible changes (although the ΔCFI was slightly above .01, the ΔRMSEA and ΔSRMR values were below their proposed values) in the descriptive fit indices when more constraints were applied. However, for invariance across time and group, one intercept had to be freed (Item 1: "Take care of small children") that would otherwise be invariant between boys and girls. Although ΔCFI = 0.021 between the configural and partial strong measurement invariance model was above the proposed value, the accompanying ΔRMSEA and ΔSRMR values were below the proposed values. This indicates that for the Time × Gender interaction, only partially strong measurement invariance applies.

For Enterprising interest, strong measurement invariance applied over time, with adequate overall fit for the strong measurement invariance model and negligible changes (although the ΔCFI was above .01, the ΔRMSEA and ΔSRMR values were below their

proposed values) in the descriptive fit indices when more constraints were applied. However, for invariance across time and group, two intercepts had to be freed (Item 1: “Sell things to others”; Item 2: “Plan an event”) that would otherwise be invariant between boys and girls. In addition to a ΔCFI = 0.021 between the configural and partial strong measurement invariance model, the value of the ΔSRMR was slightly higher than the proposed values. However, because of the adequate overall model fit, we decided to free only two of the respective intercepts overall. This indicates that for the Time × Gender interaction, only partially strong measurement invariance applies.

For Conventional interest, strong measurement invariance applied over time, with adequate overall fit for the strong measurement invariance model and negligible changes (although the ΔCFI was above .01, the ΔRMSEA and ΔSRMR values were below their proposed values) in the descriptive fit indices when more constraints were applied. However, for invariance across time and group, two intercepts had to be freed (Item 1: “Add numbers”; Item 2: “Make a list”) that would otherwise be invariant between boys and girls. In addition to a ΔCFI = 0.035 between the configural and partial strong measurement invariance model, the value for the ΔSRMR was slightly higher than the proposed values. However, because of the adequate overall model fit, we decided to free only two of the respective intercepts overall. This indicates that for the Time × Gender interaction, only partially strong measurement invariance applies.

Table C1

Measurement Models with Different Types of Invariance Over Time and Group

Int.	Model	CFI	TLI	RMSEA	SRMR	χ^2	*df*
R	MICT	0.964	0.958	0.031	0.030	1005.68*	240
	MIWT	0.962	0.959	0.030	0.033	1060.13*	255
	MIST	0.957	0.956	0.032	0.035	1187.51*	270
	MICG[a]	0.951	0.944	0.032	0.039	1321.98*	480
	MIWG	0.947	0.944	0.032	0.044	1418.07*	515
	MISG	0.928	0.928	0.036	0.051	1779.43*	550
	MISGp	0.939	0.938	0.034	0.047	1597.88*	548
I	MICT	0.903	0.889	0.045	0.049	1885.78*	240
	MIWT	0.901	0.893	0.044	0.050	1939.03*	255
	MIST	0.893	0.890	0.045	0.052	2095.91*	270
	MICG[a]	0.901	0.886	0.045	0.053	2147.53*	480
	MIWG	0.895	0.888	0.045	0.055	2273.33*	515
	MISG	0.865	0.864	0.049	0.060	2814.75*	550
	MISGp	0.880	0.880	0.046	0.057	2522.36*	548
A	MICT	0.921	0.909	0.040	0.048	1521.34*	240
	MIWT	0.918	0.911	0.039	0.050	1586.54*	255
	MIST	0.896	0.893	0.043	0.055	1959.32*	270
	MICG[a]	0.925	0.914	0.036	0.047	1535.95*	480
	MIWG	0.918	0.912	0.036	0.055	1667.14*	515
	MISG	0.817	0.817	0.052	0.079	3116.86*	550
	MISGp	0.886	0.885	0.041	0.059	2143.51*	547
S	MICT	0.938	0.929	0.038	0.034	1391.22*	240
	MIWT	0.937	0.931	0.037	0.036	1435.63*	255
	MIST	0.923	0.921	0.040	0.037	1706.29*	270
	MICG[a]	0.932	0.922	0.037	0.039	1610.26*	480
	MIWG	0.929	0.924	0.037	0.043	1694.94*	515
	MISG	0.900	0.899	0.042	0.056	2224.91*	550
	MISGp	0.911	0.910	0.040	0.046	2043.17*	549
E	MICT	0.970	0.966	0.022	0.031	651.70*	240
	MIWT	0.967	0.965	0.023	0.034	704.84*	255
	MIST	0.950	0.949	0.027	0.037	961.57*	270
	MICG[a]	0.967	0.962	0.024	0.038	935.10*	480
	MIWG	0.962	0.959	0.025	0.043	1042.99*	515
	MISG	0.915	0.914	0.036	0.055	1736.15*	550
	MISGp	0.941	0.940	0.030	0.047	1372.06*	548
C	MICT	0.972	0.968	0.023	0.033	660.57*	240
	MIWT	0.971	0.968	0.023	0.034	700.82*	255
	MIST	0.955	0.954	0.027	0.040	949.60*	270
	MICG[a]	0.976	0.972	0.021	0.035	850.06*	480
	MIWG	0.974	0.972	0.021	0.038	918.69*	515
	MISG	0.925	0.925	0.035	0.052	1706.96*	550

MISGp	0.941	0.940	0.031	0.047	1462.24*	548

Note. Int. = RIASEC interest dimensions; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; χ^2 = Chi-Square Statistic, *df* = Degrees of Freedom; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional; MICT = configural measurement invariance over time; MIWT = weak measurement invariance over time; MIST = strong measurement invariance over time; MICG = configural measurement invariance over time and gender; MIWG = weak measurement invariance over time and gender; MISG = strong measurement invariance over time and gender; MISGp = partial strong measurement invariance over time and gender.

[a]The standard errors in the model might not be trustworthy, which is most likely due to having more free parameters than clusters (the error message vanished when we did not compute cluster robust standard errors).

* $p < .01$

Table C2

Differences in Fit Indices

Int.	Model	CFI	ΔCFI	RMSEA	ΔRMSEA	SRMR	ΔSRMR
	MICT	0.964		0.031		0.030	
	MIWT	0.962	0.002	0.030	0.001	0.033	0.003
	MIST	0.957	0.007	0.032	0.001	0.035	0.002
R	MICG	0.951		0.032		0.039	
	MIWG	0.947	0.004	0.032	0.000	0.044	0.005
	MISG	0.928	0.023	0.036	0.004	0.051	0.012
	MISGp	0.939	0.012	0.034	0.002	0.047	0.008
	MICT	0.903		0.045		0.049	
	MIWT	0.901	0.002	0.044	0.001	0.050	0.001
	MIST	0.893	0.010	0.045	0.000	0.052	0.003
I	MICG	0.901		0.045		0.053	
	MIWG	0.895	0.006	0.045	0.000	0.055	0.002
	MISG	0.865	0.036	0.049	0.004	0.060	0.007
	MISGp	0.880	0.021	0.046	0.001	0.057	0.004
	MICT	0.921		0.040		0.048	
	MIWT	0.918	0.003	0.039	0.001	0.050	0.002
	MIST	0.896	0.025	0.043	0.003	0.055	0.007
A	MICG	0.925		0.036		0.047	
	MIWG	0.918	0.007	0.036	0.000	0.055	0.008
	MISG	0.817	0.118	0.052	0.016	0.079	0.032
	MISGp	0.886	0.039	0.041	0.005	0.059	0.012
	MICT	0.938		0.038		0.034	
	MIWT	0.937	0.001	0.037	0.001	0.036	0.002
	MIST	0.923	0.015	0.040	0.002	0.037	0.003
S	MICG	0.932		0.037		0.039	
	MIWG	0.929	0.003	0.037	0.000	0.043	0.004
	MISG	0.900	0.032	0.042	0.005	0.056	0.017
	MISGp	0.911	0.021	0.040	0.003	0.046	0.007
	MICT	0.970		0.022		0.031	
	MIWT	0.967	0.003	0.023	0.001	0.034	0.003
	MIST	0.950	0.020	0.027	0.005	0.037	0.006
E	MICG	0.967		0.024		0.038	
	MIWG	0.962	0.005	0.025	0.001	0.043	0.005
	MISG	0.915	0.052	0.036	0.012	0.055	0.017
	MISGp	0.941	0.026	0.030	0.006	0.047	0.011
	MICT	0.972		0.023		0.033	
	MIWT	0.971	0.001	0.023	0.000	0.034	0.001
	MIST	0.955	0.017	0.027	0.004	0.040	0.007
C	MICG	0.976		0.021		0.035	
	MIWG	0.974	0.002	0.021	0.000	0.038	0.003
	MISG	0.925	0.051	0.035	0.014	0.052	0.017

MISGp	0.941	0.035	0.031	0.010	0.047	0.012

Note. Int. = RIASEC interest dimensions; ΔCFI = Change in Comparative Fit Index; ΔTLI = Change in Tucker-Lewis Index; ΔRMSEA = Change in Root Mean Square Error of Approximation; ΔSRMR = Change in Standardized Root Mean Square Residual; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional; MICT = configural measurement invariance over time; MIWT = weak measurement invariance over time; MIST = strong measurement invariance over time; MICG = configural measurement invariance over time and gender; MIWG = weak measurement invariance over time and gender; MISG = strong measurement invariance over time and gender; MISGp = partial strong measurement invariance over time and gender.
* $p < .01$.

Fit Indices for Generalized Second-Order Growth Models

For model selection, we used the same values that were proposed for measurement invariance testing. For the overall models, when a latent factor representing quadratic growth was included, the model fit did not improve substantially according to the fit indices. Although ΔCFI values for Investigative (ΔCFI = 0.017) and Social (ΔCFI = 0.012) were higher than the proposed value of .01, both of their ΔRMSEA and ΔSRMR values were under the proposed values. Therefore, we decided to specify models with a latent factor that implied linear growth only. Models with a latent linear growth factor possessed adequate model fit according to their descriptive fit indices.

For the GSGM multigroup models, we first specified an overall model in which both groups (i.e., boys and girls) had linear as well as quadratic growth components. We then investigated whether the means of the respective quadratic growth factors were statistically significantly different from zero. If this was not the case, we specified another model in which we constrained the quadratic growth type for the respective group. If the change in the descriptive fit indices according to the values proposed by Chen (2007) was negligible, we decided to keep the constrained model.

For Realistic and Artistic interests, we kept the constrained model with quadratic growth for girls and linear growth for boys because the mean level of the latent quadratic growth factor for boys was not statistically significantly different from zero. Changes in the descriptive fit indices between the overall and constrained models were negligible and under the values proposed by Chen (2007). A similar procedure was applied for Conventional interests, with the only difference that girls had a linear growth factor, and boys had a quadratic growth factor.

For Investigative and Enterprising interests, we kept the overall model in which girls and boys both had a quadratic growth factor component. Both boys and girls had significant mean levels on the latent quadratic growth factor.

For Social interest, we specified a constrained model in which both the girls and boys had a linear growth factor. For both groups, the latent mean level of the quadratic growth factor was not statistically significantly different from zero. Although ΔCFI between the overall and the constrained value was above .01, the accompanying ΔRMSEA and ΔSRMR values were below the proposed values. We consequently chose the more parsimonious model, which was the one with linear growth factors.

Table C3

Fit Indices for Generalized Second-Order Growth Models

Int.	Model	CFI	TLI	RMSEA	SRMR	χ^2	*df*
R	GSGML	0.955	0.954	0.032	0.039	1212.21*	270
	GSGMQ	0.958	0.956	0.031	0.040	1161.68*	266
	$GSGML_{AUX}$	0.963	0.963	0.030	0.044	1228.52*	270
	$GSGMQ_{AUX}$	0.965	0.964	0.030	0.046	1183.61*	266
I	GSGML	0.901	0.899	0.043	0.079	1943.91*	270
	GSGMQ	0.918	0.915	0.039	0.072	1659.18*	266
	$GSGML_{AUX}$	0.911	0.909	0.041	0.086	1998.54*	270
	$GSGMQ_{AUX}$	0.926	0.923	0.037	0.080	1709.17*	266
A	GSGML	0.906	0.904	0.041	0.057	1793.73*	270
	GSGMQ	0.914	0.911	0.039	0.056	1655.30*	266
	$GSGML_{AUX}$	0.919	0.917	0.038	0.067	1798.12*	270
	$GSGMQ_{AUX}$	0.926	0.923	0.037	0.065	1659.15*	266
S	GSGML	0.923	0.921	0.039	0.064	1702.47*	270
	GSGMQ	0.935	0.933	0.036	0.072	1469.97*	266
	$GSGML_{AUX}$	0.934	0.933	0.037	0.068	1728.05*	270
	$GSGMQ_{AUX}$	0.945	0.943	0.034	0.080	1490.88*	266
E	GSGML	0.947	0.946	0.028	0.040	995.83*	270
	GSGMQ	0.950	0.948	0.028	0.041	951.84*	266
	$GSGML_{AUX}$	0.955	0.954	0.026	0.047	980.29*	270
	$GSGMQ_{AUX}$	0.955	0.953	0.026	0.047	975.92*	266
C	GSGML	0.960	0.959	0.026	0.049	875.04*	270
	GSGMQ	0.962	0.961	0.025	0.045	842.15*	266
	$GSGML_{AUX}$	0.966	0.965	0.024	0.061	876.90*	270
	$GSGMQ_{AUX}$	0.968	0.966	0.024	0.057	838.85*	266

Note. Int. = RIASEC interest dimensions; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; χ^2 = Chi-Square Statistic; *df* = Degrees of Freedom; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional; GSGML = GSGM with linear growth; GSGMQ = GSGM with quadratic growth; $GSGML_{AUX}$ = GSGM with linear growth and auxiliaries; $GSGMQ_{AUX}$ = GSGM with quadratic growth and auxiliaries.

* $p < .01$.

Table C4

Fit Indices for Multigroup Generalized Second-Order Growth Models

Int.	Model	Growth Type	CFI	TLI	RMSEA	SRMR	χ^2	*df*
	Ov	QQ	0.942	0.941	0.033	0.051	1541.80*	545
R	Con	QL	0.939	0.939	0.033	0.052	1591.82*	549
	Aux	QL	0.948	0.948	0.032	0.060	1629.13*	549
	Ov	QQ	0.907	0.906	0.041	0.074	2098.53*	545
I	Con	QQ	0.907	0.906	0.041	0.074	2098.53*	545
	Aux	QQ	0.918	0.917	0.040	0.084	2170.97*	545
	Ov	QQ	0.900	0.899	0.039	0.064	1946.27*	544
A	Con	QL	0.896	0.895	0.040	0.065	2010.69*	548
	Aux	QL	0.908	0.907	0.038	0.076	2052.77*	548
	Ov	QQ	0.927	0.926	0.036	0.073	1773.64*	546
S	Con	LL	0.913	0.914	0.039	0.069	2004.72*	554
	Aux	LL	0.925	0.925	0.038	0.077	2055.76*	554
	Ov	QQ	0.941	0.941	0.030	0.048	1358.64*	545
E	Con	QQ	0.941	0.941	0.030	0.048	1358.64*	545
	Aux	QQ	0.951	0.951	0.028	0.056	1347.53*	545
	Ov	QQ	0.947	0.946	0.030	0.052	1362.34*	545
C	Con	LQ	0.946	0.946	0.030	0.053	1381.48*	549
	Aux	LQ	0.956	0.956	0.028	0.064	1372.19*	549

Note. Int. = RIASEC interest dimensions; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; χ^2 = Chi-Square Statistic; *df* = Degrees of Freedom; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional; Ov. = Overall model; Con = constrained Model; Aux = Constrained model with auxiliaries; L = linear growth factor; Q = quadratic growth factor; QL = the growth factor type of the respective group, the first letter refers to girls, the second to boys.

* $p < .01$.

Parameters for Generalized Second-Order Growth Models

Table C5

Model Parameters for Generalized Second-Order Growth Models

		Intercept		Slope		
Growth Type	Int.	*M* [95% CI]	σ^2 [95% CI]	*M* [95% CI]	σ^2 [95% CI]	*r*(I,S) [95% CI]
Linear	R	3.37 [3.30, 3.43]	0.46 [0.39, 0.52]	-0.10 [-0.12, -0.08]	0.02 [0.01, 0.04]	-0.20 [-0.34, -0.07]
	I	3.64 [3.56, 3.72]	0.89 [0.79, 0.99]	-0.16 [-0.19, -0.13]	0.11 [0.09, 0.13]	-0.60 [-0.66, -0.55]
	A	3.63 [3.57, 3.70]	0.64 [0.56, 0.72]	-0.11 [-0.14, -0.09]	0.06 [0.04, 0.07]	-0.38 [-0.46, -0.31]
	S	3.19 [3.12, 3.26]	0.68 [0.57, 0.79]	0.06 [0.03, 0.08]	0.07 [0.04, 0.10]	-0.46 [-0.56, -0.36]
	E	2.87 [2.81, 2.92]	0.18 [0.13, 0.22]	-0.01 [-0.03, 0.00]	0.01 [0.00, 0.01]	-0.47 [-0.58, -0.36]
	C	3.10 [3.02, 3.19]	0.61 [0.51, 0.72]	-0.14 [-0.17, -0.10]	0.05 [0.02, 0.07]	-0.40 [-0.53, -0.28]
Quadratic	R	3.36 [3.29, 3.42]	0.58 [0.49, 0.68]	-0.01 [-0.03, 0.00]	0.04 [0.03, 0.06]	0.27 [0.17, 0.37]
	I	3.59 [3.50, 3.68]	1.05 [0.93, 1.17]	-0.04 [-0.06, -0.02]	0.10 [0.08, 0.12]	0.36 [0.28, 0.43]
	A	3.59 [3.52, 3.67]	0.73 [0.64, 0.82]	-0.03 [-0.05, -0.01]	0.05 [0.04, 0.07]	0.24 [0.16, 0.33]
	S	3.19 [3.12, 3.26]	0.82 [0.73, 0.91]	0.00 [-0.02, 0.02]	0.10 [0.08, 0.12]	0.23 [0.14, 0.32]
	E	2.82 [2.74, 2.91]	0.19 [-0.04, 0.41]	-0.02 [-0.04, 0.00]	0.01 [-0.02, 0.05]	0.10 [-0.77, 0.97]
	C*	3.12 [3.03, 3.21]	0.49 [0.25, 0.73]	0.01 [-0.01, 0.03]	0.03 [0.00, 0.06]	-0.30 [-0.95, 0.34]

Note. Int. = RIASEC interest dimensions; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional; *M* = mean of intercept or slope factor; σ^2 = variance of intercept or slope factor; *r*(I,S) = correlation between intercept and slope factor.

Table C6

Model Parameters for Multigroup Generalized Second-Order Growth Models

			Intercept		Slope		
		Int.	*M* [95% CI]	σ^2 [95% CI]	*M* [95% CI]	σ^2 [95% CI]	*r*(I,S) [95% CI]
Boys	Ov.	R	3.71 [3.64, 3.77]	0.40 [0.29, 0.51]	0.01 [-0.01, 0.03]	0.04 [0.03, 0.05]	.29 [.16, .41]
		I	3.70 [3.59, 3.80]	1.01 [0.87, 1.14]	-0.03 [-0.06, 0.00]	0.11 [0.08, 0.13]	.39 [.30, .48]
		A	3.44 [3.35, 3.53]	0.97 [0.82, 1.13]	-0.02 [-0.05, 0.01]	0.09 [0.06, 0.12]	.30 [.17, .43]
		S	2.97 [2.88, 3.06]	0.83 [0.70, 0.96]	-0.01 [-0.04, 0.03]	0.11 [0.09, 0.14]	.23 [.12, .34]
		E	2.92 [2.83, 3.01]	0.34 [0.16, 0.52]	-0.03 [-0.05, -0.01]	0.02 [0.01, 0.04]	.11 [-.32, .53]
		C	3.01 [2.90, 3.12]	0.52 [0.27, 0.78]	0.03 [0.00, 0.06]	0.04 [0.002, 0.08]	-.15 [-.73, .43]
	Con	R_L	3.66 [3.59, 3.72]	0.31 [0.23, 0.39]	-0.07 [-0.09, -0.05]	0.02 [0.01, 0.04]	-.38 [-.53, -.22]
		I_Q	3.68 [3.58, 3.78]	1.01 [0.88, 1.15]	-0.04 [-0.07, -0.01]	0.10 [0.08, 0.12]	.34 [.23, .45]
		A_L	3.47 [3.39, 3.56]	0.82 [0.68, 0.96]	-0.17 [-0.20, -0.13]	0.08 [0.05, 0.10]	-.54 [-.63, -.45]
		S_L	2.99 [2.91, 3.07]	0.69 [0.56, 0.82]	0.02 [-0.01, 0.05]	0.07 [0.04, 0.11]	-.53 [-.63, -.42]
		E_Q	3.00 [2.91, 3.10]	0.28 [0.18, 0.39]	-0.02 [-0.04, -0.01]	0.01 [-0.01, 0.04]	-.11 [-.93, .72]
		C_Q	2.99 [2.89, 3.09]	0.51 [0.29, 0.73]	0.03 [0.01, 0.06]	0.04 [0.01, 0.07]	-.21 [-.76, .33]
Girls	Ov.	R	2.97 [2.89, 3.04]	0.44 [0.36, 0.53]	-0.04 [-0.06, -0.02]	0.04 [0.03, 0.06]	.27 [.14, .40]
		I	3.50 [3.40, 3.61]	1.09 [0.94, 1.23]	-0.04 [-0.07, -0.01]	0.10 [0.08, 0.12]	.32 [.21, .43]
		A	3.95 [3.87, 4.04]	0.62 [0.49, 0.75]	-0.04 [-0.06, -0.02]	0.05 [0.03, 0.07]	.32 [.19, .45]
		S	3.43 [3.34, 3.52]	0.72 [0.60, 0.85]	0.01 [-0.02, 0.04]	0.09 [0.06, 0.11]	.26 [.14, .39]
		E	3.00 [2.89, 3.11]	0.29 [0.18, 0.40]	-0.02 [-0.05, 0.00]	0.02 [-0.01, 0.05]	.04 [-.84, .91]
		C	2.93 [2.84, 3.03]	0.46 [0.20, 0.71]	-0.02 [-0.05, 0.00]	0.01 [-0.03, 0.04]	-.86 [-3.92, 2.20]
	Con	R_Q	2.96 [2.88, 3.03]	0.41 [0.31, 0.51]	-0.03 [-0.05, -0.01]	0.03 [0.02, 0.05]	.28 [.10, .47]
		I_Q	3.48 [3.37, 3.59]	1.09 [0.95, 1.24]	-0.04 [-0.07, -0.01]	0.10 [0.08, 0.13]	.38 [.29, .47]
		A_Q	3.94 [3.85, 4.04]	0.57 [0.43, 0.70]	-0.04 [-0.06, -0.02]	0.04 [0.02, 0.06]	.32 [.15, .48]
		S_L	3.41 [3.33, 3.50]	0.58 [0.46, 0.70]	0.11 [0.08, 0.15]	0.07 [0.04, 0.09]	-.53 [-.65, -.41]
		E_Q	2.93	0.37	-0.03	0.02	.11

	[2.85, 3.02]	[0.18, 0.55]	[-0.05, -0.01]	[0.00, 0.04]	[-.31, .52]
C_L	2.95	0.59	-0.18	0.06	-.44
	[2.87, 3.04]	[0.46, 0.72]	[-0.22, -0.15]	[0.03, 0.08]	[-.60, -.28]

Note. Ov. = Overall model; Con = constrained Model; Int. = RIASEC interest dimensions; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional; *M* = mean of intercept or slope factor; σ^2 = variance of intercept or slope factor; *r*(I,S) = correlation between intercept and slope factor.

Table C7

Consistency and Occasion Specificity per Time Point for Generalized Second-Order Growth Models with Linear Growth

Int.	TL	CO [95% CI]	Δ [95% CI]	OS [95% CI]	Δ [95% CI]	ε [95% CI]	Δ [95% CI]	ΔCOOS [95% CI]
	1	.29 [.26, .32]		.19 [.16, .22]		.52 [.51, .54]		.10 [.05, .15]
	2	.32 [.30, .34]	.03 [.01, .06]	.18 [.16, .21]	-.01 [-.04, .03]	.50 [.48, .52]	-.03 [-.05, .00]	.14 [.10, .18]
R	3	.36 [.33, .38]	.04 [.02, .05]	.16 [.14, .19]	-.02 [-.05, .01]	.48 [.46, .50]	-.01 [-.04, .01]	.19 [.15, .24]
	4	.41 [.37, .44]	.05 [.02, .08]	.17 [.14, .21]	.01 [-.03, .05]	.42 [.40, .44]	.06 [.04, .08]	.23 [.16, .30]
	14		.12 [.08, .16]		-.02 [-.06, .03]		-.10 [-.13, -.08]	
	1	.26 [.23, .28]		.19 [.16, .21]		.55 [.53, .58]		.07 [.03, .11]
	2	.21 [.19, .23]	-.05 [-.07, -.03]	.24 [.22, .27]	.05 [.03, .08]	.55 [.53, .57]	-.01 [-.03, .02]	-.04 [-.07, .00]
I	3	.22 [.19, .24]	.01 [-.01, .02]	.23 [.20, .25]	-.02 [-.04, .01]	.55 [.54, .58]	.01 [-.02, .03]	-.01 [-.06, .03]
	4	.27 [.24, .30]	.06 [.04, .07]	.26 [.22, .29]	.03 [.00, .06]	.47 [.45, .50]	.09 [.06, .12]	.02 [-.04, .07]
	14		.02 [-.02, .05]		.07 [.03, .11]		-.09 [-.12, -.05]	
	1	.23 [.20, .25]		.17 [.14, .19]		.60 [.59, .63]		.06 [.02, .10]
	2	.22 [.21, .24]	.00 [-.02, .01]	.17 [.15, .20]	.01 [-.02, .04]	.61 [.58, .63]	-.01 [-.03, .02]	.05 [.02, .08]
A	3	.24 [.22, .26]	.01 [.00, .03]	.18 [.16, .21]	.01 [-.01, .04]	.58 [.56, .60]	-.03 [-.05, .00]	.05 [.02, .09]
	4	.28 [.26, .30]	.04 [.03, .06]	.15 [.13, .18]	-.03 [-.06, .00]	.57 [.54, .59]	.01 [-.01, .04]	.13 [.09, .17]
	14		.06 [.03, .08]		-.01 [-.05, .02]		-.04 [-.08, -.01]	
	1	.26 [.23, .29]		.21 [.18, .24]		.53 [.51, .56]		.06 [.00, .11]
	2	.25 [.22, .28]	-.01 [-.03, .01]	.24 [.22, .27]	.04 [.01, .07]	.51 [.49, .53]	-.03 [-.06, .00]	.01 [-.05, .06]
S	3	.27 [.24, .30]	.02 [.00, .03]	.23 [.20, .26]	-.01 [-.04, .01]	.50 [.48, .52]	-.01 [-.03, .02]	.04 [-.02, .09]
	4	.33 [.31, .36]	.07 [.04, .09]	.19 [.15, .23]	-.04 [-.08, -.01]	.48 [.44, .51]	.02 [-.01, .06]	.15 [.08, .21]
	14		.07 [.04, .11]		-.02 [-.06, .03]		-.06 [-.09, -.02]	
	1	.20 [.18, .23]		.21 [.18, .24]		.58 [.56, .61]		-.01 [-.06, .04]
	2	.22 [.20, .23]	.01 [-.01, .03]	.18 [.16, .20]	-.03 [-.07, .00]	.60 [.58, .62]	.02 [-.01, .05]	.03 [.00, .07]
E	3	.22 [.20, .23]	.00 [-.02, .02]	.17 [.15, .19]	-.01 [-.04, .02]	.62 [.59, .64]	.01 [-.01, .04]	.04 [.01, .08]
	4	.23 [.20, .27]	.02 [-.01, .05]	.17 [.14, .21]	.00 [-.04, .05]	.59 [.58, .61]	.02 [-.01, .05]	.06 [-.01, .13]
	14		.03 [-.01, .07]		-.04 [-.09, .01]		.01 [-.02, .04]	

	1	.20 [.17, .23]		.25 [.21, .28]		.56 [.53, .58]		-.05 [-.11, .01]
	2	.20 [.18, .23]	.00 [-.02, .02]	.25 [.23, .28]	.01 [-.02, .04]	.55 [.53, .56]	-.01 [-.04, .02]	-.05 [-.10, .00]
C	3	.22 [.20, .24]	.02 [.00, .03]	.22 [.20, .25]	-.03 [-.06, .00]	.56 [.54, .58]	.02 [-.01, .04]	-.01 [-.05, .04]
	4	.26 [.23, .29]	.04 [.02, .07]	.18 [.15, .21]	-.04 [-.08, -.01]	.56 [.54, .58]	.00 [-.03, .03]	.08 [.02, .14]
	14		.06 [.02, .10]		-.07 [-.11, -.02]		.01 [-.02, .04]	

Note. Int. = RIASEC interest dimensions; TL = time lag; CO = consistency coefficient; Δ = differences in the consistency, occasion specificity and residual coefficient; OS = occasion specificity coefficient; ε = residual variance; ΔCOOS = difference between consistency coefficient and occasion specificity coefficient, positive values indicate higher values for the consistency coefficient; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional; 1–4 = time points 1 to 4; 14 = lag between T1 to T4.

Table C8

Consistency and Occasion Specificity per Time Point for Generalized Second-Order Growth Models with Quadratic Growth

Int.	TL	CO [95% CI]	Δ [95% CI]	OS [95% CI]	Δ [95% CI]	ε [95% CI]	Δ [95% CI]	ΔCOOS [95% CI]
	1	.37 [.33, .40]		.11 [.07, .14]		.53 [.51, .55]		.26 [.19, .33]
	2	.36 [.34, .39]	-.01 [-.04, .03]	.16 [.13, .18]	.05 [.00, .10]	.48 [.47, .50]	-.04 [-.07, -.02]	.21 [.16, .25]
R	3	.40 [.37, .43]	.04 [.02, .06]	.13 [.10, .15]	-.03 [-.06, .00]	.47 [.45, .49]	-.01 [-.04, .01]	.27 [.23, .32]
	4	.48 [.44, .52]	.08 [.04, .13]	.09 [.05, .13]	-.04 [-.08, .01]	.42 [.40, .44]	.05 [.02, .07]	.39 [.31, .47]
	14		.12 [.07, .17]		-.01 [-.06, .04]		-.11 [-.13, -.08]	
	1	.32 [.30, .35]		.15 [.12, .17]		.53 [.51, .56]		.18 [.13, .22]
	2	.26 [.24, .28]	-.06 [-.09, -.04]	.20 [.18, .23]	.06 [.03, .08]	.54 [.51, .56]	.01 [-.03, .04]	.06 [.02, .10]
I	3	.29 [.26, .32]	.03 [.01, .05]	.18 [.15, .21]	-.02 [-.05, .00]	.53 [.51, .55]	-.01 [-.03, .02]	.11 [.06, .16]
	4	.33 [.30, .36]	.04 [.01, .06]	.21 [.19, .24]	.04 [.01, .06]	.46 [.44, .48]	.07 [.05, .10]	.11 [.06, .17]
	14		.01 [-.03, .04]		.07 [.04, .10]		-.07 [-.11, -.04]	
	1	.26 [.24, .29]		.14 [.11, .16]		.60 [.57, .62]		.13 [.08, .17]
	2	.26 [.24, .28]	.00 [-.02, .02]	.15 [.12, .17]	.01 [-.02, .03]	.59 [.57, .61]	-.01 [-.04, .03]	.12 [.08, .15]
A	3	.28 [.26, .30]	.02 [.00, .03]	.16 [.13, .18]	.01 [-.01, .04]	.56 [.54, .58]	-.03 [-.06, .01]	.12 [.08, .16]
	4	.31 [.29, .33]	.03 [.01, .05]	.14 [.12, .16]	-.02 [-.05, .00]	.55 [.53, .58]	.01 [-.02, .04]	.18 [.14, .21]
	14		.05 [.02, .08]		.00 [-.03, .03]		-.05 [-.08, .01]	
	1	.31 [.27, .34]		.17 [.13, .20]		.53 [.51, .56]		.14 [.07, .21]
	2	.32 [.29, .35]	.01 [-.01, .04]	.20 [.17, .22]	.03 [.00, .05]	.49 [.47, .51]	-.04 [-.07, -.01]	.12 [.07, .18]
S	3	.34 [.31, .38]	.02 [.00, .04]	.18 [.16, .21]	-.01 [-.03, .01]	.47 [.45, .50]	-.01 [-.03, .01]	.16 [.10, .22]
	4	.38 [.35, .41]	.04 [.01, .07]	.18 [.16, .21]	.00 [-.03, .03]	.44 [.41, .47]	.04 [.00, .07]	.20 [.15, .25]
	14		.07 [.04, .11]		.02 [-.02, .05]		-.09 [-.12, -.06]	
	1	.22 [.04, .40]		.19 [.03, .34]		.59 [.56, .63]		.03 [-.30, .37]
	2	.25 [.23, .28]	.03 [-.14, .20]	.15 [.13, .17]	-.04 [-.20, .13]	.60 [.57, .62]	.00 [-.03, .03]	.10 [.07, .14]
E	3	.26 [.24, .28]	.01 [-.01, .02]	.14 [.11, .18]	-.01 [-.05, .03]	.60 [.56, .64]	.00 [-.03, .04]	.12 [.08, .16]
	4	.28 [.07, .49]	.02 [-.18, .23]	.11 [-.10, .33]	-.03 [-.26, .21]	.60[.59, .62]	-.01 [-.05, .03]	.17 [-.26, .59]
	14		.06 [-.01, .13]		-.07 [-.17, .02]		.01 [-.03, .05]	

	1	.16 [.10, .22]		.27 [.21, .33]		.57 [.55, .59]		-.11 [-.23, .01]
	2	.25 [.21, .28]	.08 [.02, .15]	.22 [.19, .25]	-.05 [-.12, .02]	.53 [.51, .55]	-.04 [-.06, -.01]	.02 [-.04, .08]
C	3	.26 [.23, .29]	.02 [.00, .04]	.19 [.16, .22]	-.04 [-.07, .00]	.55 [.53, .57]	.02 [-.01, .04]	.08 [.02, .13]
	4	.28 [.24, .32]	.02 [-.02, .05]	.15 [.12, .19]	-.03 [-.07, .01]	.57 [.55, .59]	-.02 [-.04, .01]	.13 [.05, .20]
	14		.12 [.05, .18]		-.12 [-.18, -.05]		.00 [-.03, .03]	

Note. Int. = RIASEC interest dimensions; TL = time lag; CO = consistency coefficient; Δ = differences in the consistency, occasion specifity and residual coefficient; OS = occasion specificity coefficient; ε = residual variance; ΔCOOS = difference between consistency coefficient and occasion specificity coefficient, positive values indicate higher values for the consistency coefficient; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional; 1–4 = time points 1 to 4; 14 = lag between T1 to T4.

Table C9

Consistency and Occasion Specificity Averaged across All Time Points

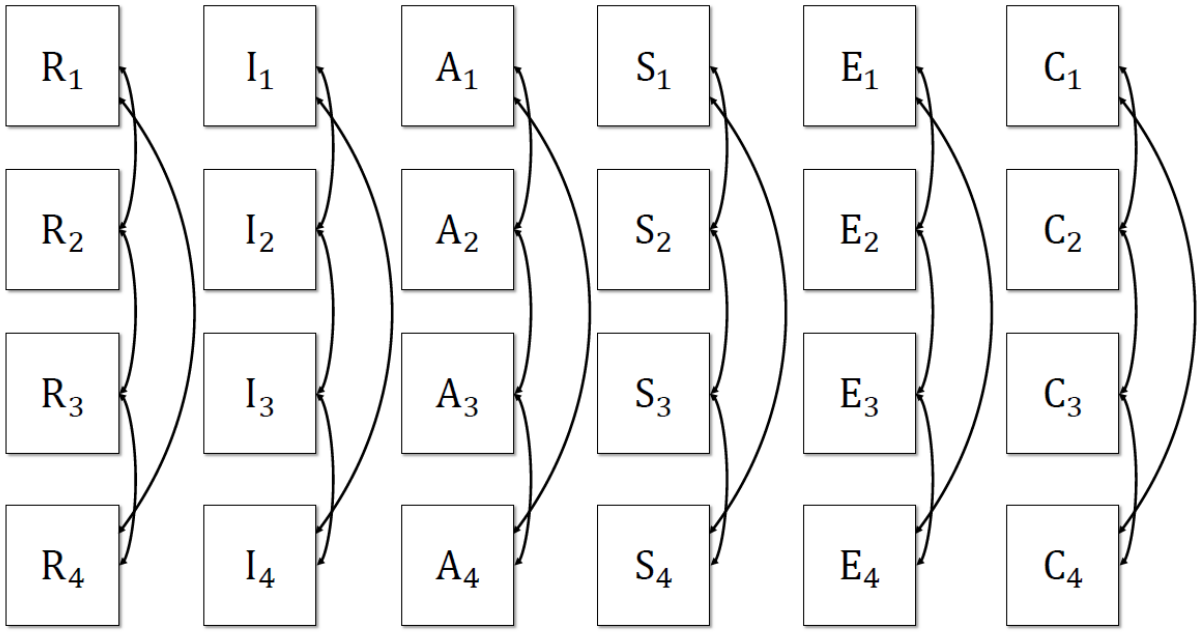
Growth Type	Int.	CO [95% CI]	OS [95% CI]	ε [95% CI]	ΔCOOS [95% CI]
Linear	R	.34 [.32, .36]	.18 [.16, .19]	.48 [.47, .49]	.17 [.13, .20]
	I	.24 [.22, .26]	.23 [.21, .25]	.53 [.52, .55]	.01 [-.03, .05]
	A	.24 [.23, .26]	.17 [.15, .18]	.59 [.58, .60]	.07 [.05, .10]
	S	.28 [.25, .30]	.22 [.19, .24]	.50 [.49, .52]	.06 [.01, .11]
	E	.22 [.20, .23]	.19 [.17, .20]	.60 [.59, .61]	.03 [.002, .06]
	C	.22 [.20, .2]	.23 [.21, .25]	.56 [.54, .57]	-.01 [-.05, .03]
Quadratic	R	.40 [.38, .42]	.12 [.10, .14]	.48 [.46, .49]	.28 [.24, .3]
	I	.30 [.28, .32]	.19 [.17, .21]	.51 [.50, .53]	.12 [.07, .16]
	A	.28 [.26, .30]	.14 [.13, .16]	.58 [.56, .59]	.14 [.11, .17]
	S	.34 [.31, .37]	.18 [.16, .21]	.48 [.47, .50]	.15 [.10, .21]
	E	.25 [.15, .36]	.15 [.06, .23]	.60 [.58, .62]	.11 [-.08, .29]
	C	.24 [.21, .27]	.21 [.18, .23]	.55 [.54, .57]	.03 [-.02, .08]

Note. Int. = RIASEC interest dimensions; CO = consistency coefficient; OS = occasion specificity coefficient; ε = residual variance; ΔCOOS = difference between consistency coefficient and occasion specificity coefficient, positive values indicate higher values for the consistency coefficient; R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional.

Supplement D: Model Depictions and Formulas

Depiction of the Saturated Path Model

Figure D1
Depiction of the saturated path model

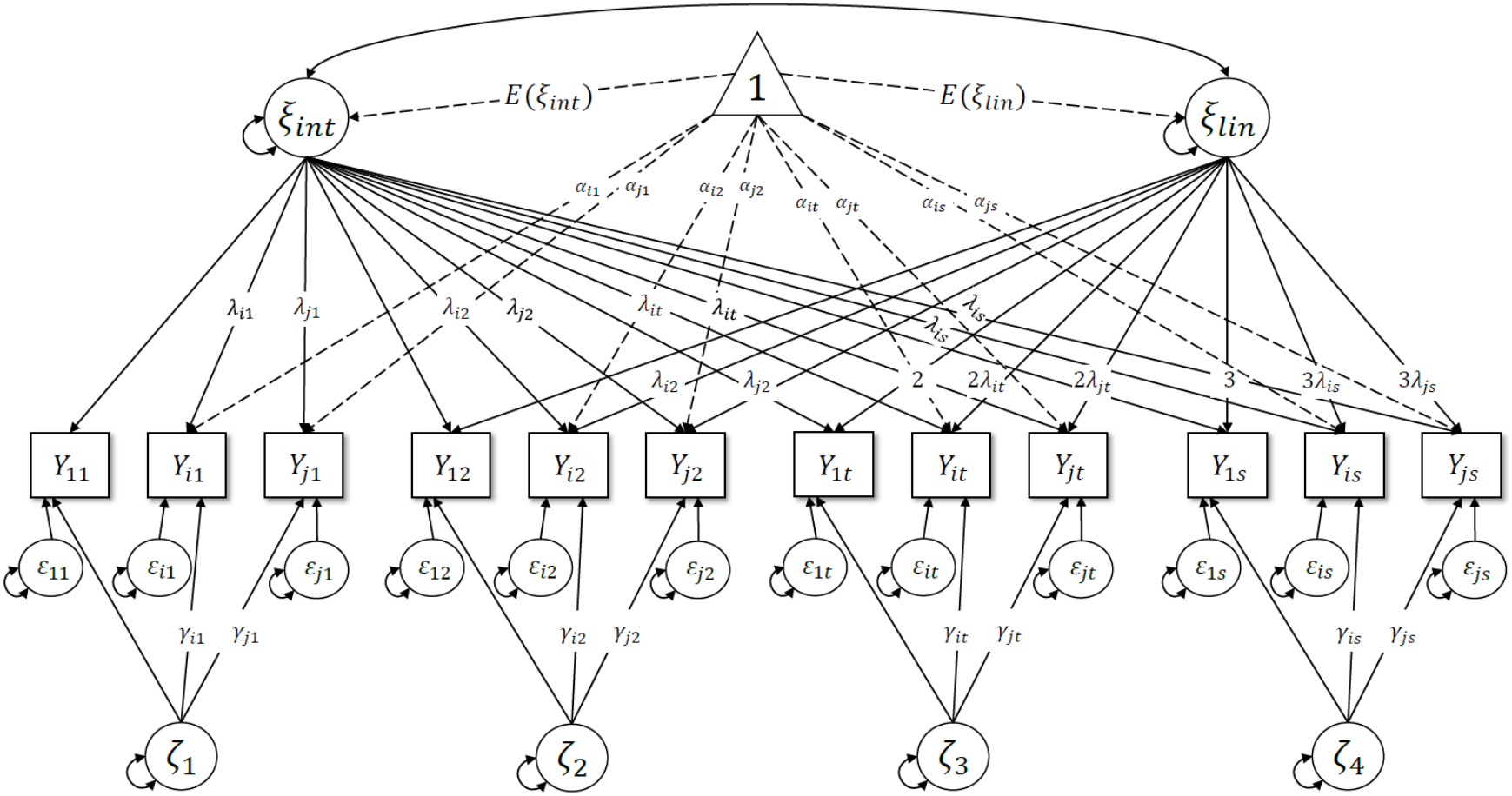


Note. Only the retest correlations are depicted. However, all possible correlations between all manifest RIASEC scales were estimated. For simplicity, the other correlations are not depicted in the model.

Depiction of the Generalized Second-Order Growth Models

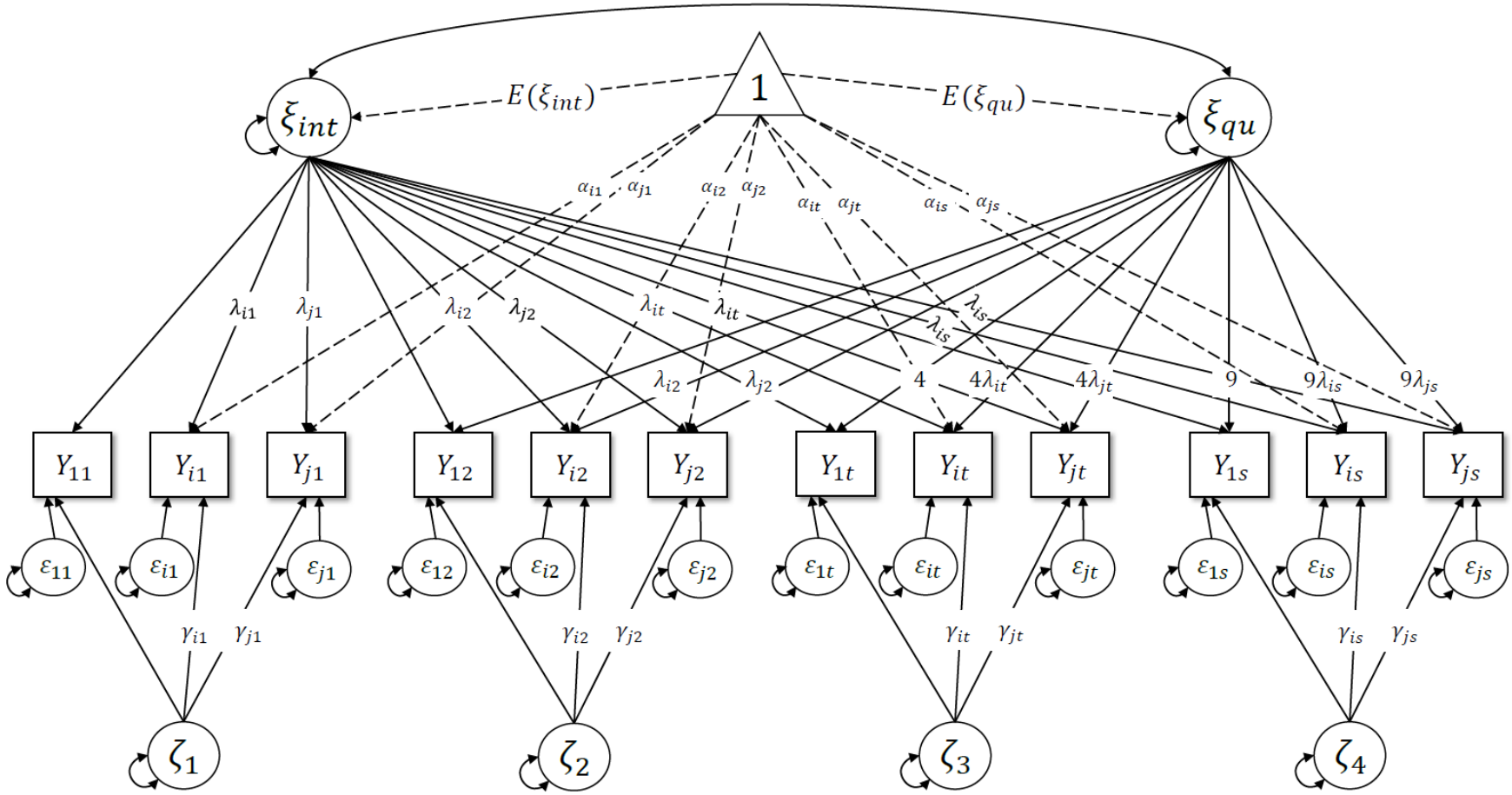
Figure D2

Depiction of the GSGM with linear growth.



Note. Correlated uniqueness is not depicted. Intercept factors were identified by fixing the first factor loading to 1 and the first intercept to 0.

Figure D3
Depiction of the GSGM with quadratic growth.



Note. Correlated uniqueness is not depicted. Intercept factors were identified by fixing the first factor loading to 1 and the first intercept to 0.

Formulas for Consistency and Occasion Specificity for the Linear Models

$$\mathrm{var}\left(Y_{it}\right) = \lambda_{it}^2\ \mathrm{var}(\xi_{\mathrm{int}}) + \lambda_{it}^2\ \mathrm{var}(\xi_{lin}) + 2\lambda_{it}^2\ \mathrm{cov}(\xi_{\mathrm{int}}, \xi_{lin}) + \gamma_{it}^2\ \mathrm{var}(\zeta_t) + \mathrm{var}(\varepsilon_{it}) \tag{S1}$$

$$CO\left(Y_{it}\right) = \frac{\lambda_{it}^2\ \mathrm{var}(\xi_{\mathrm{int}}) + \lambda_{it}^2\ \mathrm{var}(\xi_{lin}) + 2\lambda_{it}^2\ \mathrm{cov}(\xi_{\mathrm{int}}, \xi_{lin})}{\mathrm{var}\left(Y_{it}\right)} \tag{S2}$$

$$OS\left(Y_{it}\right) = \frac{\gamma_{it}^2\ \mathrm{var}(\zeta_t)}{\mathrm{var}\left(Y_{it}\right)} \tag{S3}$$

Formulas for Consistency and Occasion Specificity for the Quadratic Models

$$\begin{aligned}&\mathrm{var}\left(Y_{it}\right) = \lambda_{it}^2\ \mathrm{var}(\xi_{\mathrm{int}}) + \lambda_{it}^2\ \mathrm{var}(\xi_{lin}) + \Upsilon_{it}^2\ \mathrm{var}(\xi_{quad}) + \\ &2\lambda_{it}^2\ \mathrm{cov}(\xi_{\mathrm{int}}, \xi_{lin}) + 2(\lambda_{it}\Upsilon_{it}\ \mathrm{cov}(\xi_{\mathrm{int}}, \xi_{quad})) + 2(\lambda_{it}\Upsilon_{it}\ \mathrm{cov}(\xi_{lin}, \xi_{quad})) + \\ &\gamma_{it}^2\ \mathrm{var}(\zeta_t) + \mathrm{var}(\varepsilon_{it})\end{aligned} \tag{S4}$$

$$CO\left(Y_{it}\right) = \frac{\begin{array}{l}\lambda_{it}^2\ \mathrm{var}(\xi_{\mathrm{int}}) + \lambda_{it}^2\ \mathrm{var}(\xi_{lin}) + \Upsilon_{it}^2\ \mathrm{var}(\xi_{quad}) + \\ 2\lambda_{it}^2\ \mathrm{cov}(\xi_{\mathrm{int}}, \xi_{lin}) + 2(\lambda_{it}\Upsilon_{it}\ \mathrm{cov}(\xi_{\mathrm{int}}, \xi_{quad})) + 2(\lambda_{it}\Upsilon_{it}\ \mathrm{cov}(\xi_{lin}, \xi_{quad}))\end{array}}{\mathrm{var}\left(Y_{it}\right)} \tag{S5}$$

$$OS\left(Y_{it}\right) = \frac{\gamma_{it}^2\ \mathrm{var}(\zeta_t)}{\mathrm{var}\left(Y_{it}\right)} \tag{S6}$$

Formula for Cohen's d for Repeated Measures

$$Cohen's\ d_{rm} = \frac{M_{diff}}{\sqrt{SD_1^2 + SD_1^2 + 2rSD_1SD_2}}\sqrt{2(1-r)} \tag{S7}$$

4 STUDY 2

WHY DO SOME PEOPLE HAVE MORE STABLE VOCATIONAL INTEREST PROFILES THAN OTHERS? RESULTS FROM FOUR LONGITUDINAL STUDIES IN DIFFERENT AGE GROUPS

Gfrörer, T., Stoll, G., Rieger, S., Nagengast, B., & Trautwein, U. (to be submitted). Why Do Some People Have More Stable Vocational Interest Profiles Than Others? Results From Four Longitudinal Studies in Different Age Groups.

Abstract

A central aim of interest-based career guidance is to help young people to choose occupations or study majors that match their interests (Holland, 1997). For that purpose, the stability of vocational interest profiles is key, as job and study major suggestions are derived from them. If interest profiles undergo significant changes from one year to the next, they would be an unreliable guide for prospective career choices (Low et al., 2005; Strong, 1931). In the present investigation we examined the stability of vocational interest profiles and its predictors based on reanalyses of data sets from four separate longitudinal studies. We operationalized profile stability of vocational interests based on the Pearson correlation *r* between two profiles. The results of the four studies suggest that during the phases of vocational orientation (mean age: 11.24 to 14.24; $r = .43$), vocational training (mean age: 14.37 to 15.36; $r = .65$), college major orientation and college major choice (mean age: 16.71 to 23.34; $r = .64$), as well as workforce entrance (mean age: 21.55 to 33.55; $r = .74$), vocational interest profiles were moderately to highly stable. We found a consistent effect of gender, with females having more stable profiles than males, whereas personality traits and cognitive abilities had small and inconsistent effects on profile stability. Practical implications for career guidance are discussed.

Keywords: vocational interests, profile stability, gender, personality traits, cognitive abilities, life course

Why Do Some People Have More Stable Vocational Interest Profiles Than Others? Results From Four Longitudinal Studies in Different Age Groups

A central aim of interest-based career guidance is to help young people to choose occupations or study majors that match their interests (Holland, 1997). This is an important societal task because a good fit between (vocational) interests and work or study environments predicts beneficial outcomes such as higher work performance (Nye et al., 2012, 2017), income (Neumann et al., 2009), work satisfaction (Tsabari et al., 2005), and college major persistence (Allen & Robbins, 2008; Le et al., 2014) as well as lower levels of counter-productive work behavior (Nye et al., 2017).

Schools, universities, and employment offices that apply interest-based career guidance use inventories to measure students' interests and aggregate them into profiles that are matched with suitable occupations or study majors (Hanna & Rounds, 2020; Hoff et al., 2019; Holland, 1997). Suggestions about jobs or study majors are usually derived based on the expression of the respective interest profile. For example, jobseekers with an interest in business, artistic, and social activities would receive different suggestions compared to jobseekers with an interest in scientific, technological, and administrative activities. Because individual interest profiles are key in career guidance, their stability across longer periods of time is important for the predictive validity of interest inventories (Low et al., 2005). If interest profiles undergo significant changes from one year to the next, they would be an unreliable guide for prospective career choices (Low et al., 2005; Schomburg & Tokar, 2003; Strong, 1931; Swanson, 1999).

However, despite the widespread application of interest inventories (Hansen, 2019), still little is known about the stability of vocational interest profiles within different career stages. Only a few studies examined the stability of interest profiles (see Low et al., 2005), and investigations in younger age groups are especially scarce. In addition, the existing studies report significant interindividual differences in profile stabilities (e.g., Swanson & Hansen, 1988, 1988; Xu & Tracey, 2016), raising the question of why some people have more stable interest profiles than others. Finding constructs that predict unstable configurations of vocational interest profiles would facilitate the integration of such additional information in the career guidance process (Ackerman & Beier, 2003; Dawis, 1992; Lowman, 1991).

In the present study, we aim to provide new insights about vocational interest profile stabilities by investigating potential predictors that might be associated with individual differences in profile stability. We focused on predictors that are likely to be available in career guidance contexts (Ackerman & Beier, 2003; Dawis, 1992): personality traits, cognitive

abilities, and gender. As career guidance is important in multiple age groups, we investigated profile stability and its predictors in four life phases that confront people with actual or imminent career transitions. To this end, we reanalyzed data from four large-scale longitudinal research projects that covered these life phases and provided information on vocational interests, personality traits, cognitive abilities, and gender.

Relevant Life Phases for Career Guidance

Interest-based career guidance is frequently utilized during various life phases. In the United States alone, every year about four million high school students as well as half a million postsecondary students and adults fill out the ACT interest inventory to support their career choice process (American College Testing Program, 2009). Career guidance usually becomes important during phases in which educational or occupational transitions occur that urge people to make career-related decisions (Gati et al., 2019). This is often intertwined with the respective educational system, which determines when transitions, such as entrance to the job market, take place. Career guidance becomes relevant not only during explicit transitions, such as the completion of compulsory education, but also during more subtle transitions that occur within a respective school track, for example before the choice of advanced school courses. As educational systems usually provide students the possibility to pursue different career paths (e.g., obtaining a university degree vs. finishing vocational educational training), several life phases can be relevant for career guidance.

Although career guidance seems less important during younger age groups at first glance, adolescents begin early to attend school-based vocational orientation events that inform them about occupations, assist them in exploring their abilities and goals and provide them further help in planning their future career paths (Gfrörer et al., 2021; Gysbers & Lapan, 2001; Kracke, 1997; Noack et al., 2010). In economies where vocational educational training is common, this process usually begins during lower secondary education (about age 11 to 15; Eurostat, 2021a). For example, in Germany, adolescents often begin their vocational apprenticeship as early as the age of 15 (i.e., approximately 55,000 adolescents each year; BIBB, 2018) indicating that they need career guidance before this age (Eurostat, 2021a). Vocational educational training is often chosen by lower and middle track students who do not pursue a higher school track. For those students, vocational orientation is explicitly emphasized in the school curriculum (Maaz et al., 2008).

At the end of lower secondary education and the beginning of upper secondary education in Germany (about age 15 to 16), students have the possibility to leave school for

vocational educational training (Eurostat, 2021a). The training is characterized either by solely attending a vocational school or by attending a vocational school and simultaneously working in the respective occupation (Maaz et al., 2008). This career decision is especially common in Europe, as approximately half of the students in upper secondary education choose such a career path (Eurostat, 2021a). In Germany, students who finish lower secondary education also have the possibility to choose an academic school track that prepares them for university, instead of vocational educational training (Maaz et al., 2008). Career guidance is highly relevant during that time, as both choices (i.e., vocational educational training or the academic school track) constitute a long-term commitment (i.e., two to three years) which often sets the track for an individual's prospective career.

At the end of upper secondary education (about age 19), students are confronted with career decisions concerning tertiary education (Maaz et al., 2008). After finishing upper secondary school, students who followed an academic school track and did not choose to follow vocational educational training begin to search for a suitable study major and pursue a university degree. In 2018 about 60% of the students in tertiary education in Europe were studying for bachelor's degrees (Eurostat, 2021b). Similarly high enrollment was found in the United States, where about 20 million students were enrolled in university in the year of 2017 (National Center for Education Statistics, 2019). Career guidance concerning study major choices are often implemented by schools and universities.

The end of tertiary education, which falls into the time of young adulthood (about age 18 to 26), is characterized by students finishing their studies and obtaining a university degree. In Europe and the United States, this can be fulfilled at the undergraduate or graduate level. The end of tertiary education is also accompanied by entering the job market and finding a suitable occupation. Career guidance during that life phase is offered by government institutions such as employment offices, but also partly by universities. Students who expect to undergo such transition periods often receive career counseling that entails the application of interest inventories (Holland, 1997). Although there are numerous inventories that were designed to measure vocational interests (see Hansen, 2019), many of them allow counselors to categorize a student's vocational interests with Holland's (1997) vocational interest model (Hansen, 2019).

The Stability of Vocational Interest Profiles

Holland's (1997) RIASEC model is the most widely used taxonomy for vocational interests (Rounds & Su, 2014). According to this model, a person's vocational interests can be characterized by six interest dimensions: Realistic, Investigative, Artistic, Social, Enterprising,

and Conventional (RIASEC; for further information see Table 1). The scores of a person on the six dimensions constitute the interest profile of this person (Holland, 1997). For career guidance, the ordering of the dimensions within an interest profile—the profile shape—is particularly decisive (Holland, 1997; Prediger, 1998; Xu & Li, 2020). The profile shape constitutes the intraindividual expression of vocational likes and dislikes (Etzel & Nagy, 2021; Perera & McIlveen, 2018; Rounds et al., 1987; Xu & Li, 2020) that predict occupational group membership (Prediger, 1998). In contrast, profile characteristics such as elevation (i.e., mean level of the dimensions within a profile) are often treated as a nuisance (Tracey, 2012), as they are not practically relevant in predicting occupational group membership (Prediger, 1998).

For most interest inventories that are based on Holland's (1997) model, the matching between interest profiles and job or study major suggestions is based on the ordering of the three dimensions for which the person expresses the highest interests (Hansen, 2019). This approach is used because of its practicability and flexibility (Hansen, 2019). However, the approach also has its difficulties, as it depends on the stability of vocational interests. For example, jobseekers with highest scores on the dimensions Social, Investigative, and Artistic—descending, in that exact order—would receive job suggestions such as Mental Health Counselor, History Teacher, and Speech-Language Pathologist (O*NET, 2021). But jobseekers with highest scores on Social, Investigative and Enterprising—descending, in that exact order—would receive different suggestions including Law Teacher, Community Health Worker, and Instructional Coordinator (see Figure 1; O*NET, 2021). This illustrates that even small differences in the shape of an interest profile can have practical consequences and that the stability of the profile shape is decisive for career guidance.

Although stability of vocational interests has been frequently investigated (see Low et al., 2005), still little is known about profile stability during phases that are relevant for career guidance. Evidence suggests that vocational interest profiles are relatively stable over longer periods of time (e.g., 30-year period, $r = .54$, Rottinghaus et al., 2007; 12-year period, average $r = .73$, Zytowski, 1976). When comparing across different life phases, profile stability has been found to be lower during phases in adolescence, including the transition from lower to upper secondary education (e.g., 1-year period between 9th and 10th grade, average $r = .78$, Xu & Tracey, 2016) and higher during phases in adulthood, including the transition from upper secondary to tertiary education (e.g., 2-year period between age 21 and age 23, $r = .84$, Stoll et al., 2020). To date, empirical findings about profile stability during earlier phases of career guidance, for example at the beginning of lower secondary education, are scarce.

Table 1
Reliability Coefficients

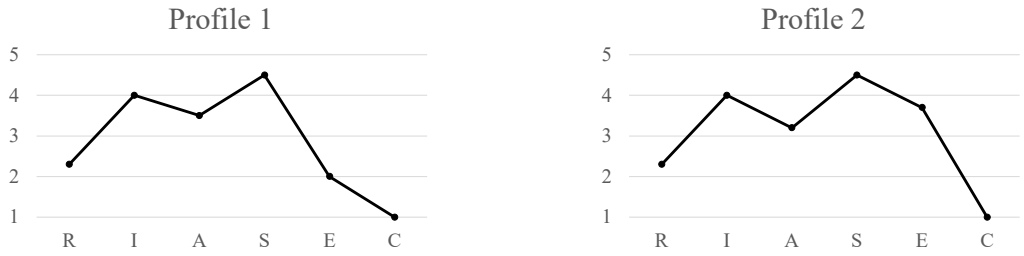
	Description	Item Examples	Study 1 ω (T1, T2)	Study 2 ω (T1, T2)	Study 3 ω (T1, T2)	Study 4 ω (T1, T2)
Vocational Interests						
Realistic	Practical activities	"Building something"	.85, .89	.88, .91	.89, .88	.87, .87
Investigative	Problem solving/analytical thinking	"Experimenting in a lab"	.81, .84	.81, .83	.83, .84	.84, .85
Artistic	Creative activities	"Drawing pictures"	.78, .79	.78, .80	.84, .86	.86, .85
Social	Teaching, caring, and informing	"Helping others"	.83, .85	.86, .87	.90, .88	.90, .88
Enterprising	Manipulating and leading others	"Leading a group"	.81, .79	.77, .78	.87, .87	.88, .88
Conventional	Structured tasks (e.g., organizing/sorting)	"Organizing things"	.82, .82	.79, .80	.85, .86	.86, .85
Personality Traits						
Openness	Need for variety/novelty/change	"I often try out new things"	.63	.67	.68	.76
Openness (w/o r)			.81	.80	.57	.63
Conscientiousness	Strong sense of purpose/high aspiration	"I am thorough"	.42	.64	.36	.84
Conscientiousness (w/o r)			.77	.79	.24	.78
Extraversion	Preference for companionship	"I like to be with people"	.35	.71	.81	.80
Extraversion (w/o r)			.72	.77	.70	.75
Agreeableness	Willingness to defer to others	"Most people like me"	.20	.48	.53	.74
Agreeableness (w/o r)			.67	.65	-	.56
Neuroticism	Tendency to experience dysphoric affect	"I often feel nervous"	.36	.55	.67	.87
Neuroticism (w/o r)			.71	.68	.68	.82
Cognitive Abilities						
Quantitative	Numerical cognitive abilities	"Complete: 3 6 12 24 X"	-	-	-	-
Figural	Figural cognitive abilities	"Rotate figure X"	-	-	-	-
Verbal	Verbal cognitive abilities	"What do bricklayers do?"	-	-	-	-

Note. ω = omega total (see McNeish, 2018); (w/o r) = implies that scales did not include the reverse coded items.

Existing findings also indicate that there is substantial between-person variation in the stability of interest profiles. Swanson (1999) stated that significant between-person variations in profile stability are well known as they have been consistently reported in studies since the 1930s (Schomburg & Tokar, 2003). For example, Swanson and Hansen (1988) showed ranges in profile correlation for vocational interests over a 12-year time period from -.27 to .97, indicating that profiles of some participants changed substantially over time (i.e., -1 implies a completely reversed profile pattern), whereas others nearly stayed the same (i.e., 1 implies a completely stable profile pattern). A similar range for a 30-year time period was reported by Rottinghaus et al. (2007) with profile correlations for vocational interests ranging from -.29 to .95. In addition, Xu and Tracey (2016) reported standard deviations in profile stability coefficients between persons that were substantially different from zero ($0.27 < SD < 0.37$).

Figure 1

*Depiction of Different Vocational Interest Profiles and Their Occupational Suggestions Derived From O*NET*



Matching occupations according to O*NET

Holland Code, SIA:
- History Teacher
- Mental Health Counselor
- Genetic Counselor
- Speech-Language Pathologist
- …

Holland Code, SIE:
- Law Teacher (Postsecondary School)
- Community Health Worker
- Instructional Coordinator
- …

Note. Hypothetical Vocational Interest Profiles with O*NET suggestions.

Possible Predictors of Vocational Interest Profile Stability

The substantial between-person differences in vocational interest profile stability raise the question of whether there are relatively stable individual factors that can explain these variations. As empirical studies are still missing to date, relatively little is known about predictors that influence vocational interest profile stability (see Schomburg & Tokar, 2003).

Swanson (1999) stated that differences in interest stability might occur due to individual difference variables, such as personality characteristics. His claims, however, are lacking empirical evidence so far. In the following, we therefore relied on theoretical and empirical work from the research area of individual differences to derive assumptions about possible predictors that might be associated with vocational interest profile stability. In addition, as vocational interest profile stability is also practically relevant, suggestions from the practice of career guidance were considered (e.g., the availability of personality measures in career guidance settings). On the basis of the respective considerations—theoretical, empirical, and practical—we selected a potential set of predictors to test systematically in four separate longitudinal studies.

Theoretical and Empirical Considerations About Predictors of Profile Stability

Many theories about personality define three broad areas of individual differences that summarize relatively stable dispositional characteristics: cognitive abilities, personality traits, and interests (Ackerman & Heggestad, 1997; Roberts & Wood, 2006). It is assumed that cognitive abilities, personality traits, and interests are responsible for a majority of people's behavioral patterns (Kandler et al., 2014; Roberts & Wood, 2006), their choices (e.g., Stoll, Einarsdóttir, et al., 2020; Wille et al., 2020) and their accomplishments (e.g., Cheng & Furnham, 2012; Roberts et al., 2007; Spengler et al., 2018; Stoll et al., 2017). Many studies have associated vocational interests with personality traits and cognitive abilities—theoretically (e.g., Ackerman, 1996; Ackerman & Beier, 2003; Schmidt, 2014) and empirically (Ackerman & Heggestad, 1997; Armstrong et al., 2008; Larson et al., 2002; Mount et al., 2005). It is suggested that the respective constructs are intertwined with each other (Armstrong et al., 2008) and develop in interaction with each other (Hoff et al., 2020). Therefore, variables that are associated with between-person differences in the stability of vocational interest profiles are most likely located within this cluster of constructs.

However, prior studies focused solely on associations between single dimensions of vocational interests, personality traits, and cognitive abilities (Ackerman & Heggestad, 1997; Hoff et al., 2020). To our knowledge, no study exists that theoretically describes or empirically investigates if personality traits and cognitive abilities also influence the stability of vocational interest profiles. This contradicts considerations derived from interest development theories, which suggest that personality characteristics might influence profile stability. Many of these theories describe that having experiences in various environments is key for interest solidification (Holland, 1997; Renninger & Hidi, 2011; Su et al., 2019). Holland (1997) and Su

et al. (2019) state that the solidification of (vocational) interests is fostered through an accumulation of positive experiences, where preferences for single activities are gradually rewarded and reinforced—for example, through personal satisfaction or external rewards. Personality traits and cognitive abilities influence how people encounter, process, and incorporate affective and cognitive components from different experiences (Kandler et al., 2014). In addition, they broadly predict the selection of different environments (e.g., Päßler & Hell, 2012) as well as activity diversity and quantity (e.g., Jackson et al., 2020). Consequently, differences in personality traits and cognitive abilities might result in having different experiences and ultimately influence the solidification of vocational interests and specifically vocational interest profiles. This also implies that when the meaning of single interest dimensions changes for a person due to different experiences, the rank order of the elements within a profile, as well as profile stability, also changes.

Another factor that might be associated with the stability of vocational interest profiles is gender. Meta-analytic evidence indicates that vocational interests demonstrate robust and large gender differences (Su et al., 2009). Although gender differences are found primarily in mean levels, it still raises the question if they can also be found in stability. Low et al. (2005) found no moderation effect of gender in their meta-analysis about stability. However, they investigated a joint stability coefficient, combining rank order and profile stability coefficients. More recent studies (Stoll, Rieger, et al., 2020; Xu & Tracey, 2016) found differences in vocational interest profile stability between men and women. Xu and Tracey (2016) found small differences in the stability coefficients between boys and girls during adolescence, with girls having more stable profiles than boys. Similarly, Stoll et al. (2020) found statistically significant gender differences in profile stability during young adulthood, with women having more stable profiles than men. Although in both studies differences were small in magnitude, their results indicate that gender could be a predictor of profile stability.

Practical Considerations About Predictors of Profile Stability

In line with the theoretical and empirical considerations, practitioners in career assessment have recommended measuring multiple constructs during vocational counseling sessions, focussing especially on personality traits, cognitive abilities, and vocational interests (Ackerman & Beier, 2003; Dawis, 1992; Lowman, 1991) because this integrated approach improves suggestions about future occupations (Lowman, 1991). Similarly, Ackerman and Beier (2003) recommended integrating the constructs of personality traits, cognitive abilities, and vocational interests to improve the understanding of career choice processes. They

illustrated that vocational counselors usually have more information about their clients than solely the results of an interest inventory (Ackerman & Beier, 2003).

Knowing how these constructs are associated with the stability of interest profiles could help vocational counselors to integrate this information into their counseling process (Ackerman & Beier, 2003; Hoff et al., 2019). On the basis of additional information from personality traits and cognitive abilities, vocational counselors could identify jobseekers with potentially lower profile stability (Swanson, 1999). Although adaptive counseling may not be feasible in a strict sense, the identification of unstable profile constellations could lead to an adjusted utilization of interest inventories. For example, for people with potentially low interest profile stability, counselors could use their interest inventory results to suggest a large number of occupations or study majors that might be suitable for exploration, rather than to predict fitting occupational or study major choices for the future.

The Present Investigation

In the present investigation we examined the stability of vocational interest profiles and its predictors based on reanalyses of data sets from four separate longitudinal studies. The respective studies provide new insights about vocational interest profile stability in early career stages, an overview of stability during various life phases in which career transitions occur or are imminent, and new information about the relationship between individual characteristics and vocational interest profile stability. Each of the four studies captured a different phase that included actual or imminent career transitions.

The first study investigated a 3-year period across late childhood and early adolescence (ages 11 to 14) that is associated with the start of vocational orientation. The second study examined a 1-year period during middle adolescence (ages 14 to 15), which is associated with the start of vocational training. Especially for students from lower and middle track schools, these life phases are related to career guidance. The third study investigated a 6-year period from late adolescence to young adulthood (ages 17 to 23), which is associated with the phase of college major orientation and college major choice. The fourth study examined a 12-year period across young adulthood (ages 22 to 34), which is associated with entering the workforce after ending tertiary education.

We had three overall study aims: (1) We examined the stability of vocational interest profiles during life phases in which career guidance seems relevant. We assumed that vocational interest profiles should be relatively stable over the course of the respective time periods, in accordance with previous evidence (Rottinghaus et al., 2007; Stoll, Rieger, et al., 2020; Swanson & Hansen, 1988; Xu & Tracey, 2016; Zytowski, 1976), but (2) also examined stability variation between participants (e.g., Rottinghaus et al., 2007; Swanson & Hansen, 1988; Xu & Tracey, 2016). When there was variation between participants, (3) we examined the association between personality traits, cognitive abilities, and gender and vocational interest profile stability during the respective life phases. As this is the first study that investigates the association between personality characteristics and profile stability of vocational interests, we made no assumptions about the strength of the relationships. Instead, we exploratively investigated the potential associations between profile stability and personality traits, cognitive abilities, and gender.

Methods

Educational System

The participants included in the four studies attended different school tracks in Germany, with almost all of them being in secondary school in the federal state of Baden-Württemberg. The German educational system offers an ideal framework to examine the three study aims, as it provides students the possibility to pursue various career paths (e.g., vocational educational training or university degree) that result in transitions at different age periods. In the German state of Baden-Württemberg, secondary school starts at 5th grade, after primary school, and consists of three educational tracks (low, middle, and high). The lowest track is characterized by students finishing school after 9th or 10th grade. Although they have the option of choosing different educational arrangements, many of these students in the lowest track start vocational educational training afterwards. The middle track ends after 10th grade and students usually have the option either to enter vocational educational training or to attend a higher school track to obtain university entrance qualifications. The highest school track prepares students for university and usually ends after 12th or 13th grade (for more information about the German school system see Maaz et al., 2008).

Operationalization of Constructs

Each of the four studies consisted of data from two time points and comprised participants that partook at either one of the two time points (or both time points; see Table 2 for sample characteristics). Vocational interests, measured at both time points, were used to estimate the profile stability coefficients. The predictors (i.e., personality traits, cognitive abilities, and gender) were measured at the first time point. In all four studies, vocational interests were operationalized based on Holland's (1997) RIASEC model, personality traits were operationalized based on the Big Five framework (Costa & McCrea, 1996) and cognitive abilities were operationalized to capture the processing capacity based on the Berlin intelligence structure model (Heller & Perleth, 2000), which included subscales of quantitative, figural, and verbal cognitive abilities. The following analytic steps were applied in all four of the studies.

Table 2
Overview of Sample Characteristics

Study	Period	Interval (# Years)	n_1	n_2	N	# Classes	# Schools	♀T1	♀T2	Mean Age T1	Mean Age T2	School Type	Original Study
1	Vocational Orientation	3	2,894	3,060	3,746	136	105	46%	46%	11.24 (5th Grade)	14.24 (8th Grade)	Lower, Middle & Multi-Track	TRAIN C1
2	Vocational Training	1	2,291	2,241	2,707	192	72	47%	48%	14.37 (8th Grade)	15.36 (9th Grade)	Lower, Middle & Multi-Track	TRAIN C2
3	College Orientation & Choice	6	3,047	940	3,047	116	66	50%	56%	16.71 (10th Grade)	23.43	Middle & Higher Track	TOSCA-10
4	Workforce Entrance	12	2,392	1,401	2,455	-	147	62%	61%	21.55	33.55	Higher Track	TOSCA-2002

Note. For Study 4, the information about the schools of TOSCA-2002 is retrieved from the first wave of the original study when participants were in 12th grade, not from the wave labelled as T1 in the current study.

Operationalization of Profile Stability

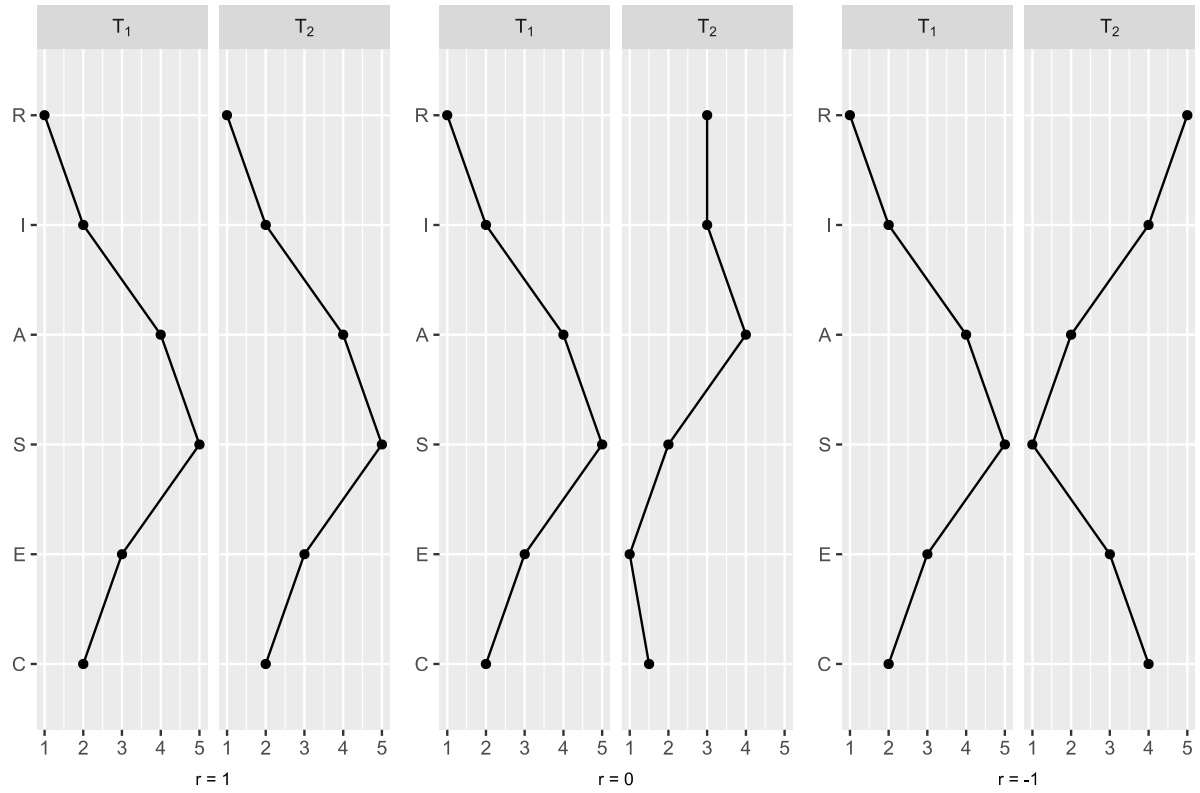
As an indicator for profile stability, we used Q-correlation—a widely used index to examine the within-profile rank-order stability (Burt, 1937; Livingston et al., 2003; Xu & Li, 2020). Q-correlation is a Pearson correlation between two variables that contain all the respective elements (i.e., the six RIASEC scores) of two distinct profiles (Livingston et al., 2003). It is sensitive to the rank order of the profile elements, and therefore to deviations from the profile shape but not to changes in profile elevation (Livingston et al., 2003). Because the relative strength (i.e., the ordering) of the RIASEC dimensions within a profile is key in career guidance contexts, the Q-correlation is the most suitable coefficient for our study (Xu & Li, 2020). Results from congruence research indicate that the Q-correlation is a profile similarity indicator with high practical relevance (Xu & Li, 2020).

Figure 2 shows three hypothetical profiles with stabilities of $r \approx -1$, $r \approx 0$ and $r \approx +1$. A profile correlation of $r = -1$ implies a completely reversed profile pattern, a profile correlation of $r = 0$ implies an orthogonal profile pattern, and a profile correlation of $r = 1$ implies a completely stable profile pattern. Higher values imply smaller changes in the shape of the profiles and higher profile stability. It is important to note that profile correlations around 0 do not suggest that the respective profile is highly unstable. As shown in Figure 2, with a profile correlation around 0, the respective profile shows changes in the rank order of only a few profile elements. In line with current empirical evidence (e.g., Rottinghaus et al., 2007), we see profile correlations between .30 and .60 as indicating moderate stability and profile correlations above .60 as indicating high profile stability (e.g., Stoll, Rieger, et al., 2020; Xu & Tracey, 2016).

We computed Q-correlation for each participant and used Mplus 8 (Muthén & Muthén, 1998-2017) to compute the mean level of the Q-correlation across all participants in each study based on an intercept-only model. We investigated the variance of profile stability by examining the variance of the Q-correlation based on the respective intercept-only model.

Figure 2

Depiction of Different Profile Stabilities According to the Q-Correlation



Note. Q-Correlations of hypothetical Vocational Interest Profiles at two different time points. Scale ranges from 1 to 5.

Investigating Predictors of Profile Stability

To investigate potential predictors, we specified multiple regression models with the Q-correlation as a dependent variable. First, we specified three models, each including personality traits (i.e., openness, conscientiousness, extraversion, agreeableness, and neuroticism), cognitive abilities (i.e., quantitative, figural, and verbal), or gender as an independent variable. Second, we specified an overall model in which personality traits, cognitive abilities, and gender jointly predict the profile stability indicators. We z-standardized the non-binary independent variables prior to the analysis. We included the remaining variables that were not used as predictors in the respective models, as auxiliary variables via the saturated correlates model approach (Graham, 2003). This was also done for the intercept-only model that investigated the mean level of the Q-correlation. The auxiliary variables were included into the analysis through correlations with variables that were part of the actual analysis to improve parameter estimation (Enders, 2008).

To estimate the model parameters, we used a maximum likelihood estimator that is robust against non-normality and nonindependence of data by choosing the analysis option

TYPE = COMPLEX in Mplus 8 (Muthén & Muthén, 1998–2017). The estimator is labeled MLR (Maximum Likelihood Robust) in Mplus 8, and its χ^2 test statistic is asymptotically equivalent to the Yuan-Bentler T_2 test statistic (Muthén & Muthén, 1998–2017; Yuan & Bentler, 2000). For significance testing, we provided 95% confidence intervals for all parameters in the subsequent models. To deal with the occurrence of missing data, we used full information maximum likelihood estimation (FIML; see e.g., Enders, 2001) in all models, which is applied by default in Mplus 8. FIML requires the missing at random (MAR) assumption. Including auxiliary variables to the respective models via the saturated correlates model approach makes the MAR assumption more plausible (Graham, 2003).

Nested Data Structure

The data sets in the four studies had a clustered data structure, with students nested in classes (or schools for Study 4). Students from the same class tend to be more similar compared to students from other classes (Raudenbusch & Bryk, 2002), and not accounting for that hierarchical data structure could lead to an underestimation of standard errors (McNeish et al., 2017). Because classes were not the focus of the analyses, we treated them as a design nuisance (McNeish et al., 2017) and therefore relied on cluster-robust standard errors to correct for a possible underestimation. For vocational interests and personality traits the ICC's across all studies were between .00 and .15; for cognitive abilities they were between .04 and .26 (see Supplement A, Table A3).

Study 1: Profile Stability During Vocational Orientation (Ages 11 to 14)

Sample

Study 1 focussed on the vocational orientation phase. We used data from the first cohort of a large-scale longitudinal multi-cohort study (Tradition and Innovation study, TRAIN; Jonkmann et al., 2013)[6] that comprised students from lower, middle, and multi-track schools (i.e., a combination of lower and middle track students, situated in the German federal state of Saxony). In the original TRAIN study, students were annually followed over four time points from 5th to 8th grade (mean age: 11.24 to 14.24). In Study 1 of the present investigation, we included students who took part at the first and/or the fourth measurement time point of the initial study. Further information about the sample such as sample size and gender composition can be found in Table 2. Attrition analyses for the study can be found in Supplement A.

Instruments

Vocational Interests

Vocational interests were assessed with a 36-item measure that comprised items from the Revised General Interest Structure-Test (AIST-R; Allgemeiner Interessen Strukturtest; Bergmann & Eder, 2005), items from the German version of the Inventory of Children's Activities (ICA; Tracey & Ward, 1998; German version [ICA-D]: von Maurice, 2006) and newly constructed items (for information about the construction, validity, and reliability of the inventory see Gfrörer et al., 2021). In the present study, reliability coefficients for the respective scales were between .78 and .89 (see Table 1). The number of items per scale, the item instruction, and the response scale can be found in Supplement B.

Personality Traits

Personality traits were measured by the German version of the Big Five Inventory (BFI; see Lang et al., 2001). Similar to Rieger et al. (2017) and Trautwein et al. (2015), we decided to exclude the negatively worded items, because reliabilities were low if the negatively worded items were included in the scale and the recoded items demonstrated negative correlations with the overall scale—indicating that the younger students had difficulties in understanding them (see Table 1). In the present study, the reliability coefficients were between .67 and .81 (see

[6] The data set used in the current study was already used for several other publications. An overview is given here: https://uni-tuebingen.de/en/faculties/faculty-of-economics-and-social-sciences/subjects/department-of-social-sciences/education-sciences-and-psychology/research/aktuelle-studien/train/. None of these studies have investigated the influence of predictors on vocational interest profile stability.

rows without reverse coded items in Table 1). The number of items per scale, the item instruction and the response scale can be found in Supplement B.

Cognitive Abilities

Cognitive abilities were measured by a short version of the revised cognitive abilities test for 4th to 12th graders (KFT 4-12+ R; Heller & Perleth, 2000), a widely used German measure for cognitive abilities. Study 1 contained the subscales quantitative, figural, and verbal cognitive abilities. Sum scores were used that captured the amount of correctly solved items of the respective subscales. The number of items per scale, the item instruction, and the scale used to score answers can be found in Supplement B.

Results

Across the 3-year time period that was investigated in Study 1 (mean age 11.24 to 14.24), the mean Q-correlation of vocational interest profiles was r = .43 [.39, .48] with a statistically significant standard deviation of SD = 0.42 [0.39, 0.44] (see Table 3 and Table A4).

Gender positively predicted profile stability, even after controlling for personality traits and cognitive abilities (see Table 4). The difference between males and females on the Q-correlation was b = 0.12 [0.05, 0.18], implying that females had more stable profiles. In the model that included only personality traits as predictors, agreeableness had a significant effect on profile stability, b = 0.08 [0.03, 0.13], indicating that higher levels of agreeableness predicted higher levels of profile stability. This effect was also significant in the model that additionally included gender and cognitive abilities. Cognitive abilities had no statistically significant effect on profile stability.

Table 3

Profile Stability of Vocational Interests According to Q-Correlation

	Study 1 (Ages 11 to 14)			Study 2 (Ages 14 to 15)			Study 3 (Ages 17 to 23)			Study 4 (Ages 22 to 34)		
	M [95% CI]	*SE*	*p*	*M* [95% CI]	*SE*	*p*	*M* [95% CI]	*SE*	*p*	*M* [95% CI]	*SE*	*p*
Q-correlation	.43 [.39, .48]	0.02	<.001	.65 [.62, .67]	0.01	<.001	.64 [.61, .67]	0.02	<.001	.74 [.72, .75]	0.01	<.001

Note. Auxiliary variables were included to improve parameter estimation.

Table 4

Predictors of Q-correlation (Study 1: Ages 11 to 14)

	Model 1			Model 2			Model 3			Model 4		
	b [95% CI]	*SE*	*p*	*b* [95% CI]	*SE*	*p*	*b* [95% CI]	*SE*	*p*	*b* [95% CI]	*SE*	*p*
Intercept	0.38 [0.32, 0.44]	0.03	<.001	0.44 [0.40, 0.48]	0.02	<.001	0.43 [0.38, 0.48]	0.02	<.001	0.39 [0.33, 0.45]	0.03	<.001
Gender	0.12 [0.05, 0.18]	0.03	<.001							0.10 [0.03, 0.16]	0.03	.005
Openness				-0.05 [-0.12, 0.01]	0.03	.103				-0.05 [-0.12, 0.01]	0.03	.098
Consc.				-0.03 [-0.08, 0.03]	0.03	.410				-0.02 [-0.08, 0.04]	0.03	.505
Extraversion				0.03 [-0.03, 0.08]	0.03	.294				0.03 [-0.02, 0.09]	0.03	.234
Agreeabl.				0.08 [0.03, 0.13]	0.03	.001				0.06 [0.02, 0.11]	0.03	.010
Neuroticism				-0.05 [-0.09, -0.01]	0.02	.027				-0.04 [-0.08, 0.00]	0.02	.073
Cog.: Quant.							-0.03 [-0.09, 0.03]	0.03	.284	-0.03 [-0.09, 0.03]	0.03	.311
Cog.: Figural							0.04 [-0.01, 0.08]	0.02	.106	0.03 [-0.01, 0.07]	0.02	.182
Cog.: Verbal							0.03 [-0.02, 0.07]	0.02	.225	0.02 [-0.03, 0.06]	0.02	.486
R^2	.02	0.01	.061	.03	0.01	.032	.01	0.01	.280	.05	0.02	.002

Note. Gender was coded as 1 = female, 0 = male. *b* = standardized regression coefficient; *SE* = standard error; *p* = *p* value. Mplus indicated that some of the standard errors from models 1 to 4 may not be trustworthy, as the predictor gender was included in the model and treated as a continuous variable. However, the respective model parameters were checked, and they seem normal

Study 2: Profile Stability During Vocational Training (Ages 14 to 15)

Sample

Study 2 focussed on the life phase immediately before the start of vocational educational training. We used data from the second cohort of the TRAIN study, which also comprised students from lower, middle, and multi-track schools from two German states, Baden-Württemberg and Saxony, who were followed from 8th to 9th grade (mean age: 14.37 to 15.36). Students who took part at the first and/or the second measurement time point of the initial study were included in the analysis sample. Further information about the sample such as sample size and gender composition can be found in Table 2. Attrition analyses for the study can be found in Supplement A.

Instruments

Vocational Interests

Vocational interests were assessed with the same inventory as in Study 1. In the present study, reliability coefficients of vocational interests were between .77 and .91 (see Table 1). The number of items per scale, the item instruction and the response scale can be found in Supplement B.

Personality Traits

Personality traits were assessed with the same inventory as in Study 1. The same procedures that were described for personality traits in Study 1 were also used in Study 2. In the present study, reliability coefficients of personality traits were between .65 and .80 (see rows without reverse coded items in Table 1). The number of items per scale, the item instruction, and the response scale can be found in Supplement B.

Cognitive Abilities

Cognitive abilities were assessed with the same inventory as in Study 1. Sum scores were used that captured the amount of correctly solved items of the respective subscales. The number of items per scale, the item instruction, and the scale used to score answers can be found in Supplement B.

Results

During the 1-year time period before the start of vocational educational training (mean age: 14.37 to 15.36), the mean Q-correlation of vocational interest profiles was r = .65 [.62,

.67] with a statistically significant standard deviation of *SD* = 0.35 [0.32, 0.37] (see Table 3 and Table A4).

Gender positively predicted profile stability, even after controlling for personality traits and cognitive abilities (see Table 5). The difference between males and females on the Q-correlation was *b* = 0.23 [0.18, 0.27], implying that females had more stable profiles. In the model that included only the personality traits, only agreeableness had a significant effect on profile stability, *b* = 0.04 [0.01, 0.06], indicating that higher levels of agreeableness predicted higher levels of profile stability. However, this effect was not significant in the model that included all the predictors. In the model that included only cognitive abilities, there was a significant effect of verbal cognitive abilities, *b* = 0.06 [0.04, 0.09]. This effect was also significant in the models that included all the predictors. Higher levels of verbal cognitive abilities therefore predicted higher levels of profile stability.

Table 5

Predictors of Q-correlation (Study 2: Ages 14 to 15)

	Model 1			Model 2			Model 3			Model 4		
	b [95% CI]	*SE*	*p*	*b* [95% CI]	*SE*	*p*	*b* [95% CI]	*SE*	*p*	*b* [95% CI]	*SE*	*p*
Intercept	0.54 [0.50, 0.58]	0.02	<.001	0.65 [0.62, 0.67]	0.01	<.001	0.65 [0.62, 0.67]	0.01	<.001	0.55 [0.51, 0.59]	0.02	<.001
Gender	0.23 [0.18, 0.27]	0.02	<.001							0.21 [0.16, 0.26]	0.02	<.001
Openness				-0.01 [-0.04, 0.03]	0.02	.694				-0.02 [-0.05, 0.01]	0.02	.128
Consc.				0.00 [-0.03, 0.03]	0.02	.993				0.02 [-0.01, 0.05]	0.02	.197
Extraversion				0.01 [-0.02, 0.04]	0.02	.704				0.00 [-0.02, 0.03]	0.01	.769
Agreeabl.				0.04 [0.01, 0.06]	0.01	.012				0.01 [-0.02, 0.03]	0.01	.484
Neuroticism				0.00 [-0.03, 0.02]	0.01	.719				0.00 [-0.02, 0.02]	0.01	.902
Cog.: Quant.							-0.02 [-0.05, 0.01]	0.02	.170	-0.02 [-0.05, 0.01]	0.02	.201
Cog.: Figural							0.03 [-0.01, 0.06]	0.02	.107	0.02 [-0.01, 0.05]	0.02	.200
Cog.: Verbal							0.06 [0.04, 0.09]	0.01	<.001	0.05 [0.03, 0.08]	0.01	<.001
R^2	.11	0.01	<.001	.01	0.07	.148	.04	0.01	.005	.13	0.01	<.001

Note. Gender was coded as 1 = female, 0 = male. *b* = standardized regression coefficient; *SE* = standard error; *p* = *p* value.

Study 3: Profile Stability During College Major Orientation and Choice (Ages 17 to 23)

Sample

Study 3 focussed on the college major orientation and college major choice phase. We used data from a cohort of the study Transformation of the Secondary School System and Academic Careers (TOSCA; Köller et al., 2004), namely TOSCA-10 (Trautwein et al., 2010).[7] The initial study investigated the educational trajectories of middle and higher track students from the German state Baden-Württemberg after their compulsory education completion in 10th grade (T1) and 6 years later (T2). Students who took part at the first and/or the second measurement time point of the initial study were included in the final analysis sample. Further information such as sample size and gender composition can be found in Table 2. Attrition analyses for the study can be found in Supplement A.

Instruments

Vocational Interests

Vocational interests were assessed with the AIST-R inventory (Bergmann & Eder, 2005), which is the most widely applied vocational interest inventory in German-speaking countries. In the present study, reliability coefficients of vocational interests were between .83 and .90 (see Table 1). The number of items per scale, the item instruction and the response scale can be found in Supplement B.

Personality Traits

Personality traits were assessed with a short version of the German Big Five Inventory (BFI-K; see Rammstedt & John, 2005). In the present study, reliability coefficients for Openness, Extraversion, and Neuroticism were between .67 and .81 (see Table 1). Conscientiousness and Agreeableness possessed low reliability, ω = .36 and ω = .53, respectively. Due to higher reliability scores in general, the reverse coded items were included in the scales. The number of items per scale, the item instruction, and the response scale can be found in Supplement B.

[7] The data set used in the current study was already used for several other publications. An overview is given here: https://uni-tuebingen.de/en/faculties/faculty-of-economics-and-social-sciences/subjects/department-of-social-sciences/education-sciences-and-psychology/research/aktuelle-studien/tosca/. One of the studies has investigated vocational interest profile stability during adulthood and gender differences in profile stability (see Stoll, Rieger, et al., 2020), but none of the studies have investigated the influence of personality traits and cognitive abilities on vocational interest profile stability.

Cognitive Abilities

Cognitive abilities were assessed with the same instrument as in Studies 1 and 2. However, in Study 3, only subscales of figural and verbal cognitive abilities were available. Sum scores were used that captured the amount of correctly solved items of the respective subscales. The number of items per scale, the item instruction, and the scale used to score answers can be found in Supplement B.

Results

During the 6-year time period in Study 3 (mean age: 16.71 to 23.43), the mean Q-correlation of vocational interest profiles was r = .64 [.61, .67] with a statistically significant standard deviation of SD = 0.33 [0.30, 0.35] (see Table 3 and Table A4).

Gender positively predicted profile stability, even after controlling for personality traits and cognitive abilities (see Table 6). The difference between males and females on the Q-correlation was b = 0.10 [0.05, 0.15], implying that females had more stable profiles. In the model that included only personality traits, only extraversion had a significant effect on profile stability, b = 0.04 [0.01, 0.06], indicating that higher levels of extraversion predicted higher levels of profile stability. In the model that included only cognitive abilities, there was a statistically significant effect of figural cognitive abilities on profile stability, b = -0.03 [-0.06, -0.01], with higher levels of figural cognitive abilities predicting lower levels of profile stability. Both effects (extraversion and figural cognitive abilities) were also significant in the model including all predictors. In addition, verbal cognitive abilities also predicted profile stability in the model including all predictors, b = 0.03 [0.00, 0.07].

Table 6

Predictors of Q-correlation (Study 3: Ages 17 to 23)

	Model 1			Model 2			Model 3			Model 4		
	b [95% CI]	*SE*	*p*	*b* [95% CI]	*SE*	*p*	*b* [95% CI]	*SE*	*p*	*b* [95% CI]	*SE*	*p*
Intercept	0.59 [0.55, 0.63]	0.02	<.001	0.64 [0.61, 0.67]	0.02	<.001	0.64 [0.61, 0.67]	0.02	<.001	0.58 [0.54, 0.63]	0.02	<.001
Gender	0.10 [0.05, 0.15]	0.03	<.001							0.11 [0.05, 0.17]	0.03	<.001
Openness				-0.02 [-0.04, 0.00]	0.01	.091				-0.03 [-0.06, -0.01]	0.01	.013
Consc.				0.02 [-0.01, 0.04]	0.01	.242				0.01 [-0.02, 0.04]	0.01	.393
Extraversion				0.04 [0.01, 0.06]	0.01	.002				0.03 [0.00, 0.05]	0.01	.023
Agreeabl.				0.01 [-0.01, 0.03]	0.01	.321				0.00 [-0.02, 0.03]	0.01	.698
Neuroticism				0.02 [-0.01, 0.04]	0.01	.201				0.00 [-0.03, 0.03]	0.01	.886
Cog.: Figural							-0.03 [-0.06, -0.01]	0.02	.024	-0.03 [-0.06, -0.01]	0.02	.020
Cog.: Verbal							0.02 [-0.01, 0.05]	0.02	.250	0.03 [0.00, 0.07]	0.02	.049
R^2	.02	0.01	.047	.02	0.01	.075	.01	0.01	.252	.05	0.02	.003

Note. Gender was coded as 1 = female, 0 = male. *b* = standardized regression coefficient; *SE* = standard error; *p* = *p* value. Mplus indicated that some of the standard errors from models 1 to 4 may not be trustworthy, as the predictor gender was included in the model and treated as a continuous variable. However, the respective model parameters were checked, and they seem normal.

Study 4: Profile Stability During Workforce Entrance (Ages 22 to 34)

Sample

Study 4 focussed on the phase of entering the workforce. We used data from the TOSCA-2002 cohort (Köller et al., 2004; Trautwein et al., 2010), which initially comprised seven measurement time points. Higher track students from the German federal state of Baden-Württemberg were included in the sample. At the first measurement time point, students were in 12th grade. Participants who took part at the second and/or the seventh measurement time point of the initial study (mean age: 21.55 to 33.55) were included in the final analysis sample. Further information about sample characteristics such as sample size and gender composition can be found in Table 2. Attrition analyses for the study can be found in Supplement A.

Instruments

Vocational Interests

Vocational interests were assessed with the same inventory as in Study 3. In the present study, reliability coefficients of all measures were between .84 and .90 (see Table 1). The number of items per scale, the item instruction, and the response scale can be found in Supplement B.

Personality Traits

Personality traits were assessed with the NEO-Five-Factor Inventory (NEO-FFI; see Borkenau & Ostendorf, 1991). In the present study, the reliability coefficients were between .65 and .87. Due to higher reliability scores in general, the reverse coded items were included in the scales. The number of items per scale, the item instruction, and the response scale can be found in Supplement B.

Cognitive Abilities

Cognitive abilities were assessed with the same inventory and subscales as in Study 3. We used sum scores that reflect the amount of correctly solved items of the respective subscales. The number of items per scale, the item instruction, and the scale used to score answers can be found in Supplement B.

Results

During the 12-year time period that was investigated in Study 4 (mean age: 21.55 to 33.55), the mean Q-correlation of vocational interest profiles was r = .74 [.72, .75] with a statistically significant standard deviation of SD = 0.26 [0.24, 0.28] (see Table 3 and Table A4).

Gender positively predicted profile stability, even after controlling for personality traits and cognitive abilities (see Table 7). The difference between males and females on the Q-correlation was b = 0.04 [0.01, 0.07], implying that females had more stable profiles. In the model that included only personality traits, only extraversion had a significant effect on profile stability, b = 0.04 [0.02, 0.06], indicating that higher levels of extraversion predicted higher levels of profile stability. This effect was also significant in the model that included all predictors. The model that included only cognitive abilities estimated no statistically significant effects. However, in the model that included all predictors, there was a statistically significant effect of verbal cognitive abilities on profile stability, b = 0.02 [0.00, 0.03], with higher levels of verbal cognitive abilities predicting higher levels of profile stability.

Table 7

Predictors of Q-correlation (Study 4: Ages 22 to 34)

	Model 1			Model 2			Model 3			Model 4		
	b [95% CI]	*SE*	*p*	*b* [95% CI]	*SE*	*p*	*b* [95% CI]	*SE*	*p*	*b* [95% CI]	*SE*	*p*
Intercept	0.71 [0.69, 0.74]	0.01	<.001	0.74 [0.72, 0.75]	0.01	<.001	0.74 [0.72, 0.75]	0.01	<.001	0.71 [0.69, 0.74]	0.01	<.001
Gender	0.04 [0.01, 0.07]	0.02	.012							0.04 [0.00, 0.07]	0.02	.043
Openness				-0.01 [-0.02, 0.01]	0.01	.244				-0.01 [-0.03, 0.00]	0.01	.082
Consc.				0.00 [-0.01, 0.02]	0.01	.816				0.00 [-0.02, 0.02]	0.01	.965
Extraversion				0.04 [0.02, 0.06]	0.01	<.001				0.04 [0.02, 0.05]	0.01	<.001
Agreeabl.				0.00 [-0.01, 0.02]	0.01	.874				0.00 [-0.02, 0.02]	0.01	.910
Neuroticism				0.01 [-0.01, 0.02]	0.01	.472				0.00 [-0.02, 0.02]	0.01	.998
Cog.: Figural							-0.01 [-0.03, 0.01]	0.01	.188	-0.01 [-0.03, 0.01]	0.01	.185
Cog.: Verbal							0.01 [-0.01, 0.02]	0.01	.337	0.02 [0.00, 0.03]	0.01	.017
R^2	.01	0.00	.208	.02	0.01	.01	.00	.00	.473	.03	0.01	.001

Note. Gender was coded as 1 = female, 0 = male. *b* = standardized regression coefficient; *SE* = standard error; *p* = *p* value.

General Discussion

Summary of the Study Results

The current investigation provided insights from four different studies on vocational interest profile stability and its predictors during different phases of career guidance. Results of the four studies suggest that vocational interest profiles were relatively stable during the investigated phases, but profile stability significantly varied from person to person, implying that some participants had more stable profiles than others. Gender was consistently related to differences in profile stability, with girls and women having more stable profiles than boys and men. In addition, in some studies, differences in extraversion and verbal cognitive abilities were related to differences in profile stability. However, the respective relationships were rather small.

Women Have More Stable Vocational Interest Profiles Than Men

The current investigation provides robust evidence that girls and women have more stable vocational interest profiles than boys and men. This is in line with previous studies that investigated the stability of vocational interest profiles (Xu & Tracey, 2016). Gender differences in profile stability were the largest during the phase of vocational training (ages 14 to 15) with $b = 0.21$ and the smallest during the phase of entering the workforce (ages 22 to 34) with $b = 0.04$. Findings of the current study suggest that gender differences are present not only in mean levels of single interest dimensions (Su et al., 2009), but also in profile stability indicators.

We came up with various explanations for why gender differences in profile stability might occur. For example, gender differences in individual career guidance efforts, such as career exploration, might influence profile stability. Vignoli et al. (2005) reported higher mean levels for girls in career exploration diversity (e.g., reading brochures, asking family members, or visiting employment offices) and career exploration frequency. Comprehensive career exploration might result in people having a better understanding about their occupational likes and dislikes, abilities, and goals (Hirschi, 2010; Holland, 1997). This improved understanding about oneself could result in more stable interest profiles for girls. However, gender differences in profile stability could also reflect differences in maturation processes. It is suggested that certain biological processes (e.g., puberty) that affect motivational brain structures of children begin to unfold earlier for girls than boys (Crone & Dahl, 2012). This could impact the within-person consistency of motivational variables such as vocational interests and would explain why gender differences were larger during middle adolescence than young adulthood.

Furthermore, gender differences in profile stability could also arise due to external influences. For example, Gottfredson (1981) proposed that adolescents evaluate the fit of their gender to the ones of occupations. Girls are assumed to prefer careers that are stereotypically seen as female, whereas boys are assumed to prefer careers that are stereotypically seen as male. At the same time, Gottfredson (1981) suggests that girls are provided with less occupational alternatives than boys. The narrowed set of options and choices could lead to more consistent occupational preferences for girls. Although the present investigation provides robust evidence for gender differences in vocational interest profile stability, possible causes of these differences can only be speculated about. Whether or not gender differences in profile stability arise due to differences in career exploration, maturation, or external factors, such as the availability of occupational alternatives, must be answered empirically in further research.

Personality Traits and Cognitive Abilities—Distant Predictors of Profile Stability?

In comparison to the influence of gender, the effects of personality traits and cognitive abilities were small and less consistent. We found positive significant effects for verbal cognitive abilities in three of the four studies and positive significant effects for extraversion in two of the four studies (see Table 8), suggesting that higher levels of verbal cognitive abilities and extraversion predicted higher levels of profile stability (i.e., $.02 < b < .05$; controlling for gender and the remaining personality traits and cognitive abilities). The current investigation therefore provides first insights on the association between the respective constructs and vocational interest profile stability.

Table 8
Overview of the Direction of the Significant Effects on Profile Stability

	r_1	r_2	r_3	r_4
Gender	+	+	+	+
Openness	.	.	−	.
Consc.	.	.	.	.
Extraversion	.	.	+	+
Agreeabl.	+	.	.	.
Neuroticism	.	.	.	.
Cog. Quant.	.	.	n/a	n/a
Cog. Fig.	.	.	−	.
Cog. Verb.	.	+	+	+

Note. Relationships were extracted from the overall model that included all predictors. The significant effect sizes of personality traits and cognitive abilities were all below .06. r_1 = profile stability Study 1; r_2 = profile stability Study 2; r_3 = profile stability Study 3; r_4 = profile stability Study 4.

Similarly to the effects of gender on profile stability, we can only speculate why extraversion and cognitive abilities are related to vocational interest profile stability. Cognitive abilities influence the accumulation of knowledge and interest in general learning (Ackerman, 1996; Schmidt, 2014). People with better cognitive abilities might therefore have better knowledge about themselves and about the activities they do and do not like. As these people are probably more consistent about the areas they are interested in (Hirschi, 2010; Holland, 1997), they could possess more stable vocational interest profiles. Extraversion could influence profile stability due to a similar process. For example, people who are more extraverted are more likely to receive support from others and are also more likely to search and demand for support from others (Swickert et al., 2002). External support from families, peers, or counseling professionals could provide an important source of information during career choice processes. This could increase the career choice readiness of adolescents and consequently the certainty of their occupational likes and dislikes. However, to answer if these processes affect profile stability, further empirical evidence is needed.

The small effect sizes of personality traits and cognitive abilities could also indicate that other variables might be more relevant in predicting differences in vocational interest profile stability. The information provided through scales of personality traits and general cognitive abilities might be too broad to capture the specific processes that influence profile stability. For example, interest theories argue that the solidification of vocational interests progresses based on experiences (Holland, 1997; Su et al., 2019), which allow people to get a better understanding of their likes and dislikes. This could imply that indicators that directly capture experiences, such as the pursuit and number of different activities, might be more closely related to profile stability. It could be debated that the personality trait openness to experience should capture that property, as people with higher degrees of openness are characterized as being curious and seeking a variety of experiences (Connelly et al., 2014). However, in the current investigation, weak empirical support was provided for the link between profile stability and openness to experience.

Practical Implications for Career Guidance

During time periods where results from interest inventories are used for prediction (e.g., prospective educational or occupational environments), vocational interest profiles should be relatively stable (Low et al., 2005; Strong, 1931). Differences in profile stability can therefore lead to distinct practical implications and distinct implementations of interest inventories in interest-based career guidance. The present investigation indicates that vocational interest

profiles were relatively stable. Even during longer time intervals (e.g., 6 years covering the college major orientation and choice phase or 12 years covering the workforce entrance phase), which are usually associated with lower construct stability (Fraley & Roberts, 2005), Q-correlations were above $r = .60$. As suggested by the variation across the four studies ($.43 < r < .74$), the magnitude of profile stability differed between life phases. In Study 1, which captured a life phase where vocational orientation is about to unfold (ages 11 to 14), profile stability was moderate (i.e., $r = .43$), whereas in Studies 2 to 4, which captured life phases including imminent or actual career decisions such as vocational training, study major choices or occupational choices, profile stability was high (i.e., $.64 < r < .74$). These results support the general reliability of vocational interests in career guidance and are similar to previous studies that investigated vocational interest profile stability (Rottinghaus et al., 2007; Stoll, Rieger, et al., 2020; Xu & Tracey, 2016; Zytowski, 1976).

In line with assumptions and findings about the stability of vocational interests in earlier life phases (Gfrörer et al., 2021; Holland, 1997; Low et al., 2005; Päßler & Hell, 2020; Su et al., 2019; Tracey, 2002), the findings of the present investigation suggest that vocational interest profiles are less stable during the phase of vocational orientation (ages 11 to 14). Although a Q-correlation of .43 can be considered as moderately stable, compared to later life phases, the stability is visibly lower (e.g., from age 20 to 32, $.67 < r < .77$; Zytowski, 1976). One way to deal with lower profile stability in practice could be to adjust the application of interest inventories. For example, vocational counselors could use the results of interest inventories to make a list of suitable occupations, which could be explored by the counselee through internships, internet research, or other arrangements. Over time, as interests become more stable, counselors could use both experiences from these activities and results from interest inventories to make more accurate predictions of suitable occupations or study majors in prospective life phases.

According to the findings of the current investigation, some personality traits and cognitive abilities were associated with higher levels of profile stability. However, the respective associations were rather small, with all standardized regression coefficients being below $b = 0.06$ (i.e., in the overall models including all the predictors). The small effect sizes raise the question if these associations are also practically relevant for the field of career guidance. Additional insights from personality inventories and ability tests might be very informative during counseling sessions (Ackerman & Beier, 2003), as they help to further restrict the possible pool of suitable occupations. However, according to the findings of the current investigation, their effect on profile stability seems rather small. To ensure the

robustness of this finding and its application in the practice of career guidance, more empirical studies are needed that investigate the relationship between personality traits, cognitive abilities, and profile stability.

For gender, the practical implications might be different. During the study that captured the life phase right before entering vocational training (ages 14 to 15), girls and boys differed in their profile stability by .21 points on the Q-correlation. This was the largest gender difference in stability that we found in the current investigation. The finding is different to previous meta-analytic evidence (Low et al., 2005), which reported no significant moderator effects of gender on stability. In addition, the effects of gender were consistent across all studies, suggesting that gender differences were quite robust. This could indicate that the application of interest inventories should be adapted with regard to gender. For example, in earlier life phases, when gender differences are larger (see Study 2), results of interest inventories from boys could be used for exploration rather than for prediction.

Limitations and Future Outlook

The major strength of the current investigation is that it shares new information on vocational interest profile stability and its predictors during various life phases. However, there are some limitations that should be mentioned. First, although vocational interests, personality traits, and cognitive abilities were operationalized based on the same theoretical framework in all four studies, the studies differed in sample composition, time interval, and measurement instrument. These variations made it difficult to investigate if changes in profile stability coefficients were related to changes in age. However, despite these limitations, each study provided comprehensive insights about profile stability and its predictors within each of the investigated life phases. Future studies could specifically focus on the investigation of age-related changes in profile stability.

Furthermore, the four studies differed in the learning environments that were provided for the respective participants. For example, the second study primarily included students from lower and middle school tracks in Germany. For these students, vocational orientation is explicitly emphasized in their curricula, which is not necessarily the case for students from higher school tracks, such as in Study 4. In addition, the relevance of the respective life phases for career guidance might also partly depend on the educational system. This raises the question of whether the investigation of interest stability during earlier career phases such as vocational training would have the same relevance for career guidance in other countries where vocational educational training is less common, such as in the United States. Perhaps not all of the life

phases examined in the current investigation are of the same importance for career guidance in other countries. Future studies should therefore replicate the findings of the current investigation based on a variety of samples from different countries and educational contexts.

Finally, the aim of the current investigation was to provide first empirical insights on the relationship between vocational interest profile stability, personality traits, cognitive abilities, and gender. In some of the studies, positive but small relationships were found between vocational interest profile stability and some of the personality traits and cognitive abilities. To ensure the robustness of these findings and to confirm consistent patterns between vocational interest profile stability, personality traits, and cognitive abilities, further studies are needed that investigate their relationships across different samples and life phases. Besides the replication of our findings, future studies should examine the underlying processes that are responsible for the differences in profile stability (e.g., differences between boys and girls). In the current discussion we suggested several mechanisms that could have led to the differences in profile stability.

References

Ackerman, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence*, *22*(2), 227–257. https://doi.org/10.1016/S0160-2896(96)90016-1

Ackerman, P. L., & Beier, M. E. (2003). Intelligence, Personality, and Interests in the Career Choice Process. *Journal of Career Assessment*, *11*(2), 205–218. https://doi.org/10.1177/1069072703011002006

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, Personality, and Interests: Evidence for Overlapping Traits. *Psychological Bulletin*, *121*(2), 219–245. https://doi.org/10.1037/0033-2909.121.2.219

Allen, J., & Robbins, S. B. (2008). Prediction of college major persistence based on vocational interests, academic preparation, and first-year academic performance. *Research in Higher Education*, *49*, 62–79. https://doi.org/10.1007/s11162-007-9064-5

American College Testing Program. (2009). *The ACT Interest Inventory Technical Manual*. https://www.act.org/content/dam/act/unsecured/documents/ACTInterestInventoryTechnicalManual.pdf

Armstrong, P. I., Day, S. X., McVay, J. P., & Rounds, J. (2008). Holland's RIASEC model as an integrative framework for individual differences. *Journal of Counseling Psychology*, *55*(1), 1–18. https://doi.org/10.1037/0022-0167.55.1.1

Bergmann, C., & Eder, F. (2005). *Allgemeiner Interessen-Struktur-Test mit Umwelt-Struktur-Test (UST-R)—Revision [General Interest Structure Test and Environmental Structure Test—Revision]*. Beltz.

BIBB (Ed.). (2018). *Datenreport zum Berufsbildungsbericht 2018. Informationen und Analysen zur Entwicklung der beruflichen Bildung* (1st ed.). Bundesinstitut für Berufsbildung.

Brown, D. (Ed.). (2002). *Career choice and development* (4th ed.). John Wiley & Sons, Inc.

Burt, C. L. (1937). Correlations between persons. *British Journal of Psychology*, *28*(1), 59–96.

Cheng, H., & Furnham, A. (2012). Childhood cognitive ability, education, and personality traits predict attainment in adult occupational prestige over 17 years. *Journal of Vocational Behavior*, *81*(2), 218–226. https://doi.org/10.1016/j.jvb.2012.07.005

Connelly, B. S., Ones, D. S., & Chernyshenko, O. S. (2014). Introducing the Special Section on Openness to Experience: Review of Openness Taxonomies, Measurement, and Nomological Net. *Journal of Personality Assessment*, *96*(1), 1–16. https://doi.org/10.1080/00223891.2013.830620

Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social–affective engagement and goal flexibility. *Nature Reviews Neuroscience*, *13*(9), 636–650. https://doi.org/10.1038/nrn3313

Dawis, R. V. (1992). The individual differences tradition in counseling psychology. *Journal of Counseling Psychology*, *39*(1), 7–19. https://doi.org/10.1037/0022-0167.39.1.7

Dunkleberger, C. J., & Tyler, L. E. (1961). Interest stability and personality traits. *Journal of Counseling Psychology*, *8*(1), 70–74. https://doi.org/10.1037/h0046021

Enders, C. K. (2001). The impact of nonnormality on full information maximum-likelihood estimation for structural equation models with missing data. *Psychological Methods*, *6*(4), 352–370. https://doi.org/10.1037//I082-989X.6.4.352

Enders, C. K. (2008). A note on the use of missing auxiliary variables in full information maximum likelihood-based structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *15*(3), 434–448. https://doi.org/10.1080/10705510802154307

Etzel, J. M., & Nagy, G. (2021). Stability and change in vocational interest profiles and interest congruence over the course of vocational education and training. *European Journal of Personality*, *35*(4), 534–556. https://doi.org/10.1177/08902070211014015

Eurostat. (2021a, January 5). *Share of students in vocational education programs in 2018.* Eurostat Statistics Explained. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Share_of_students_in_vocational_education_programmes,_2018_(%25)_ET2020.png

Eurostat. (2021b, February 5). *Tertiary education statistics in 2018.* Eurostat Statistics Explained. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Tertiary_education_statistics

Fraley, R. C., & Roberts, B. W. (2005). Patterns of Continuity: A Dynamic Model for Conceptualizing the Stability of Individual Differences in Psychological Constructs

Across the Life Course. *Psychological Review*, *112*(1), 60–74. https://doi.org/10.1037/0033-295X.112.1.60

Gati, I., Levin, N., & Landman-Tal, S. (2019). Decision-Making Models and Career Guidance. In J. A. Athanasou & H. N. Perera (Eds.), *International Handbook of Career Guidance* (2nd ed., pp. 115–145). Springer, Cham. https://doi.org/10.1007/978-3-030-25153-6_6

Gfrörer, T., Stoll, G., Rieger, S., Trautwein, U., & Nagengast, B. (2021). The Development of Vocational Interests in Early Adolescence: Stability, Change, and State-Trait Components. *European Journal of Personality*, 089020702110356. https://doi.org/10.1177/08902070211035630

Gottfredson, L. S. (1981). Circumscription and compromise: A developmental theory of occupational aspirations. *Journal of Counseling Psychology*, *28*(6), 545–579. https://doi.org/10.1037/0022-0167.28.6.545

Graham, J. W. (2003). Adding missing-data-relevant variables to FIML-based structural equation models. *Structural Equation Modeling*, *10*(1), 80–100. https://doi.org/10.1207/S15328007SEM1001_4

Gysbers, N. C., & Lapan, R. T. (2001). The Implementation and Evaluation of Comprehensive School Guidance Programs in the United States: Progress and Prospects. *International Journal for Educational and Vocational Guidance*, *1*(3), 197–208.

Hanna, A., & Rounds, J. (2020). How accurate are interest inventories? A quantitative review of career choice hit rates. *Psychological Bulletin*, *146*(9), 765–796. https://doi.org/10.1037/bul0000269

Hansen, J. C. (2019). Interest inventories. In G. Goldstein, D. N. Allen, & J. DeLuca (Eds.), *Handbook of Psychological Assessment* (4th ed., pp. 169–190). Elsevier. https://doi.org/10.1016/B978-0-12-802203-0.00006-7

Heller, K. A., & Perleth, C. (2000). *KFT 4-12+ R kognitiver Fähigkeitstest für 4. Bis 12. Klassen, Revision.* (1st ed.). Beltz Test.

Hirschi, A. (2010). Individual predictors of adolescents' vocational interest stabilities. *International Journal for Educational and Vocational Guidance*, *10*(1), 5–19. https://doi.org/10.1007/s10775-009-9171-2

Hoff, K. A., Perlus, J. G., & Rounds, J. (2019). Vocational Interests: Revisiting Assumptions About Their Development and What They Predict. In J. A. Athanasou & H. N. Perera (Eds.), *International Handbook of Career Guidance* (2nd ed., pp. 673–692). Springer, Cham. https://doi.org/10.1007/978-3-030-25153-6_31

Hoff, K. A., Song, Q. C., Einarsdóttir, S., Briley, D. A., & Rounds, J. (2020). Developmental structure of personality and interests: A four-wave, 8-year longitudinal study. *Journal of Personality and Social Psychology*, *118*(5), 1044–1064. https://doi.org/10.1037/pspp0000228

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Psychological Assessment Resources.

Hoyt, D. P., Smith, J. L., & Levy, S. (1957). A further study in the prediction of interest stability. *Journal of Counseling Psychology*, *4*(3), 228–233. https://doi.org/10.1037/h0046264

Jackson, J. J., Hill, P. L., Payne, B. R., Parisi, J. M., & Stine-Morrow, E. A. L. (2020). Linking openness to cognitive ability in older adulthood: The role of activity diversity. *Aging & Mental Health*, *24*(7), 1079–1087. https://doi.org/10.1080/13607863.2019.1655705

Jonkmann, K., Rose, N., & Trautwein, U. (2013). *Tradition und Innovation: Entwicklungsverläufe an Haupt- und Realschulen in Baden-Württemberg und Mittelschulen in Sachsen – Abschlussbericht für die Länder Baden-Württemberg und Sachsen [unpublished project report].* Universität Tübingen.

Kandler, C., Zimmermann, J., & McAdams, D. P. (2014). Core and surface characteristics for the description and theory of personality differences and development: Core and surface personality characteristics. *European Journal of Personality*, *28*(3), 231–243. https://doi.org/10.1002/per.1952

Köller, O., Watermann, R., Trautwein, U., & Lüdtke, O. (Eds.). (2004). *Wege zur Hochschulreife in Baden-Württemberg. TOSCA - Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien.* VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-80906-3

Kracke, B. (1997). Parental behaviors and adolescents' career exploration. *The Career Development Quarterly*, *45*(4), 341–350. https://doi.org/10.1002/j.2161-0045.1997.tb00538.x

Lang, F. R., Lüdtke, O., & Asendorpf, J. B. (2001). Testgüte und psychometrische Äquivalenz der deutschen Version des Big Five Inventory (BFI) bei jungen, mittelalten und alten Erwachsenen. *Diagnostica*, *47*(3), 111–121. https://doi.org/10.1026//0012-1924.47.3.111

Larson, L. M., Rottinghaus, P. J., & Borgen, F. H. (2002). Meta-analyses of Big Six Interests and Big Five Personality Factors. *Journal of Vocational Behavior*, *61*(2), 217–239. https://doi.org/10.1006/jvbe.2001.1854

Le, H., Robbins, S. B., & Westrick, P. (2014). Predicting student enrollment and persistence in college STEM fields using an expanded P-E fit framework: A large-scale multilevel study. *Journal of Applied Psychology*, *99*(5), 915–947. https://doi.org/10.1037/a0035998

Livingston, R. B., Jennings, E., Reynolds, C. R., & Gray, R. M. (2003). Multivariate analyses of the profile stability of intelligence tests: High for IQs, low to very low for subtest analyses. *Archives of Clinical Neuropsychology*, *18*(5), 487–507. https://doi.org/10.1093/arclin/18.5.487

Low, K. S. D., Yoon, M., Roberts, B. W., & Rounds, J. (2005). The stability of vocational interests from early adolescence to middle adulthood: A quantitative review of longitudinal studies. *Psychological Bulletin*, *131*(5), 713–737. https://doi.org/10.1037/0033-2909.131.5.713

Lowman, R. L. (1991). *The Clinical Practice of Career Assessment: Abilities, Interests, and Personality*. American Psychological Association.

Maaz, K., Trautwein, U., Ldtke, O., & Baumert, J. (2008). Educational Transitions and Differential Learning Environments: How Explicit Between-School Tracking Contributes to Social Inequality in Educational Outcomes. *Child Development Perspectives*, *2*(2), 99–106. https://doi.org/10.1111/j.1750-8606.2008.00048.x

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, *23*(3), 412–433. https://doi.org/10.1037/met0000144

McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods*, *22*(1), 114–140. https://doi.org/10.1037/met0000078

Mount, M. K., Barrick, M. R., Scullen, S. M., & Rounds, J. (2005). Higher Order Dimensions of the Big Five Personality Traits and the Big Six Vocational Interest Types. *Personnel Psychology*, *58*(2), 447–478. https://doi.org/10.1111/j.1744-6570.2005.00468.x

Muthén, B., & Muthén, L. K. (1998-2017). *Mplus User's Guide* (8th ed.). Muthén & Muthén.

National Center for Education Statistics. (2019). *Enrollment in elementary, secondary, and degree-granting postsecondary institutions.* Digest of Education Statistics. https://nces.ed.gov/programs/digest/d20/tables/dt20_105.30.asp

Neumann, G., Olitsky, N., & Robbins, S. (2009). Job congruence, academic achievement, and earnings. *Labour Economics*, *16*(5), 503–509. https://doi.org/10.1016/j.labeco.2009.03.004

Noack, P., Kracke, B., Gniewosz, B., & Dietrich, J. (2010). Parental and school effects on students' occupational exploration: A longitudinal and multilevel analysis. *Journal of Vocational Behavior*, *77*(1), 50–57. https://doi.org/10.1016/j.jvb.2010.02.006

Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2012). Vocational interests and performance: A quantitative summary of over 60 years of research. *Perspectives on Psychological Science*, *7*(4), 384–403. https://doi.org/10.1177/1745691612449021

Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2017). Interest congruence and performance: Revisiting recent meta-analytic findings. *Journal of Vocational Behavior*, *98*, 138–151. https://doi.org/10.1016/j.jvb.2016.11.002

O*NET. (2021, February 14). *O*NET Resource Center—Interest table.* https://www.onetcenter.org/dictionary/22.2/excel/interests.html

Päßler, K., & Hell, B. (2012). Do interests and cognitive abilities help explain college major choice equally well for women and men? *Journal of Career Assessment*, *20*(4), 479–496. https://doi.org/10.1177/1069072712450009

Päßler, K., & Hell, B. (2020). Stability and change in vocational interests from late childhood to early adolescence. *Journal of Vocational Behavior*, *121*, 103462. https://doi.org/10.1016/j.jvb.2020.103462

Perera, H. N., & McIlveen, P. (2018). Vocational interest profiles: Profile replicability and relations with the STEM major choice and the Big-Five. *Journal of Vocational Behavior*, *106*, 84–100. https://doi.org/10.1016/j.jvb.2017.11.012

Prediger, D. J. (1998). Is interest profile level relevant to career counseling? *Journal of Counseling Psychology*, *45*(2), 204–211. https://doi.org/10.1037/0022-0167.45.2.204

Rammstedt, B., & John, O. P. (2005). Kurzversion des Big Five Inventory (BFI-K): *Diagnostica*, *51*(4), 195–206. https://doi.org/10.1026/0012-1924.51.4.195

Raudenbusch, & Bryk. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). SAGE Publications.

Renninger, K. A., & Hidi, S. (2011). Revisiting the Conceptualization, Measurement, and Generation of Interest. *Educational Psychologist*, *46*(3), 168–184. https://doi.org/10.1080/00461520.2011.587723

Rieger, S., Göllner, R., Spengler, M., Trautwein, U., Nagengast, B., & Roberts, B. W. (2017). Social Cognitive Constructs Are Just as Stable as the Big Five Between Grades 5 and 8. *AERA Open*, *3*(3), 1–9. https://doi.org/10.1177/2332858417717691

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The Power of Personality: The Comparative Validity of Personality Traits, Socioeconomic Status, and Cognitive Ability for Predicting Important Life Outcomes. *Perspectives on Psychological Science*, *2*(4), 313–345. https://doi.org/10.1111/j.1745-6916.2007.00047.x

Roberts, B. W., & Wood, D. (2006). Personality Development in the Context of the Neo-Socioanalytic Model of Personality. In D. K. Mroczek & T. D. Little (Eds.), *Handbook of Personality Development* (pp. 11–39). Lawrence Erlbaum Associates Publishers. https://doi.org/10.4324/9781315805610.ch2

Rottinghaus, P. J., Coon, K. L., Gaffey, A. R., & Zytowski, D. G. (2007). Thirty-Year Stability and Predictive Validity of Vocational Interests. *Journal of Career Assessment*, *15*(1), 5–22. https://doi.org/10.1177/1069072706294517

Rounds, J., Dawis, R., & Lofquist, L. H. (1987). Measurement of person-environment fit and prediction of satisfaction in the theory of work adjustment. *Journal of Vocational Behavior*, *31*(3), 297–318. https://doi.org/10.1016/0001-8791(87)90045-5

Rounds, J., & Su, R. (2014). The nature and power of interests. *Current Directions in Psychological Science*, *23*(2), 98–103. https://doi.org/10.1177/0963721414522812

Schmidt, F. L. (2014). A General Theoretical Integrative Model of Individual Differences in Interests, Abilities, Personality Traits, and Academic and Occupational Achievement: A Commentary on Four Recent Articles. *Perspectives on Psychological Science*, *9*(2), 211–218. https://doi.org/10.1177/1745691613518074

Schomburg, A. M., & Tokar, D. M. (2003). The moderating effect of private self-consciousness on the stability of vocational interests. *Journal of Vocational Behavior*, *63*(3), 368–378. https://doi.org/10.1016/S0001-8791(03)00027-7

Spengler, M., Damian, R. I., & Roberts, B. W. (2018). How you behave in school predicts life success above and beyond family background, broad traits, and cognitive ability. *Journal of Personality and Social Psychology*, *114*(4), 620–636. https://doi.org/10.1037/pspp0000185

Stoll, G., Einarsdóttir, S., Chelsea Song, Q., Ondish, P., Sun, J.-T., & Rounds, J. (2020). The Roles of Personality Traits and Vocational Interests in Explaining What People Want out of Life. *Journal of Research in Personality*, *86*, 103939. https://doi.org/10.1016/j.jrp.2020.103939

Stoll, G., Rieger, S., Lüdtke, O., Nagengast, B., Trautwein, U., & Roberts, B. W. (2017). Vocational interests assessed at the end of high school predict life outcomes assessed 10 years later over and above IQ and Big Five personality traits. *Journal of Personality and Social Psychology*, *113*(1), 167–184. https://doi.org/10.1037/pspp0000117

Stoll, G., Rieger, S., Nagengast, B., Trautwein, U., & Rounds, J. (2020). Stability and change in vocational interests after graduation from high school: A six-wave longitudinal study. *Journal of Personality and Social Psychology*, *120*(4), 1091–1116. https://doi.org/10.1037/pspp0000359

Strong, E. K. (1931). *Change of Interests with Age*. Stanford University Press.

Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological Bulletin*, *135*(6), 859–884. https://doi.org/10.1037/a0017364

Su, R., Stoll, G., & Rounds, J. (2019). The nature of interests: Toward a unifying theory of trait-situation interest dynamics. In C. D. Nye & J. Rounds (Eds.), *Vocational interests: Rethinking their role in understanding workplace behavior and practice* (pp. 11–38). Taylor & Francis/Routledge. https://doi.org/10.4324/9781315678924-2

Swanson, J. L. (1999). Stability and change in vocational interests. In M. L. Savickas & A. R. Spokane (Eds.), *Vocational interests: Meaning, measurement, and counseling use* (pp. 135–158). Davies-Black Publishing.

Swanson, J. L., & Hansen, J.-I. C. (1988). Stability of vocational interests over 4-Year, 8-Year, and 12-Year intervals. *Journal of Vocational Behavior*, *33*(2), 185–202. https://doi.org/10.1016/0001-8791(88)90055-3

Swickert, R. J., Rosentreter, C. J., Hittner, J. B., & Mushrush, J. E. (2002). Extraversion, social support processes, and stress. *Personality and Individual Differences*, *32*(5), 877–891. https://doi.org/10.1016/S0191-8869(01)00093-9

Tracey, T. J. G. (2002). Development of interests and competency beliefs: A 1-year longitudinal study of fifth- to eighth-grade students using the ICA-R and structural equation modeling. *Journal of Counseling Psychology*, *49*(2), 148–163. https://doi.org/10.1037//0022-0167.49.2.148

Tracey, T. J. G. (2012). Problems with single interest scales: Implications of the general factor. *Journal of Vocational Behavior*, *81*(3), 378–384. https://doi.org/10.1016/j.jvb.2012.10.001

Tracey, T. J. G., & Ward, C. C. (1998). The structure of children's interests and competence perceptions. *Journal of Counseling Psychology*, *45*(3), 290–303. https://doi.org/10.1037/0022-0167.45.3.290

Trautwein, U., Lüdtke, O., Nagy, N., Lenski, A., Niggli, A., & Schnyder, I. (2015). Using individual interest and conscientiousness to predict academic effort: Additive, synergistic, or compensatory effects? *Journal of Personality and Social Psychology*, *109*(1), 142–162. https://doi.org/10.1037/pspp0000034

Trautwein, U., Neumann, M., Nagy, G., Lüdtke, O., & Maaz, K. (2010). Institutionelle Reform und individuelle Entwicklung: Hintergrund und Fragestellungen der Studie TOSCA-Repeat. In U. Trautwein, M. Neumann, G. Nagy, O. Lüdtke, & K. Maaz (Eds.), *Schulleistungen von Abiturienten: Die neu geordnete gymnasiale Oberstufe auf dem Prüfstand* (pp. 15–36). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92037-5_1

Tsabari, O., Tziner, A., & Meir, E. I. (2005). Updated Meta-Analysis on the Relationship Between Congruence and Satisfaction. *Journal of Career Assessment*, *13*(2), 216–232. https://doi.org/10.1177/1069072704273165

Vignoli, E., Croity-Belz, S., Chapeland, V., de Fillipis, A., & Garcia, M. (2005). Career exploration in adolescents: The role of anxiety, attachment, and parenting style. *Journal of Vocational Behavior*, *67*(2), 153–168. https://doi.org/10.1016/j.jvb.2004.08.006

von Maurice, J. (2006). *ICA-D. Deutschsprachige Version des Inventory of Children's Activities – Revised (ICA-R, Tracey & Ward, 1998)[unpublished instrument]*. Otto-Friedrich-Universität Bamberg.

Wille, E., Stoll, G., Gfrörer, T., Cambria, J., Nagengast, B., & Trautwein, U. (2020). It takes two: Expectancy-value constructs and vocational interests jointly predict STEM major choices. *Contemporary Educational Psychology*, *61*, 101858. https://doi.org/10.1016/j.cedpsych.2020.101858

Xu, H., & Li, H. (2020). Operationalize Interest Congruence: A Comparative Examination of Four Approaches. *Journal of Career Assessment*, *28*(4), 571–588. https://doi.org/10.1177/1069072720909825

Xu, H., & Tracey, T. J. G. (2016). Stability and change in interests: A longitudinal examination of grades 7 through college. *Journal of Vocational Behavior*, *93*, 129–138. https://doi.org/10.1016/j.jvb.2016.02.002

Yuan, K.-H., & Bentler, P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology*, *30*(1), 165–200. https://doi.org/10.1111/0081-1750.00078

Zytowski, D. G. (1976). Long-term profile stability of the Kuder Occupational Interest Survey. *Educational and Psychological Measurement*, *36*(3), 689–692. https://doi.org/10.1177/001316447603600315

Supplement Material

Supplement A: Attrition Analysis & Further Sample Characteristics

We conducted two types of attrition analyses at each time point. First, we compared participants who dropped out of the study at a respective time point with the participants who remained in the study. Second, we compared participants who joined the study at a respective time point with the participants who were already participating before that time point. The attrition analysis indicated minor to negligible differences on vocational interests, personality traits, cognitive abilities, and gender (i.e., $d \leq 0.30$).

Table A1
Differences (Cohen's d) Between Students Who Dropped Out and Continued (T1 » T2)

	Study 1 (Dropout: *N* = 686)	Study 2 (Dropout: *N* = 265)	Study 3 (Dropout: *N* = 2107)	Study 4 (Dropout: *N* = 1054)
Vocational Interests				
Realistic	0.07	-0.02	0.19	-0.06
Investigative	0.03	-0.11	-0.14	0.01
Artistic	**0.15**	-0.08	-0.11	**0.08**
Social	0.09	-0.02	-0.14	0.00
Enterprising	**0.13**	0.05	-0.03	0.07
Conventional	**0.13**	**-0.17**	-0.13	-0.04
Big Five				
Openness	**0.13**	-0.10	-0.01	**0.08**
Conscientiousness	0.07	**-0.26**	-0.12	**-0.18**
Extraversion	0.09	0.00	0.12	0.05
Agreeableness	0.09	-0.12	-0.01	**-0.07**
Neuroticism	**0.22**	**0.19**	-0.07	-0.02
Cognitive Abilities				
Quantitative	**-0.11**	**-0.29**	-	-
Figural	**-0.12**	**-0.32**	**-0.34**	**-0.17**
Verbal	**-0.13**	**-0.32**	**-0.35**	-0.07
Gender (female = 1)	OR = 0.95	**OR = 0.70**	**OR = 0.71**	OR = 0.97

Note. Bold parameters are significantly different from zero at $p < .05$, two-tailed. Positive values imply higher values for the participants who dropped out of the study. OR = Odds Ratio.

Table A2

Differences (Cohen's d) Between Students Who Joined Late and Those Who Were Already in the Study (T1 » T2)

	Study 1 (Joined: $N = 852$)	Study 2 (Joined: $N = 415$)	Study 4 (Joined: $N = 126$)
Vocational Interests			
Realistic	-0.01	-0.01	-0.02
Investigative	-0.04	-0.14	0.10
Artistic	-0.02	-0.04	-0.05
Social	0.06	0.04	-0.12
Enterprising	**0.13**	0.10	0.02
Conventional	0.02	-0.05	-0.02

Note. There were no participants in Study 3 who joined late. Vocational interests were the only constructs reported for T2, as the predictors of profile stability were selected from T1. Bold parameters are significantly different from zero at $p < .05$, two-tailed. Positive values imply higher values for the participants who joined the study late.

Table A3

Intraclass Correlations of All Outcomes

	Study 1		Study 2		Study 3		Study 4	
	T1	T2	T1	T2	T1	T2	T1	T2
Vocational Interests								
Realistic	.07	.03	.07	.02	.06	.03	.15	.14
Investigative	.09	.01	.09	.00	.04	.02	.07	.07
Artistic	.06	.01	.03	.01	.05	.05	.04	.03
Social	.08	.04	.10	.03	.03	.03	.06	.05
Enterprising	.08	.02	.07	.00	.01	.00	.02	.00
Conventional	.10	.03	.09	.00	.00	.00	.04	.03
Big Five		-		-		-		-
Openness	.05	-	.05	-	.03	-	.03	-
Conscientiousness	.05	-	.11	-	.02	-	.00	-
Extraversion	.04	-	.04	-	.00	-	.00	-
Agreeableness	.06	-	.08	-	.01	-	.01	-
Neuroticism	.07	-	.06	-	.00	-	.02	-
Cognitive Abilities								
Quantitative	.12	-	.26	-	-	-	-	-
Figural	.13	-	.22	-	.09	-	.05	-
Verbal	.15	-	.19	-	.19	-	.04	-

Note. ICC's are based on raw data. All ICC's are based on class clustering, except in Study 4, where the clustering was at the school level. Vocational interests were the only constructs reported for T2, as the predictors of profile stability were selected from T1

Table A4
Descriptive Statistics of Scales for All Studies

	M T1	*SD* T1	*N* T1	*M* T2	*SD* T2	*N* T2
Study 1: Vocational Orientation						
Q-correlation [a]	-	-	-	0.45	0.42	687
Realistic	3.09	1.17	1672	2.72	1.14	2288
Investigative	3.18	1.08	1656	2.67	0.99	2283
Artistic	3.28	1.03	1616	2.93	0.95	2273
Social	3.23	1.04	1626	3.27	0.97	2264
Enterprising	2.91	1.07	1587	2.87	0.89	2266
Conventional	2.91	1.05	1637	2.63	0.92	2251
Openness	3.54	0.80	1951	-	-	-
Conscientiousness	3.65	0.82	2073	-	-	-
Extraversion	3.46	0.82	1963	-	-	-
Agreeableness	3.58	0.85	2062	-	-	-
Neuroticism	2.84	0.88	1967	-	-	-
Quantitative	10.17	7.26	3656	-	-	-
Figural	10.34	8.24	3656	-	-	-
Verbal	6.10	4.70	3656	-	-	-
Gender	0.45	0.50	3739	-	-	-
Study 2: Vocational Training						
Q-correlation [a]	-	-	-	0.67	0.35	1006
Realistic	2.59	1.11	1685	2.61	1.13	1714
Investigative	2.63	0.95	1670	2.56	0.94	1728
Artistic	2.88	0.95	1665	2.95	0.93	1718
Social	3.23	0.99	1653	3.27	0.98	1720
Enterprising	2.87	0.88	1656	2.97	0.84	1709
Conventional	2.51	0.88	1670	2.66	0.86	1691
Openness	3.35	0.70	1640	-	-	-
Conscientiousness	3.41	0.74	1707	-	-	-
Extraversion	3.50	0.75	1672	-	-	-

Agreeableness	3.59	0.73	1696	-	-	-
Neuroticism	2.79	0.75	1651	-	-	-
Quantitative	12.04	4.60	1545	-	-	-
Figural	13.80	6.27	1857	-	-	-
Verbal	8.78	3.58	1812	-	-	-
Gender	0.47	0.50	2288	-	-	-
Study 3: College Major Orientation						
Q-correlation [a]	-	-	-	0.64	0.33	640
Realistic	2.51	0.88	2518	2.31	0.81	809
Investigative	2.62	0.81	2459	2.69	0.78	808
Artistic	2.77	0.86	2469	2.59	0.84	809
Social	3.02	0.88	2453	3.09	0.83	812
Enterprising	3.11	0.81	2411	3.18	0.78	811
Conventional	2.51	0.72	2458	2.77	0.77	792
Openness	3.70	0.67	2493	-	-	-
Conscientiousness	3.33	0.57	2494	-	-	-
Extraversion	3.79	0.80	2511	-	-	-
Agreeableness	3.68	1.06	2528	-	-	-
Neuroticism	2.11	0.65	2482	-	-	-
Quantitative	-	-	-	-	-	-
Figural	17.67	4.46	2572	-	-	-
Verbal	10.39	3.46	2574	-	-	-
Gender	0.50	0.50	3018	-	-	-
Study 4: Entering the Workforce						
Q-correlation [a]	-	-	-	0.74	0.26	1264
Realistic	2.25	0.75	2328	2.28	0.74	1388
Investigative	2.81	0.75	2328	2.81	0.76	1388
Artistic	2.95	0.87	2328	2.68	0.81	1388
Social	3.16	0.80	2328	3.10	0.75	1388
Enterprising	3.25	0.76	2328	3.13	0.75	1388

Conventional	2.63	0.71	2328	2.77	0.70	1387
Openness	2.83	0.53	2281	-	-	-
Conscientiousness	3.09	0.43	2271	-	-	-
Extraversion	2.89	0.43	2257	-	-	-
Agreeableness	3.28	0.39	2282	-	-	-
Neuroticism	2.13	0.52	2272	-	-	-
Quantitative	-	-	-	-	-	-
Figural	16.43	3.58	2449	-	-	-
Verbal	10.44	2.71	2445	-	-	-
Gender	0.62	0.49	2445	-	-	-

Note. Descriptive statistics are based on raw data, not imputed data.

[a] Stability coefficient refers to the interval between T1 and T2; however, for a better visualization it is depicted in the columns of T2.

Supplement B: Instruments

Table B1

Overview of the Number of Items, Instruction and Scale of the Instruments

Variable	Study	No. of Items	No. of Items (w/o r)	Total (w/o r)	Instruction	Likert Scale
Vocational Interests	1 2 3 4	R: 6, I: 6, A: 6, S: 6, E: 6, C:6 R: 10, I: 10, A: 10, S: 10, E: 10, C: 10	R: 6, I: 6, A: 6, S: 6, E: 6, C:6 R: 10, I: 10, A: 10, S: 10, E: 10, C: 10	36 60	"How much do you like this activity?"	1 (*not at all*) to 5 (*very*)
Personality Traits	1 2	O: 11, C: 9, E: 8, A: 8, N: 8	O: 8, C: 5, E: 5, A: 4, N: 5	44(27)	"How much do the following statements apply to you? I am someone who is …"	1 (*is absolutely not true)* to 4 *(is exactly true)*
	3	O: 5, C: 4, E: 4, A: 4, N: 4	O: 4, C: 3, E: 3, A: 1, N: 3	21(14)	"Please describe yourself based on the following statements. I ..."	1 (*very incorrect/inappropriate*) to 3 (*neither nor*) to 5 (*very correct*)
	4	O: 12, C: 12, E: 12, A: 12, N: 12	O: 5, C: 8, E: 8, A: 4, N: 8	60(32)		1 (*very incorrect*) to 5 (*very correct*)
Cognitive Abilities	1 2 3 4	Quant: 20, Figure: 20, Verbal: 25 Figure: 25, Verbal: 20	Quant: 20, Figure: 20, Verbal: 25 Figure: 25, Verbal: 20	65 45	"Below you will be asked questions. Try to answer them correctly. Some will be easy for you, others may be a bit more difficult."	0 (*incorrect*) to 1 (*correct*)

Note. (w/o r) = implies that scales did not include the reverse coded items.

5 STUDY 3

THE EFFECTS OF OUT-OF-SCHOOL ENGAGEMENT IN SCIENCE ON ADOLESCENTS' VOCATIONAL INTERESTS, OCCUPATIONAL ASPIRATIONS, COMPETENCIES, SCHOOL GRADES, AND ABILITY BELIEFS

Gfrörer, T., Stoll, G., Rieger, S., & Nagengast, B. (to be submitted). The Effects of Out-of-School Engagement in Science on Adolescents' Vocational Interests, Occupational Aspirations, Competencies, School Grades, and Ability Beliefs.

Abstract

Unstructured out-of-school time (OST) science activities, such as reading a science book, watching a science TV show, or researching on the internet about science, constitute a self-sustaining way to engage adolescents in science, technology, engineering, and mathematics (STEM). Theories about interest development and trait complexes suggest that the long-term engagement in such activities could have a broad influence on several STEM-related constructs. However, so far little is known about the impact of unstructured OST science activities. The current study investigated the effects of unstructured OST science activities on the development of vocational interests, occupational aspirations, competencies, school achievement and ability beliefs. For this purpose, we used a large longitudinal subsample (N = 2,655) from the National Educational Panel Study (NEPS; Blossfeld et al., 2011) where students in Germany were assessed in Grades 9, 11 and 12. Following the recommendations of VanderWeele et al. (2020), we used an outcome-wide longitudinal design for causal inference: Outcome-wide causal effects of a treatment were estimated while controlling for a set of joint confounders and pretest measures. Our findings show that OST science activities influence Investigative vocational interests, but not occupational aspirations, competencies, school grades, and ability beliefs. The results suggest that adolescents with similar initial interest trait levels who engage in unstructured OST science activities may develop a stronger interest towards STEM, compared to adolescents who do not engage in such activities.

Keywords: science activities, outcome-wide longitudinal design for causal inference, vocational interests, aspirations, competencies, school grades, ability beliefs

The Effects of Out-of-School Engagement in Science on Adolescents' Vocational Interests, Occupational Aspirations, Competencies, School Grades and Ability Beliefs

Climate change, biodiversity loss and disease control are some of the major challenges of the present age that require experts in science, technology, engineering, and mathematics (STEM; Fenster & Gropp, 2020). STEM graduates are in high demand (National Academy of Sciences, 2007) and are well paid (Noonan, 2017), but still too few students choose STEM careers. To counteract this, policy makers in Europe (OECD, 2008, 2016) and the U.S. (National Academy of Sciences, 2007) have sought to increase motivation and competencies of students in STEM.

Important for STEM engagement are structured learning environments within the school system, such as lessons in science subjects, enrichment programs (e.g., Golle et al., 2018), or extracurricular courses (e.g., Hirschenhauser et al., 2019). They are usually implemented by teachers, possess a defined structure, and are often integrated within the school curriculum. Independent of structured school settings, there are also unstructured out-of-school time (OST) science activities (see Dabney et al., 2012), such as reading a science book, watching a science TV show, or researching on the internet about science. They resemble "scientific pursuits in […] spare time" (Dabney et al., 2012, p. 65), occur without any instruction and constitute an informal, unstructured and self-sustaining way of engaging in science outside of schools (Maltese & Tai, 2010).

The Informal Science Education Ad Hoc Committee of the U.S. National Association of Research in Science Teaching (Dierking et al., 2003) and the U.S. National Research Council (Hein, 2009) suggest that OST science activities are an untapped resource of science engagement, which might have a wide-ranging influence on students' understanding of and interest in STEM. Engagement in unstructured OST science activities is associated with STEM outcomes (e.g., Dabney et al., 2012; Dou et al., 2019; Henriksen et al., 2015); however most empirical studies on their effects did not adequately control for potential confounders such as prior interest or achievement. This makes it difficult to draw causal conclusions about the impact of unstructured OST science activities. It cannot be ruled out that alternative mechanisms are responsible for apparent beneficial effects, such as self-selection effects of people who already favored STEM activities. It is therefore still unknown if and to what extent unstructured OST science activities foster a positive development of STEM-related constructs.

The present study investigated the impact of unstructured OST science activities on a broad range of outcomes. Based on considerations from the Trait-Situation Interest Dynamics

Model (TSID; Su et al., 2019) and Ackerman and Heggestad's (1997) trait complexes, we focused on five outcome clusters: dispositional interests, occupational aspirations, competencies, school achievement, and ability beliefs. We used a recently proposed template for outcome-wide longitudinal designs for causal inference (see VanderWeele et al., 2020). This enabled us to investigate the causal effect of unstructured OST science activities, while controlling for a set of joint confounders (e.g., pretest measures and competence tests) that were collectively selected for all outcomes. Compared to single-outcome studies, outcome-wide analyses cumulate research findings more efficiently and provide comprehensive information for researchers, policy makers, and practitioners. We investigated a longitudinal subsample (N = 2,655) from the National Educational Panel Study (NEPS; Blossfeld et al., 2011), in which students in Germany were assessed in Grades 9, 11 and 12.

Potential Effects of Unstructured Out-of-School Time Science Activities

Unstructured OST science activities are defined as spare time activities with science content (Dabney et al., 2012) that occur without any instruction and constitute an informal and unstructured way of science engagement outside of institutions like schools, universities, museums, or science centers (Dierking et al., 2003; Maltese & Tai, 2010). Although unstructured OST science activities are primarily leisure related, they may influence students' lives well beyond just leisure (Dabney et al., 2012, 2016; Dierking et al., 2003; Dou et al., 2019; Falk & Storksdieck, 2009; Maltese & Tai, 2010). For instance, unstructured OST science activities might initiate learning processes (Tal & Dierking, 2014), motivation (Uitto et al., 2006), or occupational aspirations (Henriksen et al., 2015) in STEM.

Drawing on theories about interest development (Su et al., 2019) and trait complexes (Ackerman & Heggestad, 1997) as well as on considerations from research on science engagement (Azevedo, 2015; Tal & Dierking, 2014), we derived assumptions about the potential effects of unstructured OST science activities. The three theoretical approaches indicate that unstructured OST science activities positively influence the development of STEM-related constructs, such as interests, competencies, or ability beliefs, that capture individual differences in the broad area of STEM. As reciprocal effects are expected between STEM-related constructs and unstructured OST science activities, study designs need to control for existing baseline differences in these constructs to adequately estimate the exclusive impact of activity engagement.

General Effects of Science Engagement

Tal and Dierking (2014) emphasize that STEM learning is based on an accumulation of experiences obtained not only in school, but also in structured out-of-school settings, such as museums, and unstructured out-of-school settings, such as hobbies (Tal & Dierking, 2014). Individuals are assumed to evolve their "understanding, attitudes, and behaviors" in science, based on the interaction of various real-world science experiences (Tal & Dierking, 2014, p. 252). Theories about science engagement therefore imply that the engagement in science activities has a wide-ranging impact on multiple constructs and is driven by an accumulation of real-world experiences (Tal & Dierking, 2014). These assumptions are similar to considerations from theories about trait complexes (Ackerman & Heggestad, 1997) and theories about interest development (Su et al., 2019).

Science engagement occurs on several levels and includes emotional, cognitive, and behavioral dimensions (Azevedo, 2015; Sinatra et al., 2015). For example, adolescents researching about climate change engage emotionally by thinking about the possible consequences of climate change, engage cognitively by processing the mechanisms of the greenhouse effect, and engage behaviorally by discussing the respective topic with their classmates (Sinatra et al., 2015). As science engagement consists of different dimensions (i.e., emotional, cognitive, and behavioral), science-related activities could influence the development of various constructs that each correspond to the different dimensions. This implies that, independent of the setting (i.e., structured school environment, structured out-of-school environment, or unstructured out-of-school environment), science engagement might influence not only motivational (i.e., interests and ability beliefs), but also cognitive (i.e., grades and competencies) and behavioral variables (e.g., subject, study major, or career choices).

Potential Effects on Interest Development

Interests can be impacted by the long-term engagement in activities (Su et al., 2019), while simultaneously being a major predictor for choosing these activities (Holland, 1997). Therefore, when the influence of activity engagement on interest development is investigated, the possibility of reciprocal effects must be considered. For example, if we want to investigate the influence of watching a science TV show on the development of science interest, we need to control for prior interest in science, because if we do not control for interest in science prior to watching a science TV show, differences in the development of science interest could occur not only due to differences in the amount of watching a science TV show, but also due to differences in prior science interest.

Many theories about interest development describe how engagement in activities can shape a person's interest (Hidi & Renninger, 2006). The TSID model (Su et al., 2019) assumes that stable interests are reflected by a mental representation of the object of interest, which includes corresponding affective and cognitive responses (Su et al., 2019). An accumulation of short-term situational interest gradually enriches the mental representation and ultimately leads to the evolvement of an interest disposition. Interest in a specific topic can therefore be strengthened through positive experiences—for example, novel activities that arouse curiosity, provide surprising information or are cognitively engaging (Hidi & Renninger, 2006; Renninger & Hidi, 2011)—or be weakened through negative experiences (Su et al., 2019). Unstructured OST science activities are usually accompanied by positive emotions, intrinsic value, and autonomy—properties that can facilitate short-term interest in a topic (Eccles et al., 1998; Hidi & Renninger, 2006; Ryan & Deci, 2000). Therefore, long-term engagement in these activities could lead to positive affective and cognitive appraisals of STEM and an increase in STEM interests.

People who are already interested in science are more likely to engage in unstructured OST science activities, and more engagement in unstructured OST science activities could lead to more pronounced interest in science. Therefore, to distinguish the effects of unstructured OST science activities from other possible explanations such as baseline differences, one needs to control for a wide range of confounding variables (i.e., variables that influence both activity engagement and the target outcome). This includes not only prior interest in science, but also other constructs that might influence the engagement in unstructured OST science activities, such as prior ability beliefs, aspirations, competencies, content knowledge, or personality traits. As confounders need to be measured before activity engagement, longitudinal research designs, which adhere to the temporal order of confounders, activity engagement, and target constructs, are crucial for the investigation of unstructured OST science activities.

Potential Effects on Trait Complexes

Besides influencing interest development, unstructured OST science activities could also have more wide-ranging effects. Arguing from the perspective of trait complexes (Ackerman et al., 2011; Ackerman & Heggestad, 1997), unstructured OST science activities could impact a broad cluster of STEM-related constructs. Trait complexes are combinations of abilities, interests, and personality traits (Ackerman et al., 2011) that are associated with each other and are assumed to develop in interaction with each other. Based on meta-analytic evidence and an additional review of several studies, Ackerman and Heggestad (1997)

identified a science/math trait complex, which combines science- and math-related cognitive abilities with investigative and realistic dispositional interests as well as domain-specific knowledge and ability beliefs in science and technology (Ackerman, 1997; Ackerman et al., 2001).

Trait complexes can develop as a function of the environment because different experiences can lead to different expressions of the trait complex. Ackerman and Beier (2003) showed that, depending on the study major field, students possessed other trait complexes: For example, students who graduated in the field of physical science were characterized by the science/math trait complex. This assumption is in line with recent theories about personality development, which suggest that personality traits, interests, goals, attitudes, and cognitive abilities (Roberts & Wood, 2006) could co-develop over time due to shared situational content (Wrzus & Roberts, 2017). Situations are assumed to initiate short-term changes in personality states, which might lead to long-term development of certain personality characteristics (Hoff et al., 2020; Wrzus & Roberts, 2017).

Adolescents who experience positive emotions during science activities will connect their feelings during these situations to behavioral patterns that are linked to these activities (for a similar explanation see Hoff et al., 2020). Over time, adolescents internalize these behavioral patterns so that they become stable personality characteristics. This implies that adolescents who regularly engage in unstructured OST science activities may develop more pronounced science/math trait complexes compared to adolescents who do not. As constructs within the math/science trait complex could also be predictors for the choice of unstructured OST science activities, controlling for confounding that applied to the investigation of dispositional interests also applies to the constructs located within the math/science trait complex.

Empirical Evidence on Unstructured Out-of-School Time Science Activities

In contrast to structured science-related learning environments (for meta-analyses see Aguilera & Perales-Palacios, 2020; Guzzetti et al., 1993; Therrien et al., 2011; for single studies see Bernacki et al., 2020; Friedman et al., 2017; for a What Works Clearinghouse report see What Works Clearinghouse et al., 2021), quantitative evidence on the influence of unstructured OST science activities is scarce and limited to a small range of outcomes. In addition, the existing studies used different operationalizations of unstructured OST science engagement (e.g., groups with no vs. any engagement, groups with low vs. high engagement or continuous measures of engagement) making it difficult to compare effects across studies and outcomes.

Studies Investigating the Impact of Unstructured OST Science Activities

Quantitative studies of unstructured OST science activities, defined as unstructured spare time activities with science content, demonstrate that they are associated with STEM-related constructs. For example, Dabney et al. (2012) showed in a sample of N = 6,882 U.S. university students that those students who regularly read or watched science fiction or non-fiction science content (in contrast to students who did not) were more likely to aspire to a STEM occupation. Reading or watching science fiction or non-fiction science was operationalized by a dichotomous variable (i.e., reading/watching vs. not reading/watching) based on two categorical items. The results suggest an association between the engagement in unstructured OST science activities and STEM aspirations. However, Dabney et al. (2012) did not control for confounding variables, such as science competencies or dispositional interests.

Henriksen et al. (2015) demonstrated that STEM major choice in university was associated with watching popular science shows, reading popular science books and having leisure time experiences in nature. This suggests that the engagement in unstructured OST science activities is related to study major choices. However, the analysis was based on a cross-sectional sample, in which students directly stated if certain leisure science activities influenced their study major choice. It is therefore unknown if the relationship between unstructured OST science activities and study major choice can also be found in a longitudinal sample where both constructs (i.e., activity engagement and the actual career choice) are measured separately.

Dou et al. (2019) reported similar results for the association between unstructured OST science activities and STEM interests. University students who retrospectively stated higher engagement in unstructured OST science activities during their K-4 school years had higher levels of STEM interest during university (Dou et al., 2019). Engagement was measured based on dichotomous items that indicated whether students ever pursued one of several activities at any time during their K-4 school years or not. The results suggest that STEM interest can be influenced by the engagement in unstructured OST science activities. However, the analysis was again based on cross-sectional data and retrospective measures.

Uitto et al. (2006) investigated the influence of out-of-school experiences on interest in biology in a sample of N = 3,626 Finnish secondary school students. Their results suggest that out-of-school nature experiences were highly correlated with a general interest in biology. In addition, different types of OST science experiences (e.g., technology-related or animal-related activities) were related to different areas of interest in biology. Students had to state their OST science experiences based on 61 statements on a 4-point Likert scale from “never” to “often”. Items were summarized based on factor analysis. Their results suggest that there is a

relationship between interest in specific STEM areas and corresponding unstructured OST science activities.

In summary, empirical findings on the effects of unstructured OST science activities are mostly confined to STEM interests (e.g., Dou et al., 2019) and aspirations (e.g., Dabney et al., 2012; Henriksen et al., 2015). To our knowledge, the relationship between OST science activities and achievement-related constructs (e.g., competencies, school achievement and ability beliefs) or dispositional interests (e.g., vocational interests) has not been investigated yet. Most of the current studies also implemented study designs that did not allow to distinguish between self-selection effects and effects of unstructured OST science activities. Longitudinal study designs that include multiple measurement waves and representative samples are missing. In addition, most studies did not adequately control for potential confounders.

A Need for Causal Evidence

There are two intertwined issues that limit the conclusiveness of the current findings on unstructured OST science activities. First, the majority of studies (e.g., Dabney et al., 2012; Dou et al., 2019; Henriksen et al., 2015; Uitto et al., 2006) implemented non-experimental study designs that have severe limitations with respect to clear causal conclusions. An approach to estimate causal effects in non-experimental studies is to control for all confounding variables (e.g., pretest measures, standardized competence tests, or motivational variables) that may influence both the engagement in unstructured OST science activities and the relevant STEM outcomes (VanderWeele et al., 2020). Second, because reciprocal effects between unstructured OST science activities and STEM outcomes are highly probable (Dabney et al., 2012), strong longitudinal designs are needed that measure baseline characteristics—especially pretests of the outcome, which are assessed before the respective activity and the outcome (VanderWeele et al., 2020).

Although studies (e.g., Dabney et al., 2012; Dou et al., 2019; Henriksen et al., 2015) suggest that especially motivational variables could be influenced by the engagement in unstructured OST science activities, due to their methodological shortcomings, it is unclear to what extent these effects are real or just artifacts of methodologically weak designs. As reciprocal effects are a major concern and current studies about unstructured OST science activities “do not address the role that self-selection may play” (Dabney et al., 2012, p. 65), present findings could also merely represent baseline differences in relevant STEM constructs. It is therefore a possibility that some of the findings of current studies could end up as null results as more rigorous control for confounding is applied.

Proposing an Outcome-Wide Investigation

As current studies focused only on a limited number of STEM-related outcomes, research about unstructured OST science activities would benefit from an outcome-wide investigation that examines a range of heterogenous STEM-related outcomes (i.e., interests, aspirations, competencies, school achievement, and ability beliefs). VanderWeele et al. (2020) recently provided a template for such outcome-wide analyses—the outcome-wide longitudinal design for causal inference. In this design, outcome-wide effects of a treatment are investigated while controlling for a set of joint confounders and pretests that are collectively selected for all outcomes (VanderWeele et al., 2020). In comparison to studies with a single outcome (e.g., academic performance; Meda et al., 2017), outcome-wide studies make decisions about confounder selection and statistical modelling for all outcomes simultaneously (VanderWeele et al., 2020). This is a fitting design for the investigation of unstructured OST science activities because such activities could influence several STEM-related constructs that have a joint set of confounding variables. Besides conveying more information, outcome-wide studies have several advantages. They decrease researchers' degrees of freedom (see Asendorpf et al., 2013; Wicherts et al., 2016) to choose specific model-outcome constellations that yield significant effects, they increase the reporting of null findings (see Asendorpf et al., 2013; Open Science Collaboration, 2015; Wicherts et al., 2016) because researchers are less dependent on a particular significant treatment effect, and they enable the comparison of treatment effects on different outcomes (VanderWeele et al., 2020).

The Present Study

In the present study, we investigated the effects of unstructured OST science activities on a broad range of outcomes, using an outcome-wide longitudinal design for causal inference (VanderWeele et al., 2020). Based on the reviewed literature, we investigated effects of unstructured OST science activities on five broad outcome clusters, both STEM related and unrelated: 1) dispositional interests; 2) occupational aspirations; 3) competencies in reading, math, and scientific thinking; 4) school achievement in German, math, and English; and 5) ability beliefs (i.e., self-concept) in math, German, and about one's general performance in school. We refrained from making assumptions about the effects of the current study, as some of the reported empirical findings on unstructured OST science activities could turn out as null results when more rigorous control for confounding is applied.

To fulfill the requirements of an outcome-wide longitudinal design for causal inference, we used a representative large-scale longitudinal subsample (N = 2,655) from the German National Educational Panel (NEPS) study, which allowed us to control for pretest measures of all outcome clusters (assessed in 9th grade) before the treatment (i.e., unstructured OST science activities, assessed in 11th grade) and measure all outcome variables once again after the treatment (in 12th grade). Students were representatively selected from schools in Germany and were further surveyed if they dropped out of the school system. The selection of the NEPS subsample provides the possibility to investigate the influence of unstructured OST science activities during upper secondary education, a crucial time period for the development of STEM-related constructs.

Previous studies about unstructured OST science activities varied in their operationalization of activity engagement. To ensure the robustness of our results and to investigate if differences in operationalization matter, we examined different forms of activity engagement. In our study, unstructured OST science activities were measured based on five items. Because we were interested in the overall impact, we created a composite treatment variable that separated the participants in groups of low and high engagement. However, we also investigated the influence of different activity engagement operationalizations (i.e., a variety of composite treatments and as a continuous variable) and the impact of the single activities in the form of robustness checks.

Method

Sample

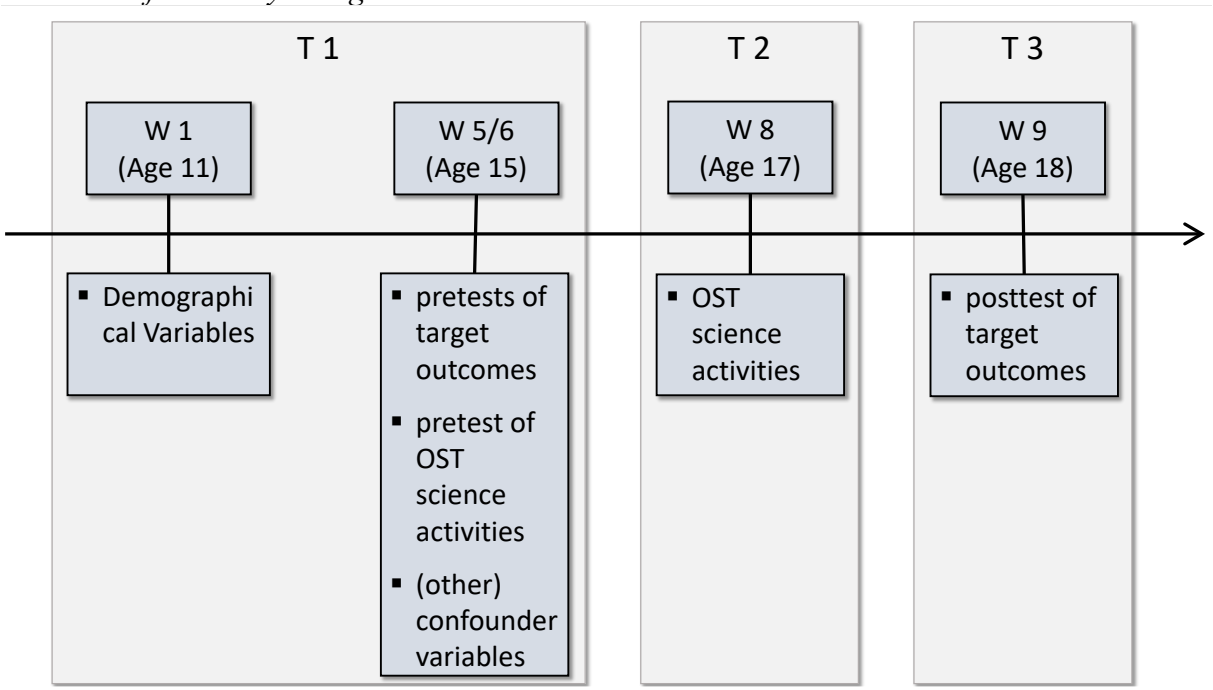
The current study used data from the National Educational Panel Study (NEPS): Starting Cohort Grade 5 (doi:10.5157/NEPS:SC3:10.0.0; for more information about the main study samples and the general study design see Blossfeld et al., 2011, and NEPS, 2021a).[8] For the present study, we selected a subsample that comprised all students who participated at each of the Waves 6 (i.e., Grade 9), 8 (i.e., Grade 11) and 9 (i.e., Grade 12). We restricted our analysis sample because the constructs that were relevant to our study were measured at these three waves. In addition, by selecting participants from these waves, we kept the necessary temporal order of confounding variables, followed by treatment, followed by target outcomes (see VanderWeele et al., 2020). The analysis sample comprised N = 2,655 participants (51% were female; 2.5% were not born in Germany) from 374 classes and 201 schools across all states of Germany (see Supplement A for more information about the subsample composition).

Study Design

Our study comprised three time points that captured different variables, namely pretests and relevant confounders (T1), treatment (T2) and target outcomes (T3; see Figure 1). Pretests and relevant confounders were taken from Waves 5 and 6 of the Starting Cohort Grade 5 Study (see NEPS, 2021b) because the two waves were only two months apart. Demographic variables that were unlikely to change or assumed to be relatively stable, such as migration background or socio-economic status, were taken from the first wave. The treatment (i.e., unstructured OST science activities) was taken from Wave 8, and the target outcomes were taken from Wave 9.

[8] From 2008 to 2013, NEPS data was collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS has been carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network.

Figure 1
Overview of the Study Design



Note. The grey areas show the time points (T) of the current study. The dark grey areas show the waves (W) of the NEPS study. A detailed overview of the respective measures can be found in Table 1.

Instruments

Table 1 provides relevant information about all instruments used in this study. For more information about the item wording see Supplement B; for further information about the NEPS-instruments see NEPS (2021a).

Treatment Operationalization

Engagement in unstructured OST science activities was measured based on five items: "Watch TV shows about natural sciences", "Borrow or buy books about natural sciences", "Visit internet sites with topics relating to natural sciences", "Read natural science magazines or articles in newspapers", and "Attend a natural science project group". The participants had to state, "How often do you do the following things?" based on the scale (1) "never", (2) "rarely", (3) "sometimes" and (4) "often". Information that is received from such items can be integrated into the analysis in different ways. Previous studies about unstructured OST science activities either investigated the influence of each item separately, created an aggregated continuous scale, or created a dichotomous variable to categorize the participants into two groups of activity engagement.

Table 1
Overview of the Instruments

Construct	Questionnaire	Item Description	Scale range	Items	Score	Reliability / Missing Rate T1	T6	T8	T9
Treatment									
OST sci.: TV	NEPS	"Watching Science TV"	1-4(+)	1	Item	-	(–/3%)	(–/39%)	-
OST sci.: Books	NEPS	"Reading Science Books"	1-4(+)	1	Item	-	(–/3%)	(–/39%)	-
OST sci.: Internet	NEPS	"Researching the Internet"	1-4(+)	1	Item	-	(–/3%)	(–/39%)	-
OST sci.: Magazine	NEPS	"Reading Science Magazine"	1-4(+)	1	Item	-	(–/4%)	(–/39%)	-
OST sci.: Project	NEPS	"Attending Science Project"	1-4(+)	1	Item	-	(–/3%)	(–/39%)	-
Outcomes									
Int.: Realistic	AIST-R/ICA-D	"building something"	1-5(+)	3	Mean	-	(.73/4%)	-	(.75/34%)
Int.: Investigative	AIST-R/ICA-D	"experimenting in a lab"	1-5(+)	3	Mean	-	(.70/4%)	-	(.75/34%)
Int.: Artistic	AIST-R/ICA-D	"drawing pictures"	1-5(+)	3	Mean	-	(.76/4%)	-	(.77/34%)
Int.: Social	AIST-R/ICA-D	"helping others"	1-5(+)	3	Mean	-	(.80/4%)	-	(.81/34%)
Int.: Enterprising	AIST-R/ICA-D	"leading a group"	1-5(+)	3	Mean	-	(.49/4%)	-	(.60/34%)
Int.: Conventional	AIST-R/ICA-D	"organizing things"	1-5(+)	3	Mean	-	(.58/5%)	-	(.60/34%)
Asp.: Realistic	NEPS	"What would be your Ideal Occupation? (ISCO 08 recoded by O*NET information)"	1-7(+)	1	Item	-	(–/19%)	-	(–/17%)
Asp.: Investigative	NEPS		1-7(+)	1	Item	-	(–/19%)	-	(–/17%)
Asp.: Artistic	NEPS		1-7(+)	1	Item	-	(–/19%)	-	(–/17%)
Asp.: Social	NEPS		1-7(+)	1	Item	-	(–/19%)	-	(–/17%)
Asp.: Enterprising	NEPS		1-7(+)	1	Item	-	(–/19%)	-	(–/17%)
Asp.: Conventional	NEPS		1-7(+)	1	Item	-	(–/19%)	-	(–/17%)

Comp.: Math	NEPS	"Mathematical literacy concept"	0-1(correct)	23	WLE	-	(.80/23%)	-	(.73/20%)
Comp.: Reading	NEPS	"Text understanding/interpr."	0-1(correct)	33	WLE	-	(.78/24%)	-	(.72/23%)
Comp.: Sc. Th.	NEPS	"Understanding sc. concepts"	0-1(correct)	32	WLE	-	-	-	(.57/57%)
Comp.: Sc. Lit.	NEPS	"Scientific literacy concept"	0-1(correct)	28	WLE	-	(.76/23%)	-	-
Grade: Math	NEPS	"State your end year grade"	1-6(+)	1	Item	-	(–/3%)	-	(–/36%)
Grade: German	NEPS		1-6(+)/	1	Item	-	(–/3%)	-	(–/36%)
Grade: English	NEPS		1-6(+)	1	Item	-	-	-	(–/39%)
Self-Con.: Math	NEPS	"I am good at math"	1-4(+)	3	Mean	-	(.89/11%)	-	(.89/34%)
Self-Con.: German	NEPS	"I learn fast in German"	1-4(+)	3	Mean	-	(.83/10%)	-	(.87/34%)
Self-Con.: School	NEPS	"I learn fast in school subjects"	1-4(+)	3	Mean	-	(.85/11%)	-	(.88/34%)
Conf.: Demogr.									
Gender	NEPS	"Male or Female"	0-1(female)	1	Item	(–/0%)	-	-	-
Migration	NEPS	"Migration of Child/Parents"	0-1(mig.)	1	Item	(–/3%)	-	-	-
SES	Ganzeboom & Treiman (1996)	"Highest ISEI Parents"	0-100(+)	1	Item	(–/22%)	-	-	-
Economic Situation	NEPS	"Parent rating of econ. situation"	1-5(+)	1	Item	(–/25%)	-	-	-
School Track	NEPS	"List of school tracks"	-	1	Item	(–/20%)	-	-	-
Conf.: Health									
Sport Frequency	NEPS	"How often do you do sport?"	1-5(+)	1	Item	-	-	(–/32%)	-
Global Self-Worth	NEPS	"Are you satisfied with yourself?"	1-5(+)	10	Mean	-	(.90/14%)	-	-
Chronic Stress	NEPS	"Stress in everyday activities"	1-5(+)	11	Mean	-	(.74/13%)	-	-
Conf.: Personality									
Conscientiousness	NEPS	"I am thorough"	1-5(+)	2	Mean	-	(.53/10%)	-	-
Extraversion	NEPS	"I am sociable"	1-5(+)	2	Mean	-	(.67/10%)	-	-
Openness	NEPS	"I am imaginative"	1-5(+)	2	Mean	-	(.49/10%)	-	-
Neuroticism	NEPS	"I am relaxed"	1-5(+)	2	Mean	-	(.47/10%)	-	-

Agreeableness	NEPS	"I am sensitive"	1-5(+)	3	Mean	-	(.39/10%)	-	-
Conf.: Cog. Ab.									
Processing Speed	NEPS	"Assign number to symbol"	0-1(correct)	93	Sum	-	(–/21%)	-	-
Logical Thinking	NEPS	"matrix tasks"	0-1(correct)	12	Sum	-	(–/22%)	-	-
Reading Speed	NEPS	"Is sentence content correct?"	0-1(correct)	51	Sum	-	(–/23%)	-	-
Conf.: Sch. Ach.									
Science Grade	NEPS	"Average of science grades"	1-6(-)	4	Mean	-	(–/4%)	-	-
Conf.: Fam. Env.									
Activities w/ Parents	NEPS	"Parent rating of act. with child"	1-7(+)	9	Mean	(.64/27%)	-	-	-
Conf.: Motivation									
Int. Mot.: Reading	NEPS	"It is fun to read books"	1-4(+)	6	Mean	-	(.91/4%)	-	-
Helplessness Math	NEPS	"I will not succeed in math"	1-4(-)	5	Mean	-	(.87/4%)	-	-
Subject Int.: Math	NEPS	"I like solving math problems"	1-4(+)	4	Mean	-	(.86/5%)	-	-
Subject Int.: German	NEPS	"I like writing"	1-4(+)	4	Mean	-	(.84/5%)	-	-
Self-Con.: Reading	NEPS	"I easily understand texts"	1-4(+)	3	Mean	-	(.83/4%)	-	-
Class Repetition	NEPS	"Did the child ever repeat a class?"	0-1(yes)	1	Item	-	(–/40%)	-	-

Note. Score describes the scale type. The reliability estimates are omega total (see McNeish, 2018) because omega total needs more than two items to identify the model without putting additional constraints on factor loadings and residual variances. We provided Cronbach's alpha for instruments with two items. Scale range describes the range of the (Likert) scales of the respective items. For example the scale for the OST science activities ranges from 1 to 4, which stands for "never", "rarely", "sometimes" and "often", respectively (see Table B1 for the wording of scales). (+) = states that higher values imply, for example, more interest, self-concept or achievement; (-) = states that higher values imply, for example, less interest, self-concept or achievement; (correct/female/mig./yes) = represent the meaning of the value 1 of the respective dichotomous scales; Conf. = confounding variable.

The approach of categorizing participants into two groups of activity engagement was the most common among previous studies (see Dabney et al., 2012; Dou et al., 2019). To be able to compare the previous results with those of our study, we focused primarily on comparing groups of low and high activity engagement to each other. Hence, we decided to build a dichotomous variable out of the respective items. However, to ensure the robustness of our results, we analyzed different operationalizations of the engagement of unstructured OST science activities as robustness checks. At first, the respective items were aggregated into a single scale by calculating their mean or sum score. Afterwards the scale was transformed into dichotomous variables that indicated either low (or none) or high (or any) engagement in unstructured OST science activities. As there are multiple ways to build dichotomous variables, we used three different approaches. We created dichotomous groups based on a median split and based on engagement frequency (i.e., no vs. any engagement and very low vs. very high engagement).

In addition, because one of the items (i.e., "Attend a natural science project group") implies a structured learning setting, we calculated all engagement operationalizations with and without this item. We therefore investigated engagement (1) on the item level, by examining the influence of the five activities separately; (2) by a continuous variable that was calculated based on the mean of the activity items (both including and excluding the project group item); and (3) based on three differently computed dichotomous variables (both including and excluding the project group item). This results in a total of five engagement operationalizations on the item level and seven engagement operationalizations on the scale level. In the following, we will explain the respective operationalizations in more detail.

Median Split. To create groups of low and high engagement in unstructured OST science activities we created a dichotomous variable out of the five items. At first, the respective items were aggregated into a single scale by calculating their mean. Afterwards, we conducted a median split of the scale. Participants who had mean scores under the median represented the low activity group (i.e., coded by a 0) and participants who had mean scores above the median represented the high activity group (i.e., coded by a 1). We focus on the median split operationalization in our study; the results of the remaining operationalizations will be depicted in the robustness checks.

No vs. Any Treatment. To create groups of no and any engagement in unstructured OST science activities we created a dichotomous variable out of the five items. At first, we calculated a sum scale out of the items and then separated the participants into the groups "no engagement" and "any engagement". Participants who indicated on all items that they "never"

engaged in any of the activities were assigned to the "no engagement" group (i.e., coded by a 0). Participants who indicated any engagement on any of the items (i.e., a score higher than "never") were assigned to the "any engagement group" (i.e., coded by a 1).

Extreme Group Comparison. To create groups of very low and very high frequencies of engagement in unstructured OST science activities we created a dichotomous variable out of the five items. At first, the respective items were aggregated into a single scale by calculating their mean. Then we z-standardized the respective scale. Afterwards we assigned participants with one standard deviation below the mean (i.e., z-scores ≤ -1) to a "very low activity group" (i.e., coded by a 0) and participants with one standard deviation above the mean (i.e., z-scores ≥ 1) to a "very high activity group" (i.e., coded by a 1). Participants who were not part of either of the two groups were excluded from the analysis.

Continuous Intensity. Besides these three dichotomous treatments we also modelled the treatment effect continuously. The respective items were aggregated into a single scale by calculating their mean.

Outcomes

Dispositional Interests. For the operationalization of dispositional interests, we used Holland's (1997) RIASEC model for vocational interests. Vocational interests are assumed to be dispositions that are relatively stable over time (Rounds & Su, 2014). Holland's (1997) model states that people's vocational interests can be characterized by six general dimensions: Realistic, Investigative, Artistic, Social, Enterprising and Conventional (RIASEC; for item descriptions see Table 1 and Table B1). The dimension Realistic is related to interest in engineering, as it captures preferences for the manipulation of objects, tools, and machines (Holland, 1997). The dimension Investigative is related to interest in science, as it captures preferences for the systemic observation and investigation of physical, cultural, and biological phenomena (Holland, 1997). The other dimensions in Holland's (1997) model are not supposed to be related to the areas of STEM.

In the NEPS study, items for the respective interest dimensions were selected from the Revised General Interest Structure Test (AIST-R; Allgemeiner Interessen Strukturtest; Bergmann & Eder, 2005) and the German version of the Inventory of Children's Activities (ICA; Tracey & Ward, 1998; German version [ICA-D]: von Maurice, 2006). For an overview of the selected items, item examples and reliability coefficients see Table 1 and Supplement B.

Occupational Aspirations. Idealistic occupational aspirations were reported by the participants in an open-ended question format. Due to reasons of data protection, the actual

answers of the participants were anonymized and available only as International Standard Classification of Occupation (ISCO-08) codes. For each participant, it was possible to transform the categorical ISCO-08 codes into a profile that consisted of continuous scores of occupational aspirations categorized based on the six RIASEC dimensions. The Occupational Information Network (O*NET), which was established by the United States Department of Labor, provides information on almost every occupation in the form of continuous scores on the six RIASEC dimensions and a crosswalk that enabled us to join these scores to the associated ISCO-08 codes. We therefore matched the ISCO-08 codes to information provided by the O*NET resource center, which in turn provides mean scores of the RIASEC dimensions (Holland, 1997) of the respective occupations (we followed the procedure of Ertl & Hartmann, 2019; see Supplement B for coding procedure). Consequently, for each participant, idealistic occupational aspirations were represented by continuous scores on Holland's (1997) six RIASEC dimensions.

Competencies. Reading and mathematical competence as well as scientific thinking were measured by standardized tests that were specifically developed for NEPS. Unfortunately, no pretest score of scientific thinking was available, therefore we used scientific literacy as an approximation. Scientific literacy was conceptualized as the use of scientific knowledge in the environmental, technological, and health contexts (Hahn et al., 2013) and was measured by a standardized test at Wave 6 that was developed for NEPS. The scientific thinking competence test entailed the areas of meta-scientific knowledge (e.g., knowledge about the scientific system), understanding of methods (e.g., knowing how knowledge is generated), and meta-scientific reflection (e.g., judging scientific processes and the generation of knowledge), which suggests an overlap between the constructs of scientific thinking and scientific literacy. For an overview of the selected items, item examples, and reliability coefficients see Table 1 and Supplement B. In addition, for detailed information about the scaling procedure see Fischer et al. (2016), NEPS (2021a) and Pohl and Carstensen (2012).

School Achievement. As an indicator for school achievement, we used end of the year grades at Wave 6 (Grade 9) and Wave 9 (Grade 12) for the school subjects English, German, and math.

Ability Beliefs. We included three constructs that represented school-related ability beliefs, namely self-concept in mathematics, German, and school. The self-concept in school represented a general belief about one's school-related abilities. Students had to rate their performance according to items such as "I get good grades in math", "I learn fast in German" or "I perform well in most of the school subjects". The questionnaire was specifically developed

for NEPS. For an overview of the selected items, item examples, and reliability coefficients see Table 1 and Supplement B.

Confounders

We adhered to the modified disjunctive cause criterion. It is one of the principles of confounder selection that was suggested by VanderWeele et al. (2020) as appropriate for outcome-wide longitudinal designs. The criterion states to "control for each covariate that is a cause of the exposure, or of the outcome, or of both … and include as a covariate any proxy for an unmeasured variable that is a common cause of both the exposure and the outcome" (VanderWeele et al., 2020, p. 441). The criterion is very inclusive concerning the selection of confounding variables (VanderWeele et al., 2020). Therefore, we controlled for variables from the following broad areas: demographics, health, personality, cognitive abilities, school achievement, family environment, and motivation (see Table 1 for detailed information about the confounders). In addition, for almost all outcomes at Grade 12 we were able to control for a pretest, or in the case of scientific thinking an approximation of the pretest, of the outcome from Grade 9. Only the outcome English school grade was without a pretest measure.

Statistical Analyses

Missing Data and Nested Data Structure

Before analyzing the data, we created multiple imputed data sets to handle missing data (for the proportion of missing data for each variable see Table 1). Because students in the NEPS study were nested within classes and schools, we conducted a multilevel multiple imputation based on the fully conditional specification approach implemented in the R package mice (van Buuren & Groothuis-Oudshoorn, 2011). To reduce model complexity, we decided to perform a two-level multiple imputation procedure and used school membership as the highest hierarchical level. We imputed $N = 100$ data sets with a number of $m = 5$ iterations. To increase predictive information, we included school cluster means to the imputation models for every continuous variable that possessed an ICC greater than .05.

All variables were imputed by predictive mean matching obtained through specifying linear mixed models (Robitzsch et al., 2021; van Buuren, 2018). Random draws were performed from the regression coefficients and random effects (van Buuren, 2018). To apply the predictive mean matching method, we used the analysis option *2l.pmm* from the mice package. The school cluster means were imputed by two-level predictive mean matching models (van Buuren, 2018). To apply that approach, we used the analysis option *2lonly.pmm* from the mice package. To

increase predictive information and restore the missing-at-random (MAR) assumption, the imputation model of every variable included all the other variables (e.g., confounders and outcomes) used in the current study.

Analysis Strategy

After data imputation, we specified multiple linear regression models for each outcome within each imputed data set. When study design (i.e., confounders, treatment, and outcomes are assessed in that exact temporal order) and control for confounding are strong, linear regression is a valid analysis option to estimate causal effects (VanderWeele et al., 2020). In the outcome-wide longitudinal design, decisions about statistical modelling are made for all outcomes simultaneously (VanderWeele et al., 2020). Consequently, all regression models consisted of the same set of independent variables. In addition, for every outcome we specified a different regression model for each treatment type (i.e., median split, no vs. any, extreme group comparison, continuous intensity, and single item treatments). Therefore, within each imputed data set, we specified $n_{outcome} * n_{treatment}$ regression models (i.e., in total 273 models, for each treatment and outcome combination).

For the outcome variables (i.e., dispositional interests, occupational aspirations, competencies, school grades, and ability beliefs) we specified multiple linear regression models. All confounding variables were included as covariates. Consequently, the multiple linear regression models consisted of the respective outcome, all confounders, and the respective treatment variable. All outcomes were z-standardized before the analysis to yield effect sizes similar to Cohen's *d* (Cohen, 1988). The continuous treatment variable (i.e., continuous intensity) was also z-standardized to yield standardized regression coefficients. The regression models were specified based on the svyglm function as implemented in the R package survey (Lumley, 2010). The results of each regression model were then pooled across the imputed data sets based on Rubin's rules (van Buuren, 2018).

Standard Error Estimation and Weighting

Because classes and schools were not the focus of our analysis, we treated them as a design nuisance (McNeish et al., 2017) and computed cluster-robust standard errors. To account for a possible underestimation of the standard errors due to the hierarchical data structure, the survey package relies on Horvitz-Thompson-type standard errors (Horvitz & Thompson, 1952) which are a generalization of the sandwich estimator (Lumley, 2010). Information about school and class membership was used from the first study wave. In addition, every multiple regression

model included the panel entry weight to make the sample representative for the study population. All information was specified within the design argument of the svyglm function from the R package survey (Lumley, 2010).

Multiple Testing Corrections and E-Values

In the current study we set the initial alpha level of the statistical tests to .05 and conducted them two-sided. In outcome-wide longitudinal designs, treatment effects are investigated on multiple outcomes simultaneously, which results in testing the identical main hypothesis multiple times. Following the recommendations of VanderWeele et al. (2020) to account for multiple testing, we therefore included the Bonferroni (see VanderWeele & Mathur, 2019) and Sidak (see Moskvina & Schmidt, 2008) corrections (i.e., number of tests were the number of outcomes, in our case 21) to adjust the initial alpha level.

In addition to the effect sizes and multiple testing metrics, it is recommended to report E-values (VanderWeele et al., 2020) as a sensitivity measure of the robustness of the causal effect of the treatment against unmeasured confounders (Ding & VanderWeele, 2016). The higher the E-value the higher unmeasured confounders have to be related to the outcome as well as the treatment, above and beyond the measured confounders, to shift the respective treatment effect to null (VanderWeele et al., 2020). The E-value is interpreted in terms of risk ratios. Risk ratios are computed when the risk that a certain event occurs in one group is compared to the risk that the event occurs in another group (e.g., risk of a medical condition in a group of males vs. females). A risk ratio is obtained by dividing the risk in one group by the risk in another group (i.e., a risk ratio of 1 indicates identical risk among groups).

In the current study we will interpret the confidence intervals of the respective E-values. If the lower bound of the confidence interval includes the value 1, it indicates that the confidence interval of the E-value includes the null. For example, the lower bound of a confidence interval of an E-value of 1.5 implies “that an unmeasured confounder that was associated with both high levels of [the treatment] and with high levels of [the outcome] by risk ratios of 1.5-fold each, above and beyond the measured covariates could suffice to shift the confidence interval to the null but weaker confounding could not” (VanderWeele et al., 2020, p. 454). This implies that the relationship between the unmeasured confounder and the treatment as well as the outcome must be higher than the respective risk ratio for the significant treatment effect to become nonsignificant when controlling for the unmeasured confounder. We computed E-values and their confidence intervals for continuous outcomes based on the approximation formula provided by Ding and VanderWeele (2016). For the approximation

formula the pooled regression coefficients and standard errors of the respective treatment effects were used.

Results

In terms of manifest re-test correlations, the outcomes were moderately stable over the 3-year time period ($.37 \leq r \leq .69$; see Table 2). In general, dispositional interests ($.48 \leq r \leq .67$), competencies ($.50 \leq r \leq .69$) and self-concepts ($.47 \leq r \leq .64$) were more stable compared to grades ($.51 \leq r \leq .52$) and occupational aspirations ($.37 \leq r \leq .49$). The effects of unstructured OST science activities operationalized with the median split treatment on all outcomes are depicted in Table 3. The results represent the mean-level differences between the groups of low (i.e., participants with mean levels below the median) and high (i.e., participants with mean levels above the median) activity engagement on each outcome after controlling for all confounding variables. As all the outcome variables were z-standardized, effects can be interpreted as differences in standard deviations. Therefore, positive values imply higher predicted mean levels on the respective outcome variables for students who had a higher engagement in unstructured OST science activities—controlling for all the confounders. Robustness checks which display the effects of all treatment operationalizations and all single activities on all outcomes can be found in Supplement C.

Effects on Dispositional Interests

Engagement in unstructured OST science activities had the strongest effect on Investigative interests. The mean-level difference between groups of low and high engagement was statistically significant with $d = 0.26$ [0.12, 0.41], implying higher predicted mean levels for students with higher engagement (see Table 3). The effect was robust against unobserved confounders. This was indicated by the lower bound of the 95% confidence interval of the E-value, which did not include the value 1. The confidence interval of the E-Value had a lower bound of 1.12, which implies that an unmeasured confounder that was associated with both high levels of unstructured OST science activities and with high levels of Investigative vocational interests by risk ratios of 1.12-fold each, above and beyond the measured covariates could suffice to shift the confidence interval to the null but weaker confounding could not. This indicates that the relationship between unmeasured confounders and the treatment as well as the outcome must be higher than the respective risk ratio (above and beyond the measured confounders) for the significant treatment effect to become nonsignificant when controlling for the unmeasured confounders.

Table 2
Descriptive Statistics and Manifest Re-Test Correlations

	M T6	*SD* T6	*N* T6	*M* T8	*SD* T8	*N* T8	*M* T9	*SD* T9	*N* T9	*r*
OST sci.: TV	-	-	-	1.99	0.91	1624	-	-	-	-
OST sci.: Books	-	-	-	1.74	0.80	1627	-	-	-	-
OST sci.: Internet	-	-	-	2.40	0.89	1625	-	-	-	-
OST sci.: Magaz.	-	-	-	1.88	0.84	1626	-	-	-	-
OST sci.: Project	-	-	-	1.44	0.75	1626	-	-	-	-
Median Treat. (No P.)	-	-	-	0.57	0.50	1620	-	-	-	-
Any Treat. (No P.)	-	-	-	0.91	0.29	1620	-	-	-	-
Extreme Treat. (No P.)	-	-	-	0.46	0.50	564	-	-	-	-
Mean Treat. (No P.)	-	-	-	2.00	0.64	1620	-	-	-	-
Median Treat.	-	-	-	0.49	0.50	1618	-	-	-	-
Any Treat.	-	-	-	0.93	0.26	1618	-	-	-	-
Extreme Treat.	-	-	-	0.45	0.50	477	-	-	-	-
Mean Treat.	-	-	-	1.89	0.55	1618	-	-	-	-
Realistic	2.88	1.00	2557	-	-	-	2.98	0.95	1747	.63
Investigative	2.82	0.97	2553	-	-	-	2.94	0.93	1746	.53
Artistic	2.51	1.01	2555	-	-	-	2.67	1.05	1749	.67
Social	3.12	0.90	2555	-	-	-	3.21	0.94	1749	.58
Enterprising	2.96	0.78	2553	-	-	-	3.16	0.82	1742	.50
Conventional	2.54	0.83	2534	-	-	-	2.72	0.84	1749	.48
R Aspirations	3.86	1.82	1972	-	-	-	3.80	1.85	1970	.48
I Aspirations	3.77	2.02	1972	-	-	-	3.84	2.06	1970	.45
A Aspirations	3.22	1.98	1972	-	-	-	2.98	1.81	1970	.45
S Aspirations	3.41	1.74	1972	-	-	-	3.71	1.79	1970	.49
E Aspirations	4.09	1.56	1972	-	-	-	4.10	1.57	1970	.37
C Aspirations	3.58	1.33	1972	-	-	-	3.57	1.30	1970	.40
Math Grade [a]	4.15	1.02	2565	-	-	-	4.38	1.09	1706	.51
German Grade [a]	4.34	0.75	2316	-	-	-	4.53	0.87	1708	.52
English Grade [a]	-	-	-	-	-	-	4.52	0.93	1629	-
Reading Comp.	1.53	1.10	2030	-	-	-	2.21	0.99	2041	.50
Math Comp.	1.87	1.15	2039	-	-	-	2.72	1.08	2112	.69
Scientific Lit.	0.18	0.88	2036	-	-	-	-	-	-	-
Scientific Think.	-	-	-	-	-	-	0.03	0.65	1149	-
Self-Con. Math	2.57	0.91	2372	-	-	-	2.50	0.98	1748	.64
Self-Con. Germ.	3.02	0.61	2381	-	-	-	2.94	0.72	1745	.47
Self-Con. School	2.94	0.59	2374	-	-	-	3.02	0.63	1742	.50

Note. Descriptive statistics are based on raw data, not imputed data. Median Treat. = Median split; Any Treat. = No vs. any treatment; Extreme Treat. = Extreme group comparison; Mean Treat. = Continuous intensity; No P. = No project group item, which indicates that the variable was computed without the project group item for OST engagement.

[a] The grades were coded in the German grading scheme of early secondary school. Grades were in the range between 1 to 6. Usually lower values indicate better grades; however, for the analysis we recoded the items so that higher values indicate better grades.

Besides Investigative interests, engagement in unstructured OST science activities had also a small effect on Realistic interests, with a mean-level difference between groups of low and high engagement of d = 0.09 [-0.05, 0.22], implying higher predicted mean levels for students with higher engagement (see Table 3). It was the third largest mean-level difference across all outcomes, besides Investigative vocational interests and Investigative occupational aspirations, according to its effect size. However, the mean-level difference was not statistically significant. In addition, because the lower bound of the 95% confidence interval of the E-value included the value 1, the effect was not robust against unobserved confounders. The remaining mean-level differences between groups of low and high engagement for Artistic, Social, Enterprising, and Conventional interests were negligible and not statistically significant. The effects reported for the median split treatment were in line with the other treatment operationalizations (see Supplement C).

Table 3

Overview of Effect Sizes, p Values, and E-Values for Median Split Treatment Effects Without Project Group Item

	d	p	α_b	α_s	E[95%CI]
Realistic	0.09 [-0.05, 0.22]	.202	ns	ns	1.38 [0.96,]
Investigative	0.26 [0.12, 0.41]	.000	*	*	1.86 [1.12,]
Artistic	0.06 [-0.07, 0.19]	.382	ns	ns	1.29 [0.94,]
Social	0.03 [-0.11, 0.18]	.649	ns	ns	1.21 [0.90,]
Enterprising	0.05 [-0.10, 0.20]	.507	ns	ns	1.27 [0.92,]
Conventional	0.06 [-0.09, 0.20]	.428	ns	ns	1.30 [0.92,]
Scientific Thinking	0.02 [-0.17, 0.20]	.876	ns	ns	1.13 [0.85,]
Math Comp.	0.01 [-0.10, 0.11]	.894	ns	ns	1.09 [0.91,]
Read. Comp.	0.03 [-0.08, 0.15]	.569	ns	ns	1.21 [0.93,]
Math Grade	0.05 [-0.10, 0.19]	.535	ns	ns	1.25 [0.91,]
German Grade	0.00 [-0.16, 0.16]	.979	ns	ns	1.05 [0.87,]
English Grade	0.05 [-0.09, 0.19]	.503	ns	ns	1.26 [0.92,]
Self-Concept Math	0.00 [-0.12, 0.12]	.981	ns	ns	1.04 [0.90,]
Self-Concept Germ.	-0.04 [-0.19, 0.11]	.612	ns	ns	1.23 [0.90,]
Self-Concept School	-0.01 [-0.15, 0.13]	.873	ns	ns	1.11 [0.89,]
Realistic Asp.	0.00 [-0.12, 0.12]	.952	ns	ns	1.06 [0.90,]
Investigative Asp.	0.11 [-0.02, 0.24]	.108	ns	ns	1.44 [0.98,]
Artistic Asp.	-0.03 [-0.14, 0.09]	.650	ns	ns	1.18 [0.92,]
Social Asp.	0.00 [-0.12, 0.12]	.999	ns	ns	1.01 [0.90,]
Enterprising Asp.	-0.09 [-0.23, 0.04]	.163	ns	ns	1.40 [0.97,]
Conventional Asp.	0.06 [-0.07, 0.19]	.355	ns	ns	1.30 [0.94,]

Note. d = depicts standardized mean-level differences in standard deviations; p = depicts the p value of the respective effect size; α_b = depicts if the respective effect size is statistically significant at the Bonferroni adjusted alpha level; α_s = depicts if the respective effect size is statistically significant at the Sidak adjusted alpha level; E[95%CI] = depicts the E-value of the respective effect sizes and its lower bound of the 95% confidence interval; ns = nonsignificant; * = significant.

Effects on Occupational Aspirations

Engagement in unstructured OST science activities had the strongest effect on Investigative occupational aspirations. It was the second highest effect size across all outcomes. There was a mean-level difference between groups of low and high engagement with $d = 0.11$ [-0.02, 0.24], implying higher predicted mean levels for students with higher engagement (see Table 3). However, the effect was not statistically significant and not robust against unobserved confounders according to the confidence interval of its E-value. Although the effect size for the mean difference on Investigative occupational aspirations was higher for the extreme group comparison operationalization, with $d = 0.26$ [-0.02, 0.53], it was still not statistically significant and not robust against unobserved confounders (see Supplement C). The remaining mean-level differences between groups of low and high engagement for Realistic, Artistic, Social, Enterprising, and Conventional occupational aspirations were not statistically significant as well. The effects reported for the median split treatment were in line with the other treatment operationalizations (see Supplement C).

Effects on Competencies

The effects of engagement in unstructured OST science activities on reading competencies, math competencies and scientific thinking were negligible. Mean-level differences between groups of low and high engagement on all variables were below $d = 0.03$ and not statistically significant. In addition, according to the confidence intervals of the E-values, all the effects were not robust against unobserved confounders. The effects reported for the median split treatment were in line with other treatment operationalizations (see Supplement C).

Effects on School Achievement

The effects of engagement in unstructured OST science activities on grades in German, math and English were negligible. Mean-level differences between groups of low and high engagement on all variables were below $d = 0.05$ and not statistically significant. In addition, according to the confidence intervals of the E-values, all the effects were not robust against unobserved confounders. According to the robustness checks, the result for the no vs. any treatment operationalization was slightly different for grades in English. There was a mean-level difference between groups of no and any engagement with $d = 0.27$ [0.05, 0.51], implying higher predicted mean levels in English grades for students with any engagement (see

Supplement C). However, it is important to note that English grade was the only variable without a pretest.

Effects on Ability Beliefs

The effects of engagement in unstructured OST science activities on ability beliefs in the subjects German and math as well as ability beliefs about general school performance were negligible. Mean-level differences between groups of low and high engagement on all variables were below $d = 0.05$ and not statistically significant. In addition, according to the confidence intervals of the E-values, all the effects were not robust against unobserved confounders. The effects reported for the median split treatment were in line with the other treatment operationalizations (see Supplement C).

Discussion

The present study investigated the influence of unstructured OST science activities on a broad range of outcomes, based on an outcome-wide longitudinal design for causal inference (VanderWeele et al., 2020). A strength of the study is the estimation of causal effects, its longitudinal research design, and its comprehensive control for confounding. It thereby overcomes methodological issues of previous research and adequately distinguishes self-selection effects from effects of unstructured OST science activities.

It is important to consider that we investigated relatively stable constructs, which typically change only slightly over multiple years (e.g., changes in vocational interests from age 14 to 18, $0.00 < d < 0.18$; Hoff et al., 2018). In addition, the development of STEM-related constructs is influenced not only by unstructured OST science activities, but also by a variety of other individual and environmental factors (e.g., school, family, or peer environments; Tal & Dierking, 2014). The multiplicity of such causes suggests that when only one of these causes is investigated, small effect sizes should be expected (Götz et al., 2021). We therefore consider even smaller effect sizes (around $d = 0.10$) as noteworthy.

In terms of effect sizes, unstructured OST science activities had the biggest impact on Investigative vocational interests. For Investigative vocational interests, the effect size was substantial with $d = 0.26$, robust against unmeasured confounders, and higher than those reported in previous studies (e.g., for STEM interest $d = 0.06$ to $d = 0.09$; Dou et al., 2019). For Realistic vocational interests and Investigative occupational aspirations, the effect sizes were smaller, with $d = 0.09$ and $d = 0.11$, respectively. Effect sizes on the ability-related constructs competencies, school achievement, and ability beliefs were the smallest and between $d = -0.04$ and $d = 0.05$. The effect sizes suggest that unstructured OST science activities have a substantial influence on Investigative vocational interests, but less of an effect on other interest dimensions, occupational aspirations, competencies, school achievement, and ability beliefs.

The Impact of Unstructured OST Science Activities on Interests and Aspirations

We used Holland's (1997) RIASEC model as a framework for dispositional interests. According to the model, the dimension Investigative represents interest in science-related activities, such as experimenting with liquids and materials, observing phenomena in nature, or solving complex problems. According to our findings, unstructured OST science activities had the biggest influence on the development of Investigative vocational interests. This indicates that adolescents who engage more in unstructured OST science activities, such as reading a popular science book, watching a science TV show, or researching on the internet about science,

develop more pronounced Investigative vocational interests, compared to adolescents who engage less in unstructured OST science activities.

Unstructured OST science activities could also influence Realistic interests, as they are closely linked to interest in engineering (e.g., technological, mechanical, and physical activities; Holland, 1997). However, the smaller effect size of $d = 0.09$ suggests that there was an attenuated influence of unstructured OST science activities on the development of Realistic interests. One explanation could be that the activities that were measured in the current study only marginally initiated situational interest in Realistic activities. The activity items focused more on general engagement in natural sciences (e.g., watching TV shows about natural sciences) and less on specific Realistic-related activities, such as working and tinkering with machines or repairing mechanical appliances. It could be assumed that activities from the field of mechanical engineering would have a bigger influence on the development of Realistic interests. However, further studies are needed to investigate that assumption.

Our findings support the assumption that unstructured OST science activities influence the development of science-related vocational interests. Previous studies that investigated the influence of unstructured OST science activities on STEM-related interests (see Dou et al., 2019; Uitto et al., 2006) reported similar results. This is also in line with considerations of interest development theories (Renninger & Hidi, 2011; Su et al., 2019). However, it is important to note that we find the most substantial and robust effect size in areas with the most content overlap, namely Investigative vocational interests. Effect sizes on Realistic interests were essentially smaller. Compared to previous studies that examined the influence of environments on the development of vocational interests (e.g., Meir & Navon, 1992; Schultz et al., 2017), we used a study design that was able to account for selection processes. Since we controlled for a variety of relevant variables temporally prior to engagement in unstructured OST science activities, we thus minimized the influence of self-selection to robustly estimate the effects of engagement.

The effect size of unstructured OST science activities on Investigative occupational aspirations was similar in magnitude to the effect size on Realistic vocational interests (both around $d = 0.10$), but much smaller than Investigative vocational interests. Although occupational aspirations and vocational interests overlap (Gottfredson, 1981), in contrast to vocational interests, occupational aspirations could depend more on objective occupational features, such as gross income, job security, and availability (Gottfredson, 1981). Unstructured OST science activities might therefore be less impactful in changing Investigative occupational aspirations in comparison to Investigative vocational interests.

Finally, it is important to note that dispositional interests such as Investigative vocational interests are strong predictors for educational decisions, such as choosing a STEM study major (Päßler & Hell, 2012; Wille et al., 2020). In addition, they also predict occupational decisions, such as choosing a STEM occupation (Holland, 1997; Lent et al., 1994); a wide range of further positive life outcomes (Stoll et al., 2017); and academic school track choices (Usslepp et al., 2020). The engagement in unstructured OST science activities could consequently be beneficial for the supply of STEM graduates, academic career pathways, and early career success. However, these assumptions need more empirical insights and should be targeted by future research.

The Impact of Unstructured OST Science Activities on Ability-Related Constructs

Based on assumptions derived from the theory of trait complexes (see Ackerman & Heggestad, 1997), we suggested that unstructured OST science activities could influence a multitude of variables, including ability-related constructs. In addition, everyday experiences such as unstructured OST science activities are assumed to initiate learning processes in STEM (Tal & Dierking, 2014). However, in comparison to Investigative vocational interests (i.e., d = 0.26), we found relatively small effect sizes (i.e., $-0.04 < d < 0.05$) of unstructured OST science activities on the development of competencies, school grades, and ability beliefs in our main analysis.

Regarding the relatively small effect sizes on competencies and school grades, it could be argued that unstructured OST science activities did not adequately initiate long-lasting and sustainable learning processes. Unstructured OST science activities are primarily leisure-related activities (Dabney et al., 2012; Dierking et al., 2003) that are, first and foremost, chosen to satisfy leisure-related needs, such as relaxation or having fun. However, to initiate learning processes that substantially accumulate knowledge or improve competencies as well as school grades, often a certain degree of effort is needed (Schneider & Stern, 2010; Stewart, 2008). For example, Schneider and Stern (2010) emphasize that the evolvement of knowledge requires hard work that contains a high degree of effort and practice. Time and effort are therefore seen as crucial determinants of knowledge development (Ericsson et al., 1993). In addition, as standardized competence tests and school grades usually assess contents that are part of the school curriculum (NEPS, 2021a), the relatively small impact on ability-related constructs could also be a result of different learning outcomes. Learning outcomes of unstructured OST science activities might differ from learning outcomes within a school curriculum, resulting in relatively small effect sizes on ability-related constructs.

We also reported relatively small effect sizes for unstructured OST science activities on ability beliefs. Ability beliefs are assumed to develop especially in situations where it is possible to compare one's abilities to those of others (Marsh et al., 2014). Such external frame of reference effects often occur in environments where abilities are evaluated due to the assignment of grades, for example in school (Marsh et al., 2014). However, as unstructured OST science activities are mostly leisure oriented, they usually provide little information and feedback about the actual performance. Furthermore, as there is often no external authority (e.g., a teacher) who judges the performance of a task, one's abilities are often self-assessed and not directly comparable to those of others. As unstructured OST science activities are primarily leisure oriented and not performance oriented, ability comparisons may not explicitly occur during these situations and students may not be aware that they are doing activities that could enhance their abilities. Furthermore, as ability beliefs of the current study focused on the school context, adolescents might not relate their engagement in unstructured OST science activities to these ability beliefs.

However, it could also be argued that the investigated period (i.e., 3-year time span) and dosage were not long and strong enough for unstructured OST science activities to initiate long-lasting learning processes, which could result in the formation of ability-related constructs that are in line with a science/math trait complex. For example, the extreme group operationalization compared groups of very low (i.e., one standard deviation below the sample mean) and very high (i.e., one standard deviation above the sample mean) engagement. Descriptively, effect sizes of the extreme group operationalization were slightly more in line with a science/math trait complex, as indicated by slightly larger effects on Realistic and Investigative interests as well as school grades in math. However, the pattern of effect sizes was mostly the same as that of the main analysis, with the largest effect sizes for Investigative vocational interests, smaller effect sizes for Realistic vocational interests and Investigative occupational aspirations, and relatively small effect sizes for ability-related constructs.

Limitations and Future Outlook

The current study was based on a nationally representative large-scale longitudinal data set that allowed controlling for a comprehensive set of confounding variables, which could influence both the treatment (i.e., unstructured OST science activities) and the outcomes, while investigating the influence of unstructured OST science activities. Due to the broad range of constructs that are available in the NEPS study, we investigated the broad influence of

unstructured OST science activities on several STEM and non-STEM related outcomes. We therefore provided evidence that is relevant to a broad audience of researchers and practitioners.

We found a few effects that were considered as noteworthy in terms of their effect size (e.g., *d* around .10), but were not statistically significant at the chosen significance level. This could indicate that the statistical power of the respective significance tests was somewhat too low. The investigation of small effect sizes usually requires large sample sizes (Cohen, 1988) to obtain adequate statistical power. However, it could be assumed that the current study might still not have achieved the required size despite its large sample of $N = 2{,}655$ participants. As our sample consisted of a clustered structure, we had to impute multilevel data, adjust standard errors for school and class clustering, and had to account for multiple testing by adjusting the chosen significance level. In combination with the small effect sizes, these necessary adjustments could have further reduced statistical power. Future studies that investigate the influence of unstructured OST science activities should therefore consider statistical power and implement designs with even larger sample sizes.

It is important to note that participants of the NEPS study were selected from different states and school forms in Germany. In Germany, school types (e.g., lower, middle, higher, or multitrack schools) or quality of schooling can differ from state to state. This implies that for students from different states and school types, school grades are probably difficult to compare. To account for that characteristic and to maximize comparability we controlled for the type of school form in all the specified regression models. In addition, because we controlled for a comprehensive set of other confounding variables in the regression models—also in the form of cluster information on the school level—the effects of different school types should be minimized and not bias our results. The results of the current study should be replicated based on samples from other countries, such as the United States, that possess a different educational system in comparison to Germany.

Unfortunately, the current study provides no viable information about the influence of unstructured OST science activities on the development of grades in science subjects such as physics, chemistry, and biology. Although we aimed to investigate that specific research question, grades in physics, chemistry, and biology were only available with a high amount of missing data (i.e., around 94%) in the NEPS data set. We decided to not impute data with that high number of missing values. Future studies about unstructured OST science activities should therefore include science-related grades beyond math to get a more differentiated look on the effects on school achievement. Unstructured OST science activities might have a bigger impact on school grades with more content overlap, such as science. The current study however

included a standardized competence test that measured scientific thinking in a broad sense, which comprises meta-scientific knowledge (e.g., knowledge about the scientific system), understanding of methods in science (e.g., knowing how knowledge is generated), and meta-scientific reflection (e.g., judging scientific processes and the generation of knowledge; NEPS, 2021a). Although not tied to a specific school subject, this achievement test is a good proxy for basic scientific competencies that are relevant in all science subjects.

Conclusion

The present study investigated the influence of unstructured OST science activities on a broad range of outcomes. We adhered to the template of an outcome-wide longitudinal design for causal inference (VanderWeele et al., 2020) by investigating the influence of a single treatment on various outcomes simultaneously and controlling for a comprehensive set of confounders. The results show that unstructured OST science activities have a robust influence on Investigative vocational interests, but less of an influence on other interest dimensions, occupational aspirations, competencies, school achievement, and ability beliefs. The findings imply that unstructured OST science activities can influence the development of dispositional interest in STEM.

References

Ackerman, P. L. (1997). Personality, self-concept, interests, and intelligence: Which construct doesn't fit? *Journal of Personality*, *65*(2), 171–205. https://doi.org/10.1111/j.1467-6494.1997.tb00952.x

Ackerman, P. L., & Beier, M. E. (2003). Intelligence, personality, and interests in the career choice process. *Journal of Career Assessment*, *11*(2), 205–218. https://doi.org/10.1177/1069072703011002006

Ackerman, P. L., Bowen, K. R., Beier, M. E., & Kanfer, R. (2001). Determinants of individual differences and gender differences in knowledge. *Journal of Educational Psychology*, *93*(4), 797–825. https://doi.org/10.1037/0022-0663.93.4.797

Ackerman, P. L., Chamorro-Premuzic, T., & Furnham, A. (2011). Trait complexes and academic achievement: Old and new ways of examining personality in educational contexts. *British Journal of Educational Psychology*, *81*(1), 27–40. https://doi.org/10.1348/000709910X522564

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, *121*(2), 219–245. https://doi.org/10.1037/0033-2909.121.2.219

Aguilera, D., & Perales-Palacios, F. J. (2020). What effects do didactic interventions have on students' attitudes towards science? A meta-analysis. *Research in Science Education*, *50*(2), 573–597. https://doi.org/10.1007/s11165-018-9702-2

Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J. A., Fiedler, K., Fiedler, S., Funder, D. C., Kliegl, R., Nosek, B. A., Perugini, M., Roberts, B. W., Schmitt, M., Van Aken, M. A. G., Weber, H., & Wicherts, J. M. (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality*, *27*(2), 108–119. https://doi.org/10.1002/per.1919

Azevedo, R. (2015). Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational Psychologist*, *50*(1), 84–94. https://doi.org/10.1080/00461520.2015.1004069

Bergmann, C., & Eder, F. (2005). *Allgemeiner Interessen-Struktur-Test mit Umwelt-Struktur-Test (UST-R)—Revision [General Interest Structure Test and Environmental Structure Test—Revision]*. Beltz.

Bernacki, M. L., Vosicka, L., & Utz, J. C. (2020). Can a brief, digital skill training intervention help undergraduates “learn to learn” and improve their STEM achievement? *Journal of Educational Psychology*, *112*(4), 765–781. https://doi.org/10.1037/edu0000405

Blossfeld, H.-P., Roßbach, H.-G., & von Maurice, J. (Eds.). (2011). *Education as a lifelong process – The German National Educational Panel Study (NEPS)*. Zeitschrift für Erziehungswissenschaft: Sonderheft 14.

Bureau of Labor Statistics. (2021). *Crosswalks between the 2010 SOC and systems used by other Federal and international statistical agencies*. Bureau of Labor Statistics. https://www.bls.gov/soc/soccrosswalks.htm

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). L. Erlbaum Associates.

Dabney, K. P., Tai, R. H., Almarode, J. T., Miller-Friedmann, J. L., Sonnert, G., Sadler, P. M., & Hazari, Z. (2012). Out-of-School time science activities and their association with career interest in STEM. *International Journal of Science Education, Part B*, *2*(1), 63–79. https://doi.org/10.1080/21548455.2011.629455

Dabney, K. P., Tai, R. H., & Scott, M. R. (2016). Informal science: Family education, experiences, and initial interest in science. *International Journal of Science Education, Part B*, *6*(3), 263–282. https://doi.org/10.1080/21548455.2015.1058990

Dierking, L. D., Falk, J. H., Rennie, L., Anderson, D., & Ellenbogen, K. (2003). Policy statement of the “informal science education” ad hoc committee. *Journal of Research in Science Teaching*, *40*(2), 108–111. https://doi.org/10.1002/tea.10066

Ding, P., & VanderWeele, T. J. (2016). Sensitivity analysis without assumptions. *Epidemiology*, *27*(3), 368–377. https://doi.org/10.1097/EDE.0000000000000457

Dou, R., Hazari, Z., Dabney, K., Sonnert, G., & Sadler, P. (2019). Early informal STEM experiences and STEM identity: The importance of talking science. *Science Education*, *103*(3), 623–637. https://doi.org/10.1002/sce.21499

Eccles, J., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (5th ed., pp. 1017–1095). John Wiley & Sons, Inc.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406.

Ertl, B., & Hartmann, F. G. (2019). The interest profiles and interest congruence of male and female students in STEM and non-STEM fields. *Frontiers in Psychology*, *10*, 897. https://doi.org/10.3389/fpsyg.2019.00897

Falk, J. H., & Storksdieck, M. (2009). Science learning in a leisure setting. *Journal of Research in Science Teaching*, *47*(2), 194–212. https://doi.org/10.1002/tea.20319

Fenster, C. B., & Gropp, R. E. (2020). On the importance of science to society: A call for government action. *BioScience*, *70*(5), 371–371. https://doi.org/10.1093/biosci/biaa047

Fischer, L., Rohm, T., Gnambs, T., & Carstensen, C. H. (2016). Linking the data of the competence tests in NEPS (NEPS Survey Paper No. 1). *Leibniz Institute for Educational Trajectories, National Educational Panel Study*. https://doi.org/10.5157/NEPS:SP01:1.0

Friedman, L. B., Margolin, J., Swanlund, A., Dhillon, S., & Liu, F. (2017). *Enhancing middle school science lessons with playground activities: A study of the impact of playground physics*. American Institutes for Research. https://files.eric.ed.gov/fulltext/ED574773.pdf

Ganzeboom, H. B. G., & Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 international standard classification of occupations. *Social Science Research*, *25*(3), 201–239. https://doi.org/10.1006/ssre.1996.0010

Golle, J., Zettler, I., Rose, N., Trautwein, U., Hasselhorn, M., & Nagengast, B. (2018). Effectiveness of a "grass roots" statewide enrichment program for gifted elementary school children. *Journal of Research on Educational Effectiveness*, *11*(3), 375–408. https://doi.org/10.1080/19345747.2017.1402396

Gottfredson, L. S. (1981). Circumscription and compromise: A developmental theory of occupational aspirations. *Journal of Counseling Psychology*, *28*(6), 545–579. https://doi.org/10.1037/0022-0167.28.6.545

Götz, F. M., Gosling, S. D., & Rentfrow, P. J. (2021). Small effects: The indispensable foundation for a cumulative psychological science. *Perspectives on Psychological Science*, 1–11. https://doi.org/10.1177/1745691620984483

Guzzetti, B. J., Snyder, T. E., Glass, G. V., & Gamas, W. S. (1993). Promoting conceptual change in science: A comparative meta-analysis of instructional interventions from reading education and science education. *Reading Research Quarterly*, *28*(2), 116–159. https://doi.org/10.2307/747886

Hahn, I., Schöps, K., Rönnebeck, S., Martensen, M., Hansen, S., Saß, S., Dalehefte, I. M., & Prenzel, M. (2013). Assessing scientific literacy over the lifespan—A description of the NEPS science framework and the test development. *Journal for Educational Research Online*, *5*(2), 110–138.

Hein, G. (2009). Learning science in informal environments: People, places, and pursuits. *Museums & Social Issues*, *4*(1), 113–124. https://doi.org/10.1179/msi.2009.4.1.113

Henriksen, E. K., Jensen, F., & Sjaastad, J. (2015). The role of out-of-school experiences and targeted recruitment efforts in Norwegian science and technology students' educational choice. *International Journal of Science Education, Part B*, *5*(3), 203–222. https://doi.org/10.1080/21548455.2014.900585

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

Hirschenhauser, K., Frigerio, D., Leithinger, V., Schenkenfelder, I., & Neuböck-Hubinger, B. (2019). Primary pupils, science and a model bird species: Evidence for the efficacy of extracurricular science education. *PLOS ONE*, *14*(7), 1–21. https://doi.org/10.1371/journal.pone.0220635

Hoff, K. A., Briley, D. A., Wee, C. J. M., & Rounds, J. (2018). Normative changes in interests from adolescence to adulthood: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *144*(4), 426–451. https://doi.org/10.1037/bul0000140

Hoff, K. A., Song, Q. C., Einarsdóttir, S., Briley, D. A., & Rounds, J. (2020). Developmental structure of personality and interests: A four-wave, 8-year longitudinal study. *Journal of Personality and Social Psychology*, *118*(5), 1044–1064. https://doi.org/10.1037/pspp0000228

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Psychological Assessment Resources.

Horvitz, D. G., & Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, *47*(260), 663–685. https://doi.org/10.1080/01621459.1952.10483446

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, *45*(1), 79–122.

Lumley, T. S. (2010). *Complex surveys: A guide to analysis using R.* John Wiley and Sons.

Maltese, A. V., & Tai, R. H. (2010). Eyeballs in the fridge: Sources of early interest in science. *International Journal of Science Education*, *32*(5), 669–685. https://doi.org/10.1080/09500690902792385

Marsh, H. W., Kuyper, H., Seaton, M., Parker, P. D., Morin, A. J. S., Möller, J., & Abduljabbar, A. S. (2014). Dimensional comparison theory: An extension of the internal/external frame of reference effect on academic self-concept formation. *Contemporary Educational Psychology*, *39*(4), 326–341. https://doi.org/10.1016/j.cedpsych.2014.08.003

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, *23*(3), 412–433. https://doi.org/10.1037/met0000144

McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods*, *22*(1), 114–140. https://doi.org/10.1037/met0000078

Meda, S. A., Gueorguieva, R. V., Pittman, B., Rosen, R. R., Aslanzadeh, F., Tennen, H., Leen, S., Hawkins, K., Raskin, S., Wood, R. M., Austad, C. S., Dager, A., Fallahi, C., & Pearlson, G. D. (2017). Longitudinal influence of alcohol and marijuana use on academic performance in college students. *PLOS ONE*, *12*(3), e0172213. https://doi.org/10.1371/journal.pone.0172213

Meir, E. I., & Navon, M. (1992). A longitudinal examination of congruence hypotheses. *Journal of Vocational Behavior*, *41*(1), 35–47. https://doi.org/10.1016/0001-8791(92)90037-Z

Moskvina, V., & Schmidt, K. M. (2008). On multiple-testing correction in genome-wide association studies. *Genetic Epidemiology*, *32*(6), 567–573. https://doi.org/10.1002/gepi.20331

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

NEPS. (2021a, March 20). *National Educational Panel Study (NEPS): Starting Cohort Grade 5. Documentation of Starting Cohort Grade 5 (SC3).* https://www.neps-data.de/Data-Center/Data-and-Documentation/Start-Cohort-Grade-5/Documentation

NEPS. (2021b, March 20). *National Educational Panel Study (NEPS): Starting Cohort Grade 5. Study Overview*. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/10-0-0/SC3_Study_Overview_W1-10_en.pdf

Noonan, R. (2017). *STEM jobs: 2017 update (ESA Issue Brief #02-17)*. https://files.eric.ed.gov/fulltext/ED594354.pdf

OECD. (2008). *Encouraging student interest in science and technology studies*. OECD Publishing. https://www.oecd-ilibrary.org/content/component/9789264040892-2-en

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

O*NET. (2021, February 14). *O*NET resource center—Interest table.* https://www.onetcenter.org/dictionary/22.2/excel/interests.html

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), 943–951. https://doi.org/10.1126/science.aac4716

Päßler, K., & Hell, B. (2012). Do interests and cognitive abilities help explain college major choice equally well for women and men? *Journal of Career Assessment*, *20*(4), 479–496. https://doi.org/10.1177/1069072712450009

Pohl, S., & Carstensen, C. H. (2012). *NEPS technical report—Scaling the data of the competence tests* (NEPS Working Paper No. 14). Leibniz Institute for Educational Trajectories, National Educational Panel Study. https://fis.uni-bamberg.de/handle/uniba/1751

Renninger, K. A., & Hidi, S. (2011). Revisiting the conceptualization, measurement, and generation of interest. *Educational Psychologist*, *46*(3), 168–184. https://doi.org/10.1080/00461520.2011.587723

Roberts, B. W., & Wood, D. (2006). Personality development in the context of the neo-socioanalytic model of personality. In D. K. Mroczek & T. D. Little (Eds.), *Handbook of personality development* (pp. 11–39). Lawrence Erlbaum Associates Publishers. https://doi.org/10.4324/9781315805610.ch2

Robitzsch, A., Grund, S., & Henke, T. (2021). *miceadds: Some additional multiple imputation functions, especially for "mice"* (3.11-6) [Computer software]. https://CRAN.R-project.org/package=miceadds

Rounds, J., & Su, R. (2014). The nature and power of interests. *Current Directions in Psychological Science*, *23*(2), 98–103. https://doi.org/10.1177/0963721414522812

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Schneider, M., & Stern, E. (2010). The cognitive perspective on learning: Ten cornerstone findings. In H. Dumont, D. Istance, & F. Benavides (Eds.), *The nature of learning: Using research to inspire practice* (pp. 69–90). OECD publishing.

Schultz, L. H., Connolly, J. J., Garrison, S. M., Leveille, M. M., & Jackson, J. J. (2017). Vocational interests across 20 years of adulthood: Stability, change, and the role of work experiences. *Journal of Research in Personality*, *71*, 46–56. https://doi.org/10.1016/j.jrp.2017.08.010

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist*, *50*(1), 1–13. https://doi.org/10.1080/00461520.2014.1002924

Stewart, E. B. (2008). School structural characteristics, student effort, peer associations, and parental involvement: The influence of school- and individual-level factors on academic achievement. *Education and Urban Society*, *40*(2), 179–204. https://doi.org/10.1177/0013124507304167

Stoll, G., Rieger, S., Lüdtke, O., Nagengast, B., Trautwein, U., & Roberts, B. W. (2017). Vocational interests assessed at the end of high school predict life outcomes assessed 10 years later over and above IQ and Big Five personality traits. *Journal of Personality and Social Psychology*, *113*(1), 167–184. https://doi.org/10.1037/pspp0000117

Su, R., Stoll, G., & Rounds, J. (2019). The nature of interests: Toward a unifying theory of trait-situation interest dynamics. In C. D. Nye & J. Rounds (Eds.), *Vocational interests: Rethinking their role in understanding workplace behavior and practice* (pp. 11–38). Taylor & Francis/Routledge. https://doi.org/10.4324/9781315678924-2

Tal, T., & Dierking, L. D. (2014). Learning science in everyday life. *Journal of Research in Science Teaching*, *51*(3), 251–259. https://doi.org/10.1002/tea.21142

Therrien, W. J., Taylor, J. C., Hosp, J. L., Kaldenberg, E. R., & Gorsh, J. (2011). Science instruction for students with learning disabilities: A meta-analysis. *Learning Disabilities Research & Practice*, *26*(4), 188–203. https://doi.org/10.1111/j.1540-5826.2011.00340.x

Tracey, T. J. G., & Ward, C. C. (1998). The structure of children's interests and competence perceptions. *Journal of Counseling Psychology*, *45*(3), 290–303. https://doi.org/10.1037/0022-0167.45.3.290

Uitto, A., Juuti, K., Lavonen, J., & Meisalo, V. (2006). Students' interest in biology and their out-of-school experiences. *Journal of Biological Education*, *40*(3), 124–129. https://doi.org/10.1080/00219266.2006.9656029

Usslepp, N., Hübner, N., Stoll, G., Spengler, M., Trautwein, U., & Nagengast, B. (2020). RIASEC interests and the Big Five personality traits matter for life success—But do they already matter for educational track choices? *Journal of Personality*, *88*(5), 1007–1024. https://doi.org/10.1111/jopy.12547

van Buuren, S. (2018). *Flexible imputation of missing data* (2nd ed.). Taylor & Francis. https://stefvanbuuren.name/fimd/sec-whyandwhen.html

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

VanderWeele, T. J., & Mathur, M. B. (2019). Some desirable properties of the Bonferroni correction: Is the Bonferroni correction really so bad? *American Journal of Epidemiology*, *188*(3), 617–618. https://doi.org/10.1093/aje/kwy250

VanderWeele, T. J., Mathur, M. B., & Chen, Y. (2020). Outcome-wide longitudinal designs for causal inference: A new template for empirical studies. *Statistical Science*, *35*(3), 437–466. https://doi.org/10.1214/19-STS728

von Maurice, J. (2006). *ICA-D. Deutschsprachige Version des Inventory of Children's Activities – Revised (ICA-R, Tracey & Ward, 1998)* [Unpublished instrument]. Otto-Friedrich-Universität Bamberg.

What Works Clearinghouse, Institute of Education Sciences, & U.S. Department of Education. (2021, May). *Science Teachers Learning through Lesson Analysis (STeLLA®)*. https://ies.ed.gov/ncee/wwc/Docs/InterventionReports/wwc_PS_STeLLA_IR_may2021.pdf

Wicherts, J. M., Veldkamp, C. L. S., Augusteijn, H. E. M., Bakker, M., van Aert, R. C. M., & van Assen, M. A. L. M. (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid p-hacking. *Frontiers in Psychology*, *7*, 1832. https://doi.org/10.3389/fpsyg.2016.01832

Wille, E., Stoll, G., Gfrörer, T., Cambria, J., Nagengast, B., & Trautwein, U. (2020). It takes two: Expectancy-value constructs and vocational interests jointly predict STEM major choices. *Contemporary Educational Psychology*, *61*, 101858. https://doi.org/10.1016/j.cedpsych.2020.101858

Wrzus, C., & Roberts, B. W. (2017). Processes of personality development in adulthood: The TESSERA framework. *Personality and Social Psychology Review*, *21*(3), 253–277. https://doi.org/10.1177/1088868316652279

Supplement Material

Supplement A: Sample Composition

The sampling procedure of the Starting Cohort Grade 5 study comprised the representative selection of students in regular and special schools in Germany, starting in fifth grade. Based on a two-stage sampling approach, first, schools were randomly selected (i.e., sampling on first stage), and then second, two classes within each school were randomly selected (i.e., sampling on second stage). All students who were in these final two selected classes per school were asked to participate in the study. After that, the participants were followed throughout secondary school within the respective educational institutions and questioned in annual waves. When participants left their respective school or the general school system altogether, they were tracked and questioned individually. To date the Starting Cohort Grade 5 study encompasses ten waves (for further information see NEPS, 2021b).

The final subsample that we selected for the current study comprised N = 2,655 participants (51% were female) from 374 classes and 201 schools. 2.5% of the participants indicated that they were not born in Germany, whereas between 14.5% to 15.1% indicated that one of their parents was not born in Germany. At the sixth wave there is information for 79.8% of the participants about their current school type. 4.8% attended the lowest school track, 23.0% attended the intermediate school track, 57.9% attended the highest school track and 13.7% attended schools with a different track system (i.e., multitrack schools). For 0.6% of the participants there was no information about which school type they attended.

Table A1
Intraclass Correlations of All Outcomes

Cluster Variable	Interests	T1	T2
Class	Realistic	.02	.03
	Investigative	.03	.01
	Artistic	.02	.03
	Social	.04	.01
	Enterprising	.02	.00
	Conventional	.03	.08
	Realistic Aspirations	.06	.03
	Investigative Aspirations	.08	.05
	Artistic Aspirations	.03	.02
	Social Aspirations	.04	.02
	Enterprising Aspirations	.04	.02
	Conventional Aspirations	.04	.03
	Math Grade	.07	.07
	German Grade	.18	.10
	English Grade	-	.14
	Reading Competencies	.31	.22
	Math Competencies	.40	.23
	Scientific Literacy	.32	-
	Scientific Thinking	-	.18
	Self-Concept Math	.01	.00
	Self-Concept German	.07	.02
	Self-Concept School	.00	.03
School	Realistic	.02	.03
	Investigative	.03	.02
	Artistic	.01	.02
	Social	.03	.01
	Enterprising	.02	.00
	Conventional	.02	.03
	Realistic Aspirations	.03	.02
	Investigative Aspirations	.05	.03
	Artistic Aspirations	.04	.03
	Social Aspirations	.04	.01
	Enterprising Aspirations	.01	.01
	Conventional Aspirations	.04	.04
	Math Grade	.05	.05
	German Grade	.17	.08
	English Grade	-	.11
	Reading Competencies	.29	.22
	Math Competencies	.39	.22
	Scientific Literacy	.29	-
	Scientific Thinking	-	.18
	Self-Concept Math	.01	.01
	Self-Concept German	.08	.03
	Self-Concept School	.00	.03

Note. ICC's are based on raw data, not imputed data.

Supplement B: Coding Procedure & Item Wording

Idealistic occupational aspirations were stated by the participants in an open-ended question format. The participants had to answer the following question, “Imagine you had all the opportunities to become what you want. What would be your ideal occupation?” Due to reasons of data protection, the actual answers of the participants were anonymized and not available for download. Therefore, we used the International Standard Classification of Occupations (ISCO-08) codes that were provided. To obtain information about the content type of the occupational aspirations (i.e., whether they are STEM related or unrelated), we matched the ISCO-08 codes to occupational information provided by the O*NET resource center, which in turn provides mean scores of the RIASEC types (Holland, 1997) of the respective occupations. The approach made it possible to assign information about RIASEC mean levels to almost every occupational aspiration in the current sample, specifically 92.88% of the ISCO-08 codes at Wave 6 and 91.30% at Wave 9.

For the matching procedure, we followed the approach of Ertl and Hartmann (2019). First, we aggregated the RIASEC mean levels of the O*NET interest table (O*NET, 2021) based on the first six digits of the O*NET-SOC code. This procedure had to be done because there are more O*NET-SOC codes than ISCO-08 codes. Second, based on the information provided by the ISCO-08 to O*NET-SOC 2010 crosswalk table of the Bureau of Labor Statistics (2021) we matched the O*NET-SOC codes to the respective ISCO-08 codes. In some cases, when the crosswalk provided multiple O*NET-SOC codes for a single ISCO-08 code, we aggregated the RIASEC mean levels within each ISCO-08 code. The coding procedure made it possible to generate information about the mean levels of the six RIASEC dimensions for almost every idealistic occupational aspiration. Consequently, for almost every participant, idealistic occupational aspirations were represented by Holland’s (1997) RIASEC dimensions.

Table B1
Wording of the Items, Their Item Strain, and Their NEPS Label

Construct	Item Strain or Item Information	Item Example	Likert Scale Wordings	NEPS Label
Treatment				
OST sci.: TV	"How often do you do the following things?"	"Watch TV shows about natural sciences"	(1) "never", (2) "rarely, (3) "sometimes", (4) "often"	t10000a
OST sci.: Books		"Borrow or buy books about natural sciences"		t10000b
OST sci.: Internet		"Visit Internet sites with topics relating to natural sciences"		t10000c
OST sci.: Magazine		"Read natural science magazines or articles in newspapers"		t10000d
OST sci.: Project		"Attend a natural science project group"		t10000e
Outcomes				
Int.: Realistic [a]	"How much are you interested in the following things?"	ICA-D 1, AIST-R 13 & 55	(1) "I have very little interest in that", (2) "I have little interest in that", (3) "I am somewhat interested in that", (4) "I am rather interested in that", (5) "I am very interested in that"	t66207a
Int.: Investigative		ICA-D 20, AIST-R 2 & 20		
Int.: Artistic		ICA-D 3, AIST-R 3 & 15		
Int.: Social		ICA-D 22, AIST-R 34 & 46		
Int.: Enterprising		ICA-D 17, AIST-R 41 & 53		
Int.: Conventional		ICA-D 18, AIST-R 42 & 54		
Asp.: Realistic	"Imagine you had all opportunities to become what you want. What would be your ideal occupation?"	ISCO 08 codes, recoded into RIASEC dimensions	-	t31060a_g3
Asp.: Investigative				
Asp.: Artistic				
Asp.: Social				
Asp.: Enterprising				
Asp.: Conventional				
Comp.: Math	"Solving the items requires recognizing and flexibly applying mathematics in realistic, mainly extra-mathematical situations (items	-	(1) task was solved correctly, (0) task was not solved correctly	mag5_sc1u

	are from different content areas and cognitive components)."			
Comp.: Reading	"Encompassed the areas of text understanding, text interpretation and multiple-choice questions concerning certain reading topics"	-	(1) task was solved correctly, (0) task was not solved correctly	reg5_sc1u
Comp.: Sc. Thinking		-	(1) task was solved correctly, (0) task was not solved correctly	stg12_sc1
Comp.: Sc. Literacy	"Conceptualized as the use of scientific knowledge in the environmental, technological and health contexts"	-	(1) task was solved correctly, (0) task was not solved correctly	scg6_sc1
Grade: Math	"What grade did you have in your last annual report card …"	"…in Math?"	(6) "very good (6)", (5) "good (5)", (4) "satisfactory (4)", (3) "passing (3)", (2) "poor (2)", (1) "failing (1)"	t724102
Grade: German		"…in German?"		t724101
Grade: English		"…in English?"		t724104
Self-Con.: Math	"How would you rate your performance at school?"	"I get good grades in math."	(1) "does not apply at all", (2) "does not really apply", (3) "applies to some extent", (4) "applies completely"	t66001a_g1
Self-Con.: German		"I learn fast in German."		t66000a_g1
Self-Con.: School		"I perform well in most of the school subjects."		t66002a_g1
Conf.: Demographics				
Gender	-	-	(0) "male", (1) "female"	tx80501
Migration	-	-	(0) "no migration background", (1) "migration background"	t400000_g1D
SES	"Please tell me what is your current occupation?" or "Please tell me what was your last occupation?"	-	"Answer coded to ISEI (Ganzeboom & Treiman, 1996)"	p731904_g5
Economic Situation	"How do you assess the economic situation of your household today?"	-	(1) "very bad", (2) "rather bad", (3) "partly good", (4) "rather good", (5) "very good"	p30300a
School Track	-	-	-	t723080_g1

Conf.: Health				
Sport Frequency	"How often do you do sports? Do NOT count the physical education classes at school."	-	(1) "never", (2) "once a month or less", (3) "several times a month or once a week", (4) "several times a week", (5) "almost daily"	t261000
Global Self-Worth	"To what extent do the following statements apply to you?"	"All in all, I am satisfied with myself."	(1) "does not apply at all", (2) "does not really apply", (3) "partially applies", (4) "applies to some extent", (5) "applies completely"	t66003a
Chronic Stress	"I would now like to talk about your personal situation in general. Please consider all areas of your life. To what extent do the following statements apply to you?"	"I feel exhausted after a normal day."	(1) "does not apply at all", (2) "does not really apply", (3) "partially applies", (4) "applies to some extent", (5) "applies completely"	t527003
Conf.: Personality				
Conscientiousness	"To what extent do the following statements apply to you?"	"I am thorough."	(1) "does not apply at all", (2) "does not really apply", (3) "partially applies", (4) "applies to some extent", (5) "applies completely"	t66800a
Extraversion		"I am out-going and sociable."		
Openness		"I do not care much about arts."		
Neuroticism		"I easily get nervous and self-conscious."		
Agreeableness		"I am considerate, sensitive."		
Conf.: Cog. Ability				
Processing Speed	"In NEPS, this is measured by the Picture Symbol Test (NEPS-BZT). This is based on an improved version of the Digit-Symbol Test (DST) from the tests of the Wechsler family."	-	(1) task was solved correctly, (0) task was not solved correctly	dgg5_sc3a
Logical Thinking	"The NEPS reasoning test (NEPS-MAT) is designed as a matrices test in the tradition of the typical reasoning tests. Each item of the	-	(1) task was solved correctly, (0) task was not solved correctly	dgg5_sc3b

Reading Speed	matrices test consists of several horizontally and vertically arranged fields in which different geometrical elements are shown." The study participants are given a total of 51 sentences which can be answered with the aid of general world knowledge, in other words no specific content-related previous knowledge is required (e.g., "mice can fly").	-	(1) task was solved correctly, (0) task was not solved correctly	rsg5_sc3
Conf.: School Ach.				
Science Grade	"What grade did you have on your last annual report card…"	"In Physics"; "In Chemistry"; "In Biology"; "In Science"	(1) "very good (1)", (2) "good (2)", (3) "satisfactory (3)", (4) "passing (4)", (5) "poor (5)", (6) "failing (6)"	t724106
Conf.: Family Env.				
Activities w/ Parents	"about things you do together as a family. In the last 12 months, how often have you taken part in the following activities?"	"How often have you made music together?"	(1) "never", (2) "once", (3) "2 to 3 times", (4) "4 to 5 times", (5) "More than five times", (6) "Once a month", (7) "Once a week or more"	p281401
Conf.: Motivation				
Int. Mot.: Reading	"What do you think about reading?"	"I enjoy reading books."	(1) "completely disagree", (2) "rather disagree", (3) "rather agree", (4) "completely agree"	td0042a
Helplessness Math	"To what extent do the following statements apply to you?"	"No matter how hard I try in math, my grades don't get any better."	(1) "does not apply at all", (2) "does not really apply", (3) "applies to some extent", (4) "applies completely"	t66005a

Subject Int.: Math	"To what extent do the following statements apply to you?"	"I enjoy puzzling over a mathematical problem."	(1) "does not apply at all", (2) "does not really apply", (3) "applies to some extent", (4) "applies completely"	t66201a
Subject Int.: German	"To what extent do the following statements apply to you?"	"It is very important for me to become more familiar with the German language and literature."	(1) "does not apply at all", (2) "does not really apply", (3) "applies to some extent", (4) "applies completely"	t66208b
Self-Con.: Reading	"How well do you read?"	"I can understand texts very well and quickly."	(1) "completely disagree", (2) "rather disagree", (3) "rather agree", (4) "completely agree"	td0043a
Class Repetition	"Did your child repeat a school year or was held back since our last interview in XX?"		(1) "yes", (2) "no"	p725000

Note. Conf. = confounding variable.

[a] The item wording was not depicted in the NEPS material, probably due to copyright reasons. Therefore, we depicted the item numbers of the original questionnaires.

Supplement C: E-Value and p Value Tables, Single Treatments & Composite Treatments

Table C1

Overview of Effect Sizes, p Values, and E-Values Across Treatment Effects Without Project Group Item for Median Treatment and Any Treatment

	Median Treatment (w/o AG)					Any Treatment (w/o AG)				
	d	p	α_b	α_s	E[95%CI]	d	p	α_b	α_s	E[95%CI]
R	0.09 [-0.05, 0.22]	.202	ns	ns	1.38 [0.96,]	0.01 [-0.19, 0.21]	.928	ns	ns	1.10 [0.84,]
I	0.26 [0.12, 0.41]	.000	*	*	1.86 [1.12,]	0.27 [0.05, 0.49]	.015	ns	ns	1.87 [1.05,]
A	0.06 [-0.07, 0.19]	.382	ns	ns	1.29 [0.94,]	0.08 [-0.10, 0.25]	.399	ns	ns	1.35 [0.91,]
S	0.03 [-0.11, 0.18]	.649	ns	ns	1.21 [0.90,]	0.00 [-0.27, 0.26]	.988	ns	ns	1.05 [0.79,]
E	0.05 [-0.10, 0.20]	.507	ns	ns	1.27 [0.92,]	0.04 [-0.25, 0.34]	.778	ns	ns	1.24 [0.79,]
C	0.06 [-0.09, 0.20]	.428	ns	ns	1.30 [0.92,]	0.01 [-0.20, 0.23]	.913	ns	ns	1.12 [0.83,]
Scien. Think.	0.02 [-0.17, 0.20]	.876	ns	ns	1.13 [0.85,]	0.08 [-0.22, 0.37]	.602	ns	ns	1.35 [0.82,]
Math Comp.	0.01 [-0.10, 0.11]	.894	ns	ns	1.09 [0.91,]	-0.03 [-0.23, 0.17]	.780	ns	ns	1.19 [0.86,]
Read. Comp.	0.03 [-0.08, 0.15]	.569	ns	ns	1.21 [0.93,]	0.19 [-0.06, 0.43]	.133	ns	ns	1.66 [0.95,]
Math Grade	0.05 [-0.10, 0.19]	.535	ns	ns	1.25 [0.91,]	0.10 [-0.10, 0.30]	.315	ns	ns	1.42 [0.92,]
Germ. Grade	0.00 [-0.16, 0.16]	.979	ns	ns	1.05 [0.87,]	0.07 [-0.14, 0.29]	.507	ns	ns	1.34 [0.88,]
Eng. Grade	0.05 [-0.09, 0.19]	.503	ns	ns	1.26 [0.92,]	0.27 [0.04, 0.51]	.025	ns	ns	1.89 [1.03,]
S.-C. Math	0.00 [-0.12, 0.12]	.981	ns	ns	1.04 [0.90,]	-0.02 [-0.23, 0.20]	.889	ns	ns	1.13 [0.84,]
S.-C. Germ.	-0.04 [-0.19, 0.11]	.612	ns	ns	1.23 [0.90,]	0.11 [-0.15, 0.37]	.405	ns	ns	1.45 [0.87,]
S.-C. School	-0.01 [-0.15, 0.13]	.873	ns	ns	1.11 [0.89,]	0.09 [-0.19, 0.36]	.525	ns	ns	1.39 [0.85,]
R Asp.	0.00 [-0.12, 0.12]	.952	ns	ns	1.06 [0.90,]	0.03 [-0.17, 0.23]	.782	ns	ns	1.19 [0.86,]
I Asp.	0.11 [-0.02, 0.24]	.108	ns	ns	1.44 [0.98,]	0.15 [-0.09, 0.38]	.222	ns	ns	1.55 [0.92,]
A Asp.	-0.03 [-0.14, 0.09]	.650	ns	ns	1.18 [0.92,]	-0.05 [-0.26, 0.16]	.632	ns	ns	1.27 [0.87,]
S Asp.	0.00 [-0.12, 0.12]	.999	ns	ns	1.01 [0.90,]	-0.05 [-0.23, 0.13]	.595	ns	ns	1.26 [0.89,]
E Asp.	-0.09 [-0.23, 0.04]	.163	ns	ns	1.40 [0.97,]	-0.12 [-0.32, 0.08]	.247	ns	ns	1.47 [0.93,]
C Asp.	0.06 [-0.07, 0.19]	.355	ns	ns	1.30 [0.94,]	0.20 [-0.03, 0.42]	.082	ns	ns	1.68 [0.98,]

Note. d = depicts standardized mean-level differences in standard deviations; p = depicts the p value of the respective effect size; α_b = depicts if the respective effect size is statistically significant at the Bonferroni adjusted alpha level; α_s = depicts if the respective effect size is statistically significant at the Sidak adjusted alpha level; E[95%CI] = depicts the E-value of the respective effect sizes and its lower bound of the 95% confidence interval; (w/o AG) = implies without project group item; ns = nonsignificant; * = significant.

Table C2

Overview of Effect Sizes, p Values, and E-Values Across Treatment Effects Without Project Group Item for Extreme Treatment and Mean Treatment

	Extreme Treatment (w/o AG)					Mean Treatment (w/o AG)				
	d	*p*	α_b	α_s	E[95%CI]	*b*	*p*	α_b	α_s	E[95%CI]
R	0.11 [-0.15, 0.38]	.407	ns	ns	1.45 [0.87,]	0.06 [-0.01, 0.13]	.097	ns	ns	1.30 [0.99,]
I	0.52 [0.24, 0.80]	.000	*	*	2.59 [1.25,]	0.19 [0.12, 0.27]	.000	*	*	1.67 [1.11,]
A	0.13 [-0.13, 0.39]	.331	ns	ns	1.50 [0.89,]	0.05 [-0.02, 0.12]	.145	ns	ns	1.28 [0.98,]
S	0.01 [-0.23, 0.26]	.912	ns	ns	1.13 [0.81,]	0.01 [-0.07, 0.10]	.774	ns	ns	1.12 [0.94,]
E	0.02 [-0.28, 0.32]	.913	ns	ns	1.14 [0.77,]	0.03 [-0.06, 0.11]	.520	ns	ns	1.19 [0.95,]
C	0.03 [-0.22, 0.28]	.804	ns	ns	1.20 [0.82,]	0.03 [-0.04, 0.11]	.368	ns	ns	1.21 [0.97,]
Scien. Think.	-0.01 [-0.38, 0.36]	.964	ns	ns	1.10 [0.72,]	0.02 [-0.09, 0.13]	.768	ns	ns	1.14 [0.92,]
Math Comp.	-0.04 [-0.27, 0.18]	.708	ns	ns	1.24 [0.85,]	0.00 [-0.06, 0.06]	.997	ns	ns	1.01 [0.95,]
Read. Comp.	0.07 [-0.18, 0.31]	.596	ns	ns	1.32 [0.85,]	0.03 [-0.04, 0.09]	.381	ns	ns	1.19 [0.97,]
Math Grade	0.14 [-0.14, 0.43]	.332	ns	ns	1.53 [0.88,]	0.05 [-0.02, 0.13]	.177	ns	ns	1.27 [0.98,]
Germ. Grade	0.06 [-0.24, 0.37]	.691	ns	ns	1.31 [0.80,]	0.03 [-0.06, 0.11]	.551	ns	ns	1.18 [0.95,]
Eng. Grade	0.20 [-0.08, 0.47]	.165	ns	ns	1.68 [0.93,]	0.06 [-0.02, 0.13]	.145	ns	ns	1.29 [0.98,]
S.-C. Math	0.02 [-0.23, 0.26]	.886	ns	ns	1.15 [0.81,]	0.02 [-0.05, 0.08]	.661	ns	ns	1.13 [0.95,]
S.-C. Germ.	-0.01 [-0.30, 0.28]	.954	ns	ns	1.10 [0.77,]	0.00 [-0.08, 0.08]	.955	ns	ns	1.05 [0.93,]
S.-C. School	0.04 [-0.23, 0.32]	.748	ns	ns	1.25 [0.81,]	0.00 [-0.08, 0.08]	.982	ns	ns	1.03 [0.93,]
R Asp.	0.05 [-0.20, 0.29]	.716	ns	ns	1.25 [0.83,]	0.00 [-0.06, 0.07]	.938	ns	ns	1.05 [0.95,]
I Asp.	0.26 [-0.02, 0.53]	.064	ns	ns	1.84 [0.99,]	0.07 [0.01, 0.14]	.035	ns	ns	1.34 [1.00,]
A Asp.	-0.12 [-0.37, 0.13]	.347	ns	ns	1.47 [0.89,]	-0.03 [-0.09, 0.03]	.352	ns	ns	1.19 [0.97,]
S Asp.	0.00 [-0.24, 0.24]	.993	ns	ns	1.03 [0.81,]	0.01 [-0.06, 0.07]	.854	ns	ns	1.08 [0.95,]
E Asp.	-0.18 [-0.43, 0.07]	.167	ns	ns	1.63 [0.94,]	-0.06 [-0.13, 0.00]	.059	ns	ns	1.31 [1.00,]
C Asp.	0.11 [-0.16, 0.37]	.442	ns	ns	1.43 [0.86,]	0.05 [-0.02, 0.12]	.159	ns	ns	1.26 [0.98,]

Note. *d* = depicts standardized mean-level differences in standard deviations; *p* = depicts the *p* value of the respective effect size; α_b = depicts if the respective effect size is statistically significant at the Bonferroni adjusted alpha level; α_s = depicts if the respective effect size is statistically significant at the Sidak adjusted alpha level; E[95%CI] = depicts the E-value of the respective effect sizes and its lower bound of the 95% confidence interval; (w/o AG) = implies without project group item; ns = nonsignificant; * = significant.

Table C3

Overview of Effect Sizes, p Values, and E-Values Across Treatment Effects With Project Group Item for Median Treatment and Any Treatment

	Median Treatment					Any Treatment				
	d	*p*	α_b	α_s	E[95%CI]	*d*	*p*	α_b	α_s	E[95%CI]
R	0.12 [-0.01, 0.25]	.079	ns	ns	1.47 [0.99,]	0.07 [-0.16, 0.29]	.556	ns	ns	1.32 [0.87,]
I	0.30 [0.15, 0.44]	.000	*	*	1.94 [1.15,]	0.36 [0.12, 0.60]	.004	ns	ns	2.11 [1.11,]
A	0.05 [-0.08, 0.17]	.446	ns	ns	1.26 [0.93,]	0.12 [-0.07, 0.31]	.210	ns	ns	1.48 [0.94,]
S	0.02 [-0.13, 0.16]	.827	ns	ns	1.14 [0.89,]	-0.11 [-0.38, 0.16]	.433	ns	ns	1.44 [0.86,]
E	0.04 [-0.12, 0.19]	.650	ns	ns	1.22 [0.90,]	-0.10 [-0.34, 0.14]	.404	ns	ns	1.43 [0.88,]
C	0.04 [-0.10, 0.18]	.547	ns	ns	1.24 [0.92,]	0.02 [-0.22, 0.26]	.885	ns	ns	1.14 [0.82,]
Scien. Think.	0.00 [-0.17, 0.18]	.975	ns	ns	1.05 [0.86,]	0.13 [-0.17, 0.43]	.401	ns	ns	1.50 [0.86,]
Math Comp.	-0.02 [-0.14, 0.10]	.712	ns	ns	1.16 [0.92,]	-0.07 [-0.25, 0.11]	.450	ns	ns	1.33 [0.91,]
Read. Comp.	0.00 [-0.12, 0.11]	.978	ns	ns	1.04 [0.90,]	0.15 [-0.14, 0.44]	.299	ns	ns	1.56 [0.88,]
Math Grade	0.04 [-0.11, 0.19]	.603	ns	ns	1.23 [0.91,]	0.14 [-0.07, 0.34]	.182	ns	ns	1.53 [0.94,]
Germ. Grade	-0.01 [-0.18, 0.15]	.883	ns	ns	1.12 [0.87,]	0.17 [-0.06, 0.40]	.138	ns	ns	1.62 [0.95,]
Eng. Grade	0.02 [-0.12, 0.17]	.753	ns	ns	1.17 [0.90,]	0.29 [0.04, 0.53]	.023	ns	ns	1.92 [1.04,]
S.-C. Math	0.03 [-0.10, 0.15]	.686	ns	ns	1.18 [0.91,]	-0.04 [-0.27, 0.19]	.722	ns	ns	1.24 [0.84,]
S.-C. Germ.	-0.02 [-0.17, 0.13]	.835	ns	ns	1.14 [0.89,]	0.16 [-0.10, 0.41]	.232	ns	ns	1.57 [0.91,]
S.-C. School	0.01 [-0.13, 0.15]	.885	ns	ns	1.11 [0.89,]	0.15 [-0.15, 0.46]	.329	ns	ns	1.56 [0.87,]
R Asp.	0.03 [-0.10, 0.16]	.648	ns	ns	1.20 [0.91,]	0.13 [-0.08, 0.33]	.235	ns	ns	1.49 [0.93,]
I Asp.	0.13 [0.00, 0.25]	.052	ns	ns	1.49 [1.00,]	0.03 [-0.20, 0.26]	.787	ns	ns	1.20 [0.84,]
A Asp.	-0.05 [-0.17, 0.06]	.366	ns	ns	1.28 [0.94,]	-0.11 [-0.34, 0.13]	.386	ns	ns	1.43 [0.89,]
S Asp.	-0.02 [-0.14, 0.10]	.771	ns	ns	1.15 [0.91,]	-0.11 [-0.33, 0.11]	.338	ns	ns	1.44 [0.90,]
E Asp.	-0.11 [-0.24, 0.02]	.101	ns	ns	1.44 [0.98,]	-0.07 [-0.30, 0.16]	.525	ns	ns	1.34 [0.87,]
C Asp.	0.06 [-0.06, 0.19]	.318	ns	ns	1.31 [0.95,]	0.14 [-0.09, 0.36]	.237	ns	ns	1.52 [0.92,]

Note. *d* = depicts standardized mean-level differences in standard deviations; *p* = depicts the *p* value of the respective effect size; α_b = depicts if the respective effect size is statistically significant at the Bonferroni adjusted alpha level; α_s = depicts if the respective effect size is statistically significant at the Sidak adjusted alpha level; E[95%CI] = depicts the E-value of the respective effect sizes and its lower bound of the 95% confidence interval; ns = nonsignificant; * = significant.

Table C4

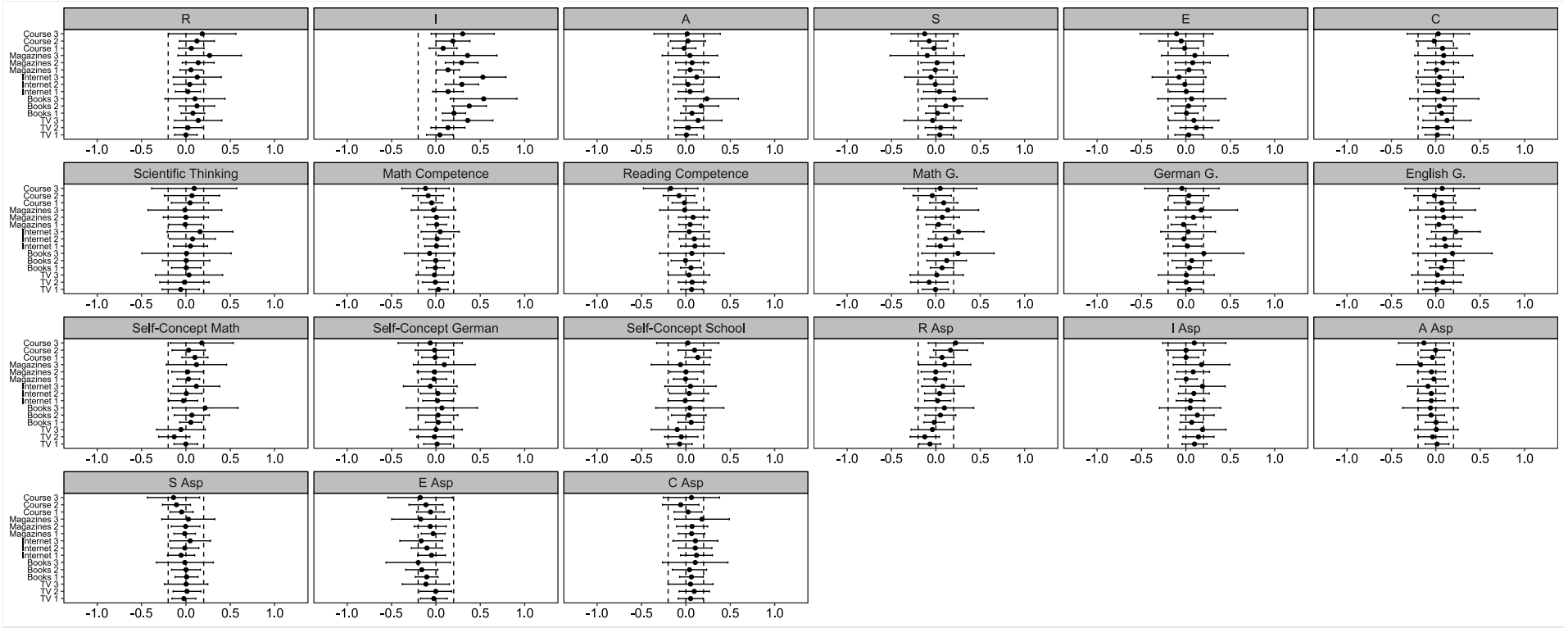
Overview of Effect Sizes, p Values, and E-Values Across Treatment Effects With Project Group Item for Extreme Treatment and Mean Treatment

	Extreme Treatment					Mean Treatment				
	d	*p*	α_b	α_s	E[95%CI]	*b*	*p*	α_b	α_s	E[95%CI]
R	0.17 [-0.10, 0.44]	.225	ns	ns	1.60 [0.91,]	0.08 [0.00, 0.15]	.038	ns	ns	1.35 [1.00,]
I	0.58 [0.29, 0.87]	.000	*	*	2.78 [1.31,]	0.21 [0.13, 0.28]	.000	*	*	1.71 [1.13,]
A	0.13 [-0.14, 0.40]	.331	ns	ns	1.51 [0.89,]	0.05 [-0.02, 0.12]	.169	ns	ns	1.27 [0.98,]
S	-0.04 [-0.32, 0.23]	.747	ns	ns	1.25 [0.81,]	0.00 [-0.08, 0.08]	.997	ns	ns	1.01 [0.93,]
E	-0.03 [-0.33, 0.27]	.854	ns	ns	1.19 [0.78,]	0.02 [-0.06, 0.10]	.676	ns	ns	1.14 [0.94,]
C	0.04 [-0.24, 0.32]	.762	ns	ns	1.24 [0.81,]	0.03 [-0.04, 0.11]	.402	ns	ns	1.21 [0.96,]
Scien. Think.	0.05 [-0.32, 0.42]	.801	ns	ns	1.26 [0.75,]	0.03 [-0.08, 0.13]	.614	ns	ns	1.18 [0.93,]
Math Comp.	-0.08 [-0.31, 0.14]	.454	ns	ns	1.37 [0.88,]	-0.01 [-0.07, 0.04]	.644	ns	ns	1.12 [0.96,]
Read. Comp.	0.02 [-0.23, 0.27]	.864	ns	ns	1.16 [0.82,]	0.01 [-0.05, 0.08]	.677	ns	ns	1.13 [0.95,]
Math Grade	0.14 [-0.16, 0.44]	.365	ns	ns	1.53 [0.86,]	0.05 [-0.03, 0.13]	.219	ns	ns	1.27 [0.97,]
Germ. Grade	0.10 [-0.22, 0.41]	.548	ns	ns	1.41 [0.82,]	0.03 [-0.06, 0.12]	.556	ns	ns	1.18 [0.95,]
Eng. Grade	0.21 [-0.09, 0.50]	.170	ns	ns	1.70 [0.92,]	0.06 [-0.02, 0.13]	.155	ns	ns	1.29 [0.98,]
S.-C. Math	0.04 [-0.23, 0.31]	.776	ns	ns	1.23 [0.81,]	0.03 [-0.05, 0.10]	.469	ns	ns	1.18 [0.96,]
S.-C. Germ.	-0.02 [-0.30, 0.27]	.908	ns	ns	1.14 [0.79,]	0.00 [-0.08, 0.08]	.971	ns	ns	1.04 [0.93,]
S.-C. School	0.07 [-0.20, 0.35]	.600	ns	ns	1.34 [0.83,]	0.01 [-0.06, 0.09]	.712	ns	ns	1.13 [0.94,]
R Asp.	0.11 [-0.13, 0.35]	.376	ns	ns	1.44 [0.89,]	0.03 [-0.04, 0.09]	.409	ns	ns	1.18 [0.97,]
I Asp.	0.18 [-0.09, 0.45]	.196	ns	ns	1.63 [0.92,]	0.07 [0.01, 0.14]	.034	ns	ns	1.33 [1.01,]
A Asp.	-0.13 [-0.39, 0.12]	.303	ns	ns	1.52 [0.90,]	-0.03 [-0.09, 0.03]	.285	ns	ns	1.21 [0.98,]
S Asp.	0.02 [-0.21, 0.26]	.860	ns	ns	1.16 [0.82,]	-0.01 [-0.07, 0.05]	.744	ns	ns	1.11 [0.96,]
E Asp.	-0.21 [-0.48, 0.06]	.132	ns	ns	1.71 [0.95,]	-0.08 [-0.14, -0.01]	.025	ns	ns	1.35 [1.01,]
C Asp.	0.10 [-0.17, 0.37]	.474	ns	ns	1.42 [0.85,]	0.04 [-0.02, 0.11]	.198	ns	ns	1.25 [0.98,]

Note. *d* = depicts standardized mean-level differences in standard deviations; *p* = depicts the *p* value of the respective effect size; α_b = depicts if the respective effect size is statistically significant at the Bonferroni adjusted alpha level; α_s = depicts if the respective effect size is statistically significant at the Sidak adjusted alpha level; E[95%CI] = depicts the E-value of the respective effect sizes and its lower bound of the 95% confidence interval; ns = nonsignificant; * = significant.

Figure 1C

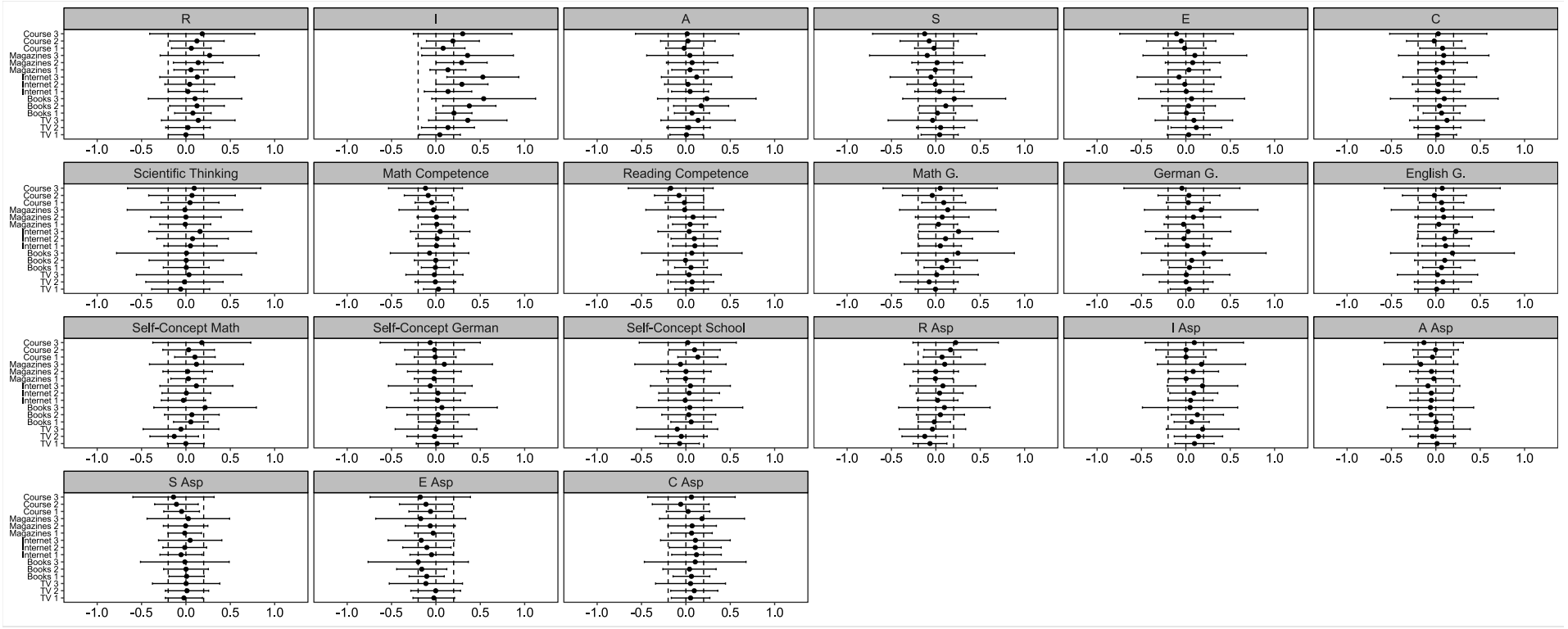
Effect Sizes of All Individual Treatment Types on All Outcome Variables With Unadjusted 95% Confidence Intervals



Note. Standardized effect sizes of all treatment effects are depicted for all outcomes, under control of all confounders mentioned in the instruments section. The right and left vertical dashed lines depict a *d* of 0.2 and -0.2, respectively. Positive values indicate a higher mean value on the outcomes for participants with more engagement in the unstructured OST science activity. 1 = indicates the dummy variable of the respective activity item that compares the groups "never" and "rarely"; 2 = indicates the dummy variable of the respective activity item that compares the groups "never" and "sometimes"; 3 = indicates the dummy variable of the respective activity item that compares the groups "never" and "often".

Figure 2C

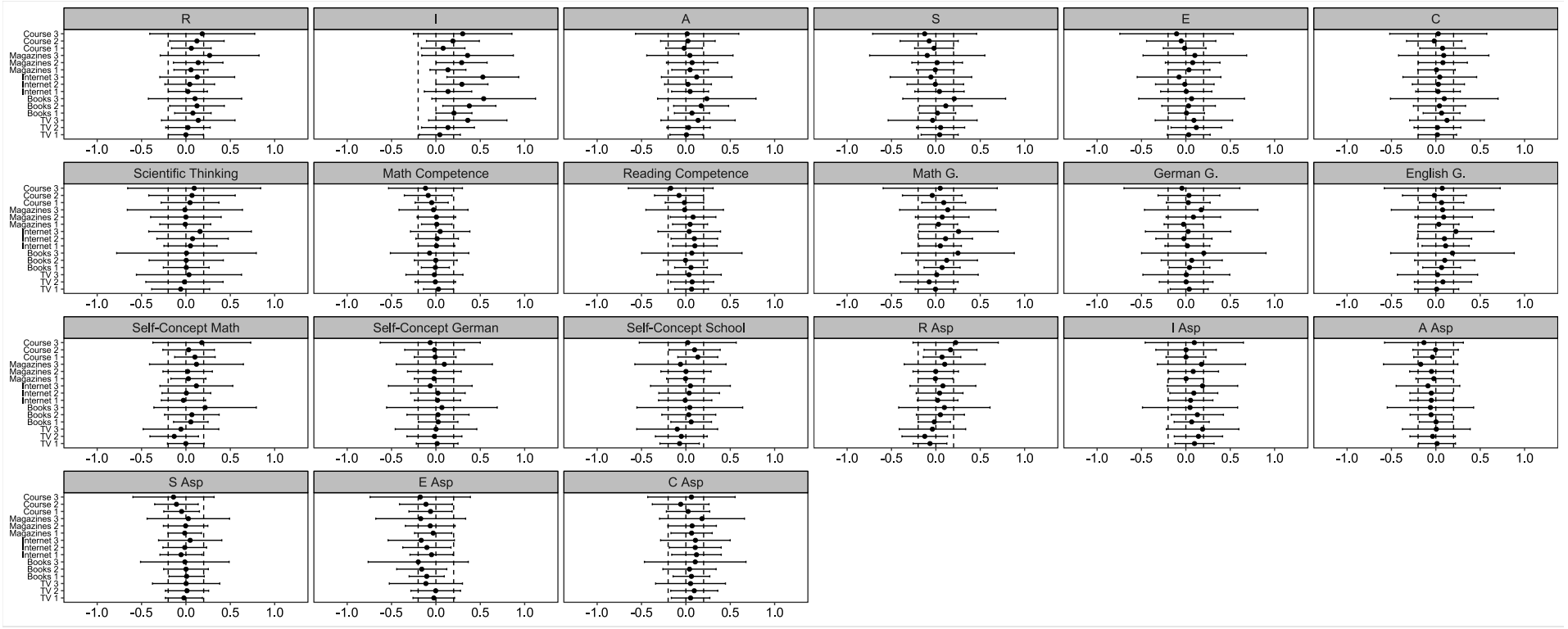
Effect Sizes of All Individual Treatment Types on All Outcome Variables With Bonferroni Adjusted 95% Confidence Intervals



Note. Standardized effect sizes of all treatment effects are depicted for all outcomes, under control of all confounders mentioned in the instruments section. The right and left vertical dashed lines depict a *d* of 0.2 and -0.2, respectively. Positive values indicate a higher mean value on the outcomes for participants with more engagement in the unstructured OST science activity. 1 = indicates the dummy variable of the respective activity item that compares the groups “never” and “rarely”; 2 = indicates the dummy variable of the respective activity item that compares the groups “never” and “sometimes”; 3 = indicates the dummy variable of the respective activity item that compares the groups “never” and “often”.

Figure 3C

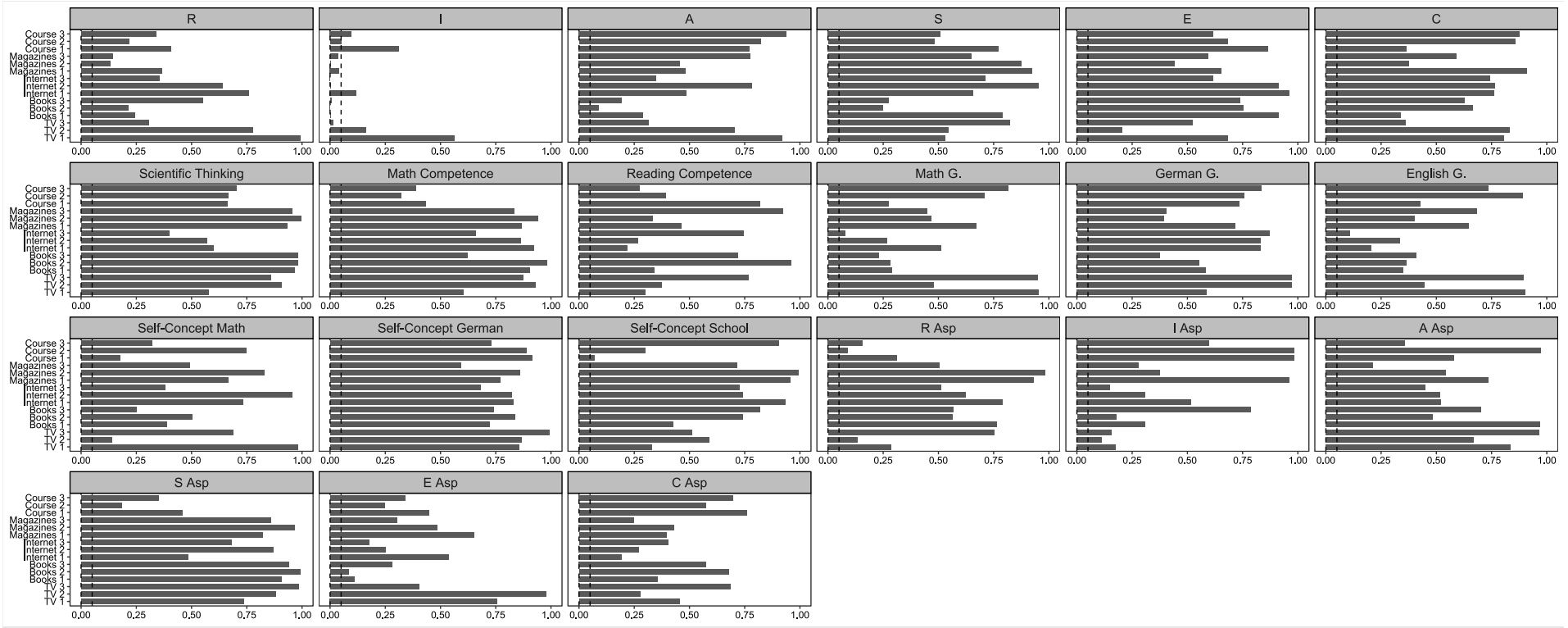
Effect Sizes of All Individual Treatment Types on All Outcome Variables With Sidak Adjusted 95% Confidence Intervals



Note. Standardized effect sizes of all treatment effects are depicted for all outcomes, under control of all confounders mentioned in the instruments section. The right and left vertical dashed lines depict a *d* of 0.2 and -0.2, respectively. Positive values indicate a higher mean value on the outcomes for participants with more engagement in the unstructured OST science activity. 1 = indicates the dummy variable of the respective activity item that compares the groups "never" and "rarely"; 2 = indicates the dummy variable of the respective activity item that compares the groups "never" and "sometimes"; 3 = indicates the dummy variable of the respective activity item that compares the groups "never" and "often".

Figure 4C

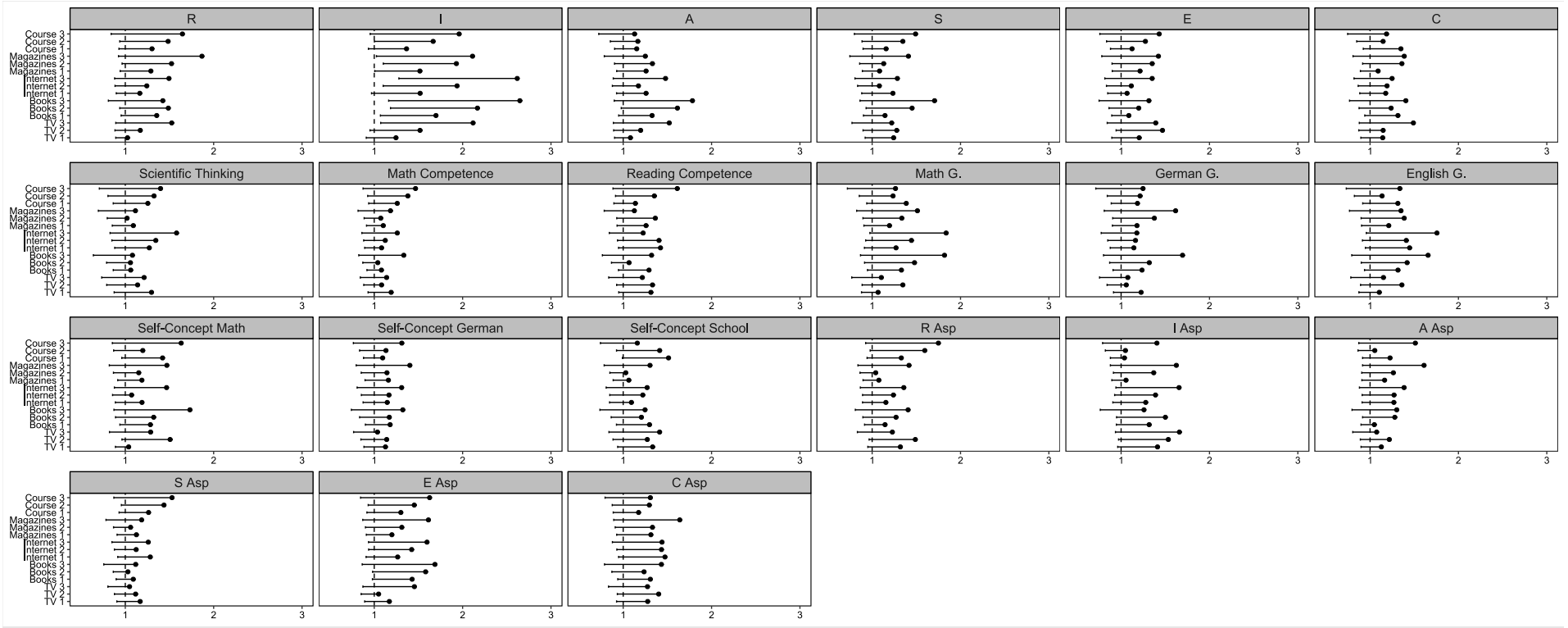
The p Values of the Effect Sizes of All Individual Treatment Types on All Outcome Variables



Note. The dashed line indicates a *p* value of 0.05; *p* values in the figure were not adjusted for multiple testing. 1 = indicates the dummy variable of the respective activity item that compares the groups "never" and "rarely"; 2 = indicates the dummy variable of the respective activity item that compares the groups "never" and "sometimes"; 3 = indicates the dummy variable of the respective activity item that compares the groups "never" and "often".

Figure 5C

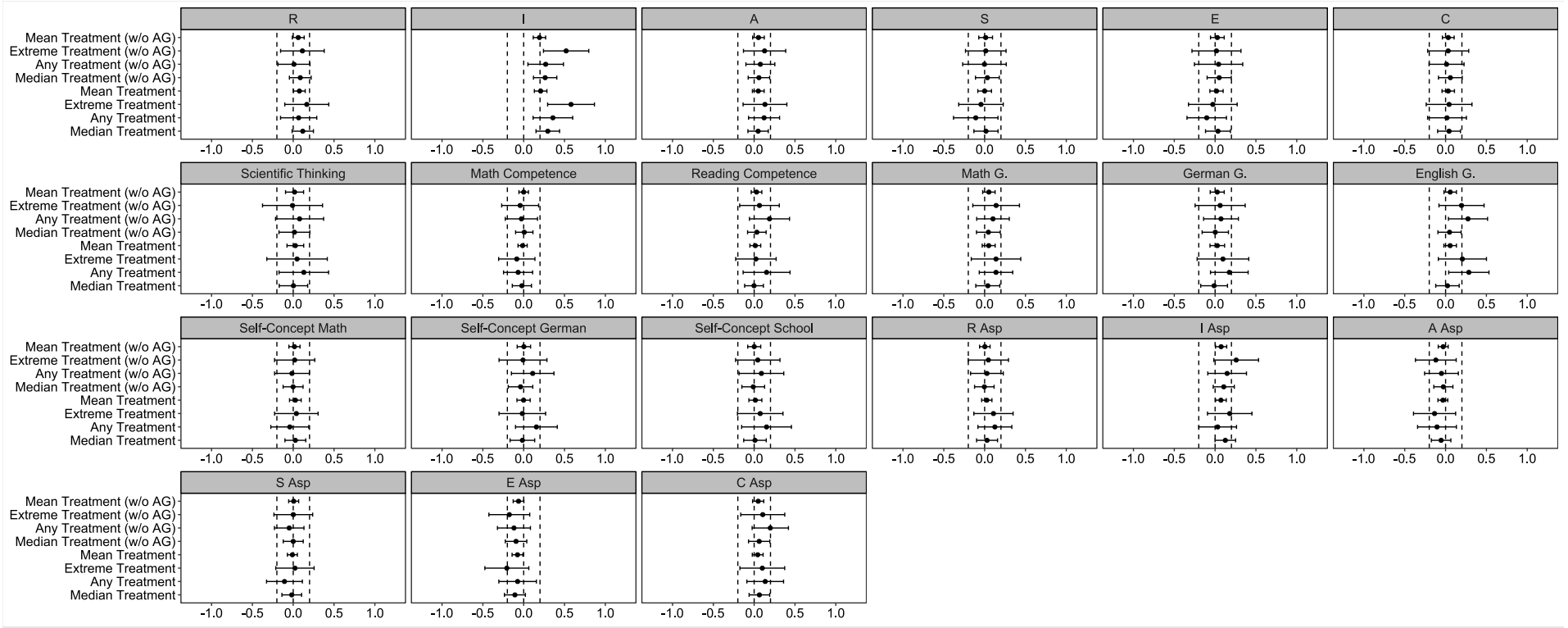
E-Values of the Effect Sizes of All Individual Treatment Types on All Outcome Variables



Note. The dashed line indicates an E-value of 1. Only the lower confidence interval is depicted. 1 = indicates the dummy variable of the respective activity item that compares the groups "never" and "rarely"; 2 = indicates the dummy variable of the respective activity item that compares the groups "never" and "sometimes"; 3 = indicates the dummy variable of the respective activity item that compares the groups "never" and "often".

Figure 6C

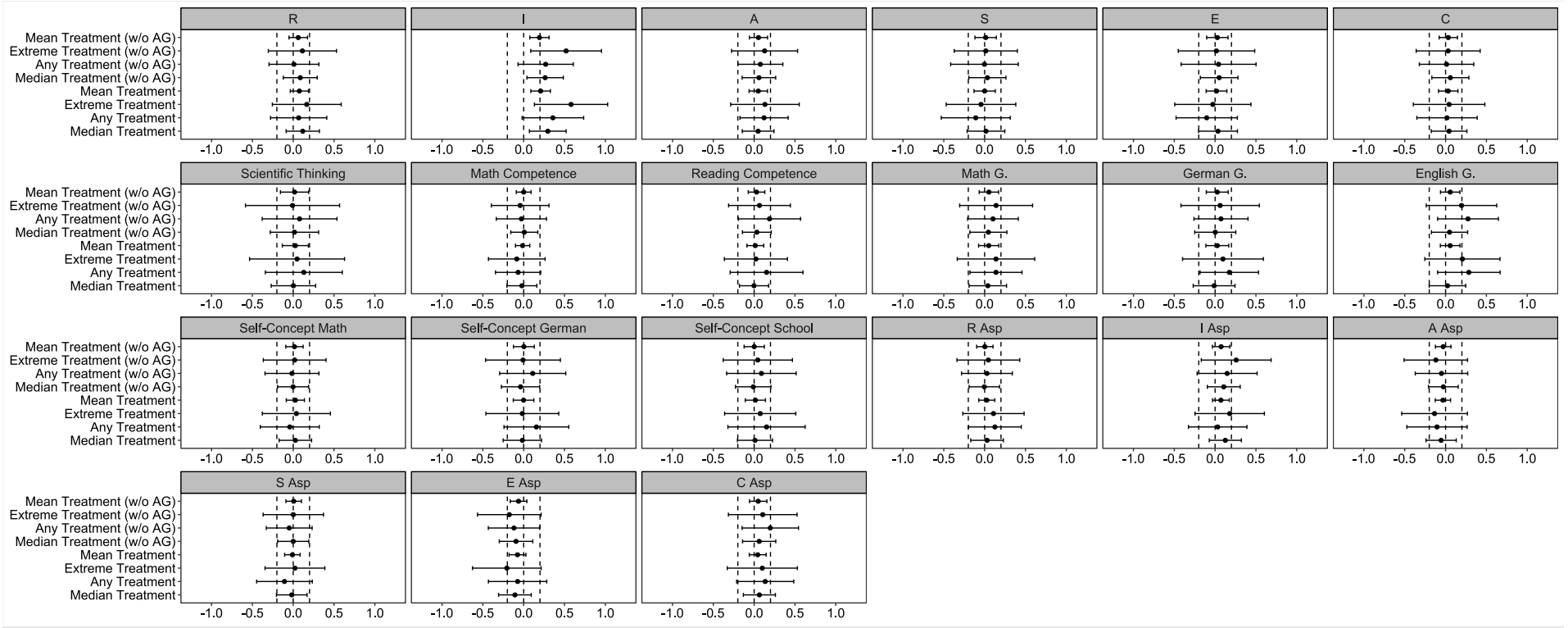
Effect Sizes of Composite Treatments on All Outcome Variables With Unadjusted 95% Confidence Intervals



Note. Standardized effect sizes of all treatment effects are depicted for all outcomes, under control of all confounders mentioned in the instruments section. The right and left vertical dashed lines depict a *d* of 0.2 and -0.2, respectively. Positive values indicate a higher mean value on the outcomes for participants with more engagement in the unstructured OST science activities. Composite treatments with the abbreviation (w/o AG) did not contain the project group item.

Figure 7C

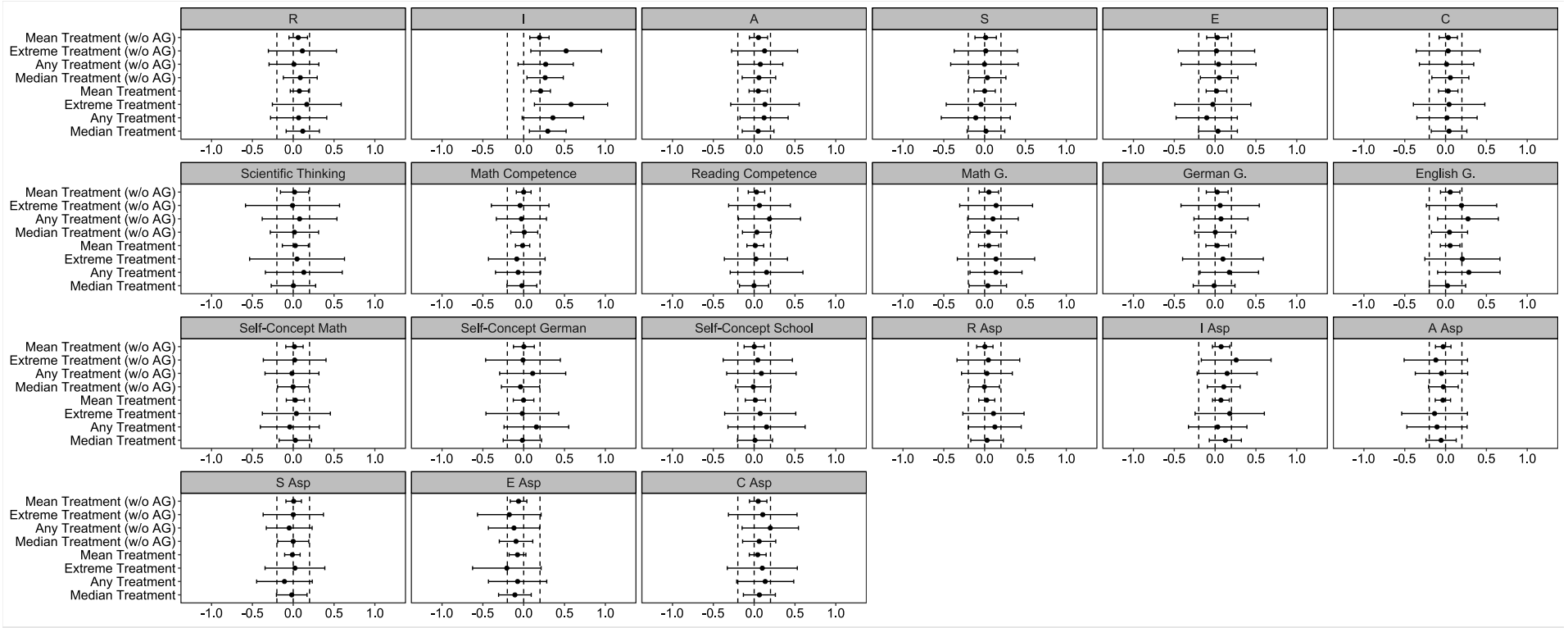
Effect Sizes of Composite Treatments on All Outcome Variables With Adjusted 95% Confidence Intervals (Bonferroni Correction)



Note. Standardized effect sizes of all treatment effects are depicted for all outcomes, under control of all confounders mentioned in the instruments section. The right and left vertical dashed lines depict a *d* of 0.2 and -0.2, respectively. Positive values indicate a higher mean value on the outcomes for participants with more engagement in the unstructured OST science activities. Composite treatments with the abbreviation (w/o AG) did not contain the project group item.

Figure 8C

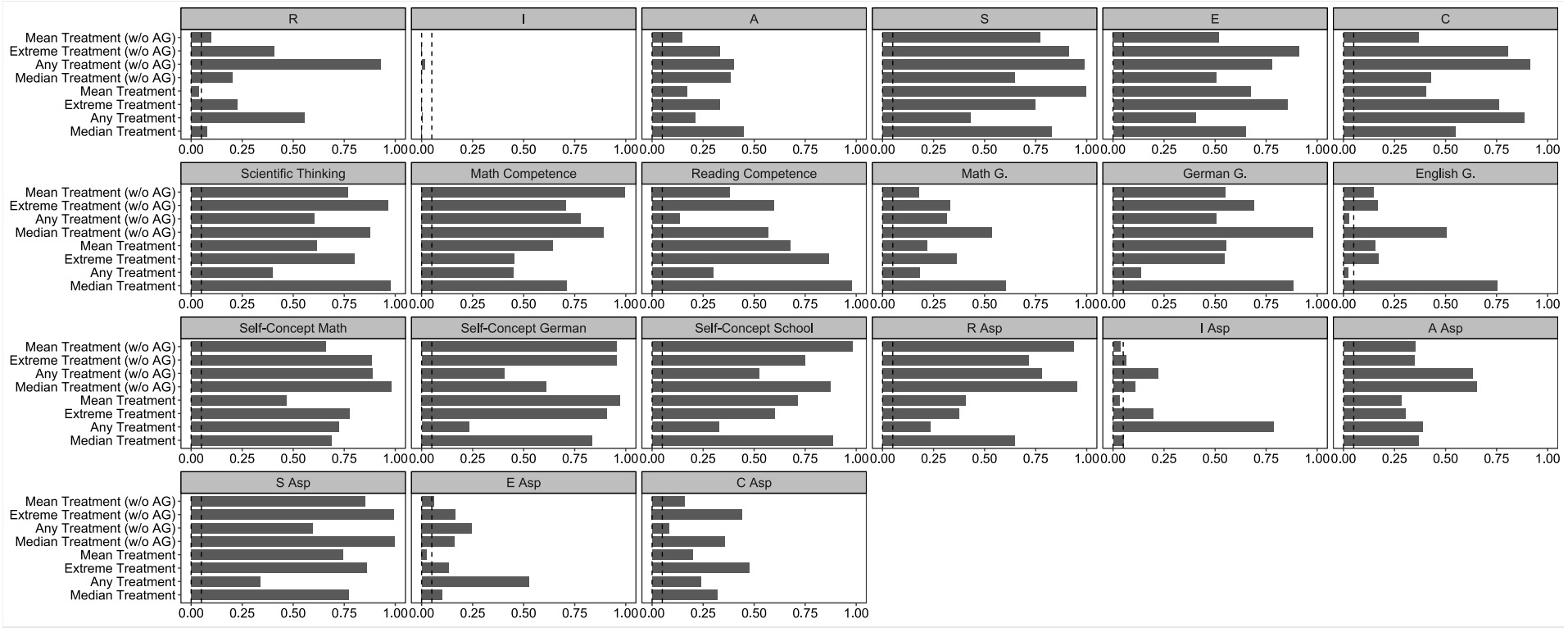
Effect Sizes of Composite Treatments on All Outcome Variables With Adjusted 95% Confidence Intervals (Sidak Correction)



Note. Standardized effect sizes of all treatment effects are depicted for all outcomes, under control of all confounders mentioned in the instruments section. The right and left vertical dashed lines depict a *d* of 0.2 and -0.2, respectively. Positive values indicate a higher mean value on the outcomes for participants with more engagement in the unstructured OST science activities. Composite treatments with the abbreviation (w/o AG) did not contain the project group item.

Figure 9C

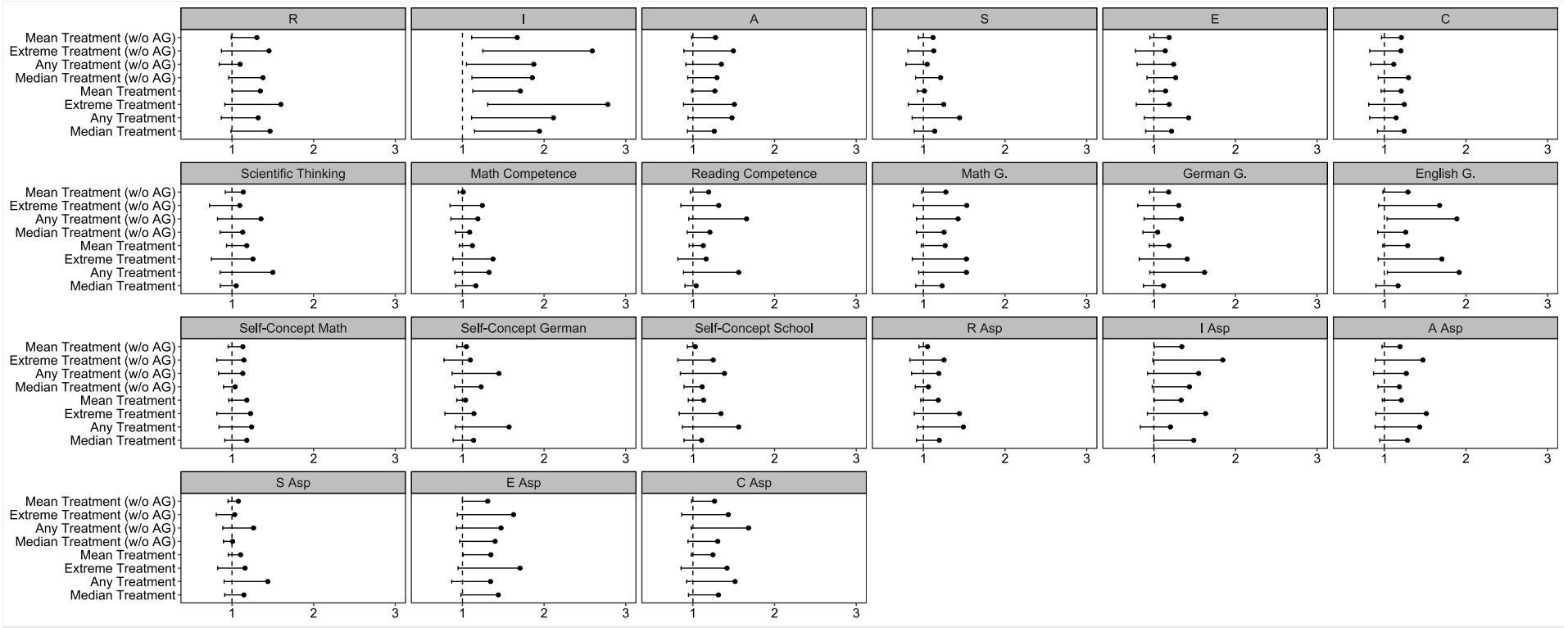
The p Values of the Effect Sizes of All Composite Treatments on All Outcome Variables



Note. The dashed line indicates a *p* value of 0.05; *p* values in the figure were not adjusted for multiple testing. Composite treatments with the abbreviation (w/o AG) did not contain the project group item. *P*-values for Investigative vocational interests were relatively small and are therefore not visible in the figure.

Figure 10C

E-Values of the Effect Sizes of All Composite Treatments on All Outcome Variables



Note. The dashed line indicates an E-value of 1. Only the lower confidence interval is depicted. Composite treatments with the abbreviation (w/o AG) did not contain the project group item.

6 GENERAL DISCUSSION

The current dissertation helps to deepen the understanding about vocational interest development over the course of adolescence. Besides comprehensive descriptive information about interest stability, intensity, and gender differentiation, insights about possible factors—individual and environmental—that influence the development of vocational interests are provided. The current dissertation describes first empirical evidence on the influence of personality traits and cognitive abilities on vocational interest profile stability. In addition, by providing evidence about the influence of leisure activities on vocational interest development (see Study 3), new and robust evidence was created about the impact of contextual factors on vocational interests. By investigating vocational interest development in adolescence based on studies with multiwave large-scale longitudinal data (see Study 1), a need in vocational interest research was filled.

The general discussion summarizes and integrates empirical findings about vocational interest development in adolescence—focusing specifically on normative change, stability, and influencing factors. In the current dissertation, the three studies operationalized vocational interests with Holland's (1997) classification model. Findings on vocational interests can therefore be described based on the six RIASEC dimensions. Chapter 6.1 summarizes the general results of the three studies, focusing on mean-level changes, re-test correlations, profile correlations, and gender differences. These findings will be integrated in Chapter 6.2 under the umbrella of the three developmental principles—evidence for, but also against the developmental principles will be discussed. The purpose of Chapter 6.2 is to describe how intensity, stability, and gender differences in vocational interests unfold over the course of adolescence. Chapter 6.3 highlights the influence of experiencing activities on the development of vocational interests. Practical implications are elaborated in Chapter 6.4. Chapter 6.5 will state the limitations of the current dissertation and conclude with an outlook for prospective research topics. Finally, Chapter 6.6 will provide a short summary of the main findings and their implications.

6.1 Summary of the Three Empirical Studies

In the following chapter the main findings of the three empirical studies will be summarized. The three empirical studies comprise evidence on normative change, stability, and influencing factors of vocational interest development in adolescence. Furthermore, as the three empirical studies also provide evidence on interest intensity, stability, and gender differences,

general results about mean levels, re-test correlations, profile correlations, and gender differences will be described. The findings of the three empirical studies will be compared to existing empirical evidence of previous studies, and novel evidence that was generated will be highlighted.

6.1.1 The Development of Vocational Interests in Early Adolescence

The first study investigated the development of vocational interests over the course of late childhood and early adolescence (ages 11 to 14). The results indicate that vocational interests were moderately to highly stable, with one-year re-test correlations between .41 and .64 and three-year re-test correlations between .32 and .49. Increases in stability were found for the dimensions Realistic, Investigative, Social and Conventional ($.06 < \Delta r < .08$), whereas the dimensions Artistic and Enterprising experienced no increases in stability over time. In comparison to previous investigations (e.g., Päßler & Hell, 2020; Tracey, 2002), the study provides robust evidence that stability of vocational interests increases over the course of late childhood and early adolescence.

In contrast to increases in stability, most interest dimensions decreased in mean levels and hence in interest intensity. Over the three-year period, the dimensions Realistic, Investigative, Artistic and Conventional decreased in mean levels ($-0.44 < \Delta d < -0.24$), whereas the dimension Social increased in mean levels ($\Delta d = 0.07$) and the dimension Enterprising had almost no change over time ($d = -0.01$). In comparison to previous (meta-analytic) findings (Hoff et al., 2018; Päßler & Hell, 2020; Tracey, 2002), the decreases reported in the study were bigger in magnitude (i.e., bigger effect sizes) and more overarching (i.e., more dimensions significantly decreased in mean levels).

Similarly, gender differences in mean levels of vocational interests were large ($-0.82 < \Delta d < 1.14$) and already present around age 11. Males were more interested in Realistic, Investigative, Enterprising and Conventional activities, whereas females were more interested in Artistic and Social activities. These results were in line with meta-analytic findings on gender differences (Hoff et al., 2018; Su et al., 2009). In addition, gender differences increased on all dimensions, except Artistic, over the course of late childhood and early adolescence. Although Hoff et al. (2018) already reported meta-analytic evidence for increases in gender differences in Realistic and Investigative interest over time, the current study was the first to generalize these findings to most of the interest dimensions.

Besides findings on interest intensity, stability, and gender differences, the study also provided evidence on the dispositional nature of vocational interests. According to latent state-trait analyses, vocational interests mainly consist of dispositional components, as indicated by the high amount of trait variance proportions. However, situational components were substantial as well, especially in late childhood, as indicated by the moderate amount of state variance proportions. These findings support the assumption of the TSID model (Su et al., 2019) that vocational interests are dispositional by nature, but also consist of situational susceptible components.

6.1.2 Predictors of Vocational Interest Profile Stability

The second study investigated vocational interest profile stability and its predictors in four different life phases. Based on a multi-study investigation, profile stability was examined in late childhood and early adolescence (Substudy 1: ages 11 to 14), middle adolescence (Substudy 2: ages 14 to 15), late adolescence to young adulthood (Substudy 3: ages 17 to 23), and young adulthood (Substudy 4: ages 22 to 34). In each of the four substudies, gender, personality traits, and cognitive abilities were investigated as possible predictors of profile stability. Profile correlations indicated that profile stability was moderate over the course of late childhood and early adolescence (r = .43) and high over the course of middle adolescence (r =. 65), late adolescence to young adulthood (r = .64), and young adulthood (r = .74). The results were in line with previous meta-analytic findings about rank-order stability (Low et al., 2005) and single studies about profile stability (e.g., Etzel & Nagy, 2021; Rottinghaus et al., 2007; Stoll, Rieger, et al., 2020; Swanson & Hansen, 1988; Zytowski, 1976), which showed that vocational interests are relatively stable, even over longer periods of time (e.g., time intervals of 12 years).

Significant variances in profile stability indicators were found in all four life phases ($0.26 < SD < 0.42$). As these indicators provide information about stability on the individual level, significant variances indicate that some people had more stable profiles than others. The most consistent predictor for these between-person differences was gender. Girls and women had significantly more stable vocational interest profiles compared to boys and men in all the respective substudies ($.04 < \Delta r < .23$). The findings on gender differences in profile stability are in line with evidence from Xu and Tracey (2016) as well as Stoll et al. (2020), which also reported more stable vocational interests for girls and women. In addition to gender, consistent and significant associations with profile stability were also found for the personality trait

extraversion and verbal cognitive abilities. Higher levels of extraversion and verbal cognitive abilities were related to higher levels of profile stability. However, in comparison to gender, the effect sizes were relatively small. The second study provides new insights about the interrelation of vocational interest profile stability, personality traits, and cognitive abilities.

6.1.3 Effects of Out of School Engagement in Science on Vocational Interests

The third study investigated the influence of unstructured out-of-school time (OST) science activities, such as reading a science book, watching a science TV show, or researching on the internet about science, on the development of vocational interests, occupational aspirations, competencies, school grades, and ability beliefs, over the course of late adolescence (ages 15 to 18). Based on an outcome-wide longitudinal design for causal inference (see VanderWeele et al., 2020), the study controlled for a comprehensive set of confounder variables to minimize the influence of self-selection. The aim was to estimate causal effects of unstructured OST science activities on the respective outcomes. According to the results of the study, only Investigative vocational interests were robustly impacted by the engagement in unstructured OST science activities.

According to the results of the third study, adolescents who engaged more in unstructured OST science activities had significantly higher Investigative vocational interests, compared to adolescents who engaged less in unstructured OST science activities (i.e., controlling for a variety of possible confounders, such as motivational variables, competencies, and personality traits). As different operationalizations of the dosage of engagement were formed, effects of unstructured OST science activities varied from $d = 0.19$ to $d = 0.58$. Descriptively, treatment operationalization that indicated a higher dosage, generally possessed bigger effect sizes. Negligible effects were found for the remaining interest dimensions. These findings are in line with previous studies, which suggest that content-related school or vocational environments could influence the development of vocational interests (e.g., Etzel & Nagy, 2021; Golle et al., 2019; Meir & Navon, 1992). Besides the main research question about the effects of unstructured OST science engagement on interest development, the third study also provides evidence on interest intensity and stability over the course of late adolescence (ages 15 to 18). According to the findings on re-test correlations, stability of vocational interests was high over the course of the three-year period ($.48 < r < .63$). According to the findings on mean levels, interest intensity descriptively increased for all the interest dimensions ($0.10 < d < 0.25$).

6.2 Evidence on the Developmental Principles in Adolescence

In the following chapter, findings from the three studies on interest intensity, stability, and gender differences are integrated under the umbrella of the three developmental principles. Findings of the three studies provide important insights about the development of vocational interests in adolescence and capture the complete period from late childhood to late adolescence (ages 11 to 18). Although the respective studies differed in properties such as sample composition, measurement instrument, or statistical analysis, they all measured vocational interests with Holland's (1997) interest model—the RIASEC framework. Their results therefore enable description of vocational interest development in terms of mean levels, re-test correlations, profile correlations, and gender differences based on the RIASEC dimensions.

6.2.1 Cumulative Learning Principle

The cumulative learning principle states that vocational interests increase in their stability over the course of adolescence. The solidification of interests is assumed to be a cumulative process that is majorly driven by experiences (Holland, 1997; Su et al., 2019). There are different considerations on why an accumulation of experiences could lead to more stable interests. For example, it is suggested that when adolescents experience more activities that are in line with their interests, they develop a mental representation of their objects of interest (Su et al., 2019). This includes getting a better sense about the activities that seem suitable for them or the activities that may match their abilities (Su et al., 2019). Over time, as adolescents reengage in activities, they become more certain about the activities they like and dislike (Su et al., 2019), which ultimately leads to more stable vocational interests. Consequently, vocational interests evolve from preferences for single activities to relatively stable dispositions as more activities are experienced (Holland, 1997; Su et al., 2019).

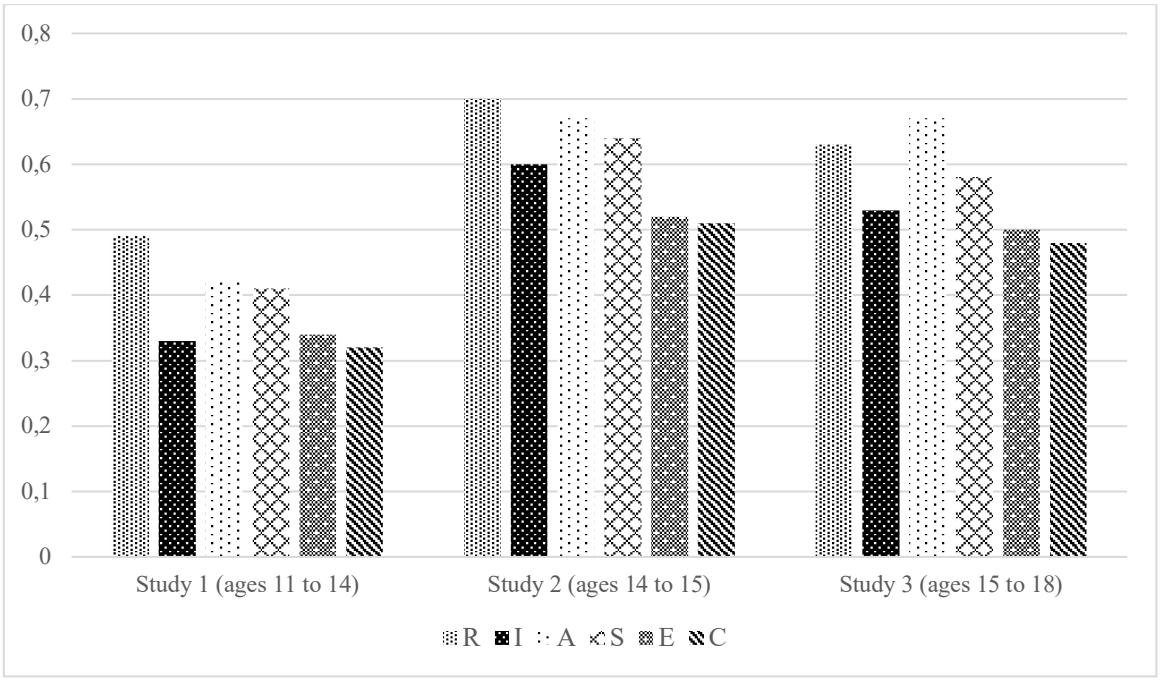
The most robust findings that are in line with the assumption of the cumulative learning principle are provided by the first study of the current dissertation. Re-test correlations as well as trait variance proportions in almost all RIASEC dimensions increased over the course of late childhood and early adolescence. This illustrates that there is an increase in stability over time. Although the reported stability increases in late childhood and early adolescence are smaller in comparison to later life phases, such as late adolescence and young adulthood (e.g., Hoff, Song, Einarsdóttir, et al., 2020), they are comprehensive as four out of six interest dimensions showed statistically significant increases. A similar trend in stability becomes apparent if stability indicators of the RIASEC dimensions from the first and the third study are compared (see

Figure 3). Vocational interests are more stable over the course of the three-year period in late adolescence (Study 3: ages 15 to 18; $.48 < r < .63$) compared to the three-year period from late childhood to early adolescence (Study 1: ages 11 to 14; $.32 < r < .49$; see Figure 3). This is also applicable to findings from the second study, where vocational interests were more stable within samples that consist of participants from later life phases in comparison to samples with participants from earlier life phases. These increases in stability over the course of adolescence are in line with theoretical considerations from Holland (1997) and Su et al. (2019) and meta-analytic evidence (Low et al., 2005).

However, it is important to note that in the first study not all RIASEC dimensions increased in their stability. This implies that there are deviations from the trends in stability that were proposed by the cumulative learning principle—some interest dimensions solidify in a different way. In the respective study (see Study 1), it was discussed that based on considerations of Holland (1997) and the TSID model (Su et al., 2019), it could be assumed that over the course of late childhood and early adolescence certain activities are experienced less often than others. As not all activities that correspond to Holland's (1997) RIASEC dimensions are accessible during late childhood and early adolescence, it could be assumed that the lack of engagement in these activities could lead to differences in the solidification process. For example, Enterprising interests did not increase over the course of late childhood and early adolescence (see Study 1). This might be because Enterprising activities, such as leading a group, selling things, and organizing events, are less prominent for children and adolescents. In comparison to Enterprising interests, Social interests increased in their stability (see Study 1). The Social dimension entails activities such as caring for others, meeting with friends, or listening to people's problems. These activities describe everyday interactions between children and adolescents, which are consequently more prominent in comparison to Enterprising activities.

Figure 3

Re-test Correlations Across the Three Studies



Note. The figure depicts descriptive re-test correlations from the three studies of the current dissertation; Re-test correlations of Study 2 were included from the second substudy, which investigated the time period of middle adolescence (age 14 to 15).

6.2.2 Restriction Initiates Growth Principle

The restriction initiates growth principle states that vocational interests decrease in their intensity from late childhood to early adolescence (ages 11 to 14) and afterwards increase from middle to late adolescence (ages 15 to 18). It is assumed that during late childhood and early adolescence, children and adolescents neglect occupations that do not fit to their developing self-concept (Gottfredson, 1981; Hoff et al., 2018; Tracey, 2002). This process is assumed to lead to overall decreases in interest intensity. After the period where adolescents have restricted their occupational choices, they begin to reengage in activities that they view as suitable in correspondence to their self-concept (Gottfredson, 1981). It was assumed that this process further deepens their vocational interests and leads to increases in interest intensity.

In line with the assumptions of the restriction initiates growth principle, the first study of the current dissertation provides evidence for decreases in vocational interest intensity over the course of late childhood and early adolescence (ages 11 to 14). Mean levels of the majority of the RIASEC dimensions (i.e., Realistic, Investigative, Artistic, and Conventional) decreased during that period. In comparison to current meta-analytic findings on mean-level development (see Hoff et al., 2018), the decreases reported in the first study of the current dissertation were

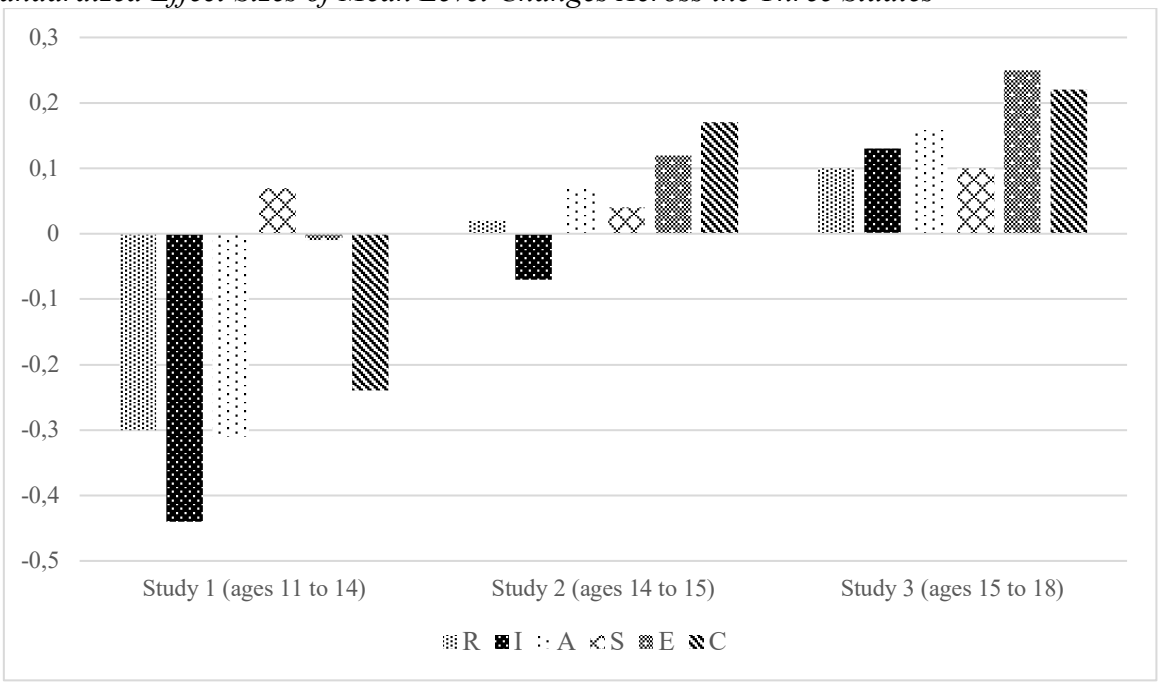
more overarching in terms of effect size magnitude ($-0.44 < \Delta d < -0.24$) and significance (i.e., four out of six interest dimensions decreased significantly). In line with the assumptions of the restriction initiates growth principle, descriptive results from the third study of the current dissertation suggest that vocational interests increase in their intensity over the course of middle to late adolescence (ages 15 to 18). In the third study, all the RIASEC dimensions descriptively increased in terms of effect sizes during the three-year period ($0.10 < \Delta d < 0.25$). This illustrates that the studies included in the current dissertation provide evidence that is in line with the restriction initiates growth principle (see Figure 4).

However, it is important to note that evidence for the two developmental patterns proposed by the restriction initiates growth principle comes from two separate studies. One of the studies investigated the period of late childhood and early adolescence (Study 1: ages 11 to 14), whereas the other study captured the period of middle to late adolescence (Study 3: ages 15 to 18). So far, there is no study that covered the whole transition from late childhood to late adolescence (ages 11 to 18). This implies that studies are still needed that confirm both developmental patterns of the restriction initiates growth principle (i.e., decreases in interests followed by increases in interests) within a sample that follows the same participants from late childhood to late adolescence.

In addition, although all of the RIASEC dimensions increased in terms of effect sizes during late adolescence, not all of the RIASEC dimensions decreased in terms of effect sizes during late childhood and early adolescence. For example, mean-level increases were reported for the interest dimension Social. This indicates that Social interests could follow a different trend than the one proposed by the restriction initiates growth principle. In the respective study (see Study 1), it was discussed that interests that are related to the Social dimension are important for successful integration in social situations and peer groups (Parkhurst & Hopmeyer, 1998). People who are characterized as Social are interested in helping others, like to engage in social situations and are described as being friendly and sociable, properties that are all beneficial in social situations (Holland, 1997). As the importance of peer groups increases over the course of late childhood and early adolescence (Berndt, 1979; Fuligni et al., 2001; Larson & Richards, 1991), Social interests could increase as a consequence of that. This assumption would again relate the development of vocational interest to experiences in social situations (see Study 1).

Figure 4

Standardized Effect Sizes of Mean Level Changes Across the Three Studies



Note. The figure depicts standardized effect sizes of mean-level changes from the three studies of the current dissertation; Mean-level changes of study 2 were included from the second sample, which investigated the time period of middle adolescence (age 14 to 15).

6.2.3 Gender Differentiation Principle

The gender differentiation principle states that gender differences in vocational interests are present during late childhood (age 11), increase over the course of late childhood and early adolescence (ages 11 to 14), and are relatively stable over the remaining course of adolescence (ages 15 and older). One assumption is that gender differences evolve due to the incorporation of gender into the developing self-concept of children and adolescents (Gottfredson, 1981). Over time, children and adolescents are supposed to reject occupations that, according to their view, do not match to their gender (Gottfredson, 1981). Another assumption is that children and adolescents gradually incorporate direct and vicarious experiences that suggest gender stereotypical activities into their developing mental representation of the object of interest (Su et al., 2019). It is assumed that this process solidifies gender differences in dispositional interests.

In line with the gender differentiation principle, the first study provides evidence for the assumption that gender differences are already present around the age of 11. The differences between boys and girls found in this study are in line with meta-analytic evidence that investigated gender differences in vocational interests (Hoff et al., 2018; Su et al., 2009). At the

age of 11, girls were more interested in Artistic and Social activities, whereas boys were more interested in Realistic and Investigative activities. According to the effect sizes these gender differences were already quite large. In addition, the first study also provides robust evidence for the assumption that gender differences increase from late childhood to early adolescence. This finding is also in line with considerations of the gender differentiation principle. Gender differences increased on all interest dimensions except Artistic. Although differences between boys and girls on Artistic interest did not increase from late childhood to early adolescence, gender differences on that variable were already large in terms of effect sizes (i.e., $d = 0.68$) in late childhood.

As gender differentiation in interest intensity was not investigated in the second and third studies of the current dissertation, they do not provide evidence on the development of mean-level differences between boys and girls. This implies that in the current dissertation, the proposed consistency of gender differences after the age of 14 was not investigated. Current meta-analyses from Su et al. (2009) and Hoff et al. (2018) suggest that gender differences are relatively stable from middle adolescence to adulthood. However, instead of mean-level differences, the results of the second study of the current dissertation suggest that gender differences in profile stability might change from middle to late adolescence. Descriptively, gender differences in profile stability were small in the sample from the life phase of late childhood and early adolescence (ages 11 to 14), moderate in the sample from the life phase of middle adolescence (ages 14 to 15), and small again in the sample from the life phase of late adolescence to young adulthood (ages 17 to 23). This finding suggests that differences in stability might increase till middle adolescence and then decrease over the course of late adolescence.

However, the question remains if the finding on gender differences on profile stability can also be applied to gender differences in mean levels. Therefore, to make robust claims about the developmental pattern that was proposed by the gender differentiation principle beyond the age of 14, further longitudinal evidence is needed. For example, as similarly proposed for the restriction initiates growth principle, multiwave longitudinal studies are needed that capture the whole span from late childhood to late adolescence (ages 11 to 18). With such data, it would be possible to investigate the development of gender differences (i.e., in interest intensity and stability) and provide empirical insights for the gender differentiation principle.

6.3 Experiencing Activities as an Initiator for Vocational Interest Development?

Based on considerations from Holland (1997) and the TSID model (Su et al., 2019), it was assumed that the development of vocational interests can be influenced by an accumulation of experiences. In line with the TSID model (Su et al., 2019), it was stated that dispositional interests are expected to develop when experiences initiate situational interest during a certain situation. Over time, as this process repeats, a mental representation of the object of interest is formed and dispositional interests start to evolve (Su et al., 2019). This process implies that if people have different experiences, they may also develop different dispositional interests. In the current dissertation, two factors were identified that could have an impact on the experiences of adolescents—individual characteristics, such as personality traits, cognitive abilities, or gender, and contextual factors, such as leisure-related environments. In line with Holland (1997) and the TSID model (Su et al., 2019), differences in individual and contextual factors could lead to differences in experiences and consequently to differences in vocational interest development. The empirical studies of the current dissertation provide first evidence that individual factors are associated to the stability of vocational interest profiles. In addition, they show that leisure activities can influence the development of vocational interests. In the following, these results are discussed.

The current dissertation provides additional evidence for the assumption that individual characteristics can influence the development of vocational interests. Adding to existing evidence on the variable level (see Hoff, Song, Einarsdóttir, et al., 2020), the second study of the current dissertation provides first information on the profile level. As personality traits, cognitive abilities and gender predict a majority of the behavioral patterns of people (Ackerman & Beier, 2003; Ackerman & Heggestad, 1997; Jackson et al., 2020; Kandler et al., 2014; Päßler & Hell, 2012; Stoll, Einarsdóttir, et al., 2020; Stoll et al., 2017), it was assumed that differences in these individual characteristics might result in different experiences over time and ultimately in different stable vocational interests. Therefore, in the second study of the current dissertation, we investigated the relationship of experiences and profile stability indirectly, by associating individual characteristics that are assumed to initiate differences in experiences to vocational interest stability. Future studies should investigate the link between experiences and profile stability with direct measures of experiences, as the second study included no information about the type and rate of experiences that the participants had. Based on the findings of the current dissertation, it could be assumed that other profile characteristics besides stability are also influenced by individual factors and, consequently, by having experiences. More studies are needed that associate profile properties of vocational interests to individual characteristics. For

example, besides stability, studies can focus on other properties such as the differentiation of the vocational interest profile or the dominant expression of the vocational profile. Profile differentiation captures the dispersion of the expressions of an interest profile (Tracey et al., 2014) and might differ in relation to some personality traits. For example, it could be assumed that people with higher scores on openness to experience have less differentiated profiles because they are characterized by having the need to experience a variety of activities (McCrea & Costa, 1999).

Besides evidence for the influence of individual characteristics, the current dissertation reports evidence for the influence of experiencing activities on the development of vocational interests (see Study 3). This is relatively new evidence as current studies that investigated contextual factors merely had descriptive information on environments, not on the activities that were provided within these environments. For example, Golle et al. (2019) investigated the influence of different career tracks on the development of vocational interests, but they did not investigate the concrete influence of different activities experienced within these environments on vocational interest development. In the current dissertation, the influence of leisure-related science activities was investigated. The results indicated that Investigative vocational interests can be developed by engaging in leisure-related science activities. It was assumed that these leisure-related science activities provided experiences that initiated situational interest. Based on that finding, and the considerations of Study 2, the development of vocational interests over the course of adolescence could be described by a fourth developmental principle—the *experiencing activities initiates development* principle. The principle could state that, over time, as adolescents experience activities that initiate situational interest and reengage in these activities, dispositional interests should develop. This principle could also offer an explanation for individual deviations from the other three developmental principles.

However, the experiencing activities initiates development principle has some assumptions. In line with the TSID model (Su et al., 2019), the principle requires that vocational interests are susceptible to situational characteristics, which can initiate situational interest. This implies that, although vocational interests are seen as dispositions that are consistent across situations (Low et al., 2005; Rounds & Su, 2014), they have to consist also of components that are susceptible to situations. Evidence for that assumption was provided by the first study of the current dissertation. Based on state-trait analyses (Geiser et al., 2015; Steyer et al., 1999, 2015), it was investigated what proportion of vocational interests can be attributed to a long-term developing trait and what proportion of vocational interests can be attributed to situational fluctuations. The proportions suggest that, over the course of late childhood and early

adolescence (ages 11 to 14), vocational interests consist of both components; however, trait components are slightly larger than situational components and increase over time. Therefore, according to the evidence of the first study of the current dissertation, vocational interests possess situational susceptible components, which is in line with the requirements of the fourth developmental principle. Future studies should investigate the state-trait variance components of vocational interests in later stages of adolescence to determine when situational susceptibility decreases and becomes substantially smaller than trait components.

In educational research the role of situational interest is frequently investigated and integrated in a variety of models (Hidi & Renninger, 2006; Reber et al., 2018; Renninger & Hidi, 2011). Recent articles from educational research propose that situational interest could facilitate the application of personalized education (Reber et al., 2018). Based on different situational properties that initiate situational interest, Reber et al. (2018) argue that personalized learning arrangements could foster the evolvement of long-term dispositional interests. For example, they state that based on elements such as content personalization or offering individual choices, educators could plan more individualized learning arrangements that trigger situational interest (Reber et al., 2018). Over time, this approach is supposed to increase dispositional interests in school subjects and the general interest in school (Reber et al., 2018). Although more evidence is needed that provides information on which components of activities initiate situational interest in vocational interests, future studies should investigate if a similar approach such as personalized education could be also applied to foster the development of vocational interests.

Finally, in the introductory section it was described that vocational interests are integrated within the broad framework of personality, implying that they are relatively stable dispositional characteristics. Empirical findings of the current dissertation suggest that vocational interests are also susceptible to situational components (see Study 1). In addition, vocational interest development can be influenced by experiencing activities (see Study 3). The idea that an accumulation of situational expressions (i.e., states) can lead to a long-term development of personality characteristics is currently discussed in personality trait research (Geukes et al., 2018; Roberts & Jackson, 2008; Wrzus & Roberts, 2017). There are personality models that imply that personality traits might change based on repeated, short-term experiences of daily behavior in situations (Geukes et al., 2018; Roberts & Jackson, 2008; Wrzus & Roberts, 2017). Empirical evidence for that assumption is provided by Quintus et al. (2021) and Roberts et al. (2017). The findings of the current dissertation suggest that similar theoretical considerations may apply to vocational interests.

6.4 Practical Implications

The topic of vocational interest development is important not only for researchers, but also for practitioners. The current dissertation provides evidence that can be used to facilitate the implementation of interventions that target vocational interests. Besides indicating which life periods might be suitable for the implementation of an intervention, insights were generated on how to design an intervention to foster interest in certain areas.

Evidence about re-test correlations and profile correlations indicate that vocational interests start to stabilize already over the course of late childhood and early adolescence (see Study 1 and Study 2). In addition, vocational interests seem to be less stable in younger life phases compared to later life phases (see Study 2). This applies to both the between-person (see Study 1) and within-person perspectives (see Study 2). Accordingly, the findings of the dissertation indicate that interventions might be more impactful during earlier life phases, such as late childhood and early adolescence, when interests are still forming, in comparison to later life phases, such as young adulthood (e.g., Stoll, Rieger, et al., 2020), when interests are relatively stable.

Information on the development of gender differences in vocational interests could be important for institutions and practitioners that want to encourage women to pursue a career in science, technology, engineering, and mathematics (STEM). Gender differences in interests are often mentioned as an influencing factor for the gender disparity in the STEM fields, as women are less likely to choose a STEM study major due to less Investigative and Realistic interests (Su et al., 2009). This is because Realistic and Investigative vocational interests are strong predictors for the choice of a study major in these areas (Päßler & Hell, 2012; Wille et al., 2020). The current dissertation provides evidence for the development of gender differences (see Study 1), which indicates that differences in interest intensity between girls and boys are already increasing over the course of late childhood and early adolescence, especially for the interest dimensions Realistic and Investigative. In addition, vocational interest profiles of girls may be more stable in comparison to the ones of boys during that period (see Study 2). This illustrates that an intervention aiming to reduce gender differences in Realistic and Investigative interests should be implemented during early life phases before gender differences start to develop and stabilize.

Besides implications for the optimal time to implement an intervention, the current dissertation also provides evidence on how to design an intervention to foster interest in certain areas. So far, there is a limited number of studies that tried to initiate vocational interests based

on interventions (e.g., Betz & Schifano, 2000; Turner & Lapan, 2005). Most intervention studies targeted the association between self-efficacy (i.e., an individual's belief in their ability to execute certain behaviors; Bandura & Walters, 1963) and vocational interests (e.g., Betz & Schifano, 2000; Lent et al., 1994; Turner & Lapan, 2005). By improving self-efficacy in a certain area, it was assumed that vocational interests increase as well (Betz & Schifano, 2000; Lent et al., 1994; Turner & Lapan, 2005), because people start to associate certain activities with positive ability beliefs. For example, Betz and Schifano (2000) tried to improve self-efficacy of men and women in Realistic activities by encouraging and supporting them while they performed Realistic activities. Although they reported no significant effects on the scale level, there were small effects on Realistic interests on the item level. Especially interests in items that directly corresponded to the activities of the intervention (e.g., screwing in a lightbulb) were influenced.

Based on the findings of the current dissertation, further implications for interventions could be derived. The dissertation provides evidence that the long-term engagement in leisure science activities can influence the development of vocational interests. Participants who engaged more in leisure science activities developed more pronounced interest in Realistic and Investigative activities, compared to participants who engaged less in leisure science activities (see Study 3). Based on processes that were proposed by the TSID model (Su et al., 2019), it was assumed that leisure activities can initiate situational interest. Over time, as engagement in leisure activities repeats, situational interest will accumulate (Su et al., 2019). This process refines the mental representation of the object of interest, leading to the development of dispositional interests in science (Su et al., 2019). Interventions aiming to increase vocational interests in science could therefore focus on activities that initiate situational interest. This could be achieved by a multiplicity of approaches, such as including "incongruous, surprising information; character identification or personal relevance" (Hidi & Renninger, 2006, p. 114); activity novelty; or utility value of the science activity.

The TSID model states that besides direct experiences also vicarious experiences can initiate situational interest (Su et al., 2019). This implies that not only conducting an activity, but also watching an activity can trigger situational interest. Based on that assumption, interest development could also be initiated through digital offers, such as videos on the internet that describe the planning, execution, and interpretation of a science experiment. Empirical findings for that assumption are provided by the current dissertation (see Study 3). For example, leisure science activities such as researching on the internet about science or watching a science TV show influenced the development of Investigative vocational interests. Both of these types of

activities produce mostly vicarious instead of direct experiences. However, it is still an open empirical question if interventions of vicarious and direct experiences have the same effect on the development of vocational interests as experiences from independently chosen activities. Future studies are needed to investigate that research question.

6.5 Limitations and Outlook

There are some limitations that should be considered when interpreting the results of the current dissertation. All the participants included in the current dissertation lived in Germany. In Germany, adolescents have the possibility to leave school around the ages 15 to 16 to enter the labor market and to pursue vocational educational training. During vocational educational training, people attend a vocational school, while simultaneously working in an occupation (BIBB, 2018). This form of education is very prominent, especially in younger age groups, as approximately 55,000 adolescents below age 16 start vocational educational training in Germany each year (BIBB, 2018). Such idiosyncrasies of the economic system of a country could influence the development of vocational interests and ultimately the three developmental principles. Theoretical processes, such as Gottfredson's (1981) phase of compromise, could be influenced by these external factors, as adolescents in Germany might be confronted with the choice of a suitable career path earlier compared to adolescents from other countries, such as the United States. Although empirical findings suggest that the developmental patterns of vocational interests are similar between adolescents from the United States (Hoff et al., 2018; Low et al., 2005) and adolescents from Germany, there might be differences in the magnitude of changes (e.g., larger decreases in mean-levels of vocational interests were reported for adolescents in Germany, as seen in the first study of the current dissertation), which could be partly explained by the earlier need of choosing a career path. However, to answer such questions empirically, future studies should compare the developmental principles across different educational and economical systems.

The current dissertation investigated three developmental principles that stated certain developmental patterns of vocational interests over the course of adolescence. All the empirical studies included in the dissertation investigated these developmental patterns by focusing on chronological age, which is defined by the number of years that a person has been alive. However, Low et al. (2005) argued that in longitudinal research, besides chronological age, also different age markers should be investigated. They suggest further operationalizations of age (e.g., biological age, social age, and psychological age; Birren & Cunningham, 1985) that account for different aspects of human development (Low et al., 2005). For example, biological age refers to age measured based on physiological markers. Social age indicates the maturation of a person on several skills that are relevant to fulfill the norms and roles in a society. Psychological age refers to the subjective age-equivalent of a person, answering the question of how old someone feels. As many of the propositions of the three developmental principles in the current dissertation orient themselves on age (e.g., phases of circumscription and

compromise; Gottfredson, 1981), their empirical investigation would benefit from a multiplicity of different age markers. Other operationalizations of age such as social age might offer different developmental patterns. As boys and girls might differ in their maturation process (e.g., the onset of puberty) over the course of adolescence (Crone & Dahl, 2012), a biological age indicator could provide additional insights for different developmental patterns. Future studies should therefore investigate the three developmental principles of vocational interest development in adolescence with different age markers.

It is important to note that in the current dissertation, vocational interests were measured with interest inventories that focused on assessing adolescents' preferences for certain activities. The first study used a self-created inventory that included items from the Revised General Interest Structure-Test (AIST-R; Allgemeiner Interessen Strukturtest; Bergmann & Eder, 2005), which is the most often used inventory in German-speaking countries; items from the German version of the Inventory of Children's Activities (ICA; Tracey & Ward, 1998; German version [ICA-D]: von Maurice, 2006); and a few newly constructed items (see Gfrörer et al., 2021). The second study used multiple inventories (i.e., the inventory that was used in the first study and the unmodified AIST-R inventory) because it investigated profile stability based on various samples during various life phases. The third study was based on data from the National Educational Panel Study (NEPS, 2021), which measured vocational interests based on a self-constructed questionnaire that included items from the AIST-R and the ICA-D. As discussed in Study 1, it was recommended by Tracey (2002) that children's vocational interest should be measured based on their preferences for "familiar activities, rather than unfamiliar occupations" (p. 149). Although research has agreed that this seems to be the appropriate way to measure vocational interests in adolescence (Tracey, 2002; Tracey & Caulum, 2015; Tracey & Ward, 1998; von Maurice, 2006), these interest inventories differ from interest inventories used in older age groups. For example, in the Self-Directed Search inventory of Holland (1994), participants have to self-report their likes and dislikes for certain activities, but also for specific occupations (Hansen, 2019). This element of stating likes and dislikes for a specific occupation is not present in the interest inventories of the current dissertation. It could be argued that stating preferences for occupations is more related to external factors, such as gross income or social status, in comparison to stating preferences for activities. Future studies should investigate if similar developmental trends in vocational interests could be expected when adolescents must state their likes and dislikes for certain occupations. Because children's view of occupations first needs to develop, this investigation might be less fruitful during late childhood (age 11).

However, it should be feasible during middle and late adolescence, when adolescents have a better sense of the world of work.

Finally, the current dissertation investigated the influence of experiencing activities on the development of vocational interests. Although there was robust evidence that engagement in activities can influence the development of vocational interests (see Study 3), relatively little is known about the mechanisms behind that process. For example, in the third study, it was assumed that leisure-related activities can initiate situational interests because of certain situational properties. To investigate if situational properties are also responsible for initiating situational interest in vocational interests, other research designs might be needed. For example, to investigate why leisure-related science activities influenced Investigative vocational interests, intensive longitudinal data based on an experience sampling design might be quite insightful. Adolescents should state and describe the characteristics of situations, the leisure-related activities that they experienced, and the feelings that they had during these activities over a period of several weeks. With such information, short-term increases or decreases in vocational interests could be associated with situational and activity-related characteristics. Based on such a design, future studies could therefore investigate which properties of activities have the biggest impact on triggering situational interest in vocational interests and initiating their long-term development.

6.6 Conclusion

The current dissertation investigated the development of vocational interests over the course of adolescence. Three empirical studies examined the development of vocational interests from late childhood to early adolescence, the profile stability of vocational interests and its predictors, as well as the influence of leisure science activities on the development of vocational interests. Evidence from the three empirical studies on vocational interest intensity, stability, and gender differentiation was integrated based on three developmental principles. In line with the cumulative learning principle, evidence suggests that vocational interests generally increased in stability over the course of adolescence. In line with the restriction initiates growth principle, evidence suggests that vocational interests generally decreased in mean levels from late childhood to early adolescence and increased from middle to late adolescence. In line with the gender differentiation principle, evidence suggests that gender differences in vocational interests increased over the course of late childhood and early adolescence. Besides these descriptive insights, robust evidence is provided for the relationship between gender and profile stability, with girls and women having more stable profiles than boys and men. Personality traits and cognitive abilities seem to be less influential regarding profile stability. It is also suggested that environmental factors, such as the engagement in leisure-related activities, can initiate the development of dispositional interests. Adolescents who engaged more in leisure-related science activities developed more pronounced Investigative interests. Experiencing activities was discussed as a fourth developmental principle that could be responsible for individual deviations from the other three developmental principles. Based on the empirical evidence, practical implications for interest interventions were derived. If vocational interests should be fostered in a certain area, interventions should occur in earlier life phases, when vocational interests are assumed to be less stable. In addition, interventions that have the purpose to develop dispositional interests in a certain area should include activities that initiate and maintain situational interest.

7 REFERENCES

Ackerman, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence*, *22*(2), 227–257. https://doi.org/10.1016/S0160-2896(96)90016-1

Ackerman, P. L. (1997). Personality, Self-Concept, Interests, and Intelligence: Which Construct Doesn't Fit? *Journal of Personality*, *65*(2), 171–205. https://doi.org/10.1111/j.1467-6494.1997.tb00952.x

Ackerman, P. L., & Beier, M. E. (2003). Intelligence, Personality, and Interests in the Career Choice Process. *Journal of Career Assessment*, *11*(2), 205–218. https://doi.org/10.1177/1069072703011002006

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, Personality, and Interests: Evidence for Overlapping Traits. *Psychological Bulletin*, *121*(2), 219–245. https://doi.org/10.1037/0033-2909.121.2.219

Allen, J., & Robbins, S. (2010). Effects of interest–major congruence, motivation, and academic performance on timely degree attainment. *Journal of Counseling Psychology*, *57*(1), 23–35. https://doi.org/10.1037/a0017267

Allen, J., & Robbins, S. B. (2008). Prediction of college major persistence based on vocational interests, academic preparation, and first-year academic performance. *Research in Higher Education*, *49*, 62–79. https://doi.org/10.1007/s11162-007-9064-5

American College Testing Program. (2009). *The ACT Interest Inventory Technical Manual*. https://www.act.org/content/dam/act/unsecured/documents/ACTInterestInventoryTechnicalManual.pdf

Anderson, M. Z., Tracey, T. J. G., & Rounds, J. (1997). Examining the Invariance of Holland's Vocational Interest Model across Gender. *Journal of Vocational Behavior*, *50*(3), 349–364. https://doi.org/10.1006/jvbe.1996.1550

Armstrong, P. I., Day, S. X., McVay, J. P., & Rounds, J. (2008). Holland's RIASEC model as an integrative framework for individual differences. *Journal of Counseling Psychology*, *55*(1), 1–18. https://doi.org/10.1037/0022-0167.55.1.1

Asendorpf, J. B. (1992). Beyond stability: Predicting inter-individual differences in intra-individual change. *European Journal of Personality*, *6*(2), 103–117. https://doi.org/10.1002/per.2410060204

Ashton, M. C., & Lee, K. (2007). Empirical, Theoretical, and Practical Advantages of the HEXACO Model of Personality Structure. *Personality and Social Psychology Review*, *11*(2), 150–166. https://doi.org/10.1177/1088868306294907

Bandura, A., & Walters, R. H. (1963). *Social learning and personality development*. Holt Rinehart and Winston.

Barrick, M. R., Mount, M. K., & Gupta, R. (2003). Meta-Analysis of the relationship between the Five-Factor model of personality and Holland's occupational types. *Personnel Psychology*, *56*(1), 45–74. https://doi.org/10.1111/j.1744-6570.2003.tb00143.x

Beltz, A. M., Swanson, J. L., & Berenbaum, S. A. (2011). Gendered occupational interests: Prenatal androgen effects on psychological orientation to Things versus People. *Hormones and Behavior*, *60*(4), 313–317. https://doi.org/10.1016/j.yhbeh.2011.06.002

Bergmann, C., & Eder, F. (2005). *Allgemeiner Interessen-Struktur-Test mit Umwelt-Struktur-Test (UST-R)—Revision [General Interest Structure Test and Environmental Structure Test—Revision]*. Beltz.

Berndt, T. J. (1979). Developmental changes in conformity to peers and parents. *Developmental Psychology*, *15*(6), 608–616. https://doi.org/10.1037/0012-1649.15.6.608

Betsworth, D. G., Bouchard, Jr., Thomas J., Cooper, C. R., Grotevant, H. D., Hansen, J.-I. C., Scarr, S., & Weinberg, R. A. (1994). Genetic and Environmental Influences on Vocational Interests Assessed Using Adoptive and Biological Families and Twins Reared Apart and Together. *Journal of Vocational Behavior*, *44*(3), 263–278. https://doi.org/10.1006/jvbe.1994.1018

Betz, N. E., & Schifano, R. S. (2000). Evaluation of an Intervention to Increase Realistic Self-Efficacy and Interests in College Women. *Journal of Vocational Behavior*, *56*(1), 35–52. https://doi.org/10.1006/jvbe.1999.1690

BIBB (Ed.). (2018). *Datenreport zum Berufsbildungsbericht 2018. Informationen und Analysen zur Entwicklung der beruflichen Bildung* (1st ed.). Bundesinstitut für Berufsbildung.

Birren, J. E., & Cunningham, W. R. (1985). Research on the psychology of aging: Principles, concepts and theory. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 3–34). Van Nostrand Reinhold Co.

Bishop, J., Geiser, C., & Cole, D. A. (2015). Modeling latent growth with multiple indicators: A comparison of three approaches. *Psychological Methods*, *20*(1), 43–62. https://doi.org/10.1037/met0000018

Bordin, E. S. (1943). A Theory of Vocational Interests as Dynamic Phenomena'. *Educational and Psychological Measurement*, *3*(1), 46–62. https://doi.org/10.1177/001316444300300105

Braun, L., Göllner, R., Rieger, S., Trautwein, U., & Spengler, M. (2020). How state and trait versions of self-esteem and depressive symptoms affect their interplay: A longitudinal experimental investigation. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspp0000295

Brown, D. (Ed.). (2002). *Career choice and development* (4th ed.). John Wiley & Sons, Inc.

Cheng, H., & Furnham, A. (2012). Childhood cognitive ability, education, and personality traits predict attainment in adult occupational prestige over 17 years. *Journal of Vocational Behavior*, *81*(2), 218–226. https://doi.org/10.1016/j.jvb.2012.07.005

Choudhury, S., Blakemore, S.-J., & Charman, T. (2006). Social cognitive development during adolescence. *Social Cognitive and Affective Neuroscience*, *1*(3), 165–174. https://doi.org/10.1093/scan/nsl024

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). L. Erlbaum Associates.

Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social–affective engagement and goal flexibility. *Nature Reviews Neuroscience*, *13*(9), 636–650. https://doi.org/10.1038/nrn3313

Darcy, M. U. A., & Tracey, T. J. G. (2007). Circumplex structure of Holland's RIASEC Interests across gender and time. *Journal of Counseling Psychology*, *54*(1), 17–31. https://doi.org/10.1037/0022-0167.54.1.17

Dewey, J. (1913). *Interest and Effort in Education.* Houghton Mifflin.

DeYoung, C. G., Weisberg, Y. J., Quilty, L. C., & Peterson, J. B. (2013). Unifying the Aspects of the Big Five, the Interpersonal Circumplex, and Trait Affiliation: Big Five and IPC. *Journal of Personality*, *81*(5), 465–475. https://doi.org/10.1111/jopy.12020

Donohue, R. (2006). Person-environment congruence in relation to career change and career persistence. *Journal of Vocational Behavior*, *68*(3), 504–515. https://doi.org/10.1016/j.jvb.2005.11.002

Dunkleberger, C. J., & Tyler, L. E. (1961). Interest stability and personality traits. *Journal of Counseling Psychology*, *8*(1), 70–74. https://doi.org/10.1037/h0046021

Einarsdóttir, S., & Rounds, J. (2009). Gender bias and construct validity in vocational interest measurement: Differential item functioning in the Strong Interest Inventory. *Journal of Vocational Behavior*, *74*(3), 295–307. https://doi.org/10.1016/j.jvb.2009.01.003

Einarsdóttir, S., & Rounds, J. (2020). A Quantitative Review of Gender Differences in Vocational Interests in Iceland: Pervasive and Persistent. *Nordic Journal of Transitions, Careers and Guidance*, *1*(1), 10–26. https://doi.org/10.16993/njtcg.29

Etzel, J. M., Lüdtke, O., Wagner, J., & Nagy, G. (2019). Similarity of vocational interest profiles within families: A person-centered approach for examining associations between circumplex profiles. *Journal of Personality*, *87*(3), 593–606. https://doi.org/10.1111/jopy.12418

Etzel, J. M., & Nagy, G. (2021). Stability and change in vocational interest profiles and interest congruence over the course of vocational education and training. *European Journal of Personality*, *35*(4), 534–556. https://doi.org/10.1177/08902070211014015

Eurostat. (2021, January 5). *Share of students in vocational education programs in 2018.* Eurostat Statistics Explained. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Share_of_students_in_vocational_education_programmes,_2018_(%25)_ET2020.png

Fleeson, W. (2004). Moving Personality Beyond the Person-Situation Debate: The Challenge and the Opportunity of Within-Person Variability. *Current Directions in Psychological Science*, *13*(2), 83–87. https://doi.org/10.1111/j.0963-7214.2004.00280.x

Fleeson, W., & Gallagher, P. (2009). The implications of Big Five standing for the distribution of trait manifestation in behavior: Fifteen experience-sampling studies and a meta-

analysis. *Journal of Personality and Social Psychology*, *97*(6), 1097–1114. https://doi.org/10.1037/a0016786

Fleeson, W., & Jayawickreme, E. (2015). Whole Trait Theory. *Journal of Research in Personality*, *56*, 82–92. https://doi.org/10.1016/j.jrp.2014.10.009

Foster, E. M. (2010). Causal inference and developmental psychology. *Developmental Psychology*, *46*(6), 1454–1480. https://doi.org/10.1037/a0020204

Fouad, N. A. (1999). Validity evidence for interest inventories. In M. L. Savickas & A. R. Spokane (Eds.), *Vocational interests: Meaning, measurement, and counseling use* (pp. 193–209). Davies-Black Publishing.

Fuligni, A. J., Eccles, J. S., Barber, B. L., & Clements, P. (2001). Early adolescent peer orientation and adjustment during high school. *Developmental Psychology*, *37*(1), 28–36. https://doi.org/10.1037/0012-1649.37.1.28

Gati, I. (1979). A hierarchical model for the structure of vocational interests. *Journal of Vocational Behavior*, *15*(1), 90–106. https://doi.org/10.1016/0001-8791(79)90021-6

Geiser, C., Keller, B. T., Lockhart, G., Eid, M., Cole, D. A., & Koch, T. (2015). Distinguishing state variability from trait change in longitudinal data: The role of measurement (non)invariance in latent state-trait analyses. *Behavior Research Methods*, *47*(1), 172–203. https://doi.org/10.3758/s13428-014-0457-z

Geiser, C., & Lockhart, G. (2012). A comparison of four approaches to account for method effects in latent state–trait analyses. *Psychological Methods*, *17*(2), 255–283. https://doi.org/10.1037/a0026977

Geukes, K., van Zalk, M., & Back, M. D. (2018). Understanding personality development: An integrative state process model. *International Journal of Behavioral Development*, *42*(1), 43–51. https://doi.org/10.1177/0165025416677847

Gfrörer, T., Stoll, G., Rieger, S., Trautwein, U., & Nagengast, B. (2021). The Development of Vocational Interests in Early Adolescence: Stability, Change, and State-Trait Components. *European Journal of Personality*, 089020702110356. https://doi.org/10.1177/08902070211035630

Golle, J., Rose, N., Göllner, R., Spengler, M., Stoll, G., Hübner, N., Rieger, S., Trautwein, U., Lüdtke, O., Roberts, B. W., & Nagengast, B. (2019). School or Work? The Choice May

Change Your Personality. *Psychological Science*, *30*(1), 32–42. https://doi.org/10.1177/0956797618806298

Gottfredson, L. S. (1981). Circumscription and compromise: A developmental theory of occupational aspirations. *Journal of Counseling Psychology*, *28*(6), 545–579. https://doi.org/10.1037/0022-0167.28.6.545

Gupta, S., Tracey, T. J. G., & Gore, P. A. (2008). Structural examination of RIASEC scales in high school students: Variation across ethnicity and method. *Journal of Vocational Behavior*, *72*(1), 1–13. https://doi.org/10.1016/j.jvb.2007.10.013

Gurtman, M. B. (1993). Constructing Personality Tests to Meet a Structural Criterion: Application of the Interpersonal Circumplex. *Journal of Personality*, *61*(2), 237–263. https://doi.org/10.1111/j.1467-6494.1993.tb01033.x

Gurtman, M. B., & Balakrishnan, J. D. (1998). Circular Measurement Redux: The Analysis and Interpretation of Interpersonal Circle Profiles. *Clinical Psychology: Science and Practice*, *5*(3), 344–360. https://doi.org/10.1111/j.1468-2850.1998.tb00154.x

Hansen, J. C. (1984). The measurement of vocational interests: Issues and future directions. In S. B. Brown & R. W. Lent (Eds.), *Handbook of counseling psychology* (pp. 99–136). John Wiley & Sons.

Hansen, J. C. (2019). Interest inventories. In G. Goldstein, D. N. Allen, & J. DeLuca (Eds.), *Handbook of Psychological Assessment* (4th ed., pp. 169–190). Elsevier. https://doi.org/10.1016/B978-0-12-802203-0.00006-7

Harris, J. A., Vernon, P. A., Johnson, A. M., & Jang, K. L. (2006). Phenotypic and genetic relationships between vocational interests and personality. *Personality and Individual Differences*, *40*(8), 1531–1541. https://doi.org/10.1016/j.paid.2005.11.024

Harter, S., Whitesell, N. R., & Kowalski, P. (1992). Individual Differences in the Effects of Educational Transitions on Young Adolescent's Perceptions of Competence and Motivational Orientation. *American Educational Research Journal*, *29*(4), 777–807. https://doi.org/10.3102/00028312029004777

Hertzog, C., & Nesselroade, J. R. (1987). Beyond autoregressive models: Some implications of the trait-state distinction for the structural modeling of developmental change. *Child Development*, *58*(1), 93–109. https://doi.org/10.2307/1130294

Hidi, S., & Renninger, K. A. (2006). The Four-Phase Model of Interest Development. *Educational Psychologist*, *41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Hirschi, A. (2010). Individual predictors of adolescents' vocational interest stabilities. *International Journal for Educational and Vocational Guidance*, *10*(1), 5–19. https://doi.org/10.1007/s10775-009-9171-2

Hoff, K. A., Briley, D. A., Wee, C. J. M., & Rounds, J. (2018). Normative changes in interests from adolescence to adulthood: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *144*(4), 426–451. https://doi.org/10.1037/bul0000140

Hoff, K. A., Perlus, J. G., & Rounds, J. (2019). Vocational Interests: Revisiting Assumptions About Their Development and What They Predict. In J. A. Athanasou & H. N. Perera (Eds.), *International Handbook of Career Guidance* (2nd ed., pp. 673–692). Springer, Cham. https://doi.org/10.1007/978-3-030-25153-6_31

Hoff, K. A., Song, Q. C., Einarsdóttir, S., Briley, D. A., & Rounds, J. (2020). Developmental structure of personality and interests: A four-wave, 8-year longitudinal study. *Journal of Personality and Social Psychology*, *118*(5), 1044–1064. https://doi.org/10.1037/pspp0000228

Hoff, K. A., Song, Q. C., Wee, C. J. M., Phan, W. M. J., & Rounds, J. (2020). Interest fit and job satisfaction: A systematic review and meta-analysis. *Journal of Vocational Behavior*, *123*, 103503. https://doi.org/10.1016/j.jvb.2020.103503

Holland, J. L. (1994). *Self-directed search form R: 1994 edition.* Psychological Assessment Resources.

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Psychological Assessment Resources.

Hong, E., Milgram, R. M., & Whiston, S. C. (1993). Leisure activities in adolescents as a predictor of occupational choice in young adults: A longitudinal study. *Journal of Career Development*, *19*(3), 221–229. https://doi.org/10.1007/BF01353280

Hoyt, D. P., Smith, J. L., & Levy, S. (1957). A further study in the prediction of interest stability. *Journal of Counseling Psychology*, *4*(3), 228–233. https://doi.org/10.1037/h0046264

Jackson, J. J., Hill, P. L., Payne, B. R., Parisi, J. M., & Stine-Morrow, E. A. L. (2020). Linking openness to cognitive ability in older adulthood: The role of activity diversity. *Aging & Mental Health*, *24*(7), 1079–1087. https://doi.org/10.1080/13607863.2019.1655705

Kandler, C., Bleidorn, W., Riemann, R., Angleitner, A., & Spinath, F. M. (2011). The Genetic Links Between the Big Five Personality Traits and General Interest Domains. *Personality and Social Psychology Bulletin*, *37*(12), 1633–1643. https://doi.org/10.1177/0146167211414275

Kandler, C., Zimmermann, J., & McAdams, D. P. (2014). Core and surface characteristics for the description and theory of personality differences and development: Core and surface personality characteristics. *European Journal of Personality*, *28*(3), 231–243. https://doi.org/10.1002/per.1952

Keijsers, L., & Poulin, F. (2013). Developmental changes in parent–child communication throughout adolescence. *Developmental Psychology*, *49*(12), 2301–2308. https://doi.org/10.1037/a0032217

Kerby, D. S., & Ragan, K. M. (2002). Activity Interests and Holland's Riasec System in Older Adults. *The International Journal of Aging and Human Development*, *55*(2), 117–139. https://doi.org/10.2190/W0G9-NBYN-H6WC-LTDN

Larson, L. M., Rottinghaus, P. J., & Borgen, F. H. (2002). Meta-analyses of Big Six Interests and Big Five Personality Factors. *Journal of Vocational Behavior*, *61*(2), 217–239. https://doi.org/10.1006/jvbe.2001.1854

Larson, R. (2000). Toward a psychology of positive youth development. *American Psychologist*, *55*(1), 170–183. https://doi.org/10.1037//0003-066X,55.1.170

Larson, R., & Richards, M. H. (1991). Daily companionship in late childhood and early adolescence: Changing developmental contexts. *Child Development*, *62*(2), 284–300. https://doi.org/10.1111/j.1467-8624.1991.tb01531.x

Le, H., Robbins, S. B., & Westrick, P. (2014). Predicting student enrollment and persistence in college STEM fields using an expanded P-E fit framework: A large-scale multilevel study. *Journal of Applied Psychology*, *99*(5), 915–947. https://doi.org/10.1037/a0035998

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, *45*(1), 79–122. https://doi.org/10.1006/jvbe.1994.1027

Leuty, M. E., Hansen, J.-I. C., & Speaks, S. Z. (2016). Vocational and Leisure Interests: A Profile-Level Approach to Examining Interests. *Journal of Career Assessment*, *24*(2), 215–239. https://doi.org/10.1177/1069072715580321

Livingston, R. B., Jennings, E., Reynolds, C. R., & Gray, R. M. (2003). Multivariate analyses of the profile stability of intelligence tests: High for IQs, low to very low for subtest analyses. *Archives of Clinical Neuropsychology*, *18*(5), 487–507. https://doi.org/10.1093/arclin/18.5.487

Low, K. S. D., & Rounds, J. (2007). Interest change and continuity from early adolescence to middle adulthood. *International Journal for Educational and Vocational Guidance*, *7*(1), 23–36. https://doi.org/10.1007/s10775-006-9110-4

Low, K. S. D., Yoon, M., Roberts, B. W., & Rounds, J. (2005). The stability of vocational interests from early adolescence to middle adulthood: A quantitative review of longitudinal studies. *Psychological Bulletin*, *131*(5), 713–737. https://doi.org/10.1037/0033-2909.131.5.713

Lubinski, D. (2000). Scientific and Social Significance of Assessing Individual Differences: "Sinking Shafts at a Few Critical Points". *Annual Review of Psychology*, *51*(1), 405–444. https://doi.org/10.1146/annurev.psych.51.1.405

McCrea, R. R., & Costa, P. T. (1999). A Five-Factor Theory of Personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 139–153). Guilford.

McGue, M., Elkins, I., Walden, B., & Iacono, W. G. (2005). Perceptions of the Parent-Adolescent Relationship: A Longitudinal Investigation. *Developmental Psychology*, *41*(6), 971–984. https://doi.org/10.1037/0012-1649.41.6.971

McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods*, *22*(1), 114–140. https://doi.org/10.1037/met0000078

Meir, E. I., & Navon, M. (1992). A longitudinal examination of congruence hypotheses. *Journal of Vocational Behavior*, *41*(1), 35–47. https://doi.org/10.1016/0001-8791(92)90037-Z

Miller, M. J. (1991). Accuracy of the Leisure Activities Finder: Expanding Holland's typology. *Journal of Vocational Behavior*, *39*(3), 362–368. https://doi.org/10.1016/0001-8791(91)90044-M

Mischel, W., Shoda, Y., & Mendoza-Denton, R. (2002). Situation-Behavior Profiles as a Locus of Consistency in Personality. *Current Directions in Psychological Science*, *11*(2), 50–54. https://doi.org/10.1111/1467-8721.00166

Moloney, D. P., Bouchard, T. J., & Segal, N. L. (1991). A Genetic and environmental analysis of the vocational interests of monozygotic and dizygotic twins reared apart. *Journal of Vocational Behavior*, *39*(1), 76–109. https://doi.org/10.1016/0001-8791(91)90005-7

Monahan, K. C., Lee, J. M., & Steinberg, L. (2011). Revisiting the Impact of Part-Time Work on Adolescent Adjustment: Distinguishing Between Selection and Socialization Using Propensity Score Matching: Revisiting the Impact of Part-Time Work on Adolescent Adjustment. *Child Development*, *82*(1), 96–112. https://doi.org/10.1111/j.1467-8624.2010.01543.x

Mount, M. K., Barrick, M. R., Scullen, S. M., & Rounds, J. (2005). Higher Order Dimensions of the Big Five Personality Traits and the Big Six Vocational Interest Types. *Personnel Psychology*, *58*(2), 447–478. https://doi.org/10.1111/j.1744-6570.2005.00468.x

Nagy, G., Marsh, H. W., Lüdtke, O., & Trautwein, U. (2009). Representing the circles in our minds: Confirmatory factor analysis of circumplex structures and profiles. In T. Teo & M. S. Khine (Eds.), *Structural equation modeling in educational research: Concepts and applications* (pp. 287–315). Sense Publishers.

Nagy, G., Trautwein, U., & Lüdtke, O. (2010). The structure of vocational interests in Germany: Different methodologies, different conclusions. *Journal of Vocational Behavior*, *76*(2), 153–169. https://doi.org/10.1016/j.jvb.2007.07.002

NEPS. (2021, March 20). *National Educational Panel Study (NEPS): Starting Cohort Grade 5. Study Overview*. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/10-0-0/SC3_Study_Overview_W1-10_en.pdf

Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2012). Vocational interests and performance: A quantitative summary of over 60 years of research. *Perspectives on Psychological Science*, *7*(4), 384–403. https://doi.org/10.1177/1745691612449021

Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2017). Interest congruence and performance: Revisiting recent meta-analytic findings. *Journal of Vocational Behavior*, *98*, 138–151. https://doi.org/10.1016/j.jvb.2016.11.002

O'Brien, & Bierman. (1988). Conceptions and Perceived Influence of Peer Groups: Interviews with Preadolescents and Adolescents. *Child Development*, *59*(5), 1360–1365. https://doi.org/10.2307/1130498

O*NET. (2021, February 14). *O*NET Resource Center—Interest table.* https://www.onetcenter.org/dictionary/22.2/excel/interests.html

Parkhurst, J. T., & Hopmeyer, A. (1998). Sociometric popularity and peer-perceived popularity: Two distinct dimensions of peer status. *The Journal of Early Adolescence*, *18*(2), 125–144. https://doi.org/10.1177/0272431698018002001

Päßler, K., & Hell, B. (2012). Do interests and cognitive abilities help explain college major choice equally well for women and men? *Journal of Career Assessment*, *20*(4), 479–496. https://doi.org/10.1177/1069072712450009

Päßler, K., & Hell, B. (2020). Stability and change in vocational interests from late childhood to early adolescence. *Journal of Vocational Behavior*, *121*, 103462. https://doi.org/10.1016/j.jvb.2020.103462

Perera, H. N., & McIlveen, P. (2018). Vocational interest profiles: Profile replicability and relations with the STEM major choice and the Big-Five. *Journal of Vocational Behavior*, *106*, 84–100. https://doi.org/10.1016/j.jvb.2017.11.012

Piaget, J. (1940). The mental development of the child. In D. Elkind (Ed.), *Six psychological studies* (pp. 3–73). Random House.

Prediger, D. J. (1982). Dimensions underlying Holland's hexagon: Missing link between interests and occupations? *Journal of Vocational Behavior*, *21*(3), 259–287. https://doi.org/10.1016/0001-8791(82)90036-7

Quintus, M., Egloff, B., & Wrzus, C. (2021). Daily life processes predict long-term development in explicit and implicit representations of Big Five traits: Testing

predictions from the TESSERA (Triggering situations, Expectancies, States and State Expressions, and ReActions) framework. *Journal of Personality and Social Psychology*, *120*(4), 1049–1073. https://doi.org/10.1037/pspp0000361

Reber, R., Canning, E. A., & Harackiewicz, J. M. (2018). Personalized Education to Increase Interest. *Current Directions in Psychological Science*, *27*(6), 449–454. https://doi.org/10.1177/0963721418793140

Renninger, K. A., & Hidi, S. (2011). Revisiting the Conceptualization, Measurement, and Generation of Interest. *Educational Psychologist*, *46*(3), 168–184. https://doi.org/10.1080/00461520.2011.587723

Rieger, S., Göllner, R., Spengler, M., Trautwein, U., Nagengast, B., & Roberts, B. W. (2017). Social Cognitive Constructs Are Just as Stable as the Big Five Between Grades 5 and 8. *AERA Open*, *3*(3), 233285841771769. https://doi.org/10.1177/2332858417717691

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, *126*(1), 3–25. https://doi.org/10.1037/0033-2909.126.1.3

Roberts, B. W., & Jackson, J. J. (2008). Sociogenomic Personality Psychology: Sociogenomic Personality Psychology. *Journal of Personality*, *76*(6), 1523–1544. https://doi.org/10.1111/j.1467-6494.2008.00530.x

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The Power of Personality: The Comparative Validity of Personality Traits, Socioeconomic Status, and Cognitive Ability for Predicting Important Life Outcomes. *Perspectives on Psychological Science*, *2*(4), 313–345. https://doi.org/10.1111/j.1745-6916.2007.00047.x

Roberts, B. W., Luo, J., Briley, D. A., Chow, P. I., Su, R., & Hill, P. L. (2017). A systematic review of personality trait change through intervention. *Psychological Bulletin*, *143*(2), 117–141. https://doi.org/10.1037/bul0000088

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. K. Mroczek & T. D. Little (Eds.), *Handbook*

of Personality Development (pp. 11–39). Lawrence Erlbaum Associates Publishers. https://doi.org/10.4324/9781315805610.ch2

Roloff Henoch, J., Klusmann, U., Lüdtke, O., & Trautwein, U. (2015). Who becomes a teacher? Challenging the "negative selection" hypothesis. *Learning and Instruction*, *36*, 46–56. https://doi.org/10.1016/j.learninstruc.2014.11.005

Rottinghaus, P. J., Coon, K. L., Gaffey, A. R., & Zytowski, D. G. (2007). Thirty-Year Stability and Predictive Validity of Vocational Interests. *Journal of Career Assessment*, *15*(1), 5–22. https://doi.org/10.1177/1069072706294517

Rounds, J., Dawis, R., & Lofquist, L. H. (1987). Measurement of person-environment fit and prediction of satisfaction in the theory of work adjustment. *Journal of Vocational Behavior*, *31*(3), 297–318. https://doi.org/10.1016/0001-8791(87)90045-5

Rounds, J., & Day, S. X. (1999). Describing, evaluating, and creating vocational interest structures. In M. L. Savickas & A. R. Spokane (Eds.), *Vocational interests: Meaning, measurement, and counseling use* (pp. 103–133). Davies-Black Publishing.

Rounds, J., & Su, R. (2014). The nature and power of interests. *Current Directions in Psychological Science*, *23*(2), 98–103. https://doi.org/10.1177/0963721414522812

Rounds, J., & Tracey, T. J. G. (1996). Cross-cultural structural equivalence of RIASEC models and measures. *Journal of Counseling Psychology*, *43*(3), 310–329. https://doi.org/10.1037/0022-0167.43.3.310

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Savickas, M. L., & Taber, B. J. (2006). Individual Differences in RIASEC Profile Similarity Across Five Interest Inventories. *Measurement and Evaluation in Counseling and Development*, *38*(4), 203–210. https://doi.org/10.1080/07481756.2006.11909782

Sawyer, S. M., Azzopardi, P. S., Wickremarathne, D., & Patton, G. C. (2018). The age of adolescence. *The Lancet Child & Adolescent Health*, *2*(3), 223–228. https://doi.org/10.1016/S2352-4642(18)30022-1

Schmidt, F. L. (2014). A General Theoretical Integrative Model of Individual Differences in Interests, Abilities, Personality Traits, and Academic and Occupational Achievement:

A Commentary on Four Recent Articles. *Perspectives on Psychological Science*, *9*(2), 211–218. https://doi.org/10.1177/1745691613518074

Schmitt, N. (2014). Personality and Cognitive Ability as Predictors of Effective Performance at Work. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*(1), 45–65. https://doi.org/10.1146/annurev-orgpsych-031413-091255

Schomburg, A. M., & Tokar, D. M. (2003). The moderating effect of private self-consciousness on the stability of vocational interests. *Journal of Vocational Behavior*, *63*(3), 368–378. https://doi.org/10.1016/S0001-8791(03)00027-7

Schultz, L. H., Connolly, J. J., Garrison, S. M., Leveille, M. M., & Jackson, J. J. (2017). Vocational interests across 20 years of adulthood: Stability, change, and the role of work experiences. *Journal of Research in Personality*, *71*, 46–56. https://doi.org/10.1016/j.jrp.2017.08.010

Silvia, P. J. (2001). Interest and Interests: The Psychology of Constructive Capriciousness. *Review of General Psychology*, *5*(3), 270–290. https://doi.org/10.1037/1089-2680.5.3.270

Silvia, P. J. (2008). Interest—The Curious Emotion. *Current Directions in Psychological Science*, *17*(1), 57–60. https://doi.org/10.1111/j.1467-8721.2008.00548.x

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology*, *100*(2), 330–348. https://doi.org/10.1037/a0021717

Soto, C. J., & Tackett, J. L. (2015). Personality Traits in Childhood and Adolescence: Structure, Development, and Outcomes. *Current Directions in Psychological Science*, *24*(5), 358–362. https://doi.org/10.1177/0963721415589345

Spear, H. J., & Kulbok, P. (2004). Autonomy and Adolescence: A Concept Analysis. *Public Health Nursing*, *21*(2), 144–152. https://doi.org/10.1111/j.0737-1209.2004.021208.x

Spengler, M., Damian, R. I., & Roberts, B. W. (2018). How you behave in school predicts life success above and beyond family background, broad traits, and cognitive ability. *Journal of Personality and Social Psychology*, *114*(4), 620–636. https://doi.org/10.1037/pspp0000185

Steinberg, L. (2005). Cognitive and affective development in adolescence. *Trends in Cognitive Sciences*, *9*(2), 69–74. https://doi.org/10.1016/j.tics.2004.12.005

Steyer, R., Mayer, A., Geiser, C., & Cole, D. A. (2015). A Theory of States and Traits—Revised. *Annual Review of Clinical Psychology*, *11*(1), 71–98. https://doi.org/10.1146/annurev-clinpsy-032813-153719

Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state-trait theory and research in personality and individual differences. *European Journal of Personality*, *13*(5), 389–408. https://doi.org/10.1002/(SICI)1099-0984(199909/10)13:5<389::AID-PER361>3.0.CO;2-A

Stoll, G., Einarsdóttir, S., Chelsea Song, Q., Ondish, P., Sun, J.-T., & Rounds, J. (2020). The Roles of Personality Traits and Vocational Interests in Explaining What People Want out of Life. *Journal of Research in Personality*, *86*, 103939. https://doi.org/10.1016/j.jrp.2020.103939

Stoll, G., Rieger, S., Lüdtke, O., Nagengast, B., Trautwein, U., & Roberts, B. W. (2017). Vocational interests assessed at the end of high school predict life outcomes assessed 10 years later over and above IQ and Big Five personality traits. *Journal of Personality and Social Psychology*, *113*(1), 167–184. https://doi.org/10.1037/pspp0000117

Stoll, G., Rieger, S., Nagengast, B., Trautwein, U., & Rounds, J. (2020). Stability and change in vocational interests after graduation from high school: A six-wave longitudinal study. *Journal of Personality and Social Psychology*, *120*(4), 1091–1116. https://doi.org/10.1037/pspp0000359

Stoll, G., & Trautwein, U. (2017). Vocational interests as personality traits. In J. Specht (Ed.), *Personality development across the lifespan* (pp. 401–417). Elsevier. https://doi.org/10.1016/B978-0-12-804674-6.00025-9

Strong, E. K. (1931). *Change of Interests with Age*. Stanford University Press.

Strong, E. K. (1943). *Vocational interests of men and women*. Stanford University Press.

Su, R. (2012). *Power of vocational interests and interest congruence in predicting career success [unpublished doctoral dissertation)*. https://www.ideals.illinois.edu/bitstream/handle/2142/34329/Su_Rong.pdf?sequence=1&isAllowed=y

Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological Bulletin*, *135*(6), 859–884. https://doi.org/10.1037/a0017364

Su, R., Stoll, G., & Rounds, J. (2019). The nature of interests: Toward a unifying theory of trait-situation interest dynamics. In C. D. Nye & J. Rounds (Eds.), *Vocational interests: Rethinking their role in understanding workplace behavior and practice* (pp. 11–38). Taylor & Francis/Routledge. https://doi.org/10.4324/9781315678924-2

Swanson, J. L. (1999). Stability and change in vocational interests. In M. L. Savickas & A. R. Spokane (Eds.), *Vocational interests: Meaning, measurement, and counseling use* (pp. 135–158). Davies-Black Publishing.

Swanson, J. L., & Hansen, J.-I. C. (1988). Stability of vocational interests over 4-Year, 8-Year, and 12-Year intervals. *Journal of Vocational Behavior*, *33*(2), 185–202. https://doi.org/10.1016/0001-8791(88)90055-3

Thorndike, E. L. (1935). The interests of adults. I. The permanence of interests. *Journal of Educational Psychology*, *26*(6), 401–410. https://doi.org/10.1037/h0058840

Thurstone, L. L. (1931). A multiple factor study of vocational interests. *Personnel Journal*, *10*, 198–205.

Tracey, T. J. G. (2001). The Development of Structure of Interests in Children: Setting the Stage. *Journal of Vocational Behavior*, *59*(1), 89–104. https://doi.org/10.1006/jvbe.2000.1787

Tracey, T. J. G. (2002). Development of interests and competency beliefs: A 1-year longitudinal study of fifth- to eighth-grade students using the ICA-R and structural equation modeling. *Journal of Counseling Psychology*, *49*(2), 148–163. https://doi.org/10.1037//0022-0167.49.2.148

Tracey, T. J. G., & Caulum, D. (2015). Minimizing gender differences in children's interest assessment: Development of the Inventory of Children's Activities-3 (ICA-3). *Journal of Vocational Behavior*, *87*, 154–160. https://doi.org/10.1016/j.jvb.2015.01.004

Tracey, T. J. G., & Rounds, J. (1993). Evaluating Holland's and Gati's Vocational-Interest Models: A Structural Meta-Analysis. *Psychological Bulletin*, *113*(2), 229–246. https://doi.org/10.1037/0033-2909.113.2.229

Tracey, T. J. G., & Ward, C. C. (1998). The structure of children's interests and competence perceptions. *Journal of Counseling Psychology*, *45*(3), 290–303. https://doi.org/10.1037/0022-0167.45.3.290

Tracey, T. J. G., Wille, B., Durr, M. R., & De Fruyt, F. (2014). An enhanced examination of Holland's consistency and differentiation hypotheses. *Journal of Vocational Behavior*, *84*(3), 237–247. https://doi.org/10.1016/j.jvb.2014.01.008

Tsabari, O., Tziner, A., & Meir, E. I. (2005). Updated Meta-Analysis on the Relationship Between Congruence and Satisfaction. *Journal of Career Assessment*, *13*(2), 216–232. https://doi.org/10.1177/1069072704273165

Turner, S. L., & Lapan, R. T. (2005). Evaluation of an intervention to increase non-traditional career interests and career-related self-efficacy among middle-school adolescents. *Journal of Vocational Behavior*, *66*(3), 516–531. https://doi.org/10.1016/j.jvb.2004.02.005

Usslepp, N., Hübner, N., Stoll, G., Spengler, M., Trautwein, U., & Nagengast, B. (2020). RIASEC interests and the Big Five personality traits matter for life success—But do they already matter for educational track choices? *Journal of Personality*, *88*(5), 1007–1024. https://doi.org/10.1111/jopy.12547

Van den Akker, A. L., Deković, M., Asscher, J., & Prinzie, P. (2014). Mean-level personality development across childhood and adolescence: A temporary defiance of the maturity principle and bidirectional associations with parenting. *Journal of Personality and Social Psychology*, *107*(4), 736–750. https://doi.org/10.1037/a0037248

Van Iddekinge, C. H., Roth, P. L., Putka, D. J., & Lanivich, S. E. (2011). Are you interested? A meta-analysis of relations between vocational interests and employee performance and turnover. *Journal of Applied Psychology*, *96*(6), 1167–1194. https://doi.org/10.1037/a0024343

VanderWeele, T. J., Mathur, M. B., & Chen, Y. (2020). Outcome-Wide Longitudinal Designs for Causal Inference: A New Template for Empirical Studies. *Statistical Science*, *35*(3), 437–466. https://doi.org/10.1214/19-STS728

Vock, M., Köller, O., & Nagy, G. (2013). Vocational interests of intellectually gifted and highly achieving young adults: *Vocational interests of the gifted*. *British Journal of*

Educational Psychology, *83*(2), 305–328. https://doi.org/10.1111/j.2044-8279.2011.02063.x

Volodina, A., Nagy, G., & Retelsdorf, J. (2015). Berufliche Interessen und der Übergang in die gymnasiale Profiloberstufe: Ihre Struktur und Vorhersagekraft für das individuelle Wahlverhalten. *Zeitschrift für Pädagogische Psychologie*, *29*(2), 89–100. https://doi.org/10.1024/1010-0652/a000154

von Maurice, J. (2006). *ICA-D. Deutschsprachige Version des Inventory of Children's Activities – Revised (ICA-R, Tracey & Ward, 1998)[unpublished instrument]*. Otto-Friedrich-Universität Bamberg.

Wille, B., De Fruyt, F., & Feys, M. (2010). Vocational interests and Big Five traits as predictors of job instability. *Journal of Vocational Behavior*, *76*(3), 547–558. https://doi.org/10.1016/j.jvb.2010.01.007

Wille, B., Tracey, T. J. G., Feys, M., & De Fruyt, F. (2014). A longitudinal and multi-method examination of interest–occupation congruence within and across time. *Journal of Vocational Behavior*, *84*(1), 59–73. https://doi.org/10.1016/j.jvb.2013.12.001

Wille, E., Stoll, G., Gfrörer, T., Cambria, J., Nagengast, B., & Trautwein, U. (2020). It takes two: Expectancy-value constructs and vocational interests jointly predict STEM major choices. *Contemporary Educational Psychology*, *61*, 101858. https://doi.org/10.1016/j.cedpsych.2020.101858

Wrzus, C., & Roberts, B. W. (2017). Processes of personality development in adulthood: The TESSERA framework. *Personality and Social Psychology Review*, *21*(3), 253–277. https://doi.org/10.1177/1088868316652279

Xu, H., & Li, H. (2020). Operationalize Interest Congruence: A Comparative Examination of Four Approaches. *Journal of Career Assessment*, *28*(4), 571–588. https://doi.org/10.1177/1069072720909825

Xu, H., & Tracey, T. J. G. (2016). Stability and change in interests: A longitudinal examination of grades 7 through college. *Journal of Vocational Behavior*, *93*, 129–138. https://doi.org/10.1016/j.jvb.2016.02.002

Zytowski, D. G. (1976). Long-term profile stability of the Kuder Occupational Interest Survey. *Educational and Psychological Measurement*, *36*(3), 689–692. https://doi.org/10.1177/001316447603600315