THE UNIVERSITY of EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Investigating Personality Differences Using Nuance-Level Traits and Links with Health

Yuzhan Hang

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy in Psychology

to

The University of Edinburgh

2021

Table of Contents

Acknowledgements

PhD is not an easy journey and I have received a lot of help from my supervisors, friends and family. I thank René Mottus, Michelle Luciano and Billy Lee for their supervision throughout my PhD. I further thank all my collaborators for their valuable feedback. I also thank my friends for their company and support. Special thanks to my friend Lydia Speyer for our endless conversions and countless online study sessions together. Finally, I cannot thank my family enough. Without my parents, I would not be who I am today.

Contributions

During my PhD studies, I have contributed to the following publications:

Publications forming parts of this thesis (contributions indicated ahead of each chapter):

Hang, Y., Soto, C. J., Speyer, L. G., Haring, L., Lee, B., Ostendorf, F., & Mõttus, R. (2021). Age Differences in the Personality Hierarchy: A Multi-Sample Replication Study across the Life Span. *Journal of Research in Personality*. doi https://doi.org/10.1016/j.jrp.2021.104121

Hang, Y., Speyer, L. G., Haring, L., Lee, B., Vainik, U., Luciano, M., & Mõttus, R. (under review at *Journal of Abnormal Psychology*). Establishing a Health Network of Personality Profiles for Adolescents and Emerging Adults. PsyArXiv pre-print: https://doi.org/10.31234/osf.io/4qh96

Hang, Y., Soto, C. J., Lee, B., Speyer, L. G., Murray, A. L., & Mõttus, R. (under review at *Journal of Personality*). Social Expectations and Abilities to Meet Them as Possible Mechanisms of Youth Personality Development. PsyArXiv pre-print: https://doi.org/10.31234/osf.io/8yp6x

Hang, Y., Speyer, L. G., Lee, B., Luciano, M., & Mõttus, R. Social Expectations and Abilities to Meet Them as Possible Mechanisms of adult Personality Development. PsyArXiv pre-print: 10.31234/osf.io/p35r2

Other Publications:

Hang, Y.*, Speyer, L. G.*, Haring, L., Murray, A. L, Mõttus, R. (Under review at *Journal of Research in Psychology*). The General P-factor of Psychopathology and its correlates with different levels of personality hierarchy. PsyArXiv pre-print: 10.31234/osf.io/g3aqh *contributed equally.

Speyer, L. G., **Hang, Y**., Hildigunnur A H., & Murray, A. L (Revised and resubmitted to *Child Development*). The Role of Harsh Parenting Practices in Early- to Middle-Childhood Socio-Emotional Development: An Examination in the Millennium Cohort Study.

Speyer, L. G., Hildigunnur A H., **Hang, Y**., Hughes, C., & Murray, A. L (2021). Within-Family Relations of Mental Health Problems across Childhood and Adolescence. *Journal of Child Psychology and Psychiatry*. In press.

Abstract

Research on personality differences has been well-established on the levels of broad personality domains such as the Big Five and their more specific facets. Over the past decade, a small number of researchers have proposed that lower-level, more specific personality traits, especially nuances (items), are also of vital importance for personality research, yet nuance-level research has not received sufficient attention. Considering that personality research has been getting broader as well as more fine-grained over time, the structured domains or the Big Five are not sufficient to fit the diversity of research purposes nowadays. Nuance-level research provides a new avenue for research by offering additional information which can supplement the Big Five focused research. This thesis aims to explore whether nuances offer additional useful information above and beyond domains and facets and whether nuance-level research can help to advance understanding of two important research areas in personality: the mechanism of personality differences over time and the association between personality and health outcomes. The data used in this thesis came from five studies: the Ained ja Arenevad Ajud (AAA; Drugs and Developing Brains), the Common-Language California Child Q-Set Study (CCQS), the Life Outcomes of Personality Replication Project (LOOPRP), the HEXACO Online Study (HEXACOS), and the German Revised NEO Personality Inventory validation study (GNEOPIRS).

Results of the empirical studies indicate that lower levels of the personality hierarchy contained substantially more unique information than higher levels which means that nuances provide unique information beyond the information captured by their shared variance that represent broader traits. Thus, aggregating individual traits might result in the loss of useful information which is only offered by nuances (chapter 2). Therefore, nuance-level research could supplement the high-

dimensional approaches (domains, and facets) focused research and provide a more complete picture of personality change. This information further allows for the testing of novel hypotheses that rely on systematic between-trait variance in age differences. Based on these findings, this thesis tested some potential mechanisms of personality differences across the life span using nuance-level traits and found that while traits' self-regulation requirements could not explain personality change, social expectations partially explained personality differences during childhood and adolescence (chapter 3). These findings are further used to test whether they support the social investment theory which has also been proposed to explain personality change based on the maturity principle. The finding that social expectations partially explain personality differences was also evident in young adulthood and results also offered some limited empirical evidence to support the social investment theory (chapter 4). In the last empirical chapter, nuances were used to create personality profiles of health-related lifestyles and outcomes, which allowed for new insights into the covariance structure of these health and wellbeing outcomes. The sixteen outcomes (health-related lifestyles and mental health problems) could be categorized into four groups based on their personality correlations (shared personality correlates): psychological distress, health awareness, emotional control and substance use. The high personality correlations between mental health problems support the hypothesis that co-occurring mental health problems can be partly explained by a general personality-based psychopathology factor, that makes individuals more vulnerable to developing any form of psychopathology.

Overall, the findings of this thesis indicate that nuance-level traits can be useful for personality research, highlighting that it is essential to develop theoretical models specifically for nuance-level personality research. Importantly, this thesis aims to illustrate the possible value of lower-level personality traits, that is nuances, to extend and diversify the field, not to criticize or negate the Big Five focused research.

Lay Summary

The Big Five taxonomy is a well-established tool to study personality, classifying personality into five domains including openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism. As personality research is getting deeper and broader, the Big Five itself is not enough to do justice to emerging topics. Therefore, more detailed analyses which take into account nuance-level personality traits are increasingly needed. Nuances are the lowest level of the personality trait hierarchy, and in practice tend to equate to individual personality questionnaire items. To date, there is still little known about this level of personality traits. This thesis aims to gain a better understanding of nuance-level traits as well as ways of using nuance-level analyses to advance personality research. In particular, this research investigated some potential mechanisms of personality development during childhood, adolescence, and early adulthood. Results showed that nuance-level traits contained unique information that was not captured by the aggregated higher-level traits (e.g., domains and facets). Overall, this thesis has highlighted an important advantage offered by nuance-level analyses – that is nuances allow for the testing of new research hypotheses that cannot be done with the traditional Big Five model, such as studying the systematic variations of traits. Results further suggested potential clinical advantages of nuance-level personality traits in terms of studying health and wellbeing. In summary, the findings of this thesis provide evidence for the value of nuance-level analyses and highlight the importance of developing a model specifically to measure nuance-level traits.

List of Tables

List of Figures

Chapter 1: Introduction

Synopsis:

This chapter provides some brief introductory background information on personality differences across the life span based on previous literature, and outlines the current research gaps that the empirical chapters of this thesis aim to address. It also introduces the aims of the thesis and provides a general overview of the following chapters.

Can we predict the future of an individual who is open to experiences, extroverted and highly organised? Will they find a decent job, form a happy family, and adopt a healthy lifestyle across their life span? Personality research has come a long way to describe and explain individual differences in people's typical thinking, feeling and behaving. While the field used to highlight the stability of personality traits (McCrae & Costa, 2008), demonstrating that personality traits are powerful predictors of a wide range of important life outcomes (Bleidorn et al., 2019; Wagner et al., 2020), research on personality development has now emphasized that it is unlikely for a person to retain the same personality traits across all developmental stages of childhood, adolescence, and adulthood (Specht, 2017).

Personality can change in various aspects, such as mean-level change, rank-order consistency, and individual differences in change (Bleidorn & Hopwood, 2019; De Fruyt et al., 2006; Roberts et al., 2008). Mean-level change refers to the extent to which a trait changes within a population. Rank-order consistency refers to whether people maintain their position and show uniform change in personality levels within a population (Roberts & DelVecchio, 2000). Individual difference in personality refers to the degree people deviate from the average population (Bleidorn, Hopwood, Back, et al., 2020b). Typically, personality traits tend to shift in a socially desirable direction, with people becoming more agreeable, conscientious, and emotionally stable as they age into adulthood, at the mean level (Allik et al., 2004; Caspi et al., 2005; Donnellan & Lucas, 2008). At the mean level, people also tend to be less extraverted and open as they age, although these changes are not strongly linked with social desirability (Wortman et al., 2012). On the other hand, personality trait variances tend to increase from childhood until mid-adolescence and plateau thereafter (de Haan et al., 2013; Kandler & Zapko-Willmes, 2017; Mõttus, Soto, et al., 2017; Van den Akker et al., 2014). Personality researchers are eager to understand the mechanisms of personality change across the life span, as well as individual differences that appear during this process – especially as a substantial number of studies show that personality differences predict many major life outcomes, such as physical health, mental health and academic performance (Hampson et al., 2007; Israel et al., 2019; Moffitt et al., 2011; Samuel & Widiger, 2008).

Prior to answering any of these research questions, it is essential to begin with a close examination of the key notion of the *personality trait*. What exactly are personality traits? Previously, the variation of behaviour has been considered as either residing within a person or within a situation. According to these beliefs, traits are the individual differences conceived analogous to static dimensions (Hunt, 1965). Later, broadly speaking, personality traits were used to refer to relatively enduring styles of behaviors, feelings, and thoughts. Alternatively, personality traits can also be understood as specific behaviors, thoughts, feelings, motives, attitudes, values, which are often referred to as personality characteristics. These personality characteristics tend to vary among people but show certain patterns (Mõttus & Allerhand, 2018) which is also the key reason for why traits are important to study. Notably, personality traits are an abstract concept, thus traits cannot be directly measured but have to be inferred from patterns of behaviour. The pattern of covariation among these traits that tends to be summarized into a few factors (e. g., the Big Five) represents the basic dimensions of personality, that is the personality structure (McCrae & Costa, 1997). For example, extraversion is a dimension encompassing multiple traits such as cheerfulness and being energetic to describe an individual as sociable (Eysenck & Eysenck, 1967).

In order to study the conceptualization and measurement of personality traits, different hierarchically structured models such as the Big Two, Big Three, Big Four, and Big Five have been widely used. In these hierarchically structured models, different personality constructs have different levels of abstraction. Specifically, moving from one or two high-level overarching constructs to specific characteristics reflected by individual test items (DeYoung, 2006; McCrae, 2015; Mõttus, Kandler, et al., 2017; Rushton et al., 2008). These hierarchically structured models solved many problems in personality assessment such as whether the same trait has been referred to by different names or different traits have been referred to by the same name (J. Block, 1996; Kelley, 1927; Thorndike, 1904) and offered a powerful framework to test individual differences in different domains (Markon, 2009). More importantly, the primary strengths of hierarchically structured models are parsimony and flexibility in describing the relations between traits from different levels (e.g., in the hierarchical model).

In particular, the Big Five (L. R. Goldberg, 1990), the Five-Factor Model (McCrae & John, 1992) and the HEXACO model (Ashton & Lee, 2020), have been most prominently used in personality research and have become a default way of operationalizing personality differences among people (Mõttus et al., 2020). Some commonly used domains in the hierarchically structured Big Five and HEXACO measures include openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism or emotional stability. These high-dimensional approaches (domains and facets) are useful in summarising how people differ through time and predicting life outcomes. Many personality researchers try to understand personality differences across the lifespan by using domain-level analysis, indicated by the large volume of research on this topic. Some findings include that a higher score on neuroticism tends to be associated with a higher risk of mental and physical health issues (Moffitt et al., 2011; Samuel & Widiger, 2008). In contrast, higher scores on openness have been found to be associated with healthy dietary habits, whilst higher scores on conscientiousness have been found to be associated with longevity (Conner et al., 2017; Roberts, Walton, et al., 2005).

However, personality traits form a hierarchy, meaning that domains can be further separated into lower-level traits such as facets and nuances. Facets are groups of items within the domains that share more content similarities and combine to form the personality traits, although the number of facets per trait is still debated. For example, each Big Five Inventory – 2 domain only accesses three facets while each NEO-PI-R domain accesses six facets. Facet-level research supplements domain-level research by adding more fine-grained insights into personality trait effects. Nuances are the lowest level of the personality trait hierarchy and in practice tend to equate to individual items (Mõttus, Sinick, et al., 2019).

Personality traits that statistically covary among people and are similar in subjective perception are combined in order to form the higher order traits. Categorizing lower-level traits into fewer and broader groups for greater parsimony are what most personality researchers have been trained to do. The simple structure of the Big Five allows for a straightforward interpretation of findings and has thus contributed to gaining a good understanding of personality differences and its associations with life outcomes on a broad level.

Despite the advantages of the hierarchically structured domains, research demonstrates that facets within the same domains often display different or opposite associations (Mõttus, Sinick, et al., 2019; Vainik, Dagher, et al., 2019). For example, Vainik et al. (2019) showed that higher BMI can be predicted by higher scores on the warmth, positive emotions and assertiveness facets but also by lower scores on the activity level facet, all within the extraversion domain. Similarly, previous research has found that nuances capture unique variance beyond the aggregated facets and domains they belong to (Mõttus, Sinick, et al., 2019; Revelle, 2019; Seeboth & Mõttus, 2018; Vainik, Dagher, et al., 2019). For example, previous studies have found a divergence of age trends in ambitiousness and working hard: in the diligence facet, the item that referred to setting ambitious goals negatively correlated with age, but items that referred to working hard positively correlated with age (Mõttus et al., 2015; Mõttus & Rozgonjuk, 2019). These findings indicate that lower-level personality traits sometimes have divergent associations with life outcomes and other factors. Simply ignoring the predictive power of lower-level personality traits, especially nuance-level traits, would lead to attenuated associations as well as lower specificity, or discriminant validity, of the observed associations. Therefore, focusing solely on the high-dimensional approaches not only reduces the potential ability to predict outcomes from personality traits but also makes it harder to differentiate between the personality correlates of different outcomes.

Since nuances particularly contain more detailed developmental information compared to domains and facets, nuance-specific associations could provide three key benefits. First, the unique information captured by nuances offers a more complete picture of personality development, either in terms of its details or general structure. Researchers could study information about which specific traits are more prone to develop, or whether the mechanisms underlying personality development are narrow and numerous or few and broadly acting. Second, nuance-level analysis could help researchers understand inconsistencies in findings. For instance, some studies found that openness is positively associated with subjective wellbeing (Ha & Kim, 2013; Wang et al., 2020), others found openness is not associated with subjective wellbeing (Hayes & Joseph, 2003), or negatively associated with subjective wellbeing (Kapure, 2017). These conflicting findings could be because these studies use different personality measurements and consequently sample different nuances. Third, nuance-level analysis allows researchers to investigate areas that could not be completely studied by the domains and facets. For example, researchers can study the systematic variation of personality traits in certain aspects of life outcomes (e.g., health) using 200 nuances which would capture a reasonable amount of variation in life outcomes, but researchers could not achieve the same goal with five domains, simply due to the limited number of traits.

As personality research becomes more detailed, focus exclusively on the high-dimensional approaches cannot be sufficient to address the diversity of research purposes (Mõttus et al., 2020). The three main purposes of personality research can be classified as description, prediction and explanation. In order to fulfil the disparate research purposes, we should choose appropriate research methods for each research aim rather than try to fit the high-dimensional approaches into every investigation and ignore that it might not be the most suitable method for some questions. For example, descriptive findings should be more detailed and nuance-level analysis could offer greater flexibility than aggregating traits (Mõttus et al., 2020). To be clear, it is not argued that it is problematical to conduct research using the high-dimensional approaches and the aim of this thesis is not primarily to criticise structured domain or facet level research. Indeed, the high-dimensional approaches are a useful tool to study personality and the parsimony of the domains and facets offers great advantages to identify and describe personality patterns. It is fair that many personality researchers traditionally aggregate numerous personality traits into more structured domains and facets, and are used to studying personality with high-dimensional approaches. Parsimony, in particular, is important to study personality traits and their effect on other outcomes, but nuance-level traits lack parsimony compared to high-dimensional approaches. Understandably so, given that hundreds of personality traits could also be overwhelming to analyse and explain. Thus, research focusing on nuance-level traits is disadvantaged by the large quantity of traits, massive results generated and difficulties in explaining complex findings. Taking all these into consideration, it is a personal preference that researchers choose to focus on a particular level of personality hierarchy for study. The goal of this thesis is to highlight the importance of nuance-level traits and draw attention to the advantages of such research. In particular, a key message of this thesis is that studies of nuance-level personality traits so far indicate much promise, and suggest that it is not always valuable *a priori* to aggregate them into broader traits.

Early history of personality traits

Although nuance-level research is still in its infancy, there is still an early history of using item-level approaches. The earliest standard personality test assumed that, while the important components of psychological functioning (attitude, emotions, and psychopathology) are held in all individuals, the differences in psychological profiles comes from the degree to which each component of psychological functioning is expressed within an individual (Buchanan, 1994). Consequently, responses on each individual items reveal the magnitude of a trait. The combination of all responses can then be used to form an individual's personality profile and this profile can be compared to the distribution of profiles from a group of participants (Buchanan, 1994). The Minnesota Multiphasic Personality Inventory (MMPI), which is one of the most widely used personality test in the world, shares a similar approach and rationale. Developed by J. Charnley McKinley and Starke R. Hathaway for use in medical or psychiatric settings, the MMPI is now used in broader settings such as for screening potential airline pilots and police personnel (Butcher & Williams, 2009). In order to develop the MMPI, Hathaway and McKinley first created an item pool based on a variety of resources such as psychiatric textbooks and previous personality scales. However, all statements taken from the prior literature were re-phrased to ensure the expression of each item was simple, frequently used and in an active voice. Thus, all items' wording should be understandable to the larger population regardless of whether one is highly educated or not. Although the items' wordings were carefully modified, the individual items were not very important (Buchanan, 1994).

Although, MMPI has been developed and revised since its first creation in 1940 (Hathaway & McKinley, 1940, 1942; McKinley et al., 1948; Meehl & Hathaway, 1946), the importance of understanding item-level traits has only been brought up during the climax of the person-situation debate. The person-situation debate is concerned with whether stable behavioral dispositions or traits exist (Epstein, 1979). In particularly, the debate over whether human behaviour is a result of an internal force and traits of the individual or the outcome of external situational influences (Mischel, 2009). Situationists argue that individual differences are essentially an error term which can be attributed to a systematic bias of reducing complex individual behavioral differences into certain limited and simple person-distinguishing underlying traits (Shweder, 1975). One of their key pieces of evidence is that personality coefficients, i.e. the correlation between behaviour under one situation and behaviour under another situation, is low, usually less than .30 (Epstein, 1979; Mischel, 1968). Situationists thus concluded that behaviour is largely depend on situational context and personality is characterized by a lack of stability in such contexts. In contrast to situationists, proponents of the view that human behavior is a result of individuals' stable traits, view the situations as the error term. They further provided evidence for the stability of personality. A series of study testing cheating behaviour of children in various situations, has been seen as influential evidence against personality stability for a long time since the results suggested that the average intercorrelation between cheating in different situations was low. However, when aggregating many tests into a single score, the reliability coefficient increased significantly, thus suggesting that the lack of stability was driven by measurement error due to situational context and in fact supporting the stability of trait personality (Hartshorne et al., 1930). Hence, the prediction will be much more reliable from when using multiple tests and averaging behaviour from multiple settings.

The trait position and the situationist position also share some similarities. The trait position aims to study consistent behavioral patterns within some selected situations and the situationist position aims to study the effects of situations on a population. In either case, measurement error could emerge if the result only rely on an individual observation since any situation could contain its uniqueness. Averaging a sufficient sample size of items is essential to reduce the measurement error to further ensure the reliability, replicability and generalizability of results (Epstein, 1979). Here then, the importance of having enough personality items and item-level personality traits became more visible. Finally, the interactionist position, which tries to find a compromise between the situationist and the trait position, believes that behavioral stability only exist within situational constraints and proposes that the relationship between the situation and the person does not have to be an either-or dilemma (Epstein, 1979).

Although the importance of personality items have been noticed at the time, there is ample evidence to show that personality has a hierarchical structure that can integrate high-dimensional traits within a common structural framework, thus suggesting hierarchy is an intrinsic and pervasive feature of trait structure (Markon, 2009; Markon et al., 2005). Thus, if any traits can be split into narrower traits and then even more narrower traits and so on, it is difficult to decide which combination of traits should be focused on. Empirical explanatory power such as the reliability or criterion-related validity also cannot provide a clear result of whether higher-level traits or lower-level traits are better. Higher-order traits or superordinate traits tend to be more reliable, show better generalization or cross-validation and better predict external criteria (Morey et al., 2007; Ones & Viswesvaran, 1996). These higher-order traits also show good descriptive utility in terms of describing certain properties of human personality, despite evidence such as heritability, cross-rater agreement, and high rank-order stability potentially not reflecting the reality of these traits (Mõttus & Allerhand, 2018). Moreover, considering why the covariation of traits can be fairly easily summarized into distinct personality factors/domains, one possible explanation could be that the clustered personality characteristics reflect some levels of shared underlying etiology (Mõttus & Allerhand, 2018). From a practical point of view, a few big dimensions are simpler to analyze and analytically/theoretically allow for better parsimony. Especially, during the initial stages of personality research, when research topics were neither as specific nor as broad as nowadays, a simple and parsimonious structure was more useful for both research and educational purposes, since these structures help personality researchers to better summarize and describe personality related topics (Mõttus & Allerhand, 2018). In contrast, lower-level traits or subordinate traits often afford greater predictive validity than higher-order traits (Paunonen, 1998; Reynolds & Clark, 2001). Especially, item-level personality traits also provide an opportunity to reduce measurement errors, improve reliability of results and create new research methods which benefits the field now that research topics are becoming more specific as well as broader. In sum, focusing on item-level traits would have been too complex, time consuming and the results of such studies would not have been easily understood in the early stage of personality research. Especially since hierarchically structured models could answer the majority research questions at the time, research focusing on hierarchically structured models became the mainstream. This facilitated the development of the personality field in terms of easily spreading these finding to different audiences for different purposes.

Personality differences and possible mechanisms

Research on personality development has indicated that people are less likely to retain the same personality traits across developmental stages (Specht, 2017). For example, a person's personality at 10 years old would likely be different to their personality at 30 years old. What are the underlying mechanisms of personality differences across time? A broad array of literature has found that both personality and social situations have important influences on behaviour (Specht, 2017). Personality is expected to have higher influences on behaviour under situations of low social pressure (Hurlock, 1994). For example, students show more diverse behaviours after school compared with in the classroom. Teachers and parents expect children to gain new knowledge and respect teachers in the classroom, in turn, these expectations or pressures drive students to behave in certain ways. In contrast, teachers and parents have less expectations for children's behaviour after school, thus children perceive less pressure to behave a certain way after school and can choose how to behave according to their own interests. According to gene-environment transactions (Mõttus, Briley, et al., 2019) or gene-environment interactions (Kandler & Zapko-Willmes, 2017), children are increasingly becoming less alike because self-selected/created environments amplify their predisposed traits and reinforce individual differences (Caspi et al., 2005). Therefore, individual differences are more likely to occur in situations that are free and with fewer social expectations on how to behave. If social expectations are a potential cause of personality maturation/change across time, and there are consistent social expectations for children to behave in a certain way, then normative personality development should be at least partially driven by the changing social expectations.

In addition to social expectations, self-regulatory abilities, which are the abilities that allow individuals to meet social expectations, are another potential explanation for personality differences across time (Denissen et al., 2013). If children's close contacts such as their parents and teachers or even their society expect children to be kind and well-behaved, children might adjust their behaviours gradually to meet these standards. This would lead to a mean-level increase in the relevant personality traits and cause children to behave in a more socially desirable way.

However, if children’s abilities have not sufficiently developed to meet these standards, that is, they do not have sufficient self-regulatory abilities, then the discrepancy between expectations and children’s trait levels might be even broader (Denissen et al., 2013; Soto, 2016; Soto et al., 2011; Van den Akker et al., 2014). This thesis tests whether social expectations and self-regulatory ability are potential mechanisms to explain personality change.

Critical periods in personality development: adolescence and emerging adulthood

Throughout the lifespan, adolescence is one of the most crucial developmental stages of personality maturation. Adolescence is a developmental period of great psychosocial and neural development, and according to the WHO lasts until age 19 years (Blakemore & Mills, 2014; Pringle et al., 2016; World Health Organization, 2021). During this period, adolescents face dramatic changes physically (e.g. puberty), mentally (e.g. want to be independent) and environmentally (e.g. leave state education), which leads to unique challenges and changes in personality (Call et al., 2002). One of the defining features of adolescence is that individuals start exploring their identities and social roles and spend less time with their family and more time in new contexts with peers, and colleagues, as well as entering into romantic relationships (Ellis et al., 2012). Experiencing new contexts and being independent can cause psychological distress and potentially increases risks of developing mental health issues (Call et al., 2002). The changes in physical and mental conditions greatly contribute to personality development during this period. Understanding the driving factors that lead to individual differences in adolescents’ personality may enable better support for this crucial period.

While adolescence is a crucial period for the establishment of social roles, this process is unlikely to be completed by age 19 years. Over the last century, social roles have changed drastically and most individuals do not settle on specific social identities until their late 20s (Schwartz et al., 2013). People in their 20s are not adolescents anymore, but they are not considered to be adults yet either since they have not taken on adult role identities (L. J. Nelson & Barry, 2005). This developmental period has been referred to as emerging adulthood (Arnett, 2000), a stage of the lifespan between adolescence and adulthood encompassing the late teens to mid-20s (approximately 18 to 25 years), characterized by instability and an increase in exposure to new environments (Moreira et al., 2015; Pusch et al., 2019). There are many overlaps between adolescence and emerging adulthood such as the common age range (about the beginning of 20s) and definitions (the transitioning period of exploring role identities). This makes it essential to include emerging adulthood in investigating individual differences in personality change. Furthermore, many studies also show that young adults may be particularly vulnerable to personality change (Bleidorn, 2015; Bleidorn, Hopwood, Back, et al., 2020b; Schwaba & Bleidorn, 2018). Therefore, the present thesis focuses on understanding personality differences and the impact of personality during these important developmental stages.

Summary

In sum, personality differences throughout the life span and the degree of change in personality traits could help to predict many life outcomes. However, existing research has mostly relied on the high-dimensional approaches and has underestimated the value of nuance-level traits. This thesis aims to examine effects of lower-level personality traits, particularly nuances, and explore possible ways to use nuance-level traits to advance understanding of personality differences over time, and associations between personality traits and life outcomes.

Thesis aim

Nuance-level research has been, until recently, relatively neglected and consequently, no fine-grained findings for many key personality research topics are available. These include, but are not limited to, the mechanisms of personality differences across time, and the associations between life outcomes and personality traits at different levels. Therefore, this thesis has the following aims:

a. Investigate the value of lower-level traits, especially nuances, and possible methods of using them.

b. Explore potential mechanisms of personality differences during critical developmental stages.

c. Gain insight into whether nuance-level research has real life implications in terms of understanding health and wellbeing.

Overview of chapters

Chapter 2 exemplifies the value of nuances, by investigating the predictive accuracies of all levels of the personality hierarchy for age. Nuance-level research provides unique information that is not shared by the domains and facets. Five datasets with different personality measurements, and participants from more than 48 countries, were tested to investigate the extent of age related information captured by domains, facets and nuances. This was to gain more evidence on how nuance-level research can supplement the high-dimensional approaches focused research and to contribute to a better understanding of personality differences. This chapter also provides an overview of all datasets used in this thesis: the Ained ja Arenevad Ajud (AAA; Drugs and Developing Brains), the Common-Language California Child Q-Set Study (CCQS), the Life Outcomes of Personality Replication Project (LOOPRP), the HEXACO Online Study (HEXACOS), and the German Revised NEO Personality Inventory validation study (GNEOPIRS). Results from chapter 2 emphasised the importance of nuance-level research in general and how nuance-level research can be used to examine mechanisms of personality differences across time. These results provide further evidence for the value of developing psychometrically sound scales for measuring nuances in future research.

Chapter 3 tests whether social expectations and self-regulation are some of the key mechanisms to explain personality differences during childhood and adolescence using nuance-level analysis. Mean-level personality traits, personality variances, socially expected level of traits and self-regulatory requirements for achieving the desired trait levels were measured by the Common-Language California Child Q-Set. Due to the characteristics of nuances, a unique method which is to study the systematic variations of personality traits in different aspects rather than the change of different people was adopted in this chapter.

Chapter 4 expands on chapter 3 by extending the age range to early adulthood to further study another critical developmental period. Results in chapter 3 conclude that self-regulation could not explain personality differences with age, however social expectations provide a partial explanation. Therefore, chapter 4 focuses solely on social expectations. Social investment theory suggests that personality differences might be due to role related social expectations changing with age. The same new approach (to focus on nuance-level traits rather than people) was applied in this chapter to investigate empirical evidence for social investment theory.

Chapter 5 highlights a practical application of nuance-level research. This chapter aims to gain a deeper understanding of health-related issues through personality traits. Specifically, personality profiles of different lifestyle aspects and mental health outcomes were created to explore how they connect with each other. In addition, this chapter also demonstrates that nuance-level research may be important for explaining associations between personality and other life outcomes such as health.

Chapter 6, the final chapter, summarises the results from Chapter 2 to 5 and contextualises the findings. It also discusses the implications and limitations of the empirical chapters and provides some directions for future research.

Chapter 2: Analysing Age Differences in each Personality Hierarchy

Synopsis:

This chapter explores why excluding nuance-level analyses in personality research is tantamount to scientific neglect by investigating how much unique age-sensitive information is captured by nuances compared to domains and facets. Aggregating personality traits into domains and facets inevitably discards a substantial amount of information that is only contained in the personality nuances. Five different datasets with participants from more than 48 counties were tested to investigate the extent of age-related information captured by domains, facets and nuances. Results indicate that across samples and measures, lower trait hierarchy levels tended to contain substantially more age-sensitive information than higher levels and nuances capture unique age-sensitive information that is not shared with domains and facets. Moreover, findings indicate that higher numbers of personality items provide better predictive accuracy.

Dissemination status:

This chapter has been published in the *Journal of Research in Personality* entitled “Age Differences in the Personality Hierarchy: A Multi-Sample Replication Study across the Life Span”. https://doi.org/10.1016/j.jrp.2021.104121

Age Differences in the Personality Hierarchy: A Multi-Sample Replication Study across the Life Span

Yuzhan Hang[a*], Christopher Soto[b], Lydia Gabriela Speyer[a], Liina Haring[c], Billy Lee[a], Fritz Ostendorf[d], René Mõttus[a]

Affiliations: [a]Department of Psychology, University of Edinburgh, EH8 9JZ, Edinburgh, United Kingdom, [b]Department of Psychology, Colby College, 04901, Maine, United States, [c]Department of Psychiatry (ARPS), University of Tartu, 50411, Tartu, Estonia, [d]Faculty of Psychology and Sport Science, Bielefeld University, 33615, Bielefeld, Germany

Abstract

This replication and extension of Mõttus and Rozgonjuk (2019) compared the extents of age-related information captured by different levels of the personality trait hierarchy (domains, facets and nuances, indexed by individual items) in several samples (N = 51,524) of different age ranges and cultural backgrounds, and tested with different instruments. Across samples and measures, lower trait hierarchy levels (especially nuances) tended to contain substantially more age-sensitive information than higher levels; most of the unique age-sensitive information was in nuances. Besides showing the need for more nuanced personality (development) research, the findings suggest ways of testing novel hypotheses that rely on systematic between-trait variance in age differences.

Keywords: personality development, age differences, the Big Five domains, facets, nuances

Introduction

Among other goals, personality research aims to describe and understand how personality – patterns of thinking, behaving, feeling, and motivation – change as individuals grow older. Besides rank-order stability (Roberts & DelVecchio, 2000) and associations between personality traits and individual-level variables such as life events (Bleidorn, Schwaba, et al., 2021), a commonly studied aspects of personality change is age differences mean trait levels, often described using the broad Five-Factor Model (FFM) or the Big Five trait domains: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience. On average, mean scores of the personality domains gradually shift in a socially desirable direction, with people becoming more agreeable, conscientious, and emotionally stable as they age in adulthood (Allemand et al., 2008; Caspi et al., 2005; Donnellan & Lucas, 2008; McCrae et al., 1999); however, the general maturation trend tends to temporarily stall or even reverse in adolescence (Denissen et al., 2013; Soto & Tackett, 2015).

Although the Big Five domains provide a useful way to summarize individual differences in behaviour, there is more to personality traits than these broad domains, as they can be split into narrower traits – personality traits form a hierarchy (Mõttus et al., 2020). Facets, narrower traits below the Big Five in the trait hierarchy, contain unique personality variance above and beyond domains (Jang et al., 1998). It is therefore not surprising that facet level analyses have shown more varied developmental age trajectories compared to their domains (Lucas & Donnellan, 2009; Mõttus et al., 2015; Mõttus, Briley, et al., 2019; Mõttus & Rozgonjuk, 2019; Soto et al., 2011; Terracciano et al., 2005). For example, Soto and John (2012) investigated the mean-level age trends of Big Five domains and facets in adulthood and reported longitudinal decreases in the Rumination and Depression facets of Neuroticism but increases for the Anxiety and Irritability facets. Similarly, Ashton and Lee (2016) examined age trends in HEXACO-PI-R self-reports and found that the Unconventionality facet of Openness slightly decreased over adulthood, the Inquisitiveness facet showed a modest upward trend during adulthood and the Aesthetic Appreciation facet remained stable during middle age and then increased. Mõttus and Rozgonjuk (2019) reported that facets contain over 50% more age-related variance than the Big Five domains. As summarized by Soto and colleagues (2011, p. 342), a "growing body of findings indicates that conceptualizing traits at the level of Big Five facets is necessary for a full understanding of life-span age differences in personality; research at the domain level can provide a rough sketch of these differences, but not a complete picture".

Although facet traits provide a more-specific picture of personality development than the Big Five domains, they are not the lowest level of the personality hierarchy. That is, facets themselves can be split into still-narrower traits, referred to as nuances, that represent specific personality variance not shared with facets (McCrae, 2015). Nuances can be indexed by single personality questionnaire items or bundles of highly similar items and are currently the lowest level of the trait hierarchy typically assessed by personality measures (Condon et al., 2020; McCrae, 2015; Mõttus, Kandler, et al., 2017). It is important to realize that nuances are not merely indicators, or stylistic expressions, of domain-level and facet-level traits, but instead display distinctive properties of traits such as cross-method agreement, rank-order stability, and heritability (Mõttus et al., 2014; Mõttus, Kandler, et al., 2017). For example, in Mõttus and colleagues (2017; 2019), even when the 240 items of the NEO Personality Inventory (NEO-PI-R) (Costa & McCrae, 1992) were residualized for the 30 facets and five domains assessed by that inventory, the remaining variance (representing unique nuances) of most items showed these properties. Nuances may also show age trends different from their domains and facets and, hence, capture unique developmental information (McCrae, 2015; Mõttus, Kandler, et al., 2017).

Indeed, Mõttus and Rozgonjuk (2019) “predicted” (in a statistical sense) age from the FFM domains, their 30 facets and 300 items to systematically compare the extents of age-related information captured by different levels of the personality trait hierarchy. The study used a large sample (*N* = 24,000), aged between 18 and 50 years with six evenly distributed and sex balanced age groups. Participants were tested with a 300-items personality questionnaire (IPIP-NEO) (W. A. Goldberg, 1999), which included a shorter 120-items version of the IPIP (the 120 items were chosen from the 300 items). The 300 items collectively captured over 40% more age-sensitive information than their facets, and over 130% more information than the five broad domains, and they also contained 20% more age-relevant information than the 120 items of the shorter version of the questionnaire. Moreover, residualizing the items for their facets (and thereby also the FFM domains) had virtually no effect on how much age-sensitive information they contained, suggesting that the variability of behaviour, thinking and affect with age was mostly driven by narrow personality characteristics better captured by single items than by broader trait constructs represented by the items’ shared variance (the residuals also out-predicted the domains and facets; see Table A1 and A2). These findings indicated that items of the same facets and domains often varied in their age differences: indeed, Mõttus and Rozgonjuk (2019) even found that for half of the facets, items had significant correlations with age in both negative and positive directions.

Nuances thus contain more detailed developmental information compared to the Big Five domains and facets. In fact, it may even be that much of the age differences in the domains and facets themselves can be accounted for by the nuances that happen to be included in them. Documenting nuance-specific associations could help researchers address some currently unanswered questions (e.g. inconsistent findings across studies may result from instruments sampling different nuances) and provide them with a more complete picture of personality development, either in terms of its details (e.g., information about which specific traits develop) or general architecture (e.g., by suggesting whether the mechanisms underlying personality development are narrow and numerous or few and broadly-acting).

We realize that some readers may be sceptical about describing age differences in terms of nuances. For example, there is currently no taxonomy for nuances, but merely evidence for their existence that has resulted from various item-level analyses (Mõttus et al., 2020). We think of nuances as traits beyond the few dozens of facets that have currently been outlined for the Big Five or HEXACO domains. Among others, these additional traits include those that have already received considerable attention in literature and are commonly referred to in everyday language (e.g., jealousy, competitiveness, humour, loneliness, procrastination, or gratitude) but simply have not been explicitly included among the facets of the Big Five or HEXACO, even if many of them are tapped into by the individual items of these facets. Given this, there should be little revolutionary in the idea that personality trait hierarchy could be extended below the existing few dozens of facets (Condon et al., 2020): often it would mean no more than explicitly outlining many of the traits that are already measured by various inventory items and / or have already been addressed in the literature. If so, *not* paying attention to possible age differences in these traits would seem like a scientific neglect.

But our current analyses do not intend to comprehensively outline how nuances vary with age, although we hope the eventual need for this to be obvious. This is exactly because there is no systematic taxonomy for nuances yet, likely because researchers never saw a need for this. As a result, our current analyses intend to further strengthen the case for the pursuit of developing a taxonomy for nuances and thereby also being able to properly study their development. The more robust and replicable evidence there is that facets do not capture all of the valid and potentially useful information about individual differences, including age differences, the more obvious the need to start the hard work of clearly mapping out the traits below facets becomes (Condon et al., 2020). That is, to motivate the research community to move towards a taxonomy of nuances in the first place, the need for it has to be shown with as much evidence and rigour as possible. The present paper is one key part of this effort.

We also realize that some readers may feel overwhelmed by the outlook of describing age differences in personality using dozens of traits. The level of details individual researchers are most comfortable with is obviously a matter of personal preference (Mõttus et al., 2020; Yarkoni, 2020). But it is also a matter of what research questions researchers are focusing on. For example, regardless of the number of traits used to describe how people psychologically vary with age, examining personality development at the nuance level would also allow for testing novel kinds of hypotheses that a) pertain to general principles rather than specific details and b) therefore result in even *more* parsimonious findings. Specifically, if many nuance-level traits show distinctive age trends, then it will be possible to quantitatively explore *systematic differences between these traits* in their age trajectories and link these with other properties of the traits (Mõttus et al., 2020; Mõttus & Rozgonjuk, 2019). For example, researchers could (finally) formally test the hypothesis that personality development reflects psychosocial maturation (Caspi et al., 2005) by quantifying a large and diverse sample of traits in terms of both their distinctive age trends and their links to psychosocial maturity, and then testing the association between these two trait-level properties. Similarly, quantifying the extent to which different nuance traits reflect behavioural, affective, cognitive, or motivational aspects of personality would allow researchers to test how much of personality development is driven by each of these four aspects (Wilt & Revelle, 2015). In short, examining systematic differences between traits in their age differences could be informative about *general principles of personality development that cut across particular traits*, broad and few or narrow and specific. This cannot be done when considering personality trait variation with age only along a few dimensions such as the Big Five.

Some readers may also think, justifiably, that increasing the number of predictors in a model is inevitably bound to increase the amount of variance that the model can explain in the sample it is fitted to, but the same result would not apply to a different sample even within the same population. This is known as over-fitting, a serious concern in statistical modelling (Seeboth & Mõttus, 2018; Yarkoni & Westfall, 2017). Thus, it may seem like nuances out-predicting higher-level traits is a given. In order to address this, the present study used a combination of regularized regression and testing the models in samples other than the samples used for constructing them: this approach has been shown to give more complex model no artefactual advantage – if anything, the opposite (Seeboth & Mõttus, 2018).

To our knowledge, Mõttus and Rozgonjuk (2019) have provided the only previous systematic test comparing the extent of age-related information captured by different levels of the personality trait hierarchy (e.g., domains, facets, and nuances). Therefore, the present study aims to test whether the key findings of Mõttus and Rozgonjuk (2019)—that personality nuances capture substantially more age-related information than facet-level traits, and that facets in turn capture more information than broad domains like the Big Five—will replicate and generalize to several different samples with a) different age ranges and b) different cultural backgrounds that c) were tested with different instruments and using different rater perspectives.

The present research thus extended that of Mõttus and Rozgonjuk (2019) in four key aspects. First, it examined both youths and adults. It is conceivable that the pattern observed in adults by Mõttus and Rozgonjuk does not generalize to youths; for example, it is possible that development becomes more nuanced (i.e., differentiated) only in adulthood, whereas personality may vary with age only along a few dimensions among children and adolescents. If so, there may be less of a need for developing nuanced designs to study adolescents, as opposed to adults. Second, we analysed both self-reports and parent-reports, allowing us to test whether previous findings derived from self-reports will generalize to informant-reports (Kööts-Ausmees et al., 2020; Rohrer et al., 2018). It is important to study key questions of personality science using multiple methods, because any one method contains considerable degree of method-specific variance which can bias the results (Costa et al., 2019; McCrae & Mõttus, 2019; Mõttus et al., 2020). Third, we tested whether our findings would generalize across different frameworks for operationalizing personality traits. This included measuring the Big Five domains with either the thirty facets of the Revised NEO Personality Inventory (NEO-PI-R; 240 items) (Costa & McCrae, 1992) and the Estonian Personality Item Pool (EE.PIP-NEO; 240 items) (Mõttus, 2006) or the fifteen facets of the Big Five Inventory–2 (BFI-2; 60 items), (Soto & John, 2017) as well as measuring the six HEXACO domains (Honesty/Humility, Emotionality, Extraversion, Agreeableness, Conscientiousness, and Openness) (Ashton & Lee, 2020) with 25 facets (100 items) (Ashton & Lee, 2007). Beyond these hierarchically structured Big Five and HEXACO measures, we also used a less-structured measure: the common-language California Child Q-Set (CCQ) (J. H. Block & Block, 1980; Caspi et al., 1992). Fourth, we tested whether our findings would generalize across different cultural contexts such as the United States, Germany and Estonia.

We expected that (a) personality nuances would contain more age-related information than facet-level traits, (b) facets, in turn, would contain more age-related information than domain-level traits, and (c) these findings would generalize across different developmental periods, rating perspectives, personality inventories, and cultural contexts. That is, we expected the original findings to replicate across the study designs and rule out the possibility that they had been driven by something questionnaire-specific, sampling or other methodological choices. We used a similar data analytic plan than the original study.

Method

Participants

Participants for measuring personality traits were drawn from five different studies: Ained ja Arenevad Ajud (AAA; Drugs and Developing Brains), the Common-Language California Child Q-Set Study (CCQS), the Life Outcomes of Personality Replication Project (LOOPRP), the HEXACO Online Study (HEXACOS), and the German Revised NEO Personality Inventory validation study (GNEOPIRS). Similarly to Mõttus and Rozgonjuk (2019), we attempted to create a relatively uniform age and sex distribution within each sample to facilitate the comparison of results across samples. This often meant limiting the age range of the sample and dropping participants who did not fit into the age range or sampling quotas (see below).

The AAA studied drug use and mental health conditions in Estonian youths, as well as their personality traits. Participants were from Estonia and data were collected online over two years between 2018 and 2019. From among the 4,005 participants in the initial sample, (2,514 females, 1,489 males, 2 did not indicate their sex, mean age = 21.42 years), we selected participants aged between 16 and 19 years old (N = 2,269, mean age = 17.41 years; 1,304 females; 963 males, 2 did not indicate their sex, participants outside this age range were scattered across many ages; some data was dropped due to concerns with its quality)[1]. Compared to most other datasets, the questionnaire for the AAA dataset was long, consisting of a total of 240 personality questions which could have impacted participant's willingness for carefully complete the whole questionnaire. While the GNEOPIRS questionnaire (described below) was also 240 items long, it was completed in person, thus increasing the likelihood of people completing the whole questionnaire with care.

The CCQS aimed to describe personality development from early childhood to the end of adolescence: participants were parents of 16,000 children aged from 3 to 20 years, with each parent rating one or more of their children in terms of their personality traits (Soto & John, 2014). At each age from 3 to 17, participants included exactly 500 boys and 500 girls, whereas ages 18 to 20 were combined to achieve the target of 500 boys and 500 girls. The target children were drawn from different countries with an estimated 83% from the United States, 7% from the United Kingdom or Ireland, 6% from Canada, and 4% from Australia or New Zealand. Approximately 78% of the target children were of White/Caucasian, 4% of Black/African American, 4% of Hispanic/Latino, 3% of Asian/Asian American, 1% of Native American/American Indian, 8% of mixed race/ethnicity, and 2% of another race/ethnicity. The data were collected online from a non-commercial website over ten years between 2004 and 2013 and participants volunteered to

[1]Data from 364 participants with more than 40 missing personality responses were dropped (N = 3,641). Then, we limited age from 16 to 19 years, because most participants with valid data were in this age range (N = 2,269). There were 306, 919, 825, 219 participants in each age, respectively, whereas other ages had 78 people at most. We replaced missing values with the median of all values.

complete this survey in exchange for automatically generated feedback about their child's personality. In the present study, target children were divided into three age groups: childhood (ages 3 to 8 years), adolescence (ages 9 to 14 years) and emerging adulthood (ages 15 to 20 years), so that non-linear age trends and age differences could be described.

The participants of the HEXACO online study (HEXACOS) completed an anonymous self-report questionnaire at the hexaco.org website between October 19, 2014 and October 18, 2018 (Lee & Ashton, 2020). There were 370,857 participants in the initial sample (mean age = 30.22 years, 57.3% were male, 41.5% were female and 1.2% did not indicate their sex). Participants were drawn from about 48 different countries (e.g. US, Germany, Australia, Japan). The HEXACOS was also divided into two age groups. In one group, we selected participants who aged between 18 and 50 years old (N = 24,000, mean age = 35.15 years)[2] to compare with the original study. In the other group, we selected participants aged between 16 and 19 years old (N = 8,000, mean age = 17.50 years) in order to study the same age range as the AAA, for comparison. Note that there was overlap in participants' ages and some participants may have been sampled into both the younger and older samples.

The LOOPRP estimated the replicability of the personality-outcome literature. Participants were collected from the Qualtrics Online Sample service in 2017 and quota sampling was used to ensure

[2] To ensure an even distribution of age and sex across the two groups, for the analysis of age differences between 18 and 50 years we sampled 4,000 participants (2,000 women, 2,000 men) based on six age levels (aged 18-25, 26-30, 31-35, 36-40, 41-45, 46-50), amounting to 24,000 participants in total; we also sampled 2,000 participants to each of four age levels (aged 16, 17, 18, 19) for the analyses of age differences between 16 and 19 years, which amounted to 8,000 participants in total. We applied a similar procedure in LOOPRP and the GNEOPIRS, trying to sample participants, so that the age ranges would be as equal as possible, although due to smaller participant numbers we could not sample equal numbers to each age level and we could not ensure sex balance.

that the samples were representative of the United States population in terms of sex, race, and ethnicity (Soto, 2019). The 6,126 participants in the initial sample were selected into two age groups: between 18 and 50 years (N = 1,662, mean age = 35.09 years) and between 18 and 25 years old (N = 3,459, mean age = 21.89 years). There was again some overlap in participants' ages, and therefore the samples partly overlapped.

The GNEOPIRS data came from the German Revised NEO Personality Inventory validation sample (Olaru et al., 2018; Ostendorf & Angleitner., 2004). The data were collected in person over several years (1992-2002; estimated mean: 1999) in more than 50 individual studies. There were a total 12,003 participants (in non-clinical subsample: 35.99% were male, in clinical subsample: 50.2% were male) with ages ranging from 16 to 91 years (mean age = 30.8 years) in the initial sample. To ensure this sample was representative of the general population, it also included a clinical subsample (N = 279). Participants were from three different countries with approximately 94% from Germany, 5% from Austria and 1% from Switzerland. We separated the GNEOPIRS into two groups, with 3,324 participants aged between 18 and 50 years old in one group (mean age = 34.48) and 5,796 participants aged between 18 and 25 years old in the other group (mean age = 21.64). There was also overlap in participants' ages in the GNEOPIRS, hence the younger and older sample partly overlapped.

We conducted a post-hoc power analysis for all included datasets using the pwr R package (Champely et al., 2018) in order to evaluate whether we had enough power to detect any true effects (Cohen, 1988), here correlations between predicted and observed ages. Results showed that sample sizes for each of our datasets had a power of .80 for detecting correlations of r = .10 or lower with two-tailed alpha set to .001.

Measures

We used five different questionnaires to measure personality traits. In the AAA dataset, participants completed a 240-item Estonian version of the International Personality Item Pool NEO (EE.PIP-NEO) (Mõttus et al., 2006). It mimics the structure of the Revised NEO Personality Inventory (NEO-PI-R) in measuring the Big Five domains and 30 facets, with each facet assessed by eight items (Costa & McCrae, 2010). The EE.PIP-NEO was comparable to the NEO-PI-R in terms of relevant psychometric properties (Mõttus et al., 2006), but the EE.PIP-NEO is linguistically simpler. Participants were required to rate the items on a 5-point Likert scale ranging from 1 (wrong/ does not agree at all) to 5 (correct/completely agree).

In the CCQS, parents rated their children on the 100-item common-language California Child Q-Set (CCQ) (J. H. Block & Block, 1980; Caspi et al., 1992), which includes 94 items assessing personality traits (Soto & John, 2014). Sixty seven of these 94 personality-relevant items can be aggregated to measure six broad trait domains representing the "Little Six": the Big Five plus activity level (Soto, 2016). However, because the CCQ items were developed to be individually informative and non-redundant (J. H. Block & Block, 1980), they can also be analyzed at the item level. Participants were asked to rate each CCQ item on a 9-point Likert scale ranging from 1 (extremely uncharacteristic) to 9 (extremely characteristic).

In the LOOPRP, personality was measures using the Big Five Inventory-2 (BFI-2) (Soto & John, 2017). The BFI-2 assesses the Big Five domains and 15 more-specific facets using 60 items written as short, easy-to-understand phrases, with four items per facet. Participants rated each item on a 5-point Likert scale ranging from 1 (disagree strongly) to 5 (agree strongly).

The HEXACO-PI, used to measure personality traits in HEXACOS, was developed to measure six major dimensions that have been found in several previous lexical studies of personality structure (Ashton et al., 2004; Lee & Ashton, 2004). Each HEXACO-PI domain subsumes four facets, with one additional, interstitial facet (Altruism) not scored on any of the six domains. Each facet scale includes four items, and participants rated each of the 100 HEXACO-PI items on a five-point scale from 1 (strongly disagree) to 5 (strongly agree).

In the GNEOPIRS, personality was measured with the German adaptation of the NEO-PI-R (Olaru et al., 2018; Ostendorf & Angleitner., 2004). NEO-PI-R measures the Big Five domains and 30 facets, with each facet assessed by eight items. Participants rated the 240 German NEO-PI-R items on a 5-point scale from 0 (strongly disagree) to 4 (strongly agree).

We emphasize that none of the measures was designed to capture personality nuances, perhaps apart from CCQ. However, based on existing evidence we considered it is likely that the items of each of the measure capture a broader range of traits than they have been designed to capture — nuances — although some more than others and none of them likely as comprehensively as an instrument designed to measure nuances would do.

For simplicity, all datasets will be labelled as the name of their measurement instruments rather than the dataset names for the rest of the study. Therefore, the AAA dataset will be referred to EE.PIP-NEO, the CCQS dataset will be referred to CCQ, the LOOPRP dataset will be referred to BFI-2, the HECACOS dataset will be referred to HEXACO, and the GNEOPIRS dataset will be referred to NEO-PI-R.

Data analysis

Statistical analyses were carried out in R (R Development Core Team). The EE.PIP-NEO, CCQ, BFI-2 data and the R scripts are publicly available at [the Online Supplemental Material: https://osf.io/jh6q8/?view_only=e1ff6aa708454cb3b120db0be157a840]. Given the conditions of research ethics approval, the HEXACO data cannot be made publicly available, but it is available for researchers from Kibeom Lee and Michael Ashton. The NEO-PI-R dataset is available for researchers from Fritz Ostendorf.

To strictly test the incremental information provided by lower-level personality traits, we analysed personality nuances in terms of both raw and residualized scores. The residuals-based analyses could show how much age-related information in nuances was unique and how much was due to facets and domains they had been designed to measure. In EE.PIP-NEO, we residualized each of the 240 items for their 30 facets (and thereby the Big Five domains) using a linear regression model (with the item being residualized removed from its facet at the time). In CCQ, we residualized both the complete set of 94 personality-relevant items and the subset of 67 items assessing the Little Six domains for these domains in a similar way. In BFI-2, we residualized the 60 items for their 15 facets (and thereby the Big Five domains). In HEXACO, we also residualized the 100 items for their 25 facets and in NEO-PI-R, we residualized the 240 items for their 30 facets in a similar manner[3].

Following Mõttus and Rozgonjuk (2019), our analyses treated age as the outcome variable (in a statistical sense) and the personality traits (raw or residualized domains, facets, and items) as the predictor variables. Specifically, age was predicted from either (a) the Big Five domains, 30 facets, 240 items, and the residuals of the EE.PIP-NEO items; (b) the Little Six domains, 94 or 67 CCQ items and their residuals; (c) the Big Five domains, 15 facets, 60 items, and residuals of the BFI-2 items; (d) the six domains, 25 facets, 100 items and residuals of the HEXACO items; or (e) the

[3] Note that residualizing each item for all of the facet scores does not simply reorganize the variance captured by the raw item responses. Instead, it removes some information from the items: the common variance that is shared across items. Importantly, this common variance is the main focus of classical test theory and latent trait models; thus, according to those approaches our residualizing procedure should remove all of the interesting variance from the items. However, our results show that the residualized items still retain considerable information.

Big Five domains, 30 facets, 240 items, and residuals of the NEO-PI-R items. Since the inclusion of a high number of predictors can lead to over-fitting with complex model having an a priori advantage, (Yarkoni & Westfall, 2017; Zou & Hastie, 2005), the models were trained using (linear) elastic net regressions and applied for prediction in separate sample partitions, with a 75% - 25% split, respectively. A simple regression model trained and applied for prediction in the same sample (as is customary in most psychological modelling) would lead to inflated (over-fit) coefficients, but the current approach of using penalized regression coefficients by a) shrinking them towards zero and b) separating model training from its testing effectively mitigates this problem (Seeboth & Mõttus, 2018). Elastic net regressions were implemented in the glmnet package (Friedman, Hastie, Simon, & Tibshirani, 2016), with 10-fold cross-validation and the regularization parameter ("lambda min") minimizing the cross-validation error. This reduces the chance of capturing more variance in the outcome purely based on the inclusion of many predictors and thus helps estimating a model that is likely to also generalize to other populations (Yarkoni & Westfall, 2017). In each dataset, we repeated the training-validation procedure for 500 random sample splits. In the validation samples, the predicted age was then correlated (using Spearman's rho) with the actual ages. This approach yielded 2,000 correlations for each type of analysis in the EE.PIP-NEO, BFI-2, HEXACO, and NEO-PI-R —500 for domains, 500 for facets, 500 for items, and 500 for the residuals of the items—and 1,500 correlations for each analysis in the CCQ (because there were no facets in this dataset). We then computed the mean and standard deviation for each type of correlation in each sample, and report these summary statistics as our main results. Using this procedure, if the associations between personality traits and age were primarily due to domains and facets, then the mean correlation between predicted and actual age would be stronger when age was predicted from the domains or facets than when it was predicted from the items (Seeboth & Mõttus, 2018).

We note that age was predicted from these sets of predictors *not* because we believed age to be causally influenced by personality *nor* because age is inherently something worthwhile to predict from other pieces of information (although sometimes it could be, for example when age is unknown but of interest). Instead, these statistical models allowed us to estimate and compare how much age-sensitive information each set of personality traits (domains, facets, and nuances) contained. That is, prediction was used in a purely statistical sense to quantify and compare information captured across trait hierarchy levels, with no causal implications.

Results

Do Personality Nuances Capture More Age-Related Information than Domain and Facet Traits?

In the combined age group of both the 67 items-based and 94 items-based CCQ, both age groups of the HEXACO and NEO-PI-R, the predictive accuracy of domains, facets and nuances followed the same pattern that lower-level personality traits containing more age-sensitive information than higher-order traits (see Table A1). This pattern was again displayed in the broader age group (18 to 50 years) of BFI-2. However, this pattern did not appear in the narrower age group (18 to 25 years): the accuracy of the domain-based predictions was only somewhat lower; the accuracy of the item-based predictions and the accuracy of the facet-based predictions were almost the same. Similarly, the pattern also did not appear in the EE.PIP-NEO. The accuracy of the domain-based predictions and facet-based predictions were similar while item-based predictions was higher than the domain and facet-based predictions.

Table A-1.Correlations between Actual Age and Age Predicted by Elastic Net Models Based on the Big Five Domains, Facets, Items, and Their Residuals.

Datasets	Age range	Parameter	Big five	Facets	Items	Residuals of items
M & R (2019)	aged 18-50	Mean	.28	.44	.65	.65
		SD	< .01	< .01	< .01	< .01
EE.PIP-NEO	aged 16-19	Mean	.11	.11	.16	.09
		SD	.04	.03	.03	.03
		95%CI	.06 - .17	.04 - .17	.10 - .21	.03 - .14
CCQ (67 items)	aged 2-20	Mean	.44	-	.68	.67
		SD	.01	-	< .01	< .01
		95%CI	.43 - .46	-	.66 - .69	.66 - .69
CCQ (94 items)	aged 2-20	Mean	-	-	.73	.73
		SD	-	-	< .01	< .01
		95%CI	-	-	.72 - .74	.72 - .74
BFI-2	aged 18-25	Mean	.10	.11	.11	.07
		SD	.02	.03	.02	.02
		95%CI	.06 - .16	.05 - .16	.07 - .17	.02 - .11
BFI-2	aged 18-50	Mean	.13	.21	.24	.21
		SD	.03	.03	.03	.03
		95%CI	.07 - .22	.14 - .27	.17 - .30	.12 - .28
HEXACO	aged 16-19	Mean	.13	.19	.24	.24
		SD	.02	.02	.02	.02

		95%CI	.11 - .17	.16 - .22	.23 - .29	.23 - .29
HEXACO	aged 18-50	Mean	.24	.36	.50	.50
		SD	.01	.01	< .01	< .01
		95%CI	.21 - .25	.34 - .38	.49 - .51	.49 - .52
NEO-PI-R	aged 18-25	Mean	.18	.25	.37	.36
		SD	.02	.02	.02	.02
		95%CI	.14 - .22	.21 - .29	.33 - .41	.32 - .41
NEO-PI-R	aged 18-50	Mean	.33	.49	.66	.65
		SD	.02	.02	.02	.02
		95%CI	.30 - .39	.45 - .53	.62 - .69	.62 - .69

Note. The mean and standard deviation (SD) are across 500 replications with the training sample of 75% of the total sample. The M & R dataset indicated the results from the original study.

We also compared the predictive ability of the Big Five domains, items, and their residuals in narrower age groups within the CCQ. Specifically, we divided the CCQ into three narrow age ranges— ages 3 to 8 years, 9 to 14 years, and 15 to 20 years—and then analysed 500 random sample partitions with 75% of the age group used for training and 25% used for validation in each case. The purpose of analysing these narrower age ranges was to test whether domain, facet, and nuance-level traits could detect the non-linear trends in age differences that are often observed in childhood and adolescence (Denissen et al., 2013; Soto & Tackett, 2015). The results (Table A2) indicated curvilinear trends in the predictability of age and hence in the associations of personality traits with age: that overall amount of age-related information captured by personality traits was highest during childhood and then gradually decreased into adolescence and early adulthood, consistent with age-differences often being non-linear.

Table A-2. Correlations between Actual Age and Age Predicted by Elastic Net Models Based on the Big Five Domains, Items, and Their Residuals in Four Age Groups in CCQ.

CCQ	Parameter		Little Six / Big Five	Facets	Items	Residuals
CCQ (67 items)	*Age 3 to 8 years*	Mean	.32	-	.53	.53
		SD	.02	-	.02	.02
		95% CI	.28 - .35	-	.51 - .56	.50 - .57
	Age 9 to 13 years	Mean	.22	-	.37	.37
		SD	.02	-	.02	.02
		95% CI	.16 - .26	-	.33 - .40	.33 - .41
	Age 14 to 20 years	Mean	.08	-	.16	.16
		SD	.02	-	.02	.02
		95% CI	.04 - .13	-	.11 - .20	.12 - .20
CCQ (94 items)	*Age 3 to 8 years*	Mean	-	-	.58	.58
		SD	-	-	.01	.01
		95% CI	-	-	.55 - .60	.55 - .61
	Age 9 to 13 years	Mean	-	-	.41	.41
		SD	-	-	.02	.02
		95% CI	-	-	.38 - .44	.38 - .44
	Age 14 to 20 years	Mean	-	-	.20	.20
		SD	-	-	.02	.02
		95% CI	-	-	.17 - .24	.16 - .24

Note. The mean and standard deviation (SD) are across 500 replications with the training sample of 75% of the total sample.

These results indicate a general trend whereby lower levels of the personality hierarchy (facets and especially nuances) contain more age-sensitive information than higher levels of the hierarchy. This pattern generalized across age groups, rating perspectives, and personality inventories, with age differences in the EE.PIP-NEO and young-adult BFI-2 sample being the exceptions, thus replicating and extending the findings of Mõttus and Rozgonjuk (2019). However, the overall amount of age-relevant information captured by personality traits varied across both developmental periods (it was generally smaller when younger participants with smaller age ranges were studied) and instruments (with the BFI-2 containing less age-sensitive information than the HEXACO and NEO-PI-R, for example). Differences in the breadth of item content included on different measures could be one possible explanation for this finding. For example, the BFI-2 was developed to measure the most prototypical content within each Big Five domain, besides being the shortest of the measures. Its domain and facet scales may therefore be more narrowly focused and unidimensional than the other measures analysed, leading to less unique information at the nuance level as well as capturing less age-sensitive information at the domain level (because domains are nothing but aggregates of their nuances).

Does this Pattern Hold while Controlling for Overlap between Domain, Facet, and Nuance Traits?

Our second set of analyses tested whether the finding that nuance-level traits capture more age-relevant information than domain- and facet- level traits would remain robust even when the overlapping information shared between domains, facets, and nuances was removed. We therefore repeated the analyses described above after residualizing each inventory item for all higher-order facets (where applicable) and domains. These residualized items no longer contained the variance of any of the Big Five (Little Six, HEXACO) domains and facets; instead, the residuals only reflect the unique age-sensitive information captured by individual items.

In the CCQ, residualizing either 67 or 94 items did not impact their collective ability to capture age-sensitive information (see Table A1); item residuals still outperformed the domain-level traits. This also applied to the HEXACO and NEO-PI-R (see Table A1). Residualizing items for their facets had essentially no impact on how much age-sensitive information they contained, either in young adulthood specifically or across three decades of adult life more broadly. These results further suggest that age-related information was better captured by single items than by broader trait constructs measured from the items' shared variance: age differences are mostly in nuances in which people vary even when statistically made identical ("clones") in the domains and facets. However, in the EE.PIP-NEO and BFI-2 dataset, residualizing items for domains and facets slightly decreased their collective ability to capture age-sensitive information (see Table A1), although they still did predict age. As stated in the previous paragraph, the BFI-2 domains and facets are more narrowly focused than the other personality measured analysed in the present research, which might explain why the BFI-2 items contained less age-sensitive residual variance once the domain and facet-level information is removed. However, this hypothesis would not help explain the relatively small predictive value of residualized items in the EE.PIP-NEO.

Why items out-predicted broader traits?

Figure A1 to A4 indicated that items of the same facets (or domains) often (but not always) differed in terms of their correlations with age, which help to explain why item-level analyses captured more age-related information than domain and facet-level analyses. For example, while all NEO-PI-R items from the O1: Fantasy facet were negatively correlated with age, several facets such as the A6: Tender Mindedness facet and the C4: Achievement Striving facet contained items that correlated with age in different directions, thus reducing the facet- and domain-level personality trait–age associations.

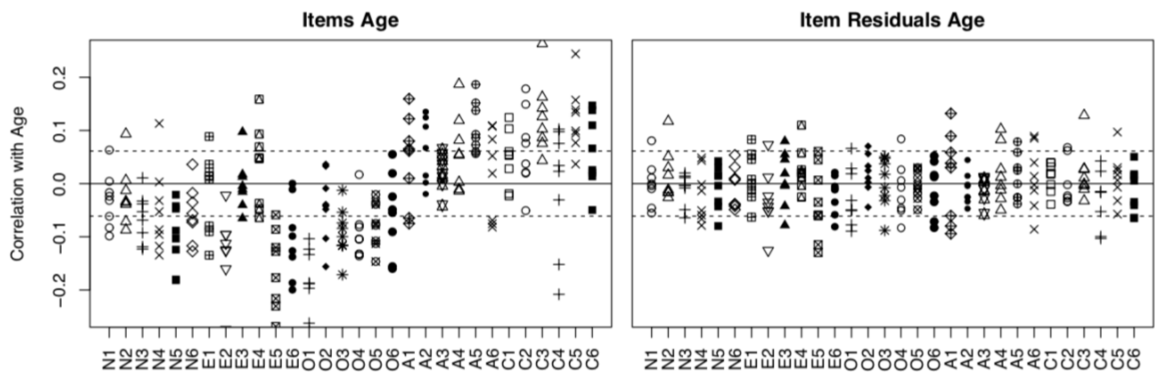


Figure A-1. Plots for the correlations of the 240 items of the NEO-PI-R and their residuals with age (with participants aged between 18 and 50). The correlations are grouped according to the Big Five domains (indicated by letter) and their facets (indicated by number). The dashed line indicates the threshold for statistical significance after correcting for multiple comparisons using Holm's correction (Holm, 1979).

As a result, nuance-level associations can provide additional information for understanding psychological development, although we note again that items of questionnaires that have not been designed to capture nuances may not be most suitable for interpreting nuance-age correlations: they only signal the presence of such associations, but for a proper interpretation of them, appropriate instruments will ultimately be required. For example for participants aged between 18 and 50 years, NEO-PI-R items in the E1: Warmth facet that referred to liking others showed positive correlations with age, but items that referred to enjoying gabbing, talking, and emotionally attaching to others decreased with age: older people might be more accepting of others, but not as talkative as younger people, on average. For the E4: Activity facet, items that referred to having a fast moving life showed negative correlations with age, but items that referred to being energetic and engaging in strenuous activities showed positive correlations with age. For the C4: Achievement Striving facet, scores of items that referred to being ambitious and wanting to get ahead slightly decreased with age, whereas scores of items that referred to being enthusiastic and working excessively showed positive correlations with age: older people might not be as ambitious as younger people, on average, and yet even more hard-working.

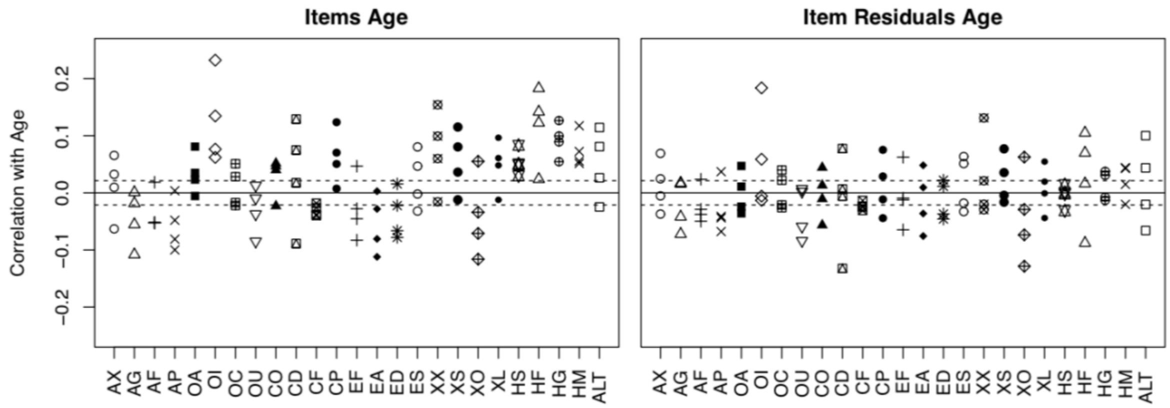


Figure A-2. Plots for the correlations of the 100 items of the HEXACO and their residuals with age (with participants aged between 18 and 50). The correlations are grouped according to the six HEXACO domains (indicated by the first letter: A for Agreeableness, O for Openness to Experience, C for Conscientiousness, E for Emotionality, X for Extraversion, H for Honesty-Humility) and their facets (indicated by the second letter: see https://osf.io/d6t37/ for facet names). The dashed line indicates the threshold for statistical significance after correcting for multiple comparisons using Holm’s correction (Holm, 1979).

For participants aged between 18 and 50 years, HEXACO items showed a similar pattern of occasional within-facets variance in age differences more generally and working hard specifically. In the Diligence facet, the item that referred to setting ambitious goals negatively correlated with age, while items that referred to working hard increased with age: older people might be more realistic in terms of setting goals than younger people, but work harder towards achieving their goals. A similar pattern for diverging age trends in ambitiousness and working hard has also been reported previously (Mõttus et al., 2015; Mõttus & Rozgonjuk, 2019). Such robustly divergent age trends can be very informative.

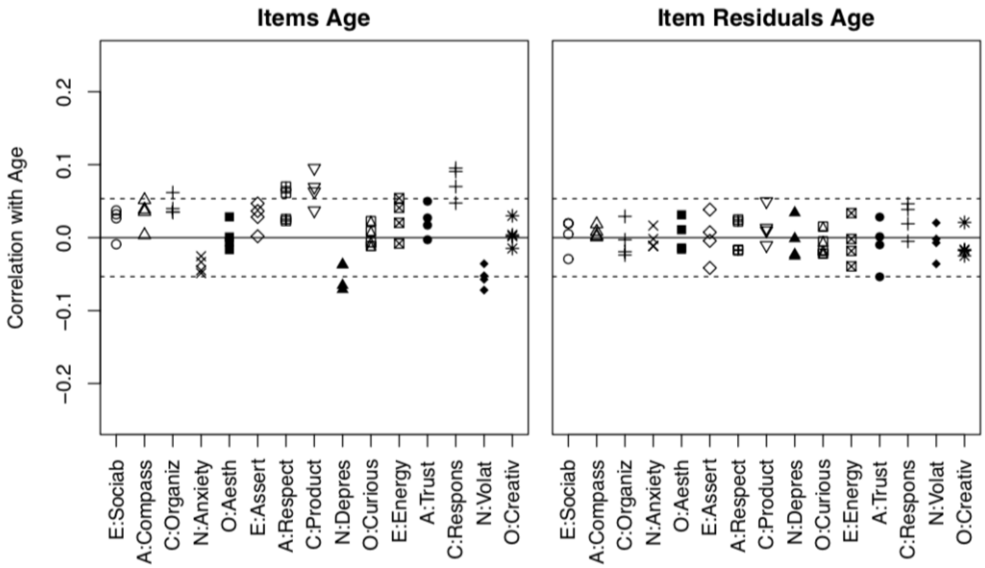


Figure A-3. Plots for the correlations of the 60 items of the BFI-2 and their residuals with age (with participants aged between 18 and 25). The correlations are grouped according to the Big Five domains (indicated by letter) and their facets (see https://osf.io/se7r4/ for facet names). The dashed line indicates the threshold for statistical significance after correcting for multiple comparisons using Holm's correction (Holm, 1979).

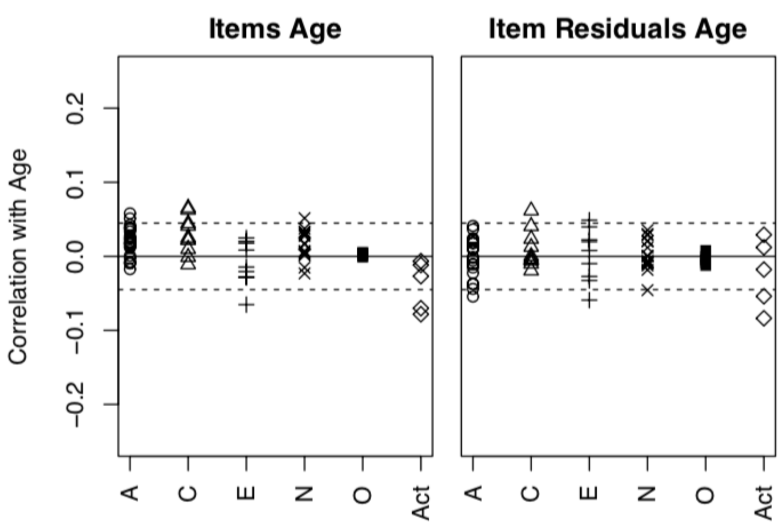


Figure A-4. Plots for the correlations of the 67 items of the CCQ and their residuals with age (for participants aged 14 to 20). The correlations are grouped according to the Little Six domains (the Big Five and activity). The dashed line indicates the threshold for statistical significance after correcting for multiple comparisons using Holm's correction (Holm, 1979).

Here we aimed to explain the reasons for the generally lower predictive accuracy in domain and facet level traits in comparison to items, and therefore highlighted item-level associations that showed opposite directions in their relations with age within the same facets. However, we want to be absolutely clear that in many cases items belonging to the same facets showed consistent correlations with age. For example, all CCQ items of the Activity domain showed negative associations with age indicating that children might become less physically active as they get older, across the nuances of the domain. Similarly, all NEO-PI-R items from the C3: Dutifulness facet and the C5: Self-discipline facet were positively correlated with age indicating that people might become more responsible and disciplined as they mature across all nuances of these facets.

Do Longer Personality Inventories Provide More Nuance-Level Information?

In their initial study comparing the age-related information captured by domain, facet, and nuance-level personality traits, Mõttus and Rozgonjuk (2019) suggested that more information in terms of more variables should allow for better prediction. To test this hypothesis, we examined the age-predictive ability for personality inventories of different lengths. Specifically, we re-ordered different personality questionnaires based on the number of items in an ascending manner: BFI-2 (60 items), CCQ (67 items), CCQ (94 items), HEXACO (100 items), NEO-PI-R (240 items), and EE.PIP-NEO (240 item). In theory, the predictive ability of personality nuances should also follow this order.

In the age range between 18 to 50 years, the results obtained with the NEO-PI-R, HEXACO and the BFI-2 indeed showed that the predictive ability of the models increased with the number of items. The predictive accuracy of the 240-item NEO-P-IR was highest among the domain, facet, and nuance models, and the predictive accuracy of the 60-item BFI-2 was the lowest among these three models (see Table A1). For samples in the age range of 14 to 25 years old, the predictive ability of the NEO-PI-R, HEXACO, CCQ (67 and 94 items) and BFI-2 mostly followed the expected order (see Table A1). However, on the domain level, the predictive ability of BFI-2 slightly outperformed the predictive ability of CCQ. The only personality inventory that did not follow the general pattern of longer inventories capturing more age-relevant information was the EE.PIP-NEO. Theoretically, the predictive accuracy of the EE.PIP-NEO should have been comparable to the NEO-PI-R because they both include 240 items. However, the predictive accuracy of EE.PIP-NEO was similar to the much-shorter BFI-2.

Another noteworthy pattern apparent from these analyses is that the predictive accuracy of groups with broader age ranges generally outperformed groups with narrower age ranges. For example, the predictive accuracy of participants aged around 18 to 50 years old in the HEXACO was substantially better than the predictive accuracy of participants aged around 16 to 19 years old in the HEXACO. This same pattern can also be observed in the BFI-2 and NEO-PI-R (see Table A1). The most likely explanation for this pattern is that greater variance in age allows for greater covariance between age and personality traits.

Discussion

We compared the amounts of age-related information captured by domain, facet, and nuance-level personality traits. We tested whether the results generalized across samples representing different developmental periods, cultural backgrounds, rating perspectives, and personality instruments.

Within each sample, we used elastic net models and random sample partitions to prevent overfitting and capitalization on chance. Our findings support four key conclusions: (a) that lower levels of the personality hierarchy (facet and especially nuance-level traits) generally contain more age-sensitive information than higher-level trait domains like the Big Five; (b) personality nuances usually retained their age-sensitive information even after controlling for overlap with their higher-order facets and domains; (c) personality inventories that can assess a larger number of personality nuances because of containing more items generally provided more age-sensitive information than do shorter measures; and (d) the amount of age-sensitive information captured by personality inventories was generally highest during childhood, then decreased into adolescence and adulthood. Moreover, the predictive advantage afforded by nuance-level traits was substantial: across all analyses reported in Table A1 that compared the three levels of the trait hierarchy, personality nuances allowed for 36% more accurate prediction of age than did facets, and 87% more accurate prediction than did domains, whereas facets allowed for 38% more accurate predictions than did domains.

These findings successfully replicate and extend previous research on age differences in personality nuances. Using a similar analytic design in the age range of 18 to 50 years, Mõttus and Rozgnonjuk (2019) found that 300 items allowed for 47% and 132% more accurate predictions of age than the Big Five facets and domains, respectively. Our results similarly indicate a pervasive pattern for facets to capture more age-relevant information than domains and item-level nuances to capture more information still. Even the unique variance in items, after controlling for the domains and facets, contained on average over 58% more age-sensitive information than the domains. This finding may seem puzzling at first: how could residualized variance, which is considered “measurement error” in classical test theory, in fact contain much more predictive information than trait scores, which are often taken to approximate error-free "true scores"? But they can in fact be explained: what varies mostly with age in personality is the unique variance of many lower-level traits that can be summarized with broader traits that aggregate these lower-level traits (as much of research does), but not accounted for by them.

This general pattern held for children, adolescents, emerging adults and adults; it also held across sample languages / cultures, rater perspectives and specific instruments. But the present results also suggest two potential boundary conditions for these findings: on average, personality nuances captured less age-sensitive information when there was less between-person variance in age (and therefore less potential for age to covary with personality nuances), and when the personality questionnaires used were relatively brief (and therefore assessed fewer personality nuances).

Implications

The present findings have important implications for understanding life-span personality differences. They suggest that age-trends in personality traits can be described along many dimensions, meaning that lower levels of the personality trait hierarchy may sometimes and possibly even often be better suited for studying age differences in personality traits than higher levels. This is because these lower-level traits contain "free" information that is typically thrown away when specific personality nuances and facets are aggregated into broad, domain-level traits; in fact, the broader traits may only be useful for summarizing but not for explaining age differences.

This insight implies at least three possible benefits. First, nuance-level analyses provide a possibility to describe age differences in personality traits in greater detail. Items within the same domains and facets can vary in their age trends (McCrae, 2015; Mõttus, Kandler, et al., 2017), so it makes sense to incorporate this information to get a more accurate picture of personality age differences and avoid undue generalizations about broader traits when some of the traits' facets and nuances show distinctive age trends. As personality research more generally matures and research questions become increasingly focused and refined, nuance-specific information may prove valuable for informing specific research questions; for example, at which life stage are people most lonely or competitive, on average. Clearly, there is still a lot to be discovered with respect to how people psychologically vary with age.

Second, representing age differences with many personality dimensions makes it possible to test new hypotheses. Specifically, nuance-level analyses will allow researchers to study systematic variations between traits in terms of how (a) nuances change with time and (b) how they intersect with developmental factors / mechanisms that characterize the nuances in different degrees. For example, some have hypothesized that personality development largely reflects social maturation (Caspi et al., 2005). To test this hypothesis, they could quantify the mean-level changes for a large pool of nuance-level traits as well as the degrees to which these traits reflect social maturity, and then correlate these two properties across all of the nuances. Similarly, some have proposed that personality maturation is largely driven by the development of self-regulation (Denissen et al., 2013). To test this hypothesis, they could correlate the degrees of age differences in a pool of personality nuances with the degrees of self-regulation required to express each nuance. The item-specific age differences that we briefly described would support the case for testing these hypotheses: it often seemed as if older age was associated with more realistic expectations and greater ability to adapt to circumstances, which are plausible signs of maturity and greater self-regulatory skills. Testing such associations of personality traits' developmental trends with the traits' other properties requires a large and diverse pool of traits, because the effective sample is the number of traits rather than the number of participants. Therefore, such tests would be best conducted when conceptualizing and measuring personality in terms of a large and diverse pool of nuance traits, rather than a much smaller set of broad traits such as the Big Five or HEXACO domains.

Third, the nuance-level analyses can help to explain why findings obtained with different instruments often do not converge (Costa et al., 2019): this is plausibly because different instruments sample different nuances. For example, age differences in Extraversion and Openness can even vary in direction, depending on which lower-level traits their measures happen to samples (Costa et al., 2019). For example, items of the diligence facet that referred to hard working positively correlated with age while scores of items that referred to having ambitious goals negatively correlated with age. This suggests that older people might have relatively more realistic goals and work hard to achieve them, while younger people might have somewhat more ambitious goals and more urge for success. Such examples of disparate correlations between items of the same facet and age were numerous.

Overall, the parsimony-focused Big Five or FFM trait models have been criticized for multiple inadequacies (e.g. lack of discriminant validity or lack of an underlying theoretical model), but they have provided the most widely used traits for operationalizing personality variance and, by virtue of allowing research findings to accumulate, they have led to fundamental advances in the description of personality-related phenomena in the past decades (J. Block, 2010; Mõttus et al., 2020; Paunonen & Ashton, 2001). That is, the Big Five domains have been very instrumental for the progress of descriptive personality research. As a result, using nuances to describe personality variation and its links with other variables may go against how many personality researchers have been trained and are used to think – to categorize lower-level traits into fewer and broader groups for greater parsimony, without necessarily considering whether those lower-level traits themselves might offer unique information. However, the present study aims to illustrate the possible value in also considering the lower-level personality traits, nuances, and not always *a priori* aggregating them into broader traits. We want to make it very clear that such nuanced research is not aimed at negating the Big Five-focused research, but to extend it and qualify the finding.

Of course, before more "nuanced" personality research can really take off, it will be necessary to develop psychometrically sound scales for measuring the nuances in future research. Developing a new generation of tools for personality assessment will require a lot of work, but it is not impossible and strategies for this have already been articulated (Condon et al., 2020; Mõttus et al., 2020). However, because developing a proper model for accessing nuances is effortful and time-consuming, it is still necessary to bolster the case for this effort, as was done in the current study.

Limitations and Future Directions

Although the present findings have important implications for understanding age differences in personality traits, they should also be interpreted with five key caveats in mind. First, the idea of "predicting" age seems unusual since, implicitly, we often think of age as an explanatory variable and not an outcome. However, in the present research "prediction" was simply a tool for quantifying the amount of age-sensitive information that a set of personality traits collectively contain. Machine learning prediction models such as elastic net are particularly well suited for comparing sets of variables in how much criterion-relevant information they contain, because they allow for better control of overfitting than traditional regression models. Therefore, the present findings regarding "prediction" should only be interpreted in a statistical sense, without any implications regarding causality.

Second, one alternative explanation for the present results is that narrow personality traits within broader domains may only show distinctive age trends because they explicitly capture social expectations and roles that vary with age. However, the finding that items out-predict domains even when only closely adjacent age-levels are compared make this unlikely to be the sole explanation: social roles and expectations are unlikely to change so quickly. The findings also generalized across inventories which vary in how their items tend to be written: for example, the NEO-PI-R and HEXACO items tend be more contextualized (and thereby perhaps more sensitive to age differences in social roles) than the BFI-2 items, which are more abstract and less likely to refer to any specific social roles and expectations. Moreover, BFI-2 items are based on adjectives and many items with interstitial content that fell between multiple domains or facets were excluded (Soto & John, 2017). These adjective based items might not capture as much social role variance as contextual and behavioural items do. Moreover, the pattern also held in parent-ratings, which mitigates the possibility that the nuance-specificity of age differences may have reflected differences in how items were interpreted at different ages. Also note that personality inventories such as NEO-PI-R tend be fairly invariant across ages in how their items correlate with each other even though the items of the same traits substantially vary in how they correlate with age, suggesting that the meanings of items do not vary as much with age as the content does (Mõttus et al., 2015).In our view, it is at least as plausible that personality traits *per se*—specific patterns of thinking, feeling, and behaving—vary in highly nuanced ways across these years.

A third caveat is that although the present study analysed data from multiple large and diverse samples of participants, all of the data were cross-sectional in nature. We therefore inferred personality change rather than directly observing it. Therefore, future research is needed to examine the development of domain, facet, and nuance-level personality traits in longitudinal samples.

Fourth, we compared personality questionnaires with different numbers of items to show that the predictive accuracy of longer personality questionnaires are better than the accuracy of shorter ones. In this way it is difficult to control for whether the predictive accuracy is solely due to the larger number of items or due to other questionnaire differences. To establish whether it indeed due to the larger number of items, it would make sense to select different sized samples from the same item pool and then compare the results. For example, we compared 67 items with 94 items in the CCQ dataset. However, we did not do the same in other datasets. Future studies could combine both methods to obtain a more reliable result.

Finally, although the present research operationalized personality nuances as individual questionnaire items, at present there are no theoretical models to justify which items should be used in nuance-level personality research. Most of the items analysed here—and in previous research on personality nuances—were chosen to measure broader trait domains and facets, and purposefully *not* to capture unique item-level information. Thus, future research would benefit from developing a stronger theoretical basis for choosing items in order to advance nuance-level analysis of age differences, as there is likely a much larger pool of nuances that vary with age than is captured by the items currently used in personality inventories. Moreover, personality traits can be conceptualized at levels of abstraction other than domains, facets, and nuances, from the very broad (e.g., meta-traits) to the very narrow (e.g., highly homogeneous item clusters). We have focused on the domain, facet, and item levels because they are the most prominent levels assessed by contemporary personality inventories. However, future research could establish a basis for extending the present results to other levels of abstraction. This basis could be established in ways similar to how researchers developed domain and facet-level personality measures (Mõttus & Rozgonjuk, 2019). In other words, researchers should consider developing new measurement models of personality that explicitly consider nuances and therefore allow for more comprehensive personality assessment. This is not unrealistic: Condon and colleagues (Condon et al., 2020) have described the concrete steps needed for the development of such models.

Conclusion

The present findings show that specific, nuance-level personality traits—as operationalized by individual personality questionnaire items—capture a great deal more of age-related personality information than broader traits do. We therefore conclude that future research can benefit from examining personality development at multiple levels of abstraction. Domain-level models like the Big Five and HEXACO are suitable for many purposes; for example, they allow researchers to efficiently describe broad-stroke patterns, which can be useful for describing personality development to the public or students (Mõttus et al., 2020). However, nuance-level analysis can allow us to characterize age differences in personality traits in much greater detail, which will be especially useful as personality research matures and the research questions become increasingly focused. As the result of this trade-off between efficiency and specificity, some studies might like to examine personality age differences at the level of broad domains; some might want to study the extra information provided by nuances; and others might choose to address personality differences using multiple levels of the personality hierarchy.

Chapter 3: Mechanisms of Personality Differences in Childhood

Synopsis:

The previous chapter discussed that nuance-level analyses capture more unique information than facets and domains and highlighted that nuance-level analyses tend to be overlooked in the current personality research literature. To extend current research on personality differences based on domain and facet level analyses, this chapter investigates mechanisms of personality development during childhood and adolescence using nuance-level personality traits. Social expectations and self-regulation were the two key mechanisms that were tested in the present chapter. Results indicate that strong social expectations partially contribute to personality differences during childhood and adolescence but, contrary to hypotheses, self-regulation did not explain personality changes.

Dissemination status:

This chapter has been under review at *Journal of Personality* entitled "Social expectations and abilities to meet them as possible mechanisms of youth personality development". A pre-print is available on *PsyArxiv*: https://doi.org/10.31234/osf.io/8yp6x

Acknowledgments:

Authors are grateful to Jaap Denissen for his extremely helpful comments throughout the study.

Social Expectations and Abilities to Meet Them as Possible Mechanisms of Youth Personality Development

Yuzhan Hang[a, *], Christopher Soto[b], Billy Lee[a], Lydia Speyer[a], Aja Murray[a], René Mõttus[a]

Affiliations:[a]Department of Psychology, University of Edinburgh, Edinburgh, United Kingdom, [b]Department of Psychology, Colby College, Maine, United States

Abstract

Rationale: Personality traits change in both mean levels and variance from childhood through mid-adolescence, but the mechanisms underlying these developmental trends remain unknown. We tested the possible roles of social pressure and self-regulation. Methods: The Common-Language California Child Q-Set was used to measure youths' mean-level personality, social expectations for youths' behaviour from multiple perspectives (parents, teachers, peers) and the self-regulatory requirements for achieving the desired trait levels. Results: There were consistent expectations for youths' traits, regardless of who described the expectations or whether these pertained to children or adolescents. Mean trait levels were moderately commensurate with social expectations, but age differences only followed these expectations in late adolescence. Traits with strong expectations showed more pronounced individual differences and increased even more in variance with age. In contrast, traits' self-regulation requirements did not predict their developmental trends. Implications: Strong social expectations may contribute to the development of individual differences, but there was little evidence for systematic development to catch up with consensually held social expectations.

Keywords: personality development; mean-level; variance; social expectations; self-regulation

Introduction

Personality traits change in both mean-levels and variances from childhood through mid-adolescence (Mõttus et al., 2016; Mõttus, Soto, et al., 2017). Although in adulthood personality traits tend to shift gradually in a socially desirable direction, with people becoming, on average, more agreeable, conscientious, and emotionally stable over time (Allemand et al., 2008; Caspi et al., 2005; Donnellan & Lucas, 2008; Van den Akker et al., 2014), the pattern is different for childhood and adolescence. Specifically, there are temporary mean-level decreases in Agreeableness, Conscientiousness, Openness to Experience and Emotional Stability (Allik et al., 2004; Denissen et al., 2013; Soto, 2016; Soto et al., 2011). Also, although there may be no systematic age-trends in the magnitudes of individual differences in adulthood (Mõttus et al., 2016), personality trait variance appears to increase from childhood until mid-adolescence and plateaus thereafter (Allik et al., 2004; de Haan et al., 2013; Kandler & Zapko-Willmes, 2017; Mõttus, Soto, et al., 2017). The mechanisms underlying these developmental trends in adolescent trait levels and variances remain unknown.

Relying on the social investment theory (SIT) (Roberts, Wood, et al., 2005), Denissen and colleagues (2013) discussed the roles of social expectations and self-regulation as possible mechanisms of youth personality development. SIT suggests that personality matures as individuals are confronted with new social expectations and alter their behaviour accordingly. Denissen and colleagues (2013) expanded on SIT by proposing that although changes in expectations drive personality change by motivating individuals to meet these expectations, regulatory skills determine how much individuals can actually change their traits. In their meta-analyses, Denissen and colleagues (2013) found that mean levels of two personality traits, conscientiousness and openness, decreased during early adolescence, followed by an increase in later adolescence. In their view, as social expectations increase with age, early adolescents might lack sufficient regulatory skills to meet these expectations which thus leads to a temporary decrease in how their personality traits are perceived by themselves and others – due to an increased gap between expectations for the youth and their actual behaviour. Late adolescents might have more developed regulatory skills which allows them to meet the expectations, which in turn leads to the rebound of rated personality traits in late adolescence – as the gap between expectations and actual behaviour narrows. The present study set out to test the key ideas of Denissen and colleagues (2013) by further investigating whether (a) social expectations for youths' behaviour, and (b) the degree of self-regulation needed to meet these expectations are among the potential mechanisms that could explain personality trait changes in both mean levels and variance.

Social expectations and normative personality differences

If there are consistent expectations from other people for children and adolescents to behave (here, broadly defined and including thinking, feeling and acting) in particular ways, then it is reasonable to hypothesize that youths will generally attempt to comply with these expectations (Roberts, 2018). If so, then normative personality development (i.e., patterns of personality differences that generalize across individuals) may in part be driven by changes in the social expectations and/or changes in youths' abilities to meet them (Denissen et al., 2013).

For example, it may be generally expected that children should consider other people's feelings and not display jealousy, and most children may indeed do this most of the time. Such expectations may become stronger with age and most children may adjust their behaviour accordingly, producing a mean-level increase in these traits. However, even if such expectations remain similar over time, children's ability to regulate their behaviour so as to meet the expectations may improve, which could also prompt mean-level changes in these characteristics. But if the expectations increase more and/or faster than youths' ability to meet them, then the gap between expectations and youths' observed trait levels may (at least temporarily) widen. Moreover, if parent-reports or self-reports of personality traits are influenced by the rater's expectations, then this gap between expectations and behaviour could produce dips in some generally desirable trait levels as reported by parents or children themselves (Denissen et al., 2013; Soto, 2016; Soto et al., 2011; Van den Akker et al., 2014).

Either increasing social expectations and/or youths' improving abilities to meet them, may be explanations for why increases in trait variance slow down and stop in mid-adolescence (Mõttus, Briley, et al., 2019; Mõttus, Soto, et al., 2017). Mechanisms such as gene-environment transactions (Mõttus, Briley, et al., 2019) or gene-environment interactions (Kandler & Zapko-Willmes, 2017) may contribute to increases in variance; for example, individuals' traits may be pulled away from their genetically influenced baseline levels as a result of self-selected/created environments amplifying the traits (Caspi et al., 2005). However, social expectations and/or youths' improving abilities to meet them may counteract these processes and constrict variance by rewarding individuals for behaving alike – in the generally expected ways. There is yet another possibility: if social expectations for trait levels rise but youths' ability to meet them is yet to catch up with the expectations, then social pressure may initially contribute to *increasing* variance as some individuals are able to meet the expectations developmentally earlier than others. If so, the increases in variance would stop and may even be reversed once slower-developing youths eventually catch up with their peers. In this paper, we explored the association between variance in personality traits and the social expectations of them, without specifying a directional hypothesis for the possible link.

Testing the hypotheses

The hypotheses were tested by quantifying a sample of diverse personality traits in terms of their a) observed mean-level and variance changes through childhood and adolescence, b) socially expected levels, c) changes in these expected levels through childhood, and d) degrees of self-regulatory ability required to meet the expected levels.

Given these data, we addressed four key research questions:

Question 1: Do children's personality traits generally comply with social expectations? We hypothesize that traits expected to have higher scores indeed tend to have higher scores across all ages (Hypothesis 1).

Question 2: Do children's personality traits change in accordance with changes in social expectations? Mean-level changes in the traits may mirror traits' expected levels and/or changes in their expected levels (Hypothesis 2a). Additionally, the association between expectations and age differences in trait levels may be moderated by age, such that the association is negative before mid-adolescence (indicating that individuals tend to move against expectations because their abilities to meet them lags behind) and positive only thereafter (as youths' abilities catch up with the competencies associated with the expectations) (Hypothesis 2b).

Question 3: Do more-demanding traits fall furthest short of expectations, especially before self-regulation abilities are more fully in place (i.e., before mid-adolescence)? Mean-level age trends may be moderated by the degree to which traits require self-regulation, with traits requiring more self-regulation displaying larger discrepancies between socially expected levels and levels actually observed in children (Hypothesis 3a). Again, this effect may be further moderated by age such that the discrepancy is linked with a trait's self-regulation demands more strongly in younger children (Hypothesis 3b).

Question 4: Do traits with stronger social expectations tend to have especially large or small variances, and are these associations moderated by age or the traits' reliance on self-regulation skills? For example, before mid-adolescence variance may increase for traits with stronger expectations due to some individuals not yet able to sufficiently self-regulate their behaviour (Hypothesis 4a). After mid-adolescence, however, more children may have sufficient self-regulatory capacities to meet the social expectations for their traits, entailing decreases in trait variances as the slower self-regulation-developers are catching up (Hypothesis 4b). For traits with lower social expectations, most children may be able to meet the expectations, thus trait variance might remain similar irrespective of age.

We emphasize that these hypotheses are tentative and do not allow clear expectations for effect sizes, because the theories on the mechanisms of personality development are still in their infancy – the very issue our paper aims to contribute to mitigating. In other words, the relative imprecision of the hypotheses is reflective of where the field is rather than our paper, whereas we hope that the present attempt will stimulate clearer thinking about the general mechanisms of personality development.

Selecting a diverse sample of traits

Addressing these specific research questions assumes that personality develops along many trait dimensions that differ in terms of their social expectations and self-regulatory requirements. Often, personality development is studied along a few broad dimensions such as the Big Five or Little Six (Shiner & DeYoung, 2013; Soto & John, 2014; Soto & Tackett, 2015). However, adequately testing our hypotheses requires comparisons between a larger number of traits. There is a wealth of evidence that personality does in fact develop along numerous dimensions. For example, specific facets of the same Big Five domains, and even specific nuances of the same facets

(McCrae, 2015), often display disparate age trends (Jackson et al., 2009; Lucas & Donnellan, 2009; Mõttus et al., 2015; Mõttus, Sinick, et al., 2019; Mõttus & Rozgonjuk, 2019; Soto et al., 2011; Terracciano et al., 2005, 2006). The present study examines personality development at the level of nuances (McCrae, 2015): specific personality traits that can be operationalized by individual questionnaire items (Condon et al., 2020; Mõttus et al., 2020). Nuances are the lowest level of the personality trait hierarchy and show the valid specific variance and distinctive properties of traits such as cross-method agreement, rank-order stability, and heritability (Condon et al., 2020; Mõttus et al., 2014; Mõttus, Kandler, et al., 2017). We capitalized on the diverse item pool of the Common-Language California Child Q-Set (CCQ) (J. H. Block & Block, 1980), which was designed deliberately to cover a broad range of youth personality characteristics with low redundancy among the items.

Quantifying social expectations

People in different social roles likely hold different expectations regarding youths' behaviour. We operationalized the social expected levels of personality nuances by separately surveying parents, teachers and students. Beyond providing each nuance with a social expectation level, this also allowed us to test variations in the expectations across the different kinds of significant others of children. Expectations may be more likely to drive personality differences when they are consistent across different social partners and thereby apply unidirectional pressure on children's behaviour, compared with children facing contradictory expectations from different people. Expectations for youths' behaviour may also change over time. For example, expectations may become greater as children mature into adolescents. We therefore also measured perceived level of expectations separately for children (8 to 10 years) and late adolescents (16 to 18 years) to allow us to assess possible changes across these two developmental stages.

Quantifying requirements of self-regulatory skills

Different behavioural traits likely vary considerably in difficulty for children and youths to enact. To quantify such differences, we also surveyed parents, teachers and students about the perceived level of effort required to enact each trait. For each nuance-trait, participants were asked to rate how much self-control it would take for a child to behave in this way when they would otherwise be inclined to behave differently. Teachers and parents are in a good position to observe youths' behaviours, and college students can draw on their own recent first-hand experiences during adolescence. We anticipated that rating the level of self-regulation skills needed for enacting different traits would be a difficult task for any single rater and that individual ratings could therefore be fairly unreliable. But we deemed it likely that the aggregate ratings of many raters would prove more reliable, also indicated by similar average ratings across different perspectives (teachers, parents, and students).

Overview of the present research

In sum, the present study examined whether social expectations for youths' behaviour and the degree of self-regulation needed to meet these expectations could be among the mechanisms to explain normative personality differences in childhood and adolescence. Specifically, we hypothesized that trait levels would positively correlate with their social expectation levels (Hypothesis 1); that traits' mean-level age trends would mirror social expectations, with age moderating this association (Hypothesis 2); that self-regulation requirements may explain gaps between social expectation levels for children's traits and their actual trait levels, especially before mid-adolescence (Hypothesis 3); and that social expectations would also intersect with age differences in personality trait variance (Hypothesis 4).

Methods

The Common-Language California Child Q-Set (CCQ)

We measured children's personality traits, social expectations for the traits and self-regulatory requirements for behaving in expected ways using the CCQ. It includes 100 items that can be used by a non-professional observer to describe a child or adolescent (J. H. Block & Block, 1980; Caspi et al., 1992), with 94 of the items focusing on personality, and the remaining six representing physical characteristics and other non-personality attributes (Soto, 2016; Soto & John, 2014). Although the CCQ items can be aggregated to measure broad traits like the Big Five or Little Six (Soto & John, 2014), they were developed to be individually informative and non-redundant (J. H. Block & Block, 1980). Due to the item-level focus of the CCQ, as well as the need to test the present hypotheses across a large and diverse set of personality traits, we analysed each of the 94 personality-focused CCQ items as representing a distinct, nuance-level personality trait (McCrae, 2015; McCrae & Mõttus, 2019). Due to its content diversity and lesser focus on *the priori* structure of items, CCQ was a more suitable personality measure for our purposes than questionnaires developed to measure particular trait models such as the Big Five.

Due to our need to collect information about children's and adolescents' mean-level personality traits, as well as social expectations and self-regulatory requirements for the traits, the personality-focused CCQ items were rated in different ways by the different samples of participants. This happens to be a strength of the present design also because it reduces the chances of over-fitting: sample-specific idiosyncrasies driving associations between multiple variables (Yarkoni & Westfall, 2017).

1) Mean-level personality traits. Parents rated their child's personality traits on each CCQ item, using a 9-point Likert scale ranging from 1 (extremely uncharacteristic) to 9 (extremely characteristic).

2) **Social expectations for personality traits.** Parent, teacher and student participants were asked to rate each CCQ item in terms of whether they would generally approve or disapprove of children thinking, feeling or behaving in the way described in the item; subsequently, students were also asked to rate how much they thought teachers would approve it. In other words, participants were instructed to rate the extent to which they perceived social pressure on children to behave or not to behave in these ways. Participants made these ratings on a 5-point Likert scale ranging from 1 (strongly disapprove) to 5 (strongly approve).

3) **Perceived self-regulation demands of personality traits.** Participants were asked to rate each CCQ item in terms of how much self-control it would require for a child or adolescent to behave in the way described in the item, also using a 5-point Likert scale ranging from 1 (almost no self-control) to 5 (an extreme amount of self-control). To make the rating process less confusing for participants, the wording of some items was modified for this condition so that all items described behaviour in a socially desirable direction (based on the previously collected expectation ratings described above). For example, the socially undesirable item 'tries to take advantage of other people' was changed to 'tries to avoid taking advantage of other people'.

Participants

Youths' mean-level personality traits. Participants for measuring children's personality traits were drawn from an initial sample of parents of 16,000 children aged from 3 to 20 years, with each parent rating their child on the CCQ items (Soto & John, 2014). The target children included 500 boys and 500 girls for each age from 3 to 17, as well as 500 boys and 500 girls from the combined ages 18 to 20 years; this was to make each age group equal in size and gender balanced (Soto & John, 2014). In the present study of later childhood and adolescence, we focused on children between ages 8 and 18-20 which accounted for 11,000 participants in total.

Social expectations for youths' personality traits. Participants for measuring perceived level of expectations for children's personality traits included college students, as well as parents and teachers of children aged between 8 and 18 years. Parents and teachers were recruited from the online research platform Prolific and each compensated with £2.50, whereas students completed the study for course credit (see Table A1). As described above, participation entailed rating each personality-focused CCQ item in terms of its most socially approved level. Each participant was allocated into one of eight rating conditions, allowing us to measure expectations in various ways. The first dimension along which the conditions differed was developmental period: in four conditions, participants had to rate expectations for children between ages 8 and 18 years (without specifying a narrower age range), whereas in four conditions the ages were more circumscribed, allowing us to study age-differences in expectations (participants rated expectations for either 8- to 10-year-olds or 16- to 18-year-olds). The second dimension was rater perspective: parents and teachers were instructed to rate expectations from their own point of view, whereas students were asked to rate how they thought children's peers and, subsequently, children's teachers would expect them to behave.

Participants who did not complete the questionnaire or failed quality checks were excluded from expectation ratings of children aged between 8 and 18 years, resulting in a sample of 25 parents, 22 teachers and 60 students. We further recruited 27 parents and 30 teachers to participate in rating expectations for either 8- to 10-year-olds or 16- to 18-year-olds. Fifteen participants were excluded because of an incomplete questionnaire or low quality answers which result in 10 parents and 10 teachers who rated expectations of children aged between 8 and 10 years; and 12 parents and 10 teachers who rated expectations of teenagers aged between 16 and 18 years (see Table B1). It is important to bear in mind that raters always rated items, *not* children actually known to them.

Table B-1. Descriptive data for eight conditions used to obtain ratings of social expectations.

Raters	Parents	Teachers	Students (rating peer expectations)	Students (rating teacher expectations)	Parents	Teachers	Parents	Teachers
N	25	22	60	60	10	10	12	10
Rater age	26-52 (*M*=39.76)	23-71 (*M*=38.73)	8-18 (*M*=13)	8-18 (*M*=13)	27-46 (*M*=37.33)	23-66 (*M*=38.90)	33-52 (*M*=42.75)	22-57 (*M*=40.40)
Rater gender (*N* of male)	9	6	30	30	8	4	8	5
Rating targets	Children aged 8 to 18 years				Children aged 8 to 10 years		Teenagers aged 16 to 18 years	
ICC	.85	.95	.92	.95	.64	.75	.73	.81

Note. ICC = Intraclass correlation of (random) average rater.

Self-regulation demands of personality traits. Participants for measuring self-regulatory skill requirements included college students, as well as parents and teachers of children aged between 8 and 18 years. Parent and teacher participants were recruited from Prolific and each compensated £2.50, college students participated for course credit. Participation required rating each personality-relevant CCQ item in terms of how much self-regulation, or effort, would be required to behave, think, or feel in a socially desirable way. Participants were allocated into five rating conditions based on target age and rater perspective. Parents and teachers were asked to rate how much self-regulation each nuance level trait would require from either 8- to 10-year-olds or 16- to 18-year-olds (see Table B2).

Table B-2. Descriptive data for five conditions used to obtain ratings of required self-regulation.

Raters	Parents	Teachers	Parents	Teachers	College students
N	14	13	13	13	54
Rater age	27-45 (M=35.36)	23-55 (M=33.85)	34-51 (M=42.69)	24-52 (M=36.54)	18-22 (M=19.12)
Rater gender (*N* of male)	5	4	6	3	19
Rating targets	Children ages 8 to 10 years old		Teenagers ages 16 to 18 years old		Youths ages 8-18 years old
ICC	.66	.45	-.03	.88	.10

Note. *ICC* = Intraclass correlation of (random) average rater.

There were 59 participants in the initial sample of rating self-regulatory requirements; six participants were excluded due to either an incomplete questionnaire or low quality answers. Of the 27 remaining parents, 14 participants rated items' self-regulation requirements for children aged between 8 and 10 years and 13 participants rated items' self-regulation requirements for teenagers aged between 16 and 18 years. Of the 26 remaining teachers, half rated items' self-regulation requirements for children aged between 8 and 10 years and half rated them for teenagers aged between 16 and 18 years (see Table B2).

Additionally, 54 college students were instructed to rate each item in terms of the effort required for children aged between 8 and 18 years. Again, obtaining ratings from different conditions allowed us to test their robustness and, when aggregated, to obtain a more reliable and generalizable assessment. Again, it is important to bear in mind that ratings were about items, not actual children.

Data analysis

The data and R code are made publicly available at Online Supplemental Material (https://osf.io/pkda3/).

To investigate the consistency of ratings within rating conditions, we used a two-way average absolute intra-class correlations (*ICCs*), as implemented in the *psych* package version 1.9.12 (Revelle, 2019). To compare the ratings across groups (i.e., rating conditions), we used eta-squares representing the proportion of variance in ratings due to group differences; we also calculated mean item-ratings within groups and compared the profiles of the 94 personality-focused CCQ item means across groups using Pearson correlations.

To investigate associations of the means and variances of parent-rated traits of actual children with the social expectations and self-regulation requirements for these traits, we constructed multi-level models (or linear mixed-effects models) as implemented in the *lme4* package version 1.1.18 (Bates et al., 2014). In these models, either items' means or variances in the parent-ratings of actual children were the dependent variable, whereas age and perceived expectations and/or self-regulation demands for these items were independent variables. Specifically, we re-arranged the data into a table with 94 (items) x 11 (age groups, from 8 to 18-20) = 1,034 rows, with the parent-rated (pertaining to actual children) mean or standard deviation of each item in the dependent variable columns, and the mean social expectations and self-regulation demands of the items in the independent variable columns, alongside age and squared age (for quadratic effects) for the particular means/variances. Parent-rated personality traits and expectations were reverse coded, so that all items were keyed in the socially desirable directions (i.e., with mean expectations at or above the scale mid-point). Specifying random intercepts and slopes for items (11 individual observations for each item, one for each age group being "nested" within the item) allowed us to test the main effects of age (and its square), expectations and self-regulation demands, as well as interactions among them, on items' means and variances while controlling for dependencies in the data (i.e., same items at different ages) and allowing items to vary in age trajectories (random slopes). Items' means, standard deviations, social expectation and self-regulation demand ratings were grand mean centred (across the 94 items and 11 ages), and age was centred at 13 (i.e., the median age) prior to computing squared age. We report fixed effects from these models, which summarize general associations across all items, while accounting for their unique deviations from the fixed effects.

Results

Did raters agree about social expectations for children's personality traits?

The *ICCs* of the ratings for social expectations for 8- to 18-year-olds from parents, teachers, peers and students ranged from .85 to .95 (Table B1). On average, the rating condition explained 6.1% of variance in the ratings, with group differences significant for 18 items after applying false discovery rate (FDR) correction for multiple testing. When the responses to each item were aggregated within the rating conditions, the four profiles of 94 average ratings correlated between .64 and .87. Also, collapsing ratings from all conditions resulted in an excellent average *ICC* of .98. Collectively these results indicate that parents, teachers and peers tended to hold fairly similarly configured expectations for youth personality traits, although there were mean-level differences for several nuances.

Did expectations for children vary with age?

Turning to parents and teachers, we also investigated the consistency of expectations as a function of whether they were rated with children (8- to 10-year-olds) or teens (16- to 18-year-olds) in mind. After correcting *p*-values for FDR, no item significantly varied in mean between the two target ages, either with or without collapsing parent-ratings and teacher-ratings for respective target ages; the average eta-squares were .06 and .02, respectively. Moreover, when average ratings were calculated for both children and adolescents (collapsing ratings from teachers and parents for each age group), the two profiles of 94 means correlated .90. These findings suggest that people generally have consistent social expectations regarding youth personality traits, not only regardless of who describes the expectations, but also regardless of whether they are asked to think of children or older adolescents. In all subsequent analyses, we therefore combined all social expectation ratings across all conditions.

Did children's personality traits generally comply with expectations?

When the means of the 94 items (from parents' ratings of their actual children) were predicted from the social expectations for these items (combined across all rating conditions) in a multi-level model, traits with higher expectations tended to have higher parent-rated means (standardized estimate *b* = .30; Table B3). Thus, supporting Hypothesis 1, the pattern of mean-level personality traits tended to comply with the pattern of social expectations.

Table B-3. Estimates of the model to test the effects social expectations and age on mean-level personality traits and their change.

	Linear model			**Quadratic model**				
	Soc	Age	Age x Soc	Soc	Age	Age^2	Age x Soc	Age^2 x Soc
Mean	.303	-.027	.000	.284	-.027	.004	.000	.002
SE	.093	.010	.010	.091	.010	.001	.010	.001
p	.002**	.007**	.976	.002**	.007**	<.001***	.976	.021*

Note. Age^2 = age squared; Soc = social expectations. * $p < .05$. ** $p < .01$. *** $p < .001$.

Did children's personality traits linearly change in accordance with expectations?

In the same model, age had a negative linear fixed effect on items' means, whereas the interaction between age and social expectations was not statistically significant (Table B3). Thus, inconsistent with Hypothesis 2a, there was no evidence that the linear trends in personality traits' developmental trajectories positively tracked the social expectations for these traits – although socially more approved traits had generally higher levels, they were no likelier to trend higher still (or no less likely to trend lower still, given the general downward trend in traits) than less approved traits.

But was the association between expectations and mean-level age trends non-linear?

When we added the quadratic term (age-squared) to the model as an additional main effect, age continued to have a negative linear fixed effect on item means, whereas age squared had a *positive* linear fixed effect on item means. When we also added the interaction between age-squared and social expectations to the model, the interaction between age and expectations remained non-significant, but the interaction between age-squared and expectations was significant (Table B3). Age trends for traits with high, medium, and low expectations (each group had a third of items) are depicted in Figure B1. It shows that: (a) items with all levels of expectations showed general declines with age, but (b) items with high expectation levels showed an uptick in later adolescence, and (c) items with low and medium expectation levels were fairly similar in their overall means, whereas most of the effects pertained to items with highest expectation levels. Taken together, these results partially support Hypothesis 2b by indicating that traits with the highest social expectations showed negative age trends from childhood through mid-adolescence, but positive trends in late adolescence.

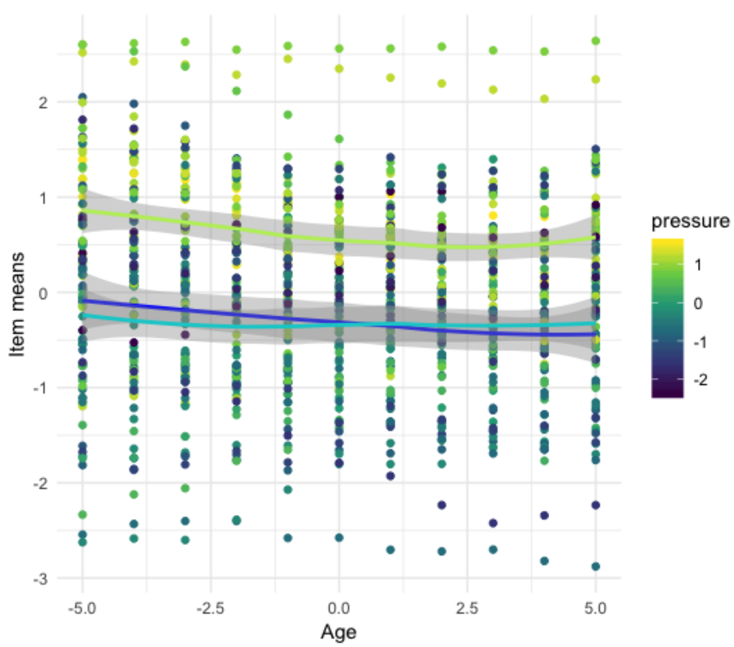


Figure B-1. The effect of different levels of social expectations and age on mean-level personality traits. The items were divided into three groups based on the high, medium and low expectations for them (lowest, middle and highest third of the items). Age was centred at 13. The y-axis represents standardized item mean scores.

Did observers agree about the self-regulation required by children's personality traits?

We estimated the consistency of self-regulation ratings within and across two age groups of targets and in a combined age group: children (8- to 10-year-olds), teens (16- to 18-year-olds) and youth generally (8- to 18-year-olds). Average rater *ICCs* ranged from -.03 for smaller rater groups rating children of more specific ages to .91 when all ratings were combined (Tables B2 and B4). When the ratings of each item were aggregated within groups, the profiles of 94 average ratings correlated from .67 to .77. These findings indicate that, when aggregated to the group level, people with different perspectives (parents, teachers, and students) hold relatively similar views of how much skills it takes children to achieve social expectations. In all subsequent analyses, we therefore averaged all self-regulation ratings across conditions. We emphasize, however, that the agreement between individual raters was often very low and it took tens of raters to achieve acceptable reliability for the averages of ratings – judging how much self-regulation it takes to display certain personality traits is hard for any one individual, but the crowd wisdom of tens of people can be sufficiently reliable.

Table B-4**.** Ratings of how much self-regulation is required: intra-class correlations when parent and teacher ratings were collapsed for children and teens, respectively, intra-class correlations for student raters and combined groups and correlations among the mean ratings across rating conditions.

ICC child	ICC teen	ICC student	ICC total	Child-Teen	Child-Student	Teen-Student
.52	.67	.88	.91	r = .67	r = .77	r = .71

Note. *ICC* = Intraclass correlation of (random) average rater.

Do more-demanding traits fall farthest short of expectations, especially before self-regulation abilities are more fully developed?

As in the analyses of social expectations, multi-level models were fitted to test whether traits' perceived self-regulation demands could help explain age differences in the gap between social expectations for traits and their mean-levels. Mean-level item scores were the outcome variable, and age, age squared, rated self-regulation demands of the items, and the interaction between them were the predictors. Because we had concluded that expectation levels for personality traits did not vary with age and expectations could be combined across all rating conditions, items' means were perfect linear transformations of their discrepancies from the expectations for them, so the means could be the outcome instead of the discrepancies. Essentially, we were still modelling whether the age trajectories in their discrepancies were moderated by the self-regulation demands for the items, as per Hypotheses 3a and 3b. (Had expectancies varied with age, we would have operationalized discrepancies as differences between expectancies for the *older* age group and mean trait levels of actual children).

Self-regulation demands did not account for mean trait levels and the interaction between age and self-regulation was also not significant (Table B5). This suggests that age-differences in the discrepancies between expected and actual trait levels of children were not moderated by the self-regulation demands of the traits. However, when we also added the interaction between age-squared and self-regulation demands, the interaction between age squared and self-regulation was significant. The results of this final model are depicted in Figure B2. Failing to support Hypothesis 3a, these results indicate that traits' self-regulation demands did not have a main effect on the generally negative age trajectories of these traits (that is, on discrepancies from expectations). However, Figure B2 also shows that traits with highest self-regulation demands—but not those with medium or low demands—showed a modest mean-level increase during late adolescence. Partially supporting Hypothesis 3b, these results suggest that the traits with the highest self-regulation demands were somewhat more likely to bounce back from the negative overall age trends.

Table B-5. Estimates of the model to test the effects self-regulation and age on mean-level personality traits and their change.

	Linear model			Quadratic model				
	SR	Age	Age x SR	SR	Age	Age^2	Age x SR	Age^2 x SR
Mean	.147	-.027	.001	.130	-.027	.004	.001	.002
SE	.097	.010	.010	.095	.010	.001	.010	.001
p	.134	.007**	.886	. 174	.007**	<.000***	.886	.038*

Note. Age ^{2} = age squared; SR = self-regulation. * $p < .05$. ** $p < .01$. *** $p < .001$.

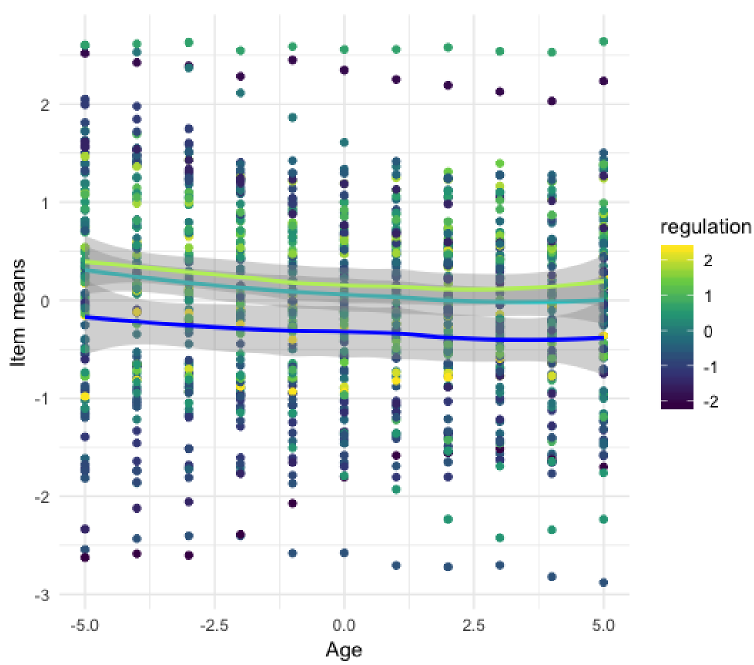


Figure B-2. The effect of different levels of self-regulation demands and age on mean-level personality traits.

Did traits with stronger expectations tend to have especially large or small increases in variance, and were these associations moderated by age or the traits' reliance on self-regulation skills?

To test whether traits with different social expectations and self-regulatory requirements showed different age trends in trait variance, we predicted traits' standard deviations from their social expectations, self-regulation and age in multi-level models; these were identical to the models described above, except that items' means were replaced with their standard deviations. For the association between traits' variance and social expectations, results showed that age had a positive linear association with personality variance (Mõttus, Soto, et al., 2017), the main effect of expectations was not significant, and the interaction between age and social expectations was significant (Table B6). When we added age-squared to the model, it had a significant negative effect on trait variance, consistent with the plateauing of the increase in trait variance (Mõttus, Soto, et al., 2017). When we also added the interaction between age-squared and expectations to the model, the interaction between age and expectation as well as the interaction between age-squared and expectations were both significant.

Table B-6. Estimates of the model to test the effects social expectations and age on personality variance and their change.

	Linear model			**Quadratic model**				
	Soc	Age	Age x Soc	Soc	Age	Age^2	Age x Soc	Age^2 x Soc
Mean	-.154	.076	.021	-.119	.076	-.009	.021	-.003
SE	.093	.008	.008	.092	.008	.001	.008	.001

p	.104	<.001***	.015*	.199	<.000***	<.000***	.015*	.005**

Note. Age ^{2} = age squared; Soc = social expectations. * $p < .05$. ** $p < .01$. *** $p < .001$.

Figure B3 depicts the results of the final model. Partially supporting Hypothesis 4a, these results indicate that, from childhood through mid-adolescence traits with high or medium expectations showed larger increases in variance than those with the lowest expectations.

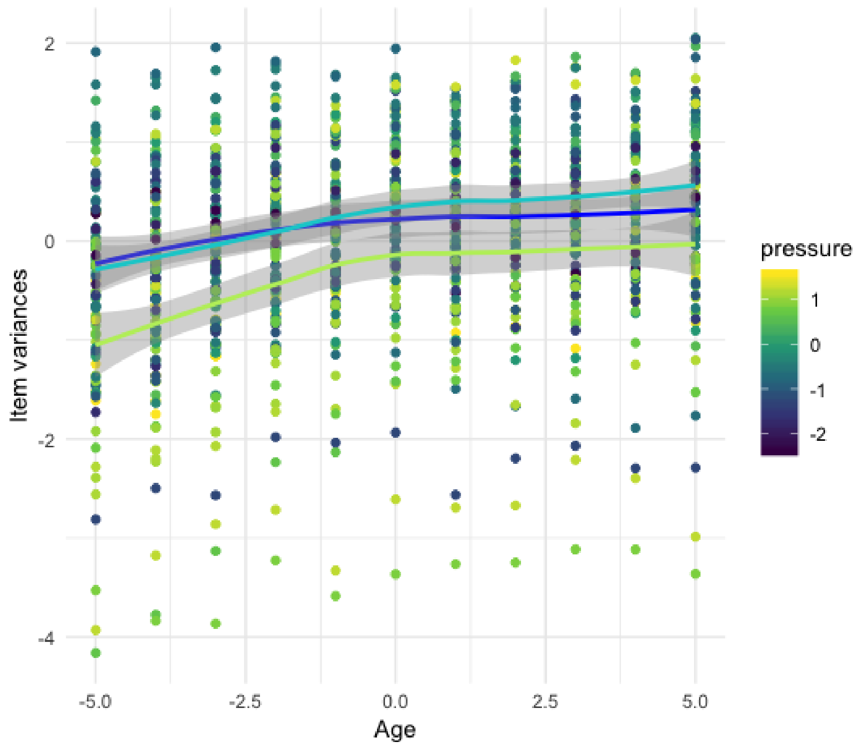


Figure B-3. The effect of different levels of social expectations and age on personality variances.

When we repeated this analysis to predict trait variance from self-regulation demands and age, self-regulation demands did not have any impact on personality variance, nor was the interaction between age and self-regulation demands significant (Table B7). When we added the quadratic age trajectory to the model and its interaction with self-regulation demands, the coefficients for age and age-squared were significant, but those for self-regulation and the interactions between age / age-squared and self-regulation were not significant. Contrary to Hypotheses 4a and 4b, these results do not provide evidence that self-regulation requirements could help explain age differences in trait variance during childhood and adolescence.

Table B-7**.** Estimates of the model to test the effects self-regulation and age on personality variance and their change.

	Linear model			**Quadratic model**				
	SR	Age	Age x SR	SR	Age	Age^2	Age x SR	Age^2 x SR
Mean	-.023	.076	-.009	-.046	.076	-.009	-.009	.002
SE	.095	.009	.009	.093	.009	.001	.009	.001
p	.806	<.000***	.301	.620	<.000***	<.000***	.301	.066

Note. Age^2 = age squared; SR = self-regulation. * $p < .05$. ** $p < .01$. *** $p < .001$.

Discussion

We addressed four questions about social expectations and self-regulation skills as possible mechanisms of personality differences during childhood and adolescence. Inspired by the influential study by Denissen and colleagues (2013), we expected nuance-level personality traits and normative changes in them to track social expectations for these traits, and these developmental shifts in the traits to be moderated by the extent to which the traits require self-regulation. Combining recent research on developmental trends in the magnitudes of individual differences (Mõttus, Briley, et al., 2019; Mõttus, Soto, et al., 2017) with the ideas of Denissen and colleagues, we also examined whether social expectations for the traits and/or their self-regulation needs would moderate age-trajectories in trait variance.

We found that: a) social expectations were linked with traits' overall levels but not with change trajectories in them apart from a slight uptick characterizing the most socially desirable traits in late adolescence, and b) traits' self-regulation needs were generally not linked with either mean trait levels or changes in them, apart from a slight uptick characterizing traits with the strongest self-regulation demands in late adolescence. We also found that traits with the strongest social expectations showed the most pronounced curvilinear increases in variance (i.e., increases in the magnitudes of individual differences prior to mid-adolescence). Therefore, findings of Denissen and colleague only partially replicated in the present study, whereas we also refined past findings regarding how children become increasingly less alike as they grow: it is traits with strongest social expectations along which high-scorers pull quicker away from low-scorers.

The role of social expectations

Our first question was whether children's personality traits generally comply with social expectations. Although the definition of social expectations has varied across research (Morgan, 2007; Roese & Sherman, 2007), they have previously been considered in terms of how they related to children's relationships and life outcomes; and how they correlate with different aspects of life in previous studies (Raag & Rackliff, 1998). For example, suitable social expectations have been suggested to contribute to children's wellbeing, whereas inappropriate social expectations may be detrimental to development (Zhou et al., 2007).

Here, we operationalized expectations thoroughly, measuring them for nearly a hundred traits and exploring their consistency across several perspectives (important others to children) and two age groups. According to our results, important others such as parents, teachers and students hold generally similar social expectations for children and adolescents. This is important to our first research question because the expectations may be more likely to systematically influence personality development if they are consistent across sources. This finding was not a given. It could have turned out that different groups of important others for children – for example, parents versus peers – hold systematically different expectations for a range of traits such as pushing boundaries, stretching rules, or enjoying physical affection.

We also found that the expectations did not vary systematically with the target's age: teachers and parents had similar personality-expectations for 8- to 10-year-olds as they did for 16- to 18-year-olds. This was surprising because it seemed plausible that adolescents would be held to higher standards than younger children. Moreover, the theory of personality development in childhood and adolescents on which we based our hypotheses (Denissen et al., 2013) postulates that personality development is driven by shifting social expectations. Our finding could indicate that people a) have already developed relatively adult-like expectations for children about 10 years old, b) hold even late adolescents to relatively low standards, c) were unable to rate the differences in their expectations across the age group (note that this would not necessarily be a methodological problem *per se* as people could also fail to perceive and apply age-graded expectations in real life), or d) internally shifted their standards of what is socially expected, but did not apply them to the ratings.

As one possibility, if the raters were unable to perceive or rate the differences in social expectations for the traits described in the personality test items, then the ratings could simply reflect the social desirability of these items. This would be partly consistent with our finding that items' average social expectations were correlated with their mean levels (as per Hypothesis 1), because items' social desirability is known to highly correlate with their mean levels (Edwards, 1953). However, in our findings the correlation between traits' mean levels and social expectations (r = .30) was much lower than has been found elsewhere (e.g., r ~ .90) – thus the social expectations ratings may not have merely reflected social desirability, after all (Konstabel et al., 2006).

Either way, a question remains: can people perceive and *apply* age-graded expectations in real life, when interacting with either younger children or late adolescents? As a possibility, even if there are some shifts in expectations with age, by and large, some traits are desirable and some not regardless of age (or the person who holds the expectations) and therefore the shifts may play a smaller role in personality development than expected, at least in the period that we considered. General expectations for how people should be may be in place and well known to children far earlier than age 8 years.

Our second key research question was whether children's personality traits change in accordance with expectations. Partially supporting our hypotheses (Hypothesis 2), we found that the linear trends in the mean-level developmental trajectories of personality traits did not track the social expectations for these traits, but there were associations between non-linear age trends and social expectations: normative personality trends from middle childhood into adolescence were generally negative, but the traits with the highest expectations showed an uptick in late adolescence. Thus, our results lend modest support to the hypothesis that personality development in childhood and adolescence is driven by social expectations.

Insofar as the mechanisms of personality development might generally be expected to be similar in childhood and adulthood, our results are (at least) not inconsistent with the broader Social Investment Theory (SIT) of personality development. This theory posits that mean-level changes are driven by people becoming committed to new social roles and the expectations that come with these roles (Bleidorn et al., 2013; Roberts, Wood, et al., 2005). Essentially, then, this theory also expects average personality traits to converge towards widely held social expectations. However, the existent pattern of empirical findings generally lends only modest support to the SIT (Bleidorn, Hopwood, Back, et al., 2020b; Denissen et al., 2019; McCrae et al., 2021) and it is in this sense that our findings do fall in line with broader literature.

The role of self-regulation skills

Our third key question (Hypothesis 3) was whether more demanding traits (i.e., those that appear to require adolescents to exert a high degree of self-control) fall farthest short of expectations, especially before self-regulation skills are more fully in place. This hypothesis was also inspired by work from Denissen and colleagues (2013) who proposed that social expectations influence personality development in conjunction with youths' ability to meet these expectations, and this requires self-regulation skills that may develop more slowly than expected. This could explain the dip in mean trait levels towards socially undesirable directions around mid-adolescence (Soto, 2016). However, we found that traits' self-regulation demands did not have a strong effect on the generally negative age trajectories of these traits: mean-level age trends were not moderated by the degree to which traits require self-regulation skills in that traits appearing to require more of them did not display larger discrepancies between socially expected levels and mean levels. However, traits with high self-regulation skills demands did show a modest mean-level increase during late adolescence, whereas traits with medium or low demands did not show the same trend. Thus, only high demanding traits fall short of expectations before mid-adolescence, which lends modest support to Hypothesis 3.

It is possible that our operationalization of the degrees to which traits require self-regulation was inappropriate: we asked panels of parents, teachers and college students (recent adolescents) to rate this, but they may not have had a very good sense of the levels of demand for youth to behave, think or feel in particular ways. Indeed, the inter-rater agreement between individual raters was sometimes very low. However, what would have been the alternative for quantifying self-regulation demands? Conceivably, we could have measured self-regulation with some experimental tasks and correlate the resultant scores with the 94 CCQ items, resulting in a score for each item in terms of how much its endorsement is more probable among youth with greater self-regulation skills. But experimental tasks are known to correlate very poorly with subjectively rated personality traits and even among themselves (Mazza et al., 2020), and experimental procedures for capturing items' self-regulation requirements could have suffered from low ecological and face validity – which were bound to be higher for subjective ratings based on human real-world experience of children and their behaviour. More importantly, the average-rater reliability of the ratings was excellent, and the average ratings from different rating conditions agreed well, suggesting that the average ratings were generally reliable and contained a substantially converging signal.

To improve the operationalisation of self-regulation, it would be helpful if youths themselves could rate the items from the self-regulation needs perspective, so that we could compare whether what youths find hard to do agrees with what adults think is hard for youths to do. The reason we did not ask youths themselves to rate the self-regulation questionnaire is that it would be quite stressful for youths, especially those age under ten years old, to answer or even understand a questionnaire with 94 relatively complex statements. As self-reported and informant-reported data reflect the same psychometric constructs (Mõttus, Allik, et al., 2019; Olino & Klein, 2015), we hoped the panel of parents, teachers and college students who are either close to children or who just past the teenager stage would provide sufficiently valid ratings.

If not self-regulation, what else could explain the dip in youth personality traits around mid-adolescence? According to Moffitt and colleagues (1996), anti-social behaviours and conduct problems are prevalent from late childhood to mid-adolescence but often stop in the early 20s. In their view, this may be an adaptive response to teens' social context: physically and cognitively mature youth increasingly desire adult privileges and see socially disapproved conduct as a way to have autonomy from parental control (Moffitt et al., 1996; Moffitt & Caspi, 2001). Possibly, our results reflect a tendency for adolescents to rebel against parental control and expectations; and this may require as much self-regulation, on average, as conforming to the expectations.

Social expectations and personality variance

Our fourth and final research question was whether youth's personality traits with stronger expectations tend to have especially large or small variance, and whether these associations are moderated by age or the traits' reliance on self-regulation skills. Partially supporting Hypothesis 4, we found that traits with high or medium expectations showed larger increases in variance than those with lowest expectations from childhood through mid-adolescence, but self-regulation demands did not have a significant effect on personality variance and age trend. On the other hand, traits with the strongest expectations increased the most in variance prior to mid-adolescence, somewhat catching up the traits with lower expectations in their variance post mid-adolescence, resulting in a non-significant main effect of expectations on trait variance. Again, we can turn to literature on adolescent antisocial behaviour: it tends to emerge around mid-adolescence when youth starts to experience a gap between their biological maturity and lack of access to adult privileges (Moffitt et al., 1996). This gap may make youth admire aggressive peers and mimic their conduct as a way to show autonomy from parents, as well as to gain affiliation with delinquent peers. Although most adolescents do not engage in antisocial behaviour and many likely to not even approve it in others, the processes underlying this tendency may increase individual difference (Moffitt & Caspi, 2001) because it could be expected that the mid-adolescent rebellion trends become particularly manifest in traits with stronger expectations: those who rebel against demands can do so most effectively for traits for which the demands are clear and strong and therefore become ever more different from their less rebellious peers exactly on these traits. Traits with low social expectations cannot reveal individual differences in rebelliousness to the same extents as high-expectation traits. If it is rebelliousness that at least in part drives trait variance, it may become less powerful a mechanism of trait development as youth mature (for a parallel, antisocial tendencies also wane by early adulthood in most of the adolescents having displayed them). In summary, the processes underlying the occurrence of the mid-adolescent antisocial trends may also increase individual difference and potentially explains why traits with high or medium social expectations showed larger increases in variance than those with the lowest expectations before mid-adolescence but not late-adolescence.

Strengths and Limitations

The present research had a number of strengths, including its use of a large parent-report sample and a personality measure—the CCQ—that was well suited to testing differences in youth personality development across an array of highly specific, nuance-level traits. Moreover, we capitalized on an underused research design by examining systematic variability among many traits in various focal properties (we studied variance between traits, not between people; for discussion, see (Mõttus et al., 2020; Mõttus & Rozgonjuk, 2019), and thoroughly measured social expectations and self-regulation skill demands of the traits, doing this in a variety of ways and testing consistency in the results.

However, a number of limitations should also be considered. First, we used a cross-sectional rather than longitudinal research design. Although the resulting data allowed us to infer changes in trait mean levels and variances across childhood and adolescence, they did not allow us to directly observe such changes. Therefore, additional research is needed to test whether the present findings extend to longitudinal data. Second, the samples used to estimate social expectations may not have been large or diverse enough to represent the general population. Moreover, the average inter-rater agreement for parents of children ages 16 to 18 years and college students were relatively low. Future research is therefore needed to further examine self-regulation in these groups. Third, the present study investigated youths aged from 8 to 18. Some of our findings showed a slight uptick in late adolescence for traits with a) the highest expectations and b) requiring the most self-regulation. These age trends may not end at age 18 and further research is needed to test whether they extend into adulthood. A fourth limitation concerns the measurement of self-regulation. We asked a panel of parents, teachers and college students to rate the degrees to which different personality traits require self-regulation. Although there was generally good agreement across raters and conditions (with the specific exceptions noted above), the perceptions of self-regulatory ability from parents, students and teachers may inaccurately reflect the actual self-regulatory demands that the traits pose to youth.

Our analyses were based on single items, which are likely to have lower reliabilities than aggregate scale scores and may have lead us to underestimate any true effects. However, although this information is not available for the CCQ, recent findings have shown that single personality test items tend to have higher reliabilities than often assumed (well over .60; e.g., (Henry & Mõttus, 2020; Mõttus, Sinick, et al., 2019)) and analyses based on comparing the properties of single items can yield meaningful results (Henry & Mõttus, 2020; Mõttus et al., 2020). Because CCQ items were designed to cover the space of personality traits more broadly than the currently dominant tests tailored to models such as the Big Five, our findings should be generalizable to a larger universe of personality traits than those reflected in the CCQ items.

Conclusions

Is personality development in childhood and adolescence driven by social expectations and youths' self-regulatory capacities? Based on the present findings, using measured rather than assumed values, we conclude that both the social expectations of significant others and children's own self-regulatory skills to meet these expectations may exert only partial and non-linear influence on youth personality development.

Chapter 4: Mechanisms of Personality Differences in Adulthood

Synopsis:

Expanding on the findings from chapter 3, this chapter examines whether social expectations are also a mechanism to explain personality differences during young adulthood. In addition, this chapter provides the first empirical evidence to support the social investment theory (SIT) in relation to personality differences based on the maturity principles. Results are partially consistent with SIT, suggesting that social expectations could be a potential mechanism to explain personality differences in young adulthood.

Dissemination status:

This chapter is in preparation for submission to the *Journal of Personality*. A pre-print is available on *PsyArxiv*: 10.31234/osf.io/p35r2

Are Social Expectations a Possible Mechanism for Adult Personality Change: Empirical Evidence for Social Investment Theory

Yuzhan Hang[a, *], Lydia Gabriela Speyer[a], Billy Lee[a], Michelle Luciano[a], René Mõttus[a]

[a]Department of Psychology, University of Edinburgh, Edinburgh, United Kingdom

Abstract

Objective: Personality traits change in both mean levels and variance across the life span but the mechanisms underlying these developmental trends remain unclear. Social investment theory is one of the leading theories trying to explain personality maturation, suggesting that social role expectations drive personality changes in adulthood. However, there is not sufficient empirical evidence for SIT, thus the present study tested whether social expectations can explain personality maturation in young adulthood. Methods: A pool of 257 personality items, which was not driven by a priori conceptual model of personality, was used to measure young adults' personality trait mean levels, personality variance and socially expected level of traits from different perspectives (friends, partners and bosses/supervisors). Results: There were consistent expectations for how young adults should think, feel and behave. Personality traits under stronger social expectations have higher mean levels and lower variances than traits under lower expectations; traits with stronger social expectations show higher means at later ages; and traits' variances increase with age. Conclusion: Our results are partially consistent with SIT, suggesting that social expectations could be a potential mechanism to explain personality changes in young adulthood.

Keywords: personality development; mean-level; variance; social expectations; social investment theory

Introduction

Personality traits change across the life span in both mean levels and variances (Mõttus et al., 2016; Mõttus, Soto, et al., 2017; Roberts & DelVecchio, 2000; Soto, 2016). Mean-level personality change, also referred to as normative personality change, is indicative of changes in average personality trait levels in a population over time (Mõttus et al., 2016; Roberts, Walton, et al., 2005; Soto, 2016; Soto et al., 2011). Specifically, mean-level changes reflect the level of changes of a personality trait among a group of same aged individuals over time (Specht et al., 2014). On average, people's personalities change in a socially desirable direction (Allemand et al., 2008; Caspi et al., 2005; Donnellan & Lucas, 2008), characterized by becoming more agreeable, conscientious, and emotionally stable across adulthood (Bleidorn et al., 2018; Roberts et al., 2006). This is referred to as the maturity principle of personality (Bleidorn & Schwaba, 2017). People also tend to become less extraverted and open to experiences with age, but these changes are not reflected in theories of the maturity principle of personality development (Wortman et al., 2012).

Although population-level patterns of personality development are most commonly explored through mean-level changes, age differences in personality trait variance offer another way to describe developmental trends (Mõttus, Briley, et al., 2019). The magnitude of individual differences in personality traits tends to increase from childhood until mid-adolescence and plateau thereafter (Kandler & Zapko-Willmes, 2017; Mõttus, 2006; Mõttus, Soto, et al., 2017). Such increases in trait variance could be caused by gene-environment transactions whereby self-selected environments magnify pre-existing individual differences (Mõttus, Briley, et al., 2019), by gene-environment interactions whereby environments moderate the expression of genetic dispositions (Kandler & Zapko-Willmes, 2017), or both. For example, extraverted people may tend to actively participate in social events and practise their social skills, which thus allows them to appear even more extraverted in those situations. In comparison, introverted people may tend to avoid social events and consequently have less opportunities to practise their social skills, thus becoming more socially insecure and, as a result, are even less likely to take part in social events in the future. A similar idea of how trait variance come about has been used to explain the Flynn effect in IQ a long time ago (Dickens & Flynn, 2001; Rowe & Rodgers, 2002).

Explaining the patterns

Besides the Five-Factor Theory, which ascribes personality development solely to intrinsic, genetically driven processes (McCrae & Costa, 2008), Social Investment Theory (Roberts, Wood, et al., 2005) is one of the leading theories to explain personality changes in terms of both mean levels and individual differences. SIT proposes social expectations as a potential explanation for the maturity principle of personality trait development: the traits mature as individuals are confronted with new social expectations, commit to new social roles and alter their behaviour accordingly, especially to accommodate major life transitions in early adulthood such as entering the work force and getting married (Denissen et al., 2013; Roberts, Wood, et al., 2005). According to SIT, there are some universal age-graded life tasks that apply to most people around similar age ranges. These social roles are associated with expectations for the individual to act in a more mature way which thus lead to mean-level changes in personality (Bleidorn, 2015; Hennecke et al., 2014). For example, if important others of a young individual expect them to be a hard-working student and a reliable and loving partner, and to become a responsible parent, it is plausible that the individual will follow these expectations and adjust their behaviours accordingly, changing in relevant personality traits as a result.

Although individuals encounter unique challenges in these transitions, which can help to sustain or even accentuate individual differences, the majority of people within different cultures may experience the similar life transitions around roughly the same time and change towards a similar pattern of expected traits (Roberts, Wood, et al., 2005). People are rewarded if they behave in socially expected ways and punished if they do not behave in the socially expected ways (Roberts, Wood, et al., 2005). If so, they should tend to become more alike over their developmental trajectories as they strive to receive rewards and avoid punishments, especially when faced with high social expectations. That is, generally shared social expectations should act to decrease the magnitudes of individual differences, counteracting or even reversing the possible increases in variance due to person-environment transactions (Mõttus, Briley, et al., 2019). So, there may be forces that tend to increase the magnitude of individual differences as well as forces that tend to decrease it.

Differential patterns across traits

But not all personality traits are alike as the behaviours associated with some traits are subject to stronger social pressures than others. This provides us with a possibility to test the hypotheses described above. Specifically, we should expect that it is especially the traits under relatively stronger social pressure (high expectations) that a) change in mean levels commensurately with the expectations and b) decrease in variance the most, or at least increase the least in variance if there is a general trend for increasing variance. To test this, we can describe developmental trends in personality traits using a diverse sample in traits that would vary in social expectations for them as well as how their means and variance change with age.

Development in the personality traits hierarchy

Most existing research on personality development has been based on the Big Five domains (Mõttus et al., 2020), but these are not the only ways to represent personality traits. Personality traits form a hierarchy and the Big Five domains can be separated into facets and nuances. Facets are groups of items that are more closely correlated with each other within the Big Five domains. Nuances are the lowest level of the personality hierarchy and can often be indexed by individual items or groups of very similar items (Condon et al., 2020). Lower levels of the personality hierarchy (facets and nuances) usually contain unique personality variance above and beyond domains (Jang et al., 1998). For example, facets show varied developmental age trajectories compared to their domains (Mõttus et al., 2015; Mõttus & Rozgonjuk, 2019; Soto et al., 2011; Terracciano et al., 2006), and nuances also show age trends different from their facets (Hang, Soto, Speyer, et al., 2021; McCrae, 2015; Mõttus, Soto, et al., 2017).

Besides potentially more refined descriptions of the population-level patterns of personality development (Hang, Soto, Speyer, et al., 2021), nuance-level analyses allow us to study systematic variations between personality traits in their developmental trends and intersection with other trait-level features such as social expectations (Mõttus et al., 2020; Mõttus & Rozgonjuk, 2019). For example, (Hang, Soto, Lee, et al., 2021) capitalised on variations between traits to investigate whether social expectations and ability to catch up with these expectations (self-regulation) explain the personality changes during childhood and adolescence. Specifically, they quantified the mean-level changes for 94 nuance-level traits and the degrees to which these traits reflect social expectations, as well as how much self-regulatory ability was required to meet the social expectations. Across the traits, levels of social expectations were significantly associated with their mean levels and individual differences were stronger for traits with strong social expectations (Hang, Soto, Lee, et al., 2021).

In the current study we used the same approach as (Hang, Soto, Lee, et al., 2021) to leverage the observation that personality traits vary with age along many dimensions, examining personality development at the level of nuances. To operationalize nuances, we used a diverse item pool of

257 items selected to measure personality as comprehensively as possible by capturing unique item-level information. Moreover, to quantify socially expected level of these personality nuances, we asked young adults to rate the items in terms of social expectations for them, from the points of views of important others for the young people (friends, intimate partners and bosses/supervisors). Since people in different social roles might have different expectations for others' behaviours, measuring young adults' perceived level of social expectations from different points of view allowed us to gain more reliable variations in the social expectations across different social relations. Arguably, the degree of social expectations only matters if individuals are affected by them, thus, measuring social expectations as perceived by their targets (young people themselves) is more suitable than measuring them in those assumed to hold the expectations (the important others of the young people). Social expectations may change most prominently during early adulthood since a majority of new social roles are established during this developmental stage (Bleidorn et al., 2013). We therefore explored the associations of social expectations with age differences in personality trait mean levels and variance for young adults (18 to 30 years).

Summary of the present research

In sum, the present study examines whether social expectations could be a potential mechanism to explain mean-level personality differences and variance during young adulthood. We hypothesise that (a) traits with higher social expectations tend to have higher mean-level scores; (b) traits under stronger social expectations will tend to vary more with age, in the direction of expectations, than those under lower expectations; (c) traits with stronger social expectations will have lower variance than traits with weaker social expectations and the variance would decline even more with age.

Method

The 257-items personality questionnaire

We used a pool of 257 items to measure young adults' personality traits and social expectations for these traits. This personality item pool was not driven by any priori conceptual model of personality, thus all items were developed to be individually informative (Henry et al., in preparation). In order to achieve this goal, all items were carefully selected in several rounds by the authors of the study and their colleagues. First, the majority of items of this personality item pool was assembled from the International Personality Item Pool (IPIP) (W. A. Goldberg, 1999) and the Synthetic Aperture Personality Assessment (SAPA) (Condon, 2018). Second, standard deviations of items were calculated, and from among highly correlated items only items with relatively higher standard deviations (calculated using the Eugene Springfield Community Sample) were selected as these items likely capture more information at the population-level (L. R. Goldberg et al., 2006). Third, items which were comparatively linguistically complex were removed (e.g. items contained more ambiguity in interpretations and items used infrequent words). Finally, items that were not covered by the Big Five or HEXACO domains measuring competitiveness, humour, sexuality, attractiveness, gratitude, and the Dark Triad were included. These items were iteratively tested for re-test reliability and items with comparatively lower re-test reliability were removed (generally those with 2-week retest correlations below .60) and other items that higher reliability were included in their place.

One reason that we did not use one of the traditional personality inventories is that there are no theoretical models to justify which items should be used in nuance-level personality research in these methods (Mõttus & Rozgonjuk, 2019). Items chosen in the high-dimensional approaches (e.g., the Big Five) aim to measure broader trait domains and facets rather than purposefully capture unique item-level information. Instead, research such as ours should strive to use a more comprehensive item pool, which does not require being a priori aligned with any hierarchical structured trait model (Condon et al., 2020).

Mean levels of personality traits and personality variances

Participants were drawn from a pre-existing dataset which were collected primarily for various research projects (including for estimating the re-test reliability of the items). These participants were recruited from the online research platform Prolific between December 2019 and March 2020 and each compensated with £4.7. There were 1,436 individuals (mean age = 27.60 years, 577 males, 844 females, and nine with unknown gender) in the initial sample and we selected participants aged between 18 and 30 years old (*N* = 1,096, mean age = 23.04 years; 446 males, 641 females, and eight with unknown gender). Participants were required to rate their own personality traits on each of the 257 item, using a 6-point Likert scale ranging from 1 (very inaccurate) to 6 (very accurate).

Socially expected personality trait levels

Participants for measuring social expectations were also recruited from the online research platform Prolific and each compensated with £3. See Table C1 for descriptive statistics. Participants were required to rate each of the 257 item in terms of whether their friends, intimate partners or bosses/supervisors would generally approve or disapprove of their thinking, feeling or behaving in the way described in these items. In other words, participants (young adults) were asked to rate each item in terms of how much social pressure they perceived to behave in these ways, from either their friends, partners or bosses/supervisors. For example, for friend's perspective, when participants answer question 'I often stop working if it becomes too hard', they need to answer to what extent their friends would approve that they stop working if it becomes too hard. Participants provided the ratings on a 5-point Likert scale ranging from 1 (strongly disapprove) to 5 (strongly approve).

Table C-1. Descriptive data for young adults aged 18 to 30 years in perspectives of friends, partners and bosses used to obtain ratings of social expectations

Raters' perspective	Friends	Partners	Bosses/supervisors
N	41	41	39
Rater age	(*M*=23.88)	(*M*=23.63)	(*M*=24.54)
Number of males	24	21	24
ICC	.95	.96	.96

Note. ICC = Intraclass correlation of (random) average rater.

Data analysis

Statistical analyses were carried out in R (R Development Core Team, 2020). The data and R code are made publicly available at Online Supplemental Material (OSF).

First, we investigated whether social expectations were consistent for young adults using a two-way average absolute random-rater intra-class correlations (*ICCs*) in the *psych* package version 2.0.9 (Revelle et al., 2020). Second, we used eta-squares to distinguish group differences (raters' perspectives) by the proportion of variance in ratings. We also collapsed the ratings within each rating condition (friends, partners, bosses) and calculated the Pearson correlations between the mean social expectations ratings (across all items).

To investigate associations of the personality trait mean levels and variances of young adults on the one hand with the social expectations for these traits on the other, we used multi-level models (or linear mixed-effects models) in the *lme4* package version 1.1-25 (Bates et al., 2014). Dependent variables were either the means or standard deviations of each item from 18 to 30 years (so there were 13 age levels with 18 being the first age level : $N = 13 * 257 = 3{,}341$), whereas independent variables were the ages to which these mean levels corresponded and the social expectations for the items to which the mean levels corresponded. We built two sets of models examining the effects of age, social expectations and their interaction on item mean and standard deviation respectively. All models included a random effect for item (allowing items to have different mean levels) and random slopes for the effects of age per item (allowing differences between items in their age trends). The interaction terms represented the key parameters: if significant, they would show that age differences in items' means/variances would depend on the expectations for these items. We reported fixed effects which summarize general associations across all items. Personality trait mean levels, variances and social expectations were standardised before building the models and age was centred on 18 (i.e. 18 was subtracted from each age to make it represent 0). Moreover, age was divided by 10 before building the model to avoid model convergence issues due to large differences in variances among individual variables. Personality traits and social expectations were reverse coded where needed, so that all items were keyed in the socially desirable directions throughout the analyses (based on the mean of the social expectation ratings, collapsed across all rating conditions).

Results

Did raters agree about social expectations for young adults' personality traits?

The *ICCs* of the ratings for social expectations were very high for the ratings conditions of friends, intimate partners and bosses/supervisors respectively (see Table C1). Also, after applying false discovery rate (FDR) correction for multiple testing, the means of only two items ('often stop working if it becomes too hard' and 'I am satisfied with my relationships') significantly varied across all three groups, suggesting that rating conditions generally produced very similar mean levels for social expectations: the rating condition explained an average of 3% of variance in the social expectation ratings. Moreover, the profiles of the 257 average ratings correlated from .90 to .94 between the three raters' perspectives. The high *ICC* indicated that young adults perceived consistent social expectations within each group of social contacts. The high correlations suggest that people with different perspectives hold relatively similar expectations for the personality traits of young adults, or at least that young adults believe that social expectations from their friends, intimate partners and bosses/supervisors share many similarities. When we collapsed ratings from all conditions, the average *ICC* increased to .98; we therefore combined raters' perspectives in all subsequent analyses.

Did young adults' personality traits generally comply with expectations?

In a multi-level model, the means of the 257 items for each age were predicted from the social expectation ratings of these items and age. Social expectations strongly predicted mean levels of personality trait ratings (see Table C2) indicating that personality traits with higher expectations were more likely to have higher means. This finding supported Hypothesis 1 that young adults' personality traits generally comply with social expectations. In the same model, the results showed that age was also a significant predictor of mean levels of personality traits, suggesting personality trait mean levels tend to be higher at later ages; since the traits had been keyed in socially desirable direction, this suggests a general normative shift towards more desirable trait levels. However, the interaction between age and social expectations was not significant (see Table C2). Thus, age differences in mean levels change irrespective of social expectations for these traits. For visualisation purposes, we rearranged the items into three groups based on whether they were subject to high, medium, or low expectations. Figure C1 shows that items subject to high social expectations show the highest personality trait mean levels, whereas items subject to low social expectations showed the lowest personality trait mean levels.

Table C-2. The effects of social expectations and age on personality trait mean levels, variances and their changes.

Parameters	Mean-level			Variance		
	Soc	Age	Age x Soc	Soc	Age	Age x Soc
Beta	.700	.117	0	-.052	.019	-.065
SE	.049	.025	.025	.045	.032	.032
p	<.001***	<.001***	.977	.001***	<.001***	.042*

Note. Soc = social expectations. * $p < .05$. ** $p < .01$. *** $p < .001$.

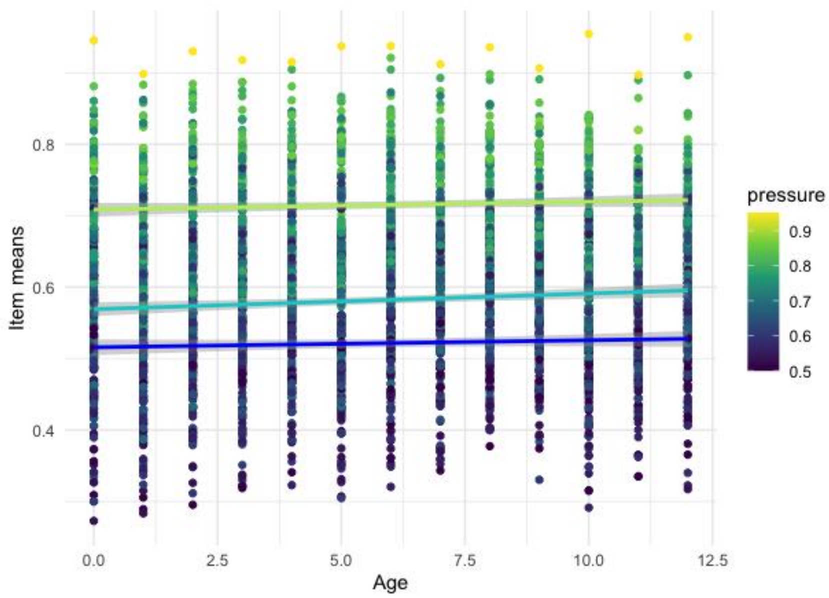


Figure C-1. The effect of different levels of social expectations and age on mean levels of personality traits. The items were divided into three groups based on the high, medium and low expectations for them (lowest (< .33), middle (>= .33, <=.66) and highest (> .66) third of the items). The y-axis represents standardized item mean scores.

Did traits with stronger expectations tend to decrease in variance and were these associations moderated by age?

In terms of predicting traits' variances from social expectations and age at which the variances were calculated, we used the same model as above but items' means were replaced with their standard deviations. There was a relatively strong negative association between traits' variance and social expectations, as well as a positive association between traits' variance and age. The interaction between age and social expectations was significant indicating that the effect of social expectations may not be as pronounced in later adulthood (see Table C2). Figure C2 shows that traits with high expectations showed the lowest level of variance and traits with low expectations showed the highest level of personality variance, Although the interaction was very weak and therefore we do not see strong support for it. Traits with medium and low levels of expectations also shows higher increases than traits with high expectations.

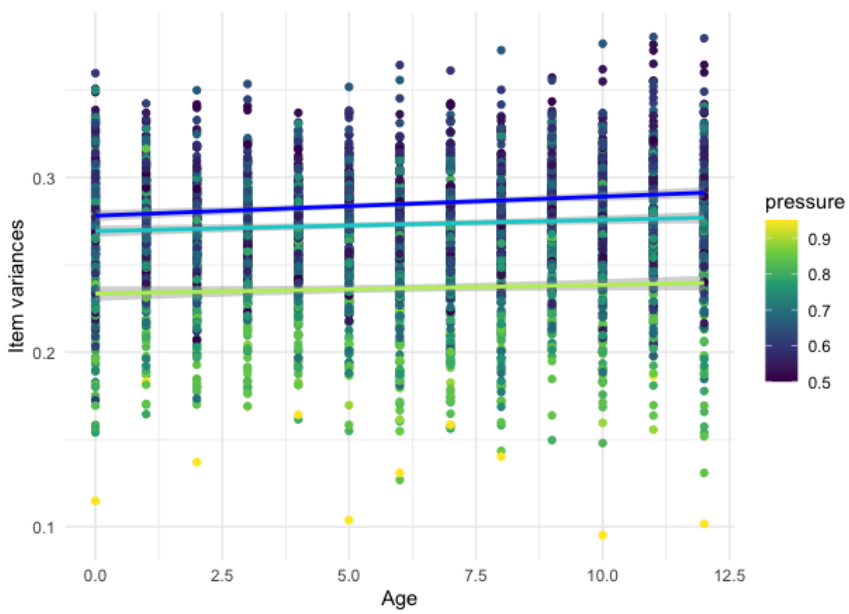


Figure C-2. The effect of different levels of social expectations and age on personality variances.

Discussion

SIT is one of the leading theories to explain personality development and is at least indirectly supported by evidence that individuals become more agreeable, conscientious and emotionally stable in adulthood after committing to different social roles (Asselmann & Specht, 2020; Hudson et al., 2012; Lodi-Smith & Roberts, 2007), although more direct evidence linking trait change with role transitions remains scarce (Bleidorn, Hopwood, Back, et al., 2020b). However, the alleged driving force of behavioural changes in accordance with committing to new age-graded life tasks is unlikely to reflect changing social roles *per se* but the expectations that come along with these social roles from one's self and close contacts (D. Wood & Roberts, 2006). Indeed, according to SIT, the social expectations attached to each social role train individuals to behave in a certain way which can lead to changes in personality if the behaviours are repeated and internalised. As yet, however, there is limited empirical evidence linking social expectations to personality changes (D. Wood & Roberts, 2006).

Comparing traits offers one yet under-explored approach to test the following hypothesis: if personality trait mean levels shift towards maturity with age, then traits that are *expected* to increase the most should in fact have the highest mean levels and even increase in them since expectations should be one of the key components in influencing personality change. Moreover, individuals should become more similar in late adulthood especially in traits that are under the strongest social pressure, since most people would tend to converge towards socially desirable levels of these traits. The present study is among the first studies to offer some evidence for social expectations being a potential mechanism for personality differences during early adulthood using nuance-level analysis.

For mean level changes, our results support SIT in two ways. First, socially desirable personality trait mean levels were higher in later adulthood than earlier in adulthood suggesting that people showed more socially desirable personality traits in older age, hence personality traits became more mature. This result supports previous literature on SIT suggesting that adults learn from their past experiences and thus mean level of personality traits change in a socially desirable direction over the adult lifespan (Denissen et al., 2013; D. Wood & Roberts, 2006). Second, the pattern of mean-level personality traits complied with the pattern of social expectations. On average, personality traits with higher expectations had higher mean levels than traits subject to lower expectations. These findings illustrate the importance of social expectations to personality changes during young adulthood by showing that, in general, traits' changes mirror the corresponding social expectations especially for traits under higher expectations. These findings further imply that high social expectations are one of the sources guiding mean-level personality change in early adulthood. Similar findings have been found in educational settings. For example teachers' expectations showed significant impacts on students' academic performances (Szumski & Karwowski, 2019; Tsiplakides & Keramida, 2010).

In addition to the two results supporting SIT, we also found a result which did not support SIT. According to SIT, we hypothesized that age differences should be moderated by the level of expectations for the traits, with traits under stronger expectations changing more in mean levels. Our result shows that the interaction between age and social expectations was not significant in predicting personality trait mean levels, thus suggesting that age differences in personality trait mean levels are not modified by the effect of social expectations. This result also implies that people have consistent social expectations for adults irrespective to their age.

Advocates of SIT many not have a clear hypothesis in relation to the age difference in individual differences but it would follow from their hypothesis that people become more alike under high social expectations since stronger awards/ punishments are naturally attached to high expectations (Bleidorn, 2015; Roberts, Wood, et al., 2005). Logically individual differences should decrease in later adulthood since people converge towards more socially desirable trait levels, especially (or at least) for traits that are under high social expectations. In the present study, we examined this hypothesis by testing whether traits with stronger expectations showed lower personality variance, and whether age moderates the association between social expectations and personality variance. On the one hand, results indicated that traits under higher social expectations have lower variances than traits under lower expectations which supports SIT. On the other hand, contrary to our hypothesis but not necessarily SIT, personality variance was higher at later ages. This could be due to gene-environment interaction strengthening predispositions, thus increasing variance at later ages (Caspi & Roberts, 2001, 2001; Mõttus, Briley, et al., 2019). Personality variance was also negatively associated with the interaction between age and social expectations indicating that age-trends in variance are moderated by the degree to which traits are subject to social expectations, so increases in variance were smaller for traits with stronger expectations than traits with lower expectations. One possible reason could be that traits' variances increases due to the environment amplifying individual differences, but the strong counteracting influence of social expectations attenuates such influences especially for traits under high social expectations. One further contributing mechanism could be that, although people's personality traits tend to become more socially desirable in later adulthood, individual differences take place even within these socially desirable traits.

If social expectations direct people to think, behave and feel in certain ways, then these expectations should be widely accepted by people with different backgrounds (D. Wood & Roberts, 2006). Social expectations take many forms (Morgan, 2007; Roese & Sherman, 2007), it was essential to test the consistency of the socially expected level of personality traits from different raters' perspectives. Therefore, another interesting finding of the current study was that there are widely accepted expectations for young adults; in particular, young adults comprehend similar social expectations from their friends, intimate partners and bosses/supervisors. Of course, most people would expect people they know to generally behave in a socially desirable manner. However, at least in principle people with different social roles could have systematically different expectations of the same person especially relating to circumstances such as 'is easy to fool' or 'is a submissive person'. For example, if person B is person A's romantic partner and person C is person A's boss, person B could have very different expectations for A than person C. But it turns out that this is not the case. One possible reason could be that the perceived social expectations we measured greatly reflect general social desirability because young adults might not have clear perceptions of what exactly their friends', partners' and bosses' are expecting, thus these expectations are similar across sources. Furthermore, as stated in the previous paragraph, people generally hold similar social expectations for young adults regardless of their age.

Strengths and Limitations

The present research had a number of strengths. First, we used an underused research design to study systematic variability among many traits in various focal properties (Mõttus et al., 2020). We studied socially expected level of personality traits, mean levels of personality traits and personality variance. Our sample is a pool of personality items rather than a group of participants. This method allowed us to use participants from different samples to measure the mean levels of personality traits and social expectations. Second, we used a large sample of item-based personality measures to test age difference in personality during young adulthood, and the measure we used was particularly designed to study item-level traits which is a more fine-grained measure compared to any hierarchical structured personality measure such as the Big Five or HEXACO. In this study's context, taking a nuance-level approach to study these associations is particularly important since focusing on the domain-level only would likely obscure social expectation-personality associations that are important for shaping individuals' personalities. Third, we measured the consistency of social expectations rather than assuming it.

However, the present study also suffered from three main caveats. First, the item pool we used is fairly unstructured yet, so there is no neat alignment with facets an domains. Second, we used cross-sectional data rather than longitudinal data; this restricts our ability to draw inferences on personality change across the life span. It would be beneficial if future studies could investigate whether the present findings extend to longitudinal data for both mean level personality, personality variance, and social expectations. Third, we only investigated young adults age between 18 to 30 years. The pattern of personality differences might be better illustrated using a broader age range since some people might engage in important social roles early or later than others. Future studies could test this idea with a broader age range and compare personality across generations in order to test how social expectations change personality traits across time. Moreover, being parents is an important social role which could also contribute to personality change. Future studies could expand the present study by including different populations such as parents.

Conclusions

Social expectations contribute to personality differences in both mean levels and variances during young adulthood. Our results show that personality traits under stronger social expectations have higher mean levels than traits under lower expectations; traits show higher means at later ages; and traits with stronger social expectations have smaller variances but variances overall increase with age. Overall, our results are only partially consistent with SIT and suggest that social expectations could be a potential mechanism to explain personality changes in young adulthood.

Chapter 5: Personality and Health Network

Synopsis:

Chapters 3 and 4 investigated the potential mechanisms of personality differences across the life span using nuance-level analysis. Expanding on the theoretical findings from the previous chapters, this chapter focuses on the practical value of nuance-level research. This chapter explores how different health-related aspects and health outcomes are linked with each other based on nuance-level personality traits. Results show that personality correlations could explain many associations between health-related behaviours and outcomes, and particularly psychopathological comorbidities.

Dissemination status:

This chapter has been under review at *Journal of Abnormal Psychology* entitled "Establishing a Health Network of Personality Profiles for Adolescents and Emerging Adults". A pre-print is available on *PsyArxiv*: https://psyarxiv.com/4qh96/

Establishing a Health Network of Personality Profiles for Adolescents and Emerging Adults

Yuzhan Hang[a*], Lydia Gabriela Speyer[a], Liina Haring[b], Billy Lee[a], Uku Vainik[c,d], Michelle Luciano[a], René Mõttus[a]

[a]Department of Psychology, University of Edinburgh, EH8 9JZ, Edinburgh, United Kingdom

[b]Department of Psychiatry (ARPS), University of Tartu, 50411, Tartu, Estonia

[c]Institute of Psychology, Faculty of Social Sciences, University of Tartu, Estonia

[d]Department of Neurology and Neurosurgery, Faculty of Medicine and Health Sciences, McGill University, Canada

Acknowledgements: Uku Vainik has been supported by Estonian Research Council's grant MOBTP94"

Abstract

Many health problems that occur later in life have their origins in behaviours and associated lifestyle habits established earlier in life. We aimed to gain new insights into the structure of health and well-being of late adolescents and emerging adults through examining a multi-dimensional network that quantitatively estimates the personality similarities (personality correlations) between sixteen different health related behaviours and outcomes. The personality correlations were based on nuance level personality traits, captured by 240 items of the EE.PIP-NEO Personality Inventory that predicted the outcomes more accurately than broader personality traits (Big Five domains and facets; $N = 2{,}269$), and analysed using Exploratory Graph Analysis. The sixteen outcomes fell into four groups based on their personality correlations: psychological distress, health awareness, emotional control and substance use. Personality correlations, quantifying the overlap among outcomes in their psychological background, can explain associations between health-related behaviours and outcomes, and psychopathological comorbidities.

Keywords: network analysis; lifestyle; mental health; personality; nuances

Introduction

Many health problems that occur later in life have their origins in behaviours and associated lifestyle habits established during adolescence. For example, experiencing new living situations and being increasingly independent can cause psychological distresses and potentially increase the risks of developing mental health issues (Pei et al., 2019). In addition, adolescents have increasing access to alcohol, tobacco and illicit drugs, but unlike adults, are cognitively under-developed to appropriately assess their potential dangers. Likewise, emerging adulthood (Arnett, 2000), a stage of life between adolescence and adulthood that spans the late teens to mid-20s (approximately from 18 to 25 years) is characterized by instability and an increase in exposure to new environments (Moreira et al., 2015; Pusch et al., 2019) and it shares many core developmental features with late adolescence. Health behaviour patterns are often established during these periods and have long lasting consequences on individuals' health and wellbeing (M. C. Nelson et al., 2008). This makes it essential to address both adolescence and emerging adulthood when investigating risk factors for adverse health.

Personality traits and health

Previous research has established that personality traits are linked with a variety of life outcomes, including health and lifestyle (Ferdosi et al., 2020; Heilmayr & Friedman, 2020; Hoffmann & Risse, 2020; Strickhouser et al., 2017). These associations are often reported at the level of broad personality domains, such as the Big Five: Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. Typically, these findings tend to conform to a pattern whereby positive (socially desirable) personality traits are linked with positive outcomes (e.g., Openness with healthy dietary habits; Conscientiousness with longevity; Agreeableness with positive mental health) and vice versa (e.g., Neuroticism with mental health disorders) (Conner et al., 2017;

Hengartner et al., 2016; Lamers et al., 2012; Roberts, Walton, et al., 2005). The ultimate health outcome, mortality, is linked with all of the Big Five traits except Openness (Graham et al., 2017). The low specificity of such broad personality findings makes it harder to understand their mechanisms or explore their possible practical uses.

Personality traits, however, can be represented as a hierarchy of narrower constructs because the Big Five domains can be broken down into *facets* and these, in turn, into *nuances* (Mõttus et al., 2020). Facets are groups of items within the Big Five domains that occur together more closely among themselves than with other items. Nuances are the lowest level of the personality trait hierarchy at which patterns of responses to personality test items have valid specific variance (Condon et al., 2020). In many cases, this will effectively equate individual items with nuances because most items have valid unique variance beyond the variance that they contribute to the measurement of facets and domains (Mõttus, Sinick, et al., 2019). Often this unique variance also drives the associations of personality traits with life outcomes (Elleman et al., 2020; Mõttus, Bates, et al., 2017; Seeboth & Mõttus, 2018).

The narrower traits, facets and nuances, may supplement the study of specific research questions since they allow for greater specificity and hence detail than the broad Big Five domains. Of course, such findings will inevitably be less parsimonious, but it is becoming increasingly clear that psychologists have to face small effect sizes and high multi-determinacy of their phenomena (Friedrich Götz, et al., 2021), and thereby accept more complex models (Mõttus et al., 2016, 2020; Yarkoni & Westfall, 2017). For example, studies have reported a consistent negative correlation between obesity and Conscientiousness (Jokela et al., 2013), but associations with the other domains are inconsistent (Arumäe et al., 2020), because facets or nuances within the same domain often vary in both the level and sometimes even direction in their associations with body mass index (BMI), which attenuates the domain-level personality trait–BMI associations (Vainik, Dagher, et al., 2019). In particular, the Warmth, Positive Emotions and Assertiveness facets of the Extraversion domain are positively linked with BMI, whereas the association with the Activity Level facet is negative (Vainik, Dagher, et al., 2019). Within the Impulsiveness facet of Neuroticism, items referring to over-eating and trouble with resisting cravings are positively correlated with BMI, while the item referring to an inability to remain calm is negatively linked with BMI (Mõttus, Sinick, et al., 2019). Similar patterns have also appeared for other outcomes such as age differences (Ashton & Lee, 2016; Hang, Soto, Speyer, et al., 2021; Mõttus & Rozgonjuk, 2019; Soto & John, 2012), or associations with health and other outcomes in adults (Mõttus et al., 2016).

Are different aspects of health and lifestyle linked to similar personality (psychological) profiles?

Facet- and especially nuance-level analyses allow for more detailed questions as to which specific traits drive the associations of personality variations with health, lifestyle and other life outcomes. In addition, they also allow investigation of broader patterns, through a systematic and quantitative examination of which outcomes are related to personality traits similarly. For example, we could quantify a large and diverse sample of nuance-level traits in terms of their links to various lifestyle aspects such as dietary behaviours, substance use, exercising and sleeping patterns and then correlate the association profiles with each other. A high correlation between two associated profiles would indicate that these lifestyle aspects are similar in terms of their personality correlates. Further, to the extent that the set of personality nuances broadly summarizes individuals' psychological profiles generally, this can show that the general psychological aetiology (or background) of these lifestyle aspects is overlapping. We can call an association between the profiles of two (health or other) outcomes a *personality correlation* or *psychological correlation*

(Seeboth & Mõttus, 2018; Vainik et al., 2020). In other terms, the nuance-level personality correlations estimate the extent to which the personality "risk" for one health outcome is associated with other health outcomes. The presence of personality correlations would indicate that personality characteristics not only account for variance in health-related lifestyles and outcomes but also their co-variance. Note that there is an instructive parallel with widely studied *genetic correlations* among observable phenomena, which quantify the extents to which they have similar genetic correlates (Neale & Maes, 1996). In these studies, a genetic correlation can reflect associations among phenomena arising from shared genetic aetiology, not necessarily due to direct causal associations between them. Likewise, a sufficiently high personality correlation could explain why healthy behaviours often go together – not because of direct causal relations between them, but because they pertain to psychologically similar people. That is, the same potentially diverse set of psychological traits may drive people to the same or different lifestyles and health outcomes. There is already evidence that vulnerability to drug use and to over-eating could be due to the same predisposing personality traits (Vainik et al., 2020).

The present study

The present study aims to gain new insights into health and well-being of adolescents and emerging adults through examining a network that links personality traits with health-promoting lifestyle choices, health-harming lifestyle choices and mental health problems. In particular, this study focused on sleeping behaviours, healthy dietary behaviours, exercise habits, troubled eating behaviours, internet use, substance use, and mental health issues. Using machine learning procedures that account for model complexity and overfitting, we first predicted each lifestyle aspect and mental health outcome from the Big Five domains, their 30 facets and numerous nuances (indexed by 240 items selected to measure the 30 facets), expecting nuances to confer the greatest predictive power in most cases. Next, we created a personality profile for each of the outcomes by correlating individual nuances (items) with them and subsequently compared these resulting correlation profiles to obtain personality correlations between the outcomes. For this, we used exploratory graph analysis, a data-driven approach that allows detecting clusters of similar variables based on their (here, personality) correlations. Finally, we compared the correlations between the observed lifestyle and mental health outcomes with their personality correlations to estimate the extent to which personality correlations could, in principle, cause the observed correlations between the lifestyle and mental health outcomes. To the extent that personality correlations notably exceed observed correlations, personality traits may confound the observed correlations, by representing a shared psychological background for them. Where observed correlations exceed personality correlations, this is less plausible and the observed correlations may be more likely to reflect direct or mediated links (Revelle et al., 2020).

Method

Participants

Participants were drawn from the Ained ja Arenevad Ajud dataset (AAA; Drugs and Developing Brains). The AAA studied drug use and mental health conditions in Estonian youths, as well as their personality traits. There were 4,005 participants in the initial sample (2,514 females, mean age = 21.42 years). A quality check was performed to check low-effort participants. First, data from participants with more than 40 missing personality responses were dropped. Then, we removed all participants who had 10 or more identical responses in a row. Finally, we correlated individuals' profiles of the 240 personality items with the average item profile across all individuals and removed participants with reverse profiles (profile correlations equal or below zero). These unusual item profiles represented either people with very unusual personalities or individuals who did not take participation seriously. As a result, 756 participants were removed because of low-effort responses (364 due to having more then 40 missing responses, 392 due to having more than 10 identical answers in a row and unusual personality profiles). After data cleaning, we selected participants aged between 16 and 25 years (N = 2,398, mean age = 18.25 years).

Measures

Personality traits

Participants completed a 240-item Estonian version of the International Personality Item Pool NEO [EE.PIP-NEO] (Mõttus et al., 2006). The EE.PIP-NEO is comparable to the NEO-PI-R in terms of relevant psychometric properties (Mõttus et al., 2006), but the EE.PIP-NEO is linguistically simpler. Participants were required to rate the items on a 5-point Likert scale ranging from 1 (wrong/ does not agree at all) to 5 (correct/completely agree).

Health-promoting lifestyles

Three health-promoting lifestyle variables were studied: healthy sleeping, dietary, and exercise habits. Sleeping behaviours were assessed by the 8-item Sleep Condition Indicator [SCI] (Espie et al., 2014). The SCI measures sleep problems, particularly insomnia disorder, according to the DSM-5 criteria. The SCI questionnaire was found to be highly reliable (Cronbach's α = .86). Participants responded to questions regarding their quality of sleep (e.g. duration of sleep problem). Participants were required to rate the items on a 5-point Likert scale ranging from 0 to 4. High scores reflected healthier sleeping habits and fewer issues with sleep.

Healthy dietary and exercise habits were measured using a 12-item questionnaire which measures individuals' lifestyles in terms of their exercise habits, dietary habits, and internet usage. The exercise habits subscale consisted of two items measure the frequency of exercise on a weekly basis and the length of exercise every time (Cronbach's $\alpha = .60$). The items were scaled, thus higher scores indicate more exercise and healthier lifestyle while lower scores indicate more sedentary lifestyle. For the healthy diet variable, we conducted a 1-factor principal component analysis (PCA) on three items measuring weekly intakes of vegetables, fruits, and soft drinks. Participants were required to rate each of the three questions on a 7-point scale ranging from 1 (1 day/week) to 7 (7days/week). The extracted principal component was characterized by positive loadings of vegetable and fruit intake and negative loadings of soft-drink intake (component loading: .85, .85, and -.40). A higher score reflects more health-promoting behaviours: more vegetable and fruit consumption and less soft-drink consumption.

Health-harming lifestyles

Internet usage, troubled eating behaviours, and substance usage were the three health-harming variables included in this study. Similarly to dietary choices and exercise habits, internet use was also assessed by the lifestyle questionnaire. We extracted the first component from a 1-factor PCA on three items measuring length of weekly internet use during weekdays, weekends and before sleep (component loading: .88, .86 and .75). Participants were required to rate each of the three questions on a 7-point scale ranging from 1 (< 1 hour) to 5 (> 5 hours). A higher sub-score indicates potentially more health-harming behaviours (e.g., addiction). While there may be positive effects of internet use (Castellacci & Tveito, 2018), here we consider greater internet use as maladaptive because of its addictive features (Greenfield, 2011) and its associations with health problems such as low vision and obesity (Bener et al., 2011).

Troubled eating behaviours were assessed by the Rapid Assessment of Reward-Related Eating [RED-X5] (Vainik, Eun Han, et al., 2019). RED-X5 is a self-report questionnaire which assesses eating behaviours in three related constructs: lack of control/over-eating, lack of satiety, and preoccupation with food. This scale includes items based on existing questionnaires (e.g. Binge Eating Scale) as well as newly developed items (Vainik, Eun Han, et al., 2019). Participants were required to rate five items on a five-point Likert scale ranging from 1 (does not agree at all) to 5 (completely agree; Cronbach's $\alpha = .84$).

Substance use was assessed by the WHO Alcohol, Smoking and Substance Involvement Screening Test [ASSIST] (Humeniuk et al., 2008). The ASSIST is a substance use instrument developed by the World Health Organization to identify lifetime and current substance use issues across many countries and cultures (McNeely et al., 2014; Onifade, 2014). The ASSIST measures ten different substances including tobacco, alcohol, cannabis, cocaine, amphetamine, inhalants, sedatives, hallucinogens, opioids, and others. We chose tobacco, alcohol and cannabis in our analysis because all other substances were rarely used and their use was unlikely to be reliably predictable (only 11% of participants used any of the other drugs) Cronbach's alphas for the ASSIST questionnaire were .84, .78 and .84, for tobacco, alcohol and cannabis use respectively.

Mental health outcomes

Mental health conditions were measured by the DSM-5 Self-Rated Level 1 Cross-Cutting Symptom Measure [CCSM] (American Psychiatric Association, 2013, 2014). It assesses mental health difficulties in multiple domains. The adult version consists of 23 questions that assess 13 psychiatric domains and the child version consists of 25 questions that assess 12 psychiatric domains (Cronbach's $\alpha = .90$) (American Psychiatric Association, 2013). Participants aged under 18 years completed the child version and participants aged above 18 years completed the adult version. We chose eight psychiatric domains (depression, anxiety, anger, mania, repetitive thoughts, somatic symptoms, suicidal ideations, and psychosis) that were common to both the adult and the child version. We discarded the sleep problems and substance abuse domains in both the adult and child assessments because we are already using the SCI questionnaire to measure sleep problems and the ASSIST questionnaire to measure substance use problems (the SCI and ASSIST contain more items and could provide a more comprehensive view of sleep problems and substance use problems compared to the CCSM). The inattention domain and irritability domain were only included in the child version, while memory, dissociation, and personality functioning domains were only included in the adult version, thus these domains were also excluded from the current analyses. A higher sub-score in each domain indicates more mental health problems.

Data analysis

Statistical analyses were carried out in R (R Development Core Team, 2019). The data and R code are publicly available at [the Online Supplemental Material: https://osf.io/jwmtn/].

The elastic net regression model

To examine the predictive ability of each level of the personality hierarchy, we predicted lifestyle aspects and mental health outcomes separately from each of these levels: domains, facets, and nuances. Thus, lifestyle and mental health were treated as the outcome variables and the personality traits were treated as the predictor variables. We trained the (linear) elastic net regression models for each outcome and validated them for prediction in separate sample partitions (75% training - 25% validation split) to avoid over-fitting and more complex models having an a priori advantage (Seeboth & Mõttus, 2018; Yarkoni & Westfall, 2017). The models were trained using the glmnet package (Friedman et al., 2010), with 10-fold cross-validation choosing the regularisation parameter that minimised cross-validation error ("lambda min"). This approach of using penalized regression coefficients by shrinking them towards zero avoids inflated (over-fit) coefficients. We repeated the training-validation procedure for 100 random sample splits and computed the mean and standard deviation for each type of correlation. Mean absolute errors of the predictive accuracy are reported in the supplementary materials (see Table S1). This analysis allowed us to determine whether associations between personality traits and different lifestyle variables and mental health outcomes were primarily due to domains, facets or nuances.

Plotting the associations

In addition to predicting lifestyle and mental health variables from personality traits, we created univariate correlation-based personality profiles of these variables and compared the profile similarities in the form of partial correlations. As our analyses (see results) showed that nuance level best predict lifestyle and mental health outcomes, we created personality profiles based on the nuance level analysis by correlating all 240 items with each lifestyle aspect and mental health outcome, yielding a profile including 240 Spearman's correlations for each. The correlations between individual items and each outcome were visualized using plots, such that the items are grouped by the facets and domains they are selected to measure: this helped us to see how items within facets vary in their correlations with outcomes, but also how facets of the same domains vary in how their items tend to be linked with the outcomes.[4]

As the next step, we correlated the personality profiles with each other to measure profile similarities, that is, personality correlations. Since some of our data was not normally distributed, we used Spearman's correlations for all tests. To assess and visualise the psychometric properties of these 16 personality profiles, we conducted an Exploratory Graph Analysis as implemented in the R package EGAnet [EGA] (H. Golino et al., 2019; H. F. Golino & Epskamp, 2017). Using graphical lasso estimation with extended Bayesian information criterion as regularization

[4] We further conducted a robustness check of the reliability of the personality profiles of lifestyle factors and mental health outcomes using cross prediction. Results were in line with the results of the elastic net model (see supplementary materials).

parameter, EGA first estimates a sparse undirected graph structure that is classified into several psychometric dimensions based on a weighted network community detection algorithm. Thus, it represents an intuitive visualization of complex dependency structures between multiple variables in the form of a graphical model, in which measured variables are represented by nodes with edges indicating relations between variable, and further allows variables to be grouped together based on their psychometric similarities (Christensen et al., 2020). We visualized the network structure using the Fruchterman-Reingold algorithm as implemented in the R package qgraph (Epskamp et al., 2012), which places more related nodes closer together in space. We only included edge weights greater than .10 in the visualisation in order to focus on interpretability of the strongest relations. Finally, using *bootEGA*, we employed bootstrapping routines (N = 1,000) to evaluate the stability of the detected communities and to assess the accuracy of edge weights using 95% confidence intervals (*CIs*). Since the EGA was conducted using LASSO regularisation, bootstrapped *CI*'s were not centred on the true parameter value making this a conservative test (Epskamp et al., 2018).

Results

Predictions

Results of the prediction analysis indicated that nuance-level personality traits explained somewhat more variance than facets, and facets explained more variance than domains, although the differences in accuracies were small. Mean accuracy for prediction at each level of the personality hierarchy are summarized in Table D1 for all lifestyle aspects and mental health outcomes. Across the 16 outcomes, the observed outcomes correlated with their predicted values on average r = .41, .47, and .51, when based on domains, facets and nuances (items), respectively. The largest increase in prediction was for troubled eating, with *r* = .41, *r* = .52, and *r* = .64 respectively for domains, facets and nuances. Facets explained on average 31% more variance in outcomes than domains (22.1% vs 16.8%), while nuances explained 18% more variance than facets (26.0% vs 22.1%). Nuances captured 55% more variance than domains.

Nuance-level correlations with each investigated lifestyle aspect and mental health outcome are visualized in the form of plots (see Figures D1, and Figures S2 to S15 in the supplementary materials). These plots show that items from the same facets, representing partly distinct nuances, often differed in their correlations with different lifestyle variables and mental health outcomes. Sometimes the items of the same facets even varied in correlation direction. For example, some items from the A6: Sympathy facet showed positive, whereas the others showed negative correlations, with the sleeping variable (see Figure D1). Consequently, if just using domain or facet level analyses, personality traits' associations with sleep phenomena might have been reduced, which is why nuance-level analyses captured more variance than domain and facet-level based analyses in many cases. Since nuance-level personality traits showed better predictions than the domain- and facet-level personality traits, subsequent analyses were focused only on nuance-level personality traits.

Table D-1. Correlations between Observed Lifestyle Aspects and Mental Health Outcomes and Their Predicted Values from Personality Traits

		Parameter	**Domains**	**Facets**	**Nuances**
Health promoting lifestyle	Sleeping	Mean	.53	.58	.60
		SD	.02	.02	.02
	Exercise	Mean	.28	.31	.35
		SD	.04	.03	.03

	Healthy diet	Mean	.32	.36	.38
		SD	.03	.03	.03
Health harming lifestyle	Internet use	Mean	.35	.39	.44
		SD	.03	.03	.03
	Troubled eating	Mean	.41	.52	.64
		SD	.03	.03	.02
	Tobacco	Mean	.30	.38	.44
		SD	.03	.03	.03
	Alcohol	Mean	.42	.48	.56
		SD	.03	.03	.02
	Cannabis	Mean	.27	.31	.36
		SD	.04	.03	.03
Mental health outcomes	Depression	Mean	.63	.73	.75
		SD	.02	.02	.01
	Anxiety	Mean	.67	.70	.74
		SD	.02	.02	.02
	Anger	Mean	.58	.63	.64
		SD	.02	.02	.02
	Mania	Mean	.27	.33	.37
		SD	.04	.03	.03
	Repetitive thoughts	Mean	.45	.48	.52
		SD	.03	.03	.03
	Somatic symptoms	Mean	.44	.47	.49
		SD	.03	.03	.03
	Suicidal ideations	Mean	.46	.55	.58
		SD	.03	.02	.02
	Psychosis	Mean	.25	.31	.35
		SD	.03	.03	.03

Note. The prediction was trained across 100 replications with the training sample of 75% of the total sample.

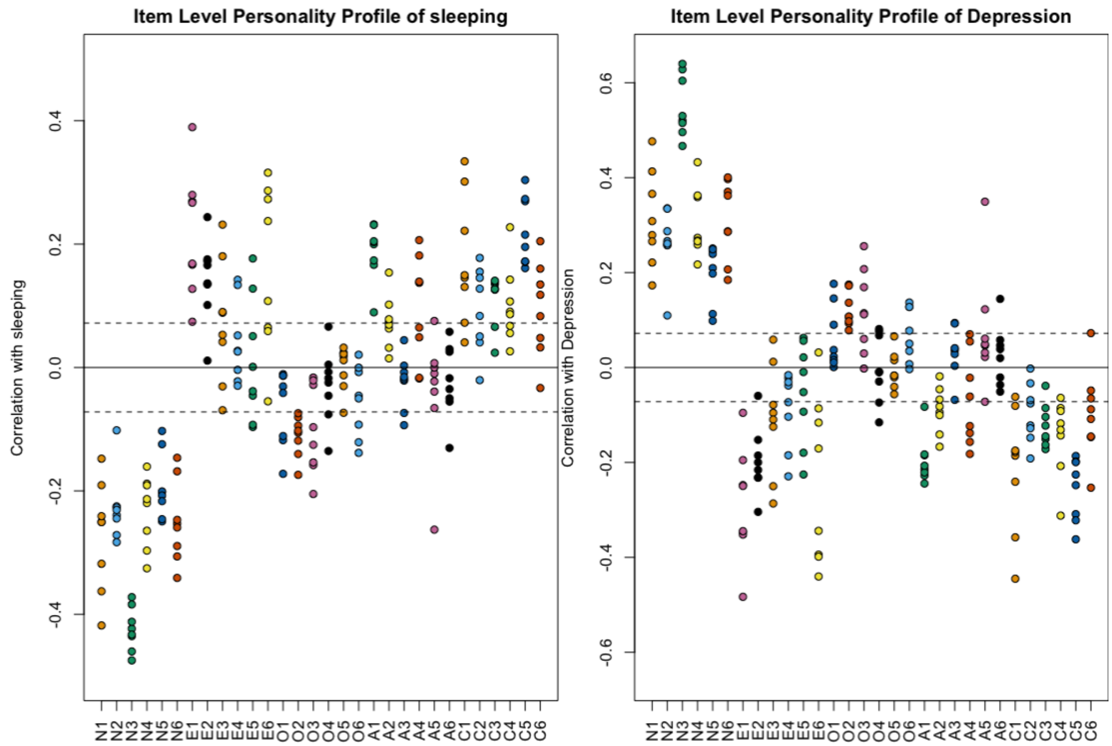


Figure D-1. Plot of the 240 EE.PIP-NEO item correlations with two criteria: sleeping quality and depression. The correlations are grouped according to the Big Five domains (indicated by letter) and their facets (indicated by number). The dashed line indicates the threshold for statistical significance after correcting for multiple comparisons using Holm's correction (Holm, 1979). Facet labels are provided in the supplementary materials.

For example, items that referred to influencing and imposing their opinion on other people's opinion from E3: Assertiveness facet, showed negative correlations with sleep, but items that referred to leading others and making decision for themselves showed positive correlations with sleep; the sleep quality might change with level of confidence. For E4: Activity level, items that referred to having many plans, having things to do all the time, acting swiftly and handling several things at once, showed positive correlations with sleep while items that referred to liking to rush and not act calmly showed negative correlations with sleep. Thus, on average, it is possible that people who are calm and have a fulfilling life might have better sleep quality than people who are irritable.

As for depression, items of the A4: Cooperation facet that referred to avoiding conflicts showed positive correlations with depression, but items that referred to being conciliatory and not being angry for long, showed negative correlations with depression. For A6: Sympathy, items that referred to sympathizing with the homeless and being concerned about other people showed positive correlations with depression, but items that referred to liking people with soft hearts, thinking about those who are in trouble, and empathizing with others showed negative correlations with depression.

Some of these item-level predictions on lifestyle aspects and health outcomes might seem circular. For example, an item of the N3: Depression facet that referred to "I get depressed sometimes" represented the most predictive nuance for the sleeping personality profile. But it is obvious that the items similar to the outcome *also* drive the predictive power of higher order traits, and removing these items would diminish their predictive power (Vainik et al., 2015). In that the sense, item-level analyses *reveal* trivialities rather than *cause* them.

Exploratory Graph Analysis

We used nuances to assess the similarities between lifestyle and mental health personality profiles – personality correlations – to explore how different lifestyle aspects and mental health outcomes overlapped in their corresponding psychological characteristics. Similarities between different personality profiles were depicted by a heatmap of correlations, which tended to be quite high in general, but also formed more specific patterns (see Figure D2).

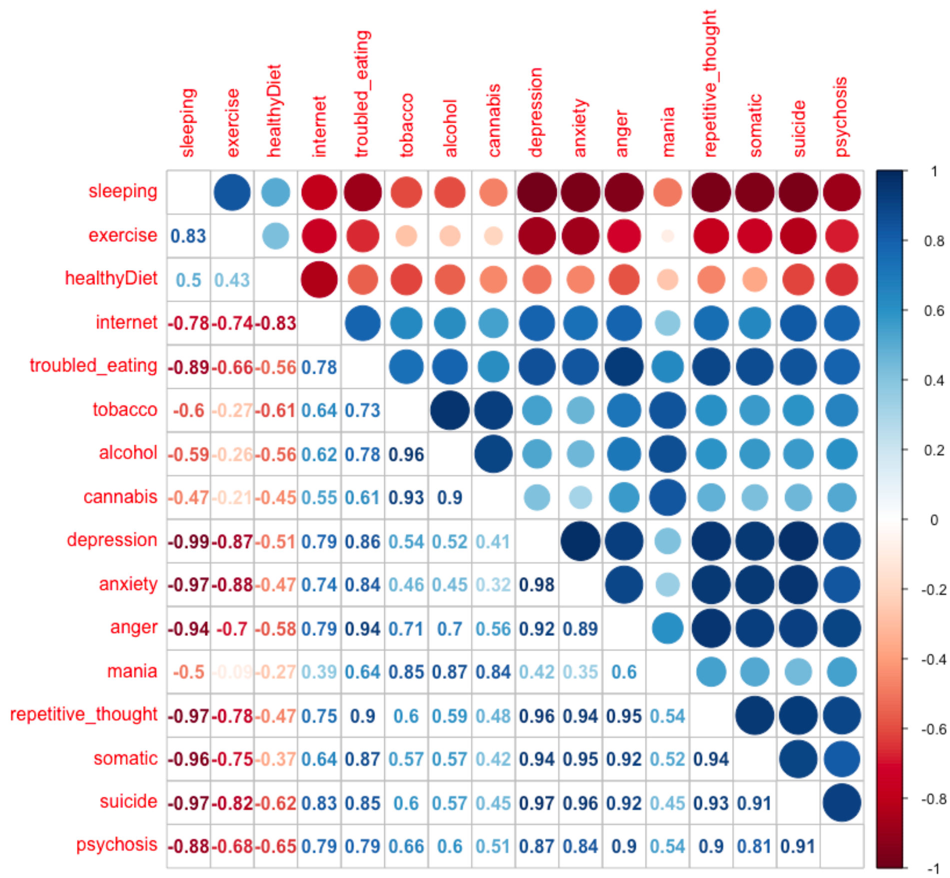


Figure D-2. Spearman correlations between 16 nuance-level personality profiles. Each personality profile was created by correlating 240 EE.PIP-NEO and different lifestyle or mental health variables (e.g. the personality profile of healthy diet lifestyle is the correlation between personality nuances and healthy diet lifestyle). Darker colours indicate higher correlations between two personality profiles which further suggest more similarities between these two profiles. The blue circles and numbers indicate positive correlations and the red circles and numbers indicate negative correlations.

According to Figure D2 some personality profiles were very highly correlated such as the depression personality profile with the sleeping personality profile (r = -.99), and the repetitive thoughts personality profile with the sleeping personality profile (r = -.97). There were also some low correlations between personality profiles such as the mania personality profile with the exercise personality profile (r = -.09).

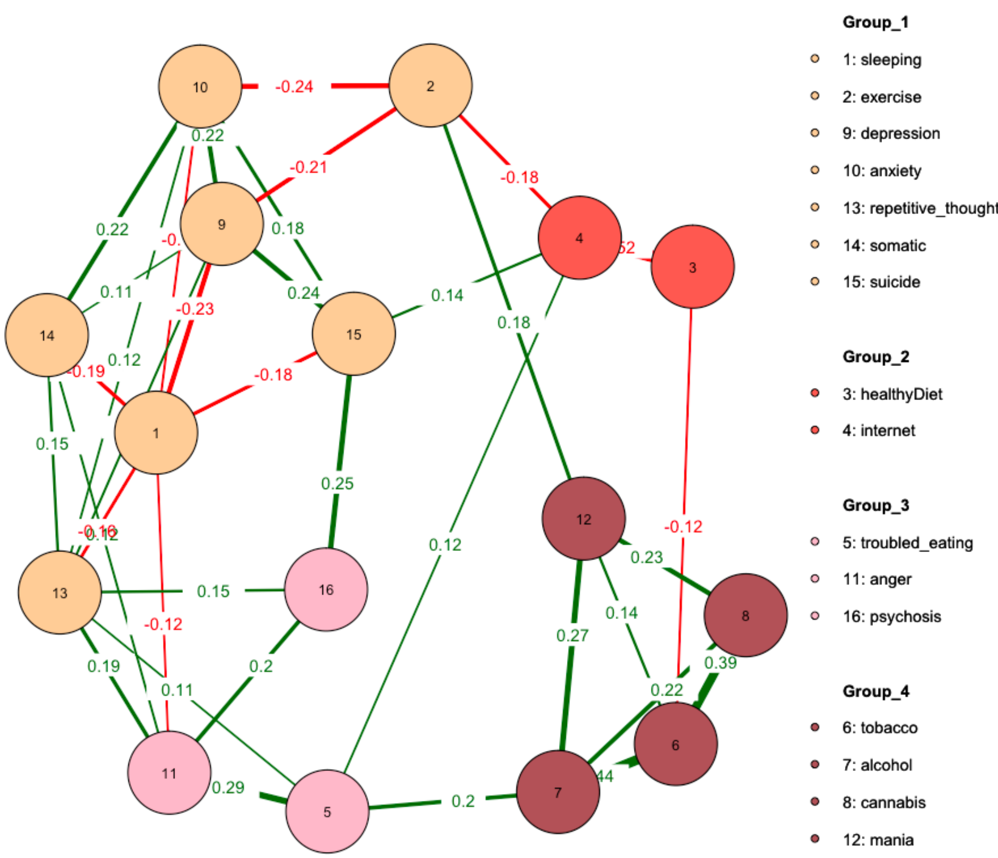


Figure D-3. Spring-embedded network graph of personality profiles visualized in the form of partial correlations using the Fruchterman-Reingold algorithm. Circles indicate personality profiles. Green edges indicate positive relationships and red edges indicate negative relationships between personality profiles. Node colour indicates group membership.

The EGA gave insights into conditional dependencies between the different personality profiles in the form of partial correlations, that is, personality correlations between lifestyle aspects and mental health outcomes controlling for all other personality correlations. These associations can be more informative than the raw correlations shown in the heatmap, because they provide a sparser network of associations. Among other things, they adjust for the overall tendency for different outcomes to have similar personality correlates, discussed in the introduction. Personality profiles were classified into four groups based on the EGA (Figure D3).

<u>Group one: Psychological Distress</u>

Five mental health problems in group one tended to overlap in terms of their broader psychological backgrounds as indicated by strong personality correlations between them (conditional on all other correlations), such as between anxiety and somatic symptoms (w = .22, CI =.17 to .26) and anxiety and depression (w = .22, CI =.20 to .25). These results suggested that mental health problems shared many personality similarities and potentially a common underlying psychological aetiology that might explain their comorbidity (e.g., psychological characteristics that pertain to higher vulnerability to depression would also pertain to higher vulnerability to anxiety). In comparison, sleeping and exercise profiles showed negative correlations with mental health profiles (sleep and depression: w = -.23, CI = -.26 to -.20; exercise and anxiety: w = -.24, CI = -.29 to -.18), but they did not overlap much with each other, net of all other profiles.

<u>Group two: Health Awareness</u>

Healthy diet and internet use (w = -.52, CI = -.58 to -.46) were the only two profiles in the second group. These two profiles showed negative correlations with each other, net of all other profiles in the network, indicating that people who tend to adopt a healthy diet would have psychologically opposite traits compared to people who spend a long time using internet. The most predictive item of healthy diet referred to enjoying meaningful and thoughtful reading, while the most predictive item of internet use referred to avoiding spending time on pointless things. Although these two items came from different facets and even domains, their contents do share some similarities in terms of enjoying/hating leisure time.

Group three: Emotional Control

Troubled eating was grouped with psychosis and anger (anger and troubled eating: w = .29, CI =.24 to .34; anger and psychosis: w = .20, CI =.15 to .25). However, there was no direct link between the troubled eating and the psychosis profile, controlling for all other profiles, therefore anger could play a mediating role between troubled eating and psychosis. The most predictive item of troubled eating referred to often eating too much which could be largely characterised by emotional eating (both positive and negative emotions) while the most predictive item of anger and psychosis referred to getting angry quickly and often being glad. These two items indicate a tendency of being emotional, which is consequently linked with troubled eating traits.

Group four: Substance Use

Substance use personality profiles were clustered together in this group. Our results suggested that all three substances (tobacco, alcohol and cannabis) showed positive personality correlations with each other, controlling for all other profiles in the network (tobacco-alcohol; w = .43, CI =.39 to .48; tobacco-cannabis: w = .38, CI =.33 to .44; alcohol-cannabis: w = .22, CI =.16 to .28). Moreover, all three substances shared personality correlations with mania (tobacco use: w = .14, CI = .09 to .20; alcohol use: w = .27, CI = .21 to .32 and cannabis use: w = .23, CI =.16 to .30). Note that alcohol use had a personality correlation with troubled eating from group three and mania with exercising from group one.

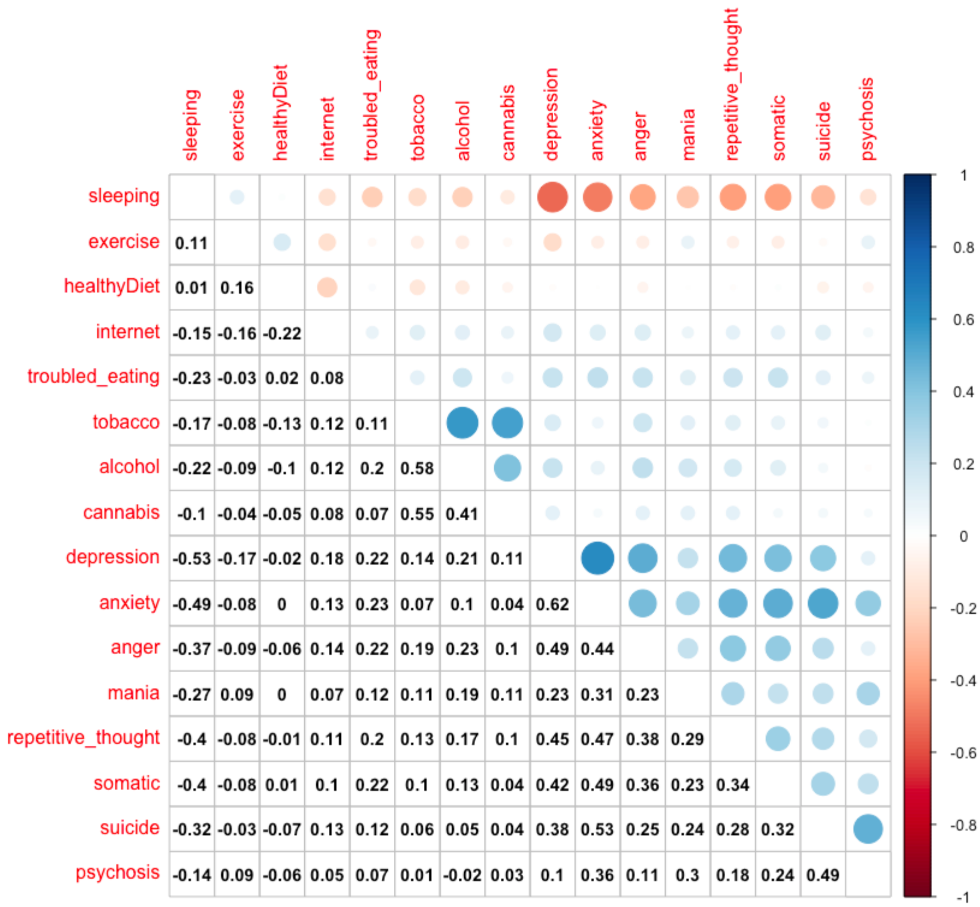


Figure D-4. Spearman correlations between 16 lifestyle and mental health variables. This heat map indicates the associations between lifestyle aspects and mental health outcomes. Darker colours indicate higher correlations between two variables. The blue circles and numbers indicate positive correlations and the red circles and numbers indicate negative correlations.

Did personality correlations track the phenotypic correlations?

The observed correlations between the (phenotypic) lifestyle aspects and mental health outcomes (*absolute mean* = .24; Figure D4) were generally weaker than the correlations between the personality profile correlations of these variables (*absolute mean* = .71; Figure D2). Within the 256 calculated correlations, 162 personality profile correlations exceeded the phenotypic correlations. This finding suggests that lifestyle aspects and mental health outcomes could often overlap because they are influenced by the same personality characteristics. In contrast, for the remaining 94 correlations, phenotypic correlations were higher than personality profile correlations, indicating that these outcomes may be characterised by more direct associations – or that they may be confounded by things other than sharing similar psychological correlates. Despite personality correlations generally exceeding the phenotypic, the relative strengths were similar: the correlations between the raw variables (lifestyle and mental health) and the personality correlations of these variables correlated at r = .83 (this is the correlation between the off-diagonal elements of the two correlation matrices), suggesting a similar trend for both personality profile correlation and phenotypical correlation.

Discussion

The present study investigated the personality profiles of health, health-related behaviour and well-being of adolescents and emerging adults. For each of the outcomes, we tested the predictive power of three levels of the personality hierarchy (domain, facet, and nuance) and created personality profiles for them based on the most informative trait level, nuances. The network which was created based on these nuance-level personality profiles comprised four groups based on personality similarities: psychological distress, health awareness, emotional control, and substance use.

Predictive power of each personality hierarchy level

Nuances, captured in the variance of individual items contained the most information in describing lifestyle behaviours and health outcomes in terms of psychological traits, compared to facets and domains. Across all different lifestyle and mental health aspects, nuances allowed for 18% more accurate prediction than facets and 55% more accurate prediction than domains. These findings align with those of Mõttus and Rozgnonjuk (2019) who found that nuances allowed for 47% more accurate prediction of age than facets, and 132% more accurate predictions than domains. Mõttus, Bates et al., (2017) also showed that items allowed for an average of 64% better prediction of a range of outcomes than facets. The incremental increase in predictive power of nuances in the present study being smaller compared to some previous studies could be due to these studies predicting objective outcomes (e.g., age and BMI). Here, we predicted scores on self-report scales and the shared method variance (social desirability, other rating biases) might have inflated domain-level associations in particular.

The figure further visualised potential reasons for the lower predictive power of higher-order traits by showing that facets within the same domain and nuances within the facet sometimes show trends in opposite directions. For example, half of the nuance-level traits in the O5: Intellect facet were positively correlated with sleeping and the other half of the nuance-level traits in the same facet were negatively correlated with sleeping. If nuances within the facets showed both positive and negative directions, the predictive abilities of higher order traits (facets and domains) would invariably be diminished (Revelle et al., 2020).

Personality profile associations

The network of profile correlations can be seen as a psychological “atlas” of the outcomes being considered, giving insights into how the lifestyle aspects and mental health outcomes are related to each other in terms of their underlying psychological associations, as well as into some more specific patterns within these associations (Vainik et al., 2020). Generally, a healthy personality profile network can be characterized by adopting healthy dietary habits, exercising regularly, sleeping well, avoiding drugs, using less internet, and having less mental health problems. The identified health network is generally congruent with the pattern that personality profiles characterized by healthy lifestyles were negatively associated with personality profiles characterized by negative lifestyles and mental health problems. This pattern is also in line with findings from previous research on identifying a healthy personality profile (Bleidorn, Hopwood, Ackerman, et al., 2020). The fact that most personality correlations were found to be stronger than phenotypic correlations further illustrates that personality similarities between lifestyle aspects and mental health outcomes might be the more likely to explain the associations among these outcomes. Thus, their phenotypic associations may often be "confounded" by a largely shared psychological background.

In particular, personality profiles of lifestyle and mental health could be separated into four groups based on their stronger similarities in personality traits characterising these profiles. The psychological distress group included five mental health profiles (depression, anxiety, repetitive thoughts, somatic complaint, and suicidal ideation) as well as the sleeping and exercise profile. For these outcomes, their personality correlations tended to exceed their phenotypic correlations, suggesting that these mental health problems might be explained by shared underlying psychological traits. This further points to the possibility that there might be a common aetiology of the comorbidity of mental health problems. This would be in line with previous research that has found that most mental health problems share common risk factors (Caspi et al., 2014). However, the personality correlations between mental health outcomes with sleep or exercise did not exceed the phenotypic correlations between them, thus illustrating that personality traits alone cannot explain these relationships.

The health awareness group only included two profiles which were the healthy diet and internet use. The most predictive items of these two profiles shared similarities in terms of the attitude towards leisure time. However, since the personality correlation between healthy diet and internet use did not exceed their phenotypic correlations, the observed direct link might better explain the association (Coughlin et al., 2015; Pollard et al., 2015).

The emotional control group included three profiles: troubled eating, psychosis and anger, with anger mediating the personality correlation between troubled eating and psychosis. The most predictive items of the troubled eating profile were over-eating and emotional eating, and the most predictive items of the anger and psychosis profiles may be characterised by getting emotional easily. Over-eating was significantly correlated with emotional eating and emotional eaters often overeat in response to both positive and negative emotions (Bongers et al., 2013; Sultson et al., 2017; Vainik, García-García, et al., 2019). Therefore, all three profiles seem to overlap in terms of their vulnerability towards emotions.

The substance use group included the three substance use profiles and the mania profile. Our results showed that the observed substance-mania associations could be due to shared personality correlations and thereby a common psychological background. Using tobacco, alcohol and cannabis share similar personality traits with being manic. Interestingly, compared with other mental health profiles, mania showed stronger correlations with substance use profiles. Substance disorders (SUD) commonly co-occur with Bipolar disorder (BD), which is marked by alternating intervals of manic and depressive episodes, with a prevalence rate of 20%-70% (Gold et al., 2018). The high comorbidities between SUD and BD might be due to the shared mechanisms of impulsivity, poor modulation of motivation and responses to rewarding stimuli, and susceptibility to behavioural sensitization (Swann, 2010). Since most of these studies focused on BD and rarely focused on mania in particular, understanding the mania-substance abuse association through personality traits could prove useful in advancing targeted interventions.

A further notable finding of the present study relates to the network structure of mental health disorders which were found to be highly connected. Using a novel approach of investigating personality correlations between a wide range of mental health disorders, we were able to point to one general underlying factor, that is a combination of specific personality traits, which might reflect an individual's propensity for developing any form of psychopathology (Caspi et al., 2014; Murray et al., 2016). This is also supported by the finding that many mental health problems share symptoms and often respond to the same interventions (Barlow et al., 2010).

Our health network was created for the specific transitioning periods of late adolescence and emerging adulthood. During these transitioning periods, people face unique challenges such as exploring new role identities, forming long-term health related habits, and experiencing physical and psychological maturation. Given these developmentally specific challenges, personality traits become an even stronger explanation of health relations, since personality shares a similar maturation process as lifestyle aspects and mental health problems (Bleidorn, Hopwood, Ackerman, et al., 2020).

Limitations

The present study has three key caveats, with the first one being the nuance-lifestyle/health associations. Although items can be seen as indicators of nuances and display trait-like properties including cross-method agreement, rank-order stability, and heritability (Mõttus, Bates, et al., 2017; Mõttus et al., 2014), no measures to capture personality nuances have been designed yet. Therefore, the present findings only show the potential of the current approach using tools not designed for it: ultimately, more nuanced explorations of how personality traits are linked with lifestyle and health, and possibly explain associations among these outcomes, will require measurement models specifically designed to capture nuances. Such tools are not unrealistic (Condon et al., 2020).

The second caveat is that our health network only focused on the personality correlations of lifestyle and mental health. This health network could benefit from incorporating more risk factors that particularly impact adolescents and emerging adults (e.g. risk-taking behaviours, vicissitudes of social life) to obtain an even more complete view of health and wellbeing during these two developmental stages. However, since the present study established a health network with a novel approach of creating personality profiles of nuance-based lifestyle aspects and mental health outcomes, researchers could easily insert more health-related variables into the network, simply by creating a new personality profile of the variable.

The third caveat is that the present study analysed cross-sectional data and only focused on late adolescence and emerging adulthood. Thus we cannot make any inferences about health networks at different developmental stages. Future studies could examine health networks during different developmental periods and investigate whether changes in the health network would follow maturation theories, be most prominent during adolescence, and more stable during adulthood.

Conclusion

This study provides new psychological insights into health-related lifestyles and mental health issues in late adolescence and emerging adulthood. By using nuance level personality traits, we maximised the predictive power and specificity for health outcomes, and explored multiple health outcomes in terms of overlapping personality profiles. The health network was made up of four groups: psychological distress, health awareness, emotional control and substance use. These were based on the stronger personality similarities characterising lifestyle and mental health profiles. Such a fine-grained personality analysis can identify people at heightened risk for certain health outcomes, and adopting the use of nuance assessment tools could be informative within clinical settings. Finally, we have confirmed that co-morbid mental health problems can be partly explained by a general personality-based psychopathology factor, that makes individuals' more vulnerable to developing any form of psychopathology.

Chapter 6: Discussion

Synopsis:

This chapter summarizes the findings from the previous empirical Chapters 2 to 5 and provides a general discussion. Overall, findings from chapter 2 underlined the importance of nuance-level analysis in the personality field. Personality research has tended to focus on the domain and facet levels because they are easier to interpret due to their categorical structure. However, nuances provide more unique information than domains and facets. Using nuance-level analyses, this thesis examined some new hypotheses, focusing on one of the key questions in the personality field: personality differences over time. Chapters 3 and 4 examined whether social expectations and self-regulation are possible mechanisms to explain personality differences from childhood to adulthood and empirically tested social investment theory. Chapter 5 further expanded on the theoretical findings of the previous chapters and applied nuance-level analyses to real life issues, showing that personality correlations can explain many associations between health-related aspects and health outcomes. Finally, this chapter summarizes the implications as well as the strengths and limitations of the empirical work presented in this thesis.

This thesis aimed to gain a clearer understanding of nuance-level personality traits and the benefits that it can bring to personality research. Further, it aimed to investigate potential mechanisms to explain personality differences across the life span using nuance-level traits, and tested the practical and clinical value of nuance-level traits as a new avenue for studying personality by examining the relation between nuance-level traits and mental health and wellbeing.

Chapter 2 began by examining the value of nuance-level personality traits in four different aspects. In particular, five datasets with samples representing different developmental periods, cultural backgrounds, rating perspectives, and personality instruments were used to compare the amounts of age-related information captured by each level of the personality hierarchy (domains, facets, and nuances). Elastic net models and random sample partitions were used to prevent overfitting and capitalization on chance. It was found that nuances could capture more age-sensitive information than domains and facets. Specifically, nuances contained more developmental information than facets, and facets contained more developmental information than domains. Furthermore, nuances captured unique information that was not shared with the collective higher level traits. Thus, it was shown that aggregating individual items into domains and facets would inevitably lead to a loss of useful information that is only contained in nuance-level traits. Third, personality inventories with a greater number of personality nuances provided more age-sensitive information than shorter personality inventories, suggesting that shorter personality inventories might be generalized and miss important information. Lastly, the amount of age-sensitive information captured by personality traits was highest during childhood and then gradually decreased into adulthood. These findings highlight that there is a clear need for personality research to move beyond the high-dimensional approaches to lower-level traits and especially nuances. While it is not surprising that many predictors would do a better job at predicting another outcome than a few, it is important to clarify that including more items and predictors does not simply benefit prediction in a statistical sense, but indeed offers researchers more diverse methods of studying personality, for instance allowing for network analyses (Mõttus & Allerhand, 2018). Thus, the purpose of using a large number of item-level traits is not simply improving prediction but also understanding personality from a new perspective.

Based on the findings from chapter 2, chapters 3 and 4 tested an innovative method of using nuance-level analysis to benefit current personality research. Given that a key interest for personality science is understanding the mechanisms of personality differences across the life span, chapters 3 and 4 uncovered potential mechanisms to explain personality differences during childhood and young adulthood within a nuance-level analysis framework. Both chapters adopted a unique research design, which was to examine the systematic variability among many traits in various focal properties. This allows personality researchers to study variance between traits, not between people, through a large number of personality traits. This research design has several advantages compared to traditional methods. First, this method requires fewer participants and consequently may benefit studies with a limited budget for recruiting participants. Second, different aspects of personality traits (here the socially expected level and self-regulatory level of traits as well as the mean levels and variance of traits) can be measured in different participants, thus, reducing participant burden as individual participants only need to complete a subset of measures required for the completion of a study.

Specifically, chapter 3 quantified a large amount of diverse personality traits in terms of their mean-level and variance changes at different ages, socially expected levels, and the extent of self-regulatory ability required to meet these expected levels. Findings suggested that only strong social expectations showed slightly increased trait mean levels, which is in line with the literature that suggests that social pressure leads to personality change (Hurlock, 1994; Specht, 2017). Traits with the strongest social expectations showed the most pronounced curvilinear increases in the magnitudes of individual differences prior to mid-adolescence. However, traits' self-regulation abilities were not associated with either mean trait levels or variance. These findings indicated that social pressure has a significant impact on children's personality differences and make them behave in a more socially desirable manner from late-adolescence. Self-regulation abilities required to meet these expectations, on the other hand, are likely not contributing to personality differences in that age group. These findings partially confirmed our hypotheses of chapter 3 that trait mean levels would track social expectations for these traits. Findings from chapter 3 further suggested that traits under the strongest social expectations make children become less alike as they grow up until mid-adolescence as indicated by larger trait variances.

Building on the research from chapter 3, chapter 4 examined the role of social expectation as a mechanism to explain personality differences during young adulthood. Self-regulatory ability required to meet social expectations were not included in this chapter as it was found from the research conducted in chapter 3 that it is unlikely to explain personality change. Social investment theory suggests that personality maturation is driven by changing social expectations which go hand in hand with different social roles (Bleidorn et al., 2013; Roberts, Wood, et al., 2005). There is also evidence suggesting that individual differences increase in environments with low social pressures and decrease in environments with high social pressures (Hurlock, 1994; Santee & Maslach, 1982; Shek & Chan, 1999; Specht, 2017). Chapter 4 provided empirical evidence to support SIT. Like chapter 3, chapter 4 also quantified a large amount of diverse personality traits in terms of their mean-level and variance changes during young adulthood as well as their socially expected levels. Findings from chapter 4 only partially supported SIT since, while social expectations were associated with changes in traits mean levels, the interaction between social expectations and age on traits mean levels was not significant. This, however, still suggested that social expectations could be a potential mechanism to explain personality changes in young adulthood: personality traits under stronger social expectations had higher mean levels than traits under lower expectations and traits with stronger social expectations had smaller variances. Hence, individuals behave in a more socially desirable way under stronger expectations and, at the same time, become more similar to each other. Findings from both chapter 3 and 4 clearly indicated that nuance-level research is a powerful tool to study personality differences in terms of obtaining the average trend of changes in traits in detail. It is impossible to study the systematic variability of traits based on the domain or facet level traits, simply because there are not enough of them. Therefore, it is essential that personality researchers work on developing nuance-level research in order to better understand personality related topics.

In addition to evaluating the theoretical value of nuance-level traits, chapter 5 intended to demonstrate the practical value of nuance-level research by creating a network between the personality profiles of health-related lifestyles and outcomes. Understanding the associations between personality and other outcomes is a key topic in the field as it may pave the way for clinical interventions. To this end, the study in chapter 5 created personality profiles of health-related lifestyles and outcomes by correlating individual nuances (items) with all outcomes. These personality profiles were compared by obtaining personality correlations between the profiles. Results showed that more than half of the correlations between the observed lifestyle and mental health outcomes were weaker than their personality correlations. This result indicated that for more than half of the outcomes, their associations were more likely due to a shared psychological background between them, rather than a direct causal relationship.

Findings from chapter 5 therefore suggested that the co-occurrence of many mental health problems might be due to a shared personality-based psychopathology factor that make people more vulnerable to develop any form of psychopathology. Chapter 5 illustrated the potential utility of using nuance-level analyses in more applied areas, here psychopathology research. This line of research is in its beginning stages with much further in-depth research required before interpreting results for clinical application. In the future, such findings may potentially allow researchers to gain practical insights into health and wellbeing and underscore the value of advancing nuance-level research. Overall, the findings of this thesis have illustrated that nuance-level traits are important to study personality, and that researchers should look to move beyond high-dimensional approaches and consider nuance-level analyses.

Implications

There are several important implications that are worth noting. First of all, nuance-level analyses allow researchers to describe various life outcomes in terms of personality traits in detail. Personality items within the same facets and domains sometimes show different or even opposite directions in terms of their associations with other life outcomes (e.g. age, lifestyle, mental health problems). Aggregating these items into broader traits would lead to a significant loss of important information and may consequently attenuate the effects of higher-level traits. Moreover, simply ignoring the conflicting associations of life outcomes and nuance-level traits within the same facets can be seen as an over-generalization of the results. It is essential to include these details to obtain a more accurate and complete picture of personality research. Furthermore, there could also be practical advantages of using nuances, especially for building prediction models for relevant outcomes. For example, knowing whether certain personality traits are associated with an increased risk of developing specific health issues will help with early prevention and may help reduce disease burden.

Additionally, chapters 3 and 4 also provided an example of how to test new hypotheses using nuance-level analyses. Researchers could study systematic variations between traits in terms of how nuances change with time and how they intersect with developmental factors / mechanisms that characterize the nuances to different degrees. This method requires using a large and diverse pool of traits as the effective sample is the number of traits rather than the number of participants. In this case, a few (e.g., five) broad traits would not be enough to test these hypotheses.

Finally, while the substantial amount of nuances can be overwhelming, they have been shown to have a higher predictive power, demonstrated in chapters 2 and 5. However, this can only be offered by a large number of nuance-level traits and, as chapter 5 showed, may (down the line) also provide insights with potential for clinical implications. Analogues to polygenic risk scores (PRS), which is the estimation of genetic liability to traits based on genome-wide association studies (GWAS), each individual nuance could serve as a potential disorder marker (Mõttus et al., 2020; Revelle et al., 2020) and personality researchers could create poly-trait risk scores based on 'nuance-wide association studies' by aggregating the very small effects of multiple personality traits into a risk score that may help identify individuals at higher risk of specific health problems. A similar idea has also been used in other studies such as using a questionnaire-wide association study design and a longitudinal experience-wide association study design (Bleidorn, Hopwood, Back, et al., 2020a; Weiss et al., 2013).

Explanations of personality differences

From chapters 3 and 4, social expectations only partially explained personality differences across the life span. So what are other explanations of personality differences? Over the past years, many studies have examined the sources of personality change using various sophisticated methods. However, there is not a convincing explanation yet, since findings are inconsistent (Wagner et al., 2020).

One of the central debates behind the causes of personality change is the degree to which individuals are shaped by nature or nurture (Bleidorn et al., 2014). The Five Factor Theory considers personality traits as “endogenous dispositions that follow intrinsic paths of development essentially independent of environmental influences” (McCrae et al., 2000, p. 173). In this case, traits are strongly influenced by genetic disposition. Life events, shared experiences or culture are not hypothesised to have an influence on traits but personality traits are thought to influence culture (Bleidorn, Hopwood, Back, et al., 2020b; Roberts, Wood, et al., 2005).

However, proponents of the nurture aspect of the debate believe that environmental influences do indeed lead to personality changes (Roberts, Wood, et al., 2005; Roberts & Wood, 2006). A substantial amount of research on how specific environmental factors change personality has been conducted in the past decades (Bleidorn, Hopwood, Back, et al., 2020b; Briley et al., 2018; Hopwood et al., 2011; Kandler et al., 2010). For example, the first romantic experience is associated with increases in emotional stability, self-esteem, and conscientiousness (Lehnart et al., 2010; Luciano & Orth, 2017; Neyer & Asendorpf, 2001; Wagner et al., 2015), and having children is also associated with increases in emotional stability (Jokela et al., 2009). Many of these changes emerge when people engage in a new social role, which supports the social investment theory.

There is also well-established evidence on the interdependence of gene and environmental influences. Indeed, behavioural-genetic studies have provided evidence that both genetic and environmental influences play an important role in shaping personality change (Bleidorn et al., 2014; Bleidorn, Hopwood, Ackerman, et al., 2020; Briley et al., 2018; Hopwood et al., 2011; Kandler et al., 2010). For example, environmentally triggered personality changes could be partially impacted by people's genetic predispositions (Shah et al., 2014). The extent to which people are predisposed to prefer certain environments and further change their personality characteristics could also depend on individuals genetic sensitivity to those influences (Borghuis et al., 2020; Byrd & Manuck, 2014). In this case, self-concept and self-regulatory process could contribute to the interaction of genetic and environmental influences. For example, social expectations may drive people to achieve certain goals but people's ability to regulate their behaviour to meet these expectations might vary (Denissen et al., 2013).

It is difficult to explain personality change solely based on either genetic or environmental influences. Social expectations are a part of the environmental influences and there are more environmental influences that might contribute to personality changes, such as life satisfaction (Specht et al., 2013). Therefore, it is essential to have a more comprehensive model which includes both genetic and environmental influences to fully explain personality change and differences. Nuance-level analysis is a promising method to help achieve such goals. In particular, twin studies may be well suited to disentangle genetic and environmental influences using nuance-level analyses.

Limitations and future directions

The contents of this thesis have important implications for understanding nuance-level traits. However, there are also some key caveats that need to be taken into account when interpreting the findings of the empirical chapters. The first caveat is that, as of yet, there is no theoretical model for selecting nuances. This thesis only operationalized personality nuances as individual questionnaire items. Although individual items are equivalent to nuances in many cases and further display trait-like properties including cross-method agreement, rank-order stability, and heritability (Mõttus et al., 2014; Mõttus, Kandler, et al., 2017), these items were chosen to measure the broad trait domains and facets, rather than to capture unique nuance-level information. Chapter 2, 3 and 5 utilized personality items from different questionnaires as an indicator of nuances but these items may only capture part of the advantages of nuances. Thus, the full potential of nuances based on adequate measures is not yet known.

For example, in chapter 2, we predicted age from personality domains, facets and nuances. There is a vast amount of psychometric personality research ensuring good reliability and validity of domain- and facet-level traits but not nuance-level traits. Nevertheless, chapter 2 used the CCQ to measure personality traits which has been designed to cover a broader space of personality traits than the Big Five. Since the findings of the other four measurements mostly corresponded with the findings of the CCQ, these findings should be fairly reliable. In chapters 3, our analysis was based on items. However, this chapter focused on the average of multiple raters rating the same items, thus rendering concerns regarding the validity of these results unlikely. Comparably, chapter 5 created the nuances and health associations using personality items chosen from EE.PIP-NEO Personality Inventory. If items, which only capture part of the unique nuance information, not designed for measuring nuances still allow for such interesting findings, it seems all the more essential to develop a theoretical model that particularly focuses on nuances. In order to achieve such a goal, personality researchers could develop a personality measure for nuances that relied on multiple items measuring the same nuance or a single item which captures the full variance of the represented nuance, thus, ensuring better reliability and making it more likely to detect the true effect. This will help to achieve the full potential of nuances and further advance lower-level personality traits research. Indeed, such measures are already in the process of being developed. Chapter 4 did use an item pool that was specifically designed for measuring nuances, however, this item pool is still fairly unstructured, thus, further methodological work is needed.

Future research could establish the measurements for nuances in a similar way as researchers developed domain and facet-level personality measures (Mõttus & Rozgonjuk, 2019). Condon and colleagues (2020) have introduced a detailed bottom-up method of developing such models. Once there are systematic measurements for nuances, researchers could even analyze the association between individual nuances with other outcomes. For example, in chapters 2 and 5, we created some plots to visualize how items from the same facets display opposite directions with age (or other outcomes). Results from chapter 5, as well as some other literatures, indicated that younger people are more ambitious in their work but that older people are more hard-working (Mõttus et al., 2015; Mõttus & Rozgonjuk, 2019). This result again highlights the importance of studying individual items as these show differential associations with different outcomes, thus, future study could benefit from investigating the association of individual nuances and life outcomes.

Moreover, the current hierarchy of domains, facets and nuances could even be an over-simplification of the structure of personality (Condon et al., 2020). For example, using a bottom-up approach, researchers could easily create sub-facets level by assessing fewer items than the facets level. Personality traits can be conceptualized at multiple levels of abstraction other than domains, facets, and nuances. However, there is currently a lack of understanding of these lower levels of the personality hierarchy, as well as any potentially unidentified levels. These unidentified levels might offer researchers a better understanding of personality and, in the future, researchers could also establish new personality inventories with different levels of abstraction from very broad to very narrow that may help shed light on the associations of personality with life outcomes.

The second caveat is that all data used in this thesis is cross-sectional rather than longitudinal. Chapters 2 to 4 aimed to study age differences in personality change and only longitudinal data allows researchers to observe such changes with certainty. Therefore, this thesis can only infer personality differences during each developmental stage. Future research is needed to re-examine all these findings in longitudinal data to obtain more accurate conclusions.

The third caveat is that this thesis only investigated certain developmental stages. Chapters 3 and 4 focused on understanding why personality differences happened from childhood to early adulthood. Although adolescence and emerging adulthood are crucial developmental period for studying personality differences (Bleidorn, 2015; Hill & Edmonds, 2017), it is still important to understand personality throughout the life span, and therefore to also incorporate the middle-aged and elderly into analyses. Chapter 5 also only focused on late adolescence and emerging adulthood population, thus, it is unable to make inferences about health networks in any other age range. Future study could test if these findings extend to mid-age and old age.

The fourth caveat is that some of the studies had insufficient information to make more generalizable conclusions. In particular, chapter 3 compared youths' self-regulatory levels of personality traits with trait mean levels and variances. However, traits' self-regulatory levels were only based on others-ratings, specifically from college students, and youths' parents and teachers. There was no self-rating for self-regulatory level of personality traits which is one of the vital limitations of this chapter. Chapter 4 aimed to understand the change in role-related social expectations during young adulthood. These role-related social expectations come from ones' close contacts. Chapter 4 included social expectations from intimate partners, friends and bosses/supervisors but is missing parents. The role of parents might not be as influential for adults than for children, but it is still one of the most influential social contacts which should have been included in this study. Chapter 5 created a health network to study the personality correlations of 16 health-related lifestyle and health outcomes. This health network did include many useful risk factors such as mental health issues and lifestyle habits but also left out other risk factors that were not measured in the studied sample such as peer problems. The more risk factors included in the network, the more complete the network of health and wellbeing. Therefore, future study should include more health-related variables into the network to advance our understanding of health and wellbeing in the context of individuals' personality.

The fifth caveat is explaining personality differences from one discrete variable without considering the complexity of life. Chapters 3 and 4 investigated whether social expectations are the mechanism to explain personality change. Like the majority of personality studies, this chapter only focused on the influence of one discrete variable (social expectations) on personality change. However, most life experiences are not isolated and instead link with each other (Bleidorn, Hopwood, Back, et al., 2020b). For example, marriage is a major life outcome that has been shown to have great impact on personality change. However, marriage is closely link to many other life events such as being parents, entering the workforce, gaining financial independence all contributing to change one's social role. It is difficult to disentangle any single variable from the network of life outcome as the sole driver of personality differences because most of them are interconnected with each other. Since this thesis only aims to set up an example of some possible ways of using nuance-level analyses, future studies should test mechanisms of personality differences through a linkage of multiple variables such as Longitudinal Experience-Wide Association Studies designs (Bleidorn, Hopwood, Back, et al., 2020b).

The sixth caveat is that some of the chapters used a self-report mono-method which means that they might suffer from a lack of convergent validity and generalizability. For example, data collected in chapter 5 largely relies on self-report measurements to assess both personality traits and health. While self-report measures are the most common way of studying personality development and personality differences mainly because they are cost-efficient, and sensitive to individual differences in self-concept (Paulhus & Vazire, 2007), an over-reliance on self-report measures could result in several problems. The first general problem is rater-specific response biases. Research participants want to make themselves look good, and thus tend to respond in a socially desirable way (Donaldson & Grant-Vallone, 2002). Second, people tend to compare themselves to a reference group when considering personality items, thus, researchers might not draw a fair estimation over the degree of changes in personality traits (Credé et al., 2010; A. M. Wood et al., 2012), resulting in self-report measures often lacking convergent validity and generalizability. Third, self-report measures have been designed to capture individual differences in personality, but they might not be the ideal method to test within-person changes or personality process (Horstmann & Ziegler, 2020). Fourth, self-report mono-method measures solely rely on people's subjective feelings and can be easily influenced by other factors which might thus reduce the validity of results (Bleidorn, Hopwood, et al., 2021). Self-report and observer ratings are the two commonly used method to assess personality traits, but a large body of literature showed that the self/other agreement are far from unity, with typical correlations of .4 to .6 (McCrae et al., 2004; McCrae, 2018). For example, neuroticism scores tend to be higher in self-reports than in observer ratings (Allik et al., 2010). From another perspective, if viewing traits as intrinsically compound, indicators can only reveal part of the traits (McCrae, 2015). Consequently, self-reports might be able to allow access to a private, intrapsychic part of neuroticism, whereas observer ratings would assess another likely more externalizing part of neuroticism. Thus, neither a measure of neuroticism that is solely based on self-reports nor one that is solely based on observer ratings can provide a complete measure of it.

Although the weaknesses of self-reports have been widely reported, many studies still choose to rely on self-reported measures thereby increasing the likelihood for mono-method bias. Mono-method bias, also known as common method bias, has long been regarded as a danger to construct validity (de Guinea et al., 2013). Mono-method bias refers to estimates of true relationships between constructs being skewed as a result of incorporating systematic variance into measures through a measurement approach (Brannick et al., 2010; Burton-Jones, 2009; Straub & Gefen, 2004). Thus, mono-method bias can lead to both Type I and Type II errors by either inflating or deflating the observed relationships between constructs (Podsakoff et al., 2003; Spector, 2006). Specifically in personality research, personality assessments tends to assess a single, substantive construct such as a Big Five trait and focuses on its reliability (the score of a respondent over different items, test forms and so on) but not on its validity (McCrae, 2018). Therefore, it is crucial to enhance the validity and generalizability of existing self-report mono method findings by methods such as a multi-trait multi-method (MTMM) approach (Bleidorn, Hopwood, et al., 2021). The MTMM matrix is an approach introduced by Campbell and Fiske in 1959 in order to assess construct validity. Although, the MTMM approach has been introduced more than fifty years ago, it has not been widely used (Ortiz de Guinea et al., 2013). In personality research, an MTMM approach could be used to evaluate the convergent and discriminant properties of different personality measurements using a confirmatory factor analysis (CFA) approach (Marsh &

Grayson, 1995; Millsap, 1995; Widaman, 1985). CFA-MTMM studies could examine the degree of trait variance contained in personality ratings from different sources (self, peer, or others, etc.) and thus help provide support for the construct validity of trait models.

There are a number of other reasons to use multi-trait and multi-method approaches. For example, adding more raters to estimate the true score variance increases the proportion of valid variance substantially. Although informant ratings are not perfect, averaging multiple raters would reduce most method biases and much random measurement error. Again suggesting that the collection of data from multiple informants is important (McCrae et al., 2013). Similarly, some personality measurements choose to only use a few item scales which could be a problem. This is because that these scales tend to include more variance which due to item specifics rather than the traits they want to measure (McCrae, 2015), whereas averaging a large number of items could be helpful in such situations. Therefore, future research should use multi-trait multi-method approaches in personality study to improve both reliability and validity. Moreover, multi-trait multi-method research can somewhat better capture the complexity and dynamism of personality in terms of its rich details. The explanatory scope of the underlying trait perspective is limited to describing static individual differences. However, personality is a dynamic system of complex interactions between humans and the environment, that is personality is influenced by how people choose, modify and react to their environment (Mõttus & Allerhand, 2018). The current trait models can only provide efficient static trait descriptions, but lack explanations for how individuals interact with each other and their environments from moment to moment, thus failing to take the comprehensiveness and dynamism of personality in real life into full considerations (Baumert et al., 2017).

The seventh caveat is that it is important to also acknowledge that treating residuals purely as unique information (not captured by domains, or facets) when predicting other outcomes, such as age in Chapter 2, is not necessarily capturing the whole complexity of measuring personality. In principle, residuals are defined as the measured difference between the observed values and their corresponding estimated quantity, thus representing the observable estimate of the unobservable statistical error. Hence, in the context of regressions, if the regression equation could perfectly predict the outcome (i.e. if all factors contributing to the outcome where perfectly measured), residuals would have a value of zero (McCrae, 2015). However, it is not usually to capture all factors that may contribute to a specific outcome. In the context of age, arguably it cannot be expected for personality traits to perfectly predict an individual's age, thus, residuals are not likely to be zero. In this context, it is also important to consider measurement error which is also captured by residuals. Measurement error stems from the fact that any measure is likely limited in their measurement of a latent variable (i.e., construct), for instance because the measurement does not ask the right question to really capture the construct of neuroticism, or because participants randomly choose an answer on a questionnaire as they are distracted by something else. Thus, when making conclusions about the additional unique information captured in residuals, it is important to consider whether this is indeed unique information specific to personality traits, or whether this information may also capture other statistical artifacts such as measurement error. With regards to predicting age, it could for instance be the case that peoples' behaviour when filling in questionnaires changes based on age, thus, the measurement error could to some extent be age specific and consequently also be predictive of age.

Further to measurement error, the more complex the analysis, the higher the risk of other regression artifacts. Regression artifacts refers to pseudo casual inferences that draw from causes other than the real cause (Campbell, 1996). Common regression artefacts among others include under-adjustment bias (Rohrer, 2018), suppression effects (Tzelgov & Henik, 1991) or collider bias

(Campbell & Kenny, 2002; Shrier & Platt, 2008; Williams et al., 2018). Results of our residual analyses may for instance be biased by under-adjustment since, given that the latent variable of neuroticism is not perfectly measured, it is possible that adjusting for the neuroticism domain and facets did not really adequately eliminate all variance related to neuroticism from the residuals. Hence, it is likely that the residuals used in this thesis as predictor of age represent a mix of multiple factors including variance not adequately captured by adjusting for domains and facets and measurement error but also unique information of nuances. Importantly however, residualizing items for domains and facets does remove the common variance that is shared across items. Notably, this common variance is the main focus of classical test theory and latent trait models; thus, according to those approaches the residualizing procedure used in this thesis should remove all of the interesting variance from the items. However, results show that the residualized items still retained considerable information, thus highlighting the value of considering the predictive abilities of residuals in more detail. For the future, it will be important to more accurately tease apart the information captured by nuances. In order to achieve this, a combination of approaches is necessary. For one, multi-item measures of nuances can help to limit the amount of measurement error represented by residual as the unique variance underlying a nuance can be more accurately captured using latent measurement models within a structural equation modelling framework. For another, simulation studies can help to more closely study the structure of residuals as they allow for the flexibility in simulating residuals under a variety of scenarios, hence, allowing for an investigation into which statistical techniques are most suitable to disaggregating any unique information of nuances from other statistical artefacts.

Last but not least, most results in this thesis are primarily based on descriptive comparisons of predictive validity, thus, we were not able to analyse whether predictive abilities of domains, facets or nuances were indeed significantly different. Future studies should investigate other indices that may help shed light on the different benefits of using nuances vs domains or facets. Also, a possible confusion is that although the primary goal of this thesis is to demonstrate the benefits of nuance-level traits, we did not extensively focus on individual items, for instance on which items are the ones doing the most in predicting another outcome. This is because nuance-level research is still at a very early stage, and there is not yet a specific measurement for nuances available. All current measurements show some degree of limitations in terms of measuring nuances, hence results drawn from these imperfect measurements might lack credibility and should not be interpreted too closely. Therefore, we only provide some examples to show the potential of individual items in relations to other major life events and health outcomes without going to deep into explaining such relationships. After creating a specific measurement for nuances, detailed research on individual nuances and the possible usage of them will be a promising next step for future research.

Conclusions

This thesis contributes to showing the value of lower-level personality traits, in particular nuances. Through nuance-level analyses, the empirical chapters of this thesis advance the current understanding of personality differences from childhood to adulthood as well as the fine-grained associations between personality and multiple life outcomes. Nuance-level analyses are a needed supplement to the traditional high-dimensional approaches as they provide unique information on personality and new ways to study it. Taken together, the empirical findings of this thesis further emphasize the importance of establishing a proper measurement tool for nuances.

References

Allemand, M., Zimprich, D., & Hendriks, A. A. J. (2008). Age differences in five personality domains across the life span. *Developmental Psychology*, *44*(3), 758–770. https://doi.org/10.1037/0012-1649.44.3.758

Allik, J., Laidra, K., Realo, A., & Pullmann, H. (2004). Personality development from 12 to 18 years of age: Changes in mean levels and structure of traits. *European Journal of Personality*, *18*(6), 445–462. https://doi.org/10.1002/per.524

Allik, J., Realo, A., Mõttus, R., Borkenau, P., Kuppens, P., & Hřebíčková, M. (2010). How people see others is different from how people see themselves: A replicable pattern across cultures. *Journal of Personality and Social Psychology*, *99*(5), 870–882. https://doi.org/10.1037/a0020963

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

American Psychiatric Association. (2014). *Online assessment measures*.

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, *55*(5), 469–480. https://doi.org/10.1037/0003-066X.55.5.469

Arumäe, K., Briley, D. A., Colodro-Conde, L., Mortensen, E. L., Jang, K. L., Ando, J., Kandler, C., Sørensen, T. I. A., Dagher, A., Mõttus, R., & Vainik, U. (2020). *Two genetic analyses to elucidate causality between body mass index and personality* [Preprint]. NutriXiv. https://doi.org/10.31232/osf.io/q8ehr

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

Ashton, M. C., & Lee, K. (2016). Age trends in HEXACO-PI-R self-reports. *Journal of Research in Personality*, *64*, 102–111. https://doi.org/10.1016/j.jrp.2016.08.008

Ashton, M. C., & Lee, K. (2020). Objections to the HEXACO Model of Personality Structure—And why those Objections Fail. *European Journal of Personality*, *34*(4), 492–510. https://doi.org/10.1002/per.2242

Ashton, M. C., Lee, K., Perugini, M., Szarota, P., de Vries, R. E., Di Blas, L., Boies, K., & De Raad, B. (2004). A Six-Factor Structure of Personality-Descriptive Adjectives: Solutions From Psycholexical Studies in Seven Languages. *Journal of Personality and Social Psychology*, *86*(2), 356–366. https://doi.org/10.1037/0022-3514.86.2.356

Asselmann, E., & Specht, J. (2020). Taking the ups and downs at the rollercoaster of love: Associations between major life events in the domain of romantic relationships and the Big Five personality traits. *Developmental Psychology*, *56*(9), 1803–1816. https://doi.org/10.1037/dev0001047

Barlow, D. H., Farchione, T. J., Fairholme, Ellard, K. K., Boisseau, C. L., Allen, L. B., & May. (2010). *Unified protocol for transdiagnostic treatment of emotional disorders: Therapist guide.* Oxford University Press.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4. *ArXiv:1406.5823 [Stat]*. http://arxiv.org/abs/1406.5823

Baumert, A., Schmitt, M., Perugini, M., Johnson, W., Blum, G., Borkenau, P., Costantini, G., Denissen, J. J. A., Fleeson, W., Grafton, B., Jayawickreme, E., Kurzius, E., MacLeod, C., Miller, L. C., Read, S. J., Roberts, B., Robinson, M. D., Wood, D., & Wrzus, C. (2017). Integrating Personality Structure, Personality Process, and Personality Development. *European Journal of Personality*, *31*(5), 503–528. https://doi.org/10.1002/per.2115

Bener, A., Al-Mahdi, H. S., Ali, A. I., Al-Nufal, M., Vachhani, P. J., & Tewfik, I. (2011). Obesity and low vision as a result of excessive Internet use and television viewing. *International Journal of Food Sciences and Nutrition*, *62*(1), 60–62. https://doi.org/10.3109/09637486.2010.495711

Blakemore, S.-J., & Mills, K. L. (2014). Is Adolescence a Sensitive Period for Sociocultural Processing? *Annual Review of Psychology*, *65*(1), 187–207. https://doi.org/10.1146/annurev-psych-010213-115202

Bleidorn, W. (2015). What Accounts for Personality Maturation in Early Adulthood? *Current Directions in Psychological Science*, *24*(3), 245–252. https://doi.org/10.1177/0963721414568662

Bleidorn, W., Hill, P. L., Back, M. D., Denissen, J. J. A., Hennecke, M., Hopwood, C. J., Jokela, M., Kandler, C., Lucas, R. E., Luhmann, M., Orth, U., Wagner, J., Wrzus, C., Zimmermann, J., & Roberts, B. (2019). The policy relevance of personality traits. *American Psychologist*, *74*(9), 1056–1067. https://doi.org/10.1037/amp0000503

Bleidorn, W., & Hopwood, C. J. (2019). *Stability and change in personality traits over the lifespan* (Vol. 1). Handbook of personality development.

Bleidorn, W., Hopwood, C. J., Ackerman, R. A., Witt, E. A., Kandler, C., Riemann, R., Samuel, D. B., & Donnellan, M. B. (2020). The healthy personality from a basic trait perspective. *Journal of Personality and Social Psychology*, *118*(6), 1207–1225. https://doi.org/10.1037/pspp0000231

Bleidorn, W., Hopwood, C. J., Back, M. D., Denissen, J. J. A., Hennecke, M., Hill, P. L., Jokela, M., Kandler, C., Lucas, R. E., Luhmann, M., Orth, U., Roberts, B. W., Wagner, J., Wrzus, C., & Zimmermann, J. (2021). Personality trait stability and change. *Personality Science*, *2*, e6009. https://doi.org/10.5964/ps.6009

Bleidorn, W., Hopwood, C. J., Back, M. D., Denissen, J. J. A., Hennecke, M., Jokela, M., Kandler, C., Lucas, R. E., Luhmann, M., Orth, U., Roberts, B. W., Wagner, J., Wrzus, C., & Zimmermann, J. (2020a). Longitudinal Experience–Wide Association Studies—A Framework for Studying Personality Change. *European Journal of Personality*, *34*(3), 285–300. https://doi.org/10.1002/per.2247

Bleidorn, W., Hopwood, C. J., Back, M., Denissen, J. J. A., Hennecke, M., Jokela, M., Kandler, C., Lucas, R. E., Luhmann, M., Orth, U., Roberts, B., Wagner, J., Wrzus, C., & Zimmermann, J. (2020b). *Longitudinal Experience-Wide Association Studies (LEWAS) – A Framework for Studying Personality Change* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/krfw9

Bleidorn, W., Hopwood, C. J., & Lucas, R. E. (2018). Life Events and Personality Trait Change: Life Events and Trait Change. *Journal of Personality*, *86*(1), 83–96. https://doi.org/10.1111/jopy.12286

Bleidorn, W., Kandler, C., & Caspi, A. (2014). The Behavioural Genetics of Personality Development in Adulthood—Classic, Contemporary, and Future Trends. *European Journal of Personality*, *28*(3), 244–255. https://doi.org/10.1002/per.1957

Bleidorn, W., Klimstra, T. A., Denissen, J. J. A., Rentfrow, P. J., Potter, J., & Gosling, S. D. (2013). Personality Maturation Around the World: A Cross-Cultural Examination of

Social-Investment Theory. *Psychological Science*, *24*(12), 2530–2540. https://doi.org/10.1177/0956797613498396

Bleidorn, W., & Schwaba, T. (2017). Personality development in emerging adulthood. In *Personality Development Across the Lifespan* (pp. 39–51). Elsevier. https://doi.org/10.1016/B978-0-12-804674-6.00004-1

Bleidorn, W., Schwaba, T., & Hopwood, C. J. (2021). Health Adversity and Value Change. *Social Psychological and Personality Science*, *12*(2), 248–257. https://doi.org/10.1177/1948550620901980

Block, J. (1996). Some jangly remarks on Baumeister and Heatherton. *Psychological Inquiry*, *7*, 28–32.

Block, J. (2010). The Five-Factor Framing of Personality and Beyond: Some Ruminations. *Psychological Inquiry*, *21*(1), 2–25. https://doi.org/10.1080/10478401003596626

Block, J. H., & Block, J. (1980). *The California child Q-set*. Consulting Psychologists Press.

Bongers, P., Jansen, A., Havermans, R., Roefs, A., & Nederkoorn, C. (2013). Happy eating. The underestimated role of overeating in a positive mood. *Appetite*, *67*, 74–80. https://doi.org/10.1016/j.appet.2013.03.017

Borghuis, J., Bleidorn, W., Sijtsma, K., Branje, S., Meeus, W. H. J., & Denissen, J. J. A. (2020). Longitudinal associations between trait neuroticism and negative daily experiences in adolescence. *Journal of Personality and Social Psychology*, *118*(2), 348–363. https://doi.org/10.1037/pspp0000233

Brannick, M. T., Chan, D., Conway, J. M., Lance, C. E., & Spector, P. E. (2010). What Is Method Variance and How Can We Cope With It? A Panel Discussion. *Organizational Research Methods*, *13*(3), 407–420. https://doi.org/10.1177/1094428109360993

Briley, D. A., Livengood, J., & Derringer, J. (2018). Behaviour Genetic Frameworks of Causal Reasoning for Personality Psychology. *European Journal of Personality*, *32*(3), 202–220. https://doi.org/10.1002/per.2153

Buchanan, R. D. (1994). The development of the Minnesota Multiphasic Personality Inventory. *Journal of the History of the Behavioral Sciences*, *30*(2), 148–161. https://doi.org/10.1002/1520-6696(199404)30:2<148::AID-JHBS2300300204>3.0.CO;2-9

Burton-Jones. (2009). Minimizing Method Bias through Programmatic Research. *MIS Quarterly*, *33*(3), 445. https://doi.org/10.2307/20650304

Butcher, J. N., & Williams, C. L. (2009). Personality Assessment with the MMPI-2: Historical Roots, International Adaptations, and Current Challenges. *Applied Psychology: Health and Well-Being*, *1*(1), 105–135. https://doi.org/10.1111/j.1758-0854.2008.01007.x

Byrd, A. L., & Manuck, S. B. (2014). MAOA, Childhood Maltreatment, and Antisocial Behavior: Meta-analysis of a Gene-Environment Interaction. *Biological Psychiatry*, *75*(1), 9–17. https://doi.org/10.1016/j.biopsych.2013.05.004

Call, K. T., Riedel, A. A., Hein, K., McLoyd, V., Petersen, A., & Kipke, M. (2002). Adolescent Health and Well-Being in the Twenty-First Century: A Global Perspective. *Journal of Research on Adolescence*, *12*(1), 69–98. https://doi.org/10.1111/1532-7795.00025

Campbell, D. T., & Kenny, D. A. (2002). *A primer on regression artifacts.* Guilford Press.

Campbell✠, D. T. (1996). Regression artifacts in time-series and longitudinal data. *Evaluation and Program Planning*, *19*(4), 377–389. https://doi.org/10.1016/S0149-7189(96)00025-0

Caspi, A., Block, J., Block, J. H., Klopp, B., Lynam, D., Moffitt, T. E., & Stouthamer-Loeber, M. (1992). A 'common-language' version of the California Child Q-Set for personality assessment. *Psychological Assessment*, *4*(4), 512–523. https://doi.org/10.1037/1040-3590.4.4.512

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., Meier, M. H., Ramrakha, S., Shalev, I., Poulton, R., & Moffitt, T. E. (2014). The p Factor: One General Psychopathology Factor in the Structure of Psychiatric Disorders? *Clinical Psychological Science*, *2*(2), 119–137. https://doi.org/10.1177/2167702613497473

Caspi, A., & Roberts, B. W. (2001). Personality Development Across the Life Course:The Argument for Change and Continuity. *Psychological Inquiry*, *12*(2), 49–66. https://doi.org/10.1207/S15327965PLI1202_01

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality Development: Stability and Change. *Annual Review of Psychology*, *56*(1), 453–484. https://doi.org/10.1146/annurev.psych.55.090902.141913

Castellacci, F., & Tveito, V. (2018). Internet use and well-being: A survey and a theoretical framework. *Research Policy*, *47*(1), 308–325. https://doi.org/10.1016/j.respol.2017.11.007

Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl, S., Anandkumar, A., Ford, C., Volcic, R., De Rosario, H., & De Rosario, M. H. (2018). *Package 'pwr'.* R package version.

Christensen, A. P., Garrido, L. E., & Golino, H. (2020). *Unique variable analysis: A novel approach for detecting redundant variables in multivariate data* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/4kra2

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdle.

Condon, D. M. (2018). *The SAPA Personality Inventory: An empirically-derived, hierarchically-organized self-report personality assessment model* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/sc4p9

Condon, D. M., Wood, D., Mõttus, R., Booth, T., Costantini, G., Greiff, S., Johnson, W., Lukaszewski, A., Murray, A. L., Revelle, W., Wright, A. G. C., Ziegler, M., & Zimmermann, J. (2020). *Bottom Up Construction of a Personality Taxonomy* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/u2n7s

Conner, T. S., Thompson, L. M., Knight, R. L., Flett, J. A. M., Richardson, A. C., & Brookie, K. L. (2017). The Role of Personality Traits in Young Adult Fruit and Vegetable Consumption. *Frontiers in Psychology*, *8*, 119. https://doi.org/10.3389/fpsyg.2017.00119

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

Costa, P. T., & McCrae, R. R. (2010). NEO Inventories professional manual. *Psychological Assessment Resources*.

Costa, P. T., McCrae, R. R., & Löckenhoff, C. E. (2019). Personality Across the Life Span. *Annual Review of Psychology*, *70*, 423–448.

Coughlin, S. S., Whitehead, M., Sheats, J. Q., Mastromonico, J., Hardy, D., & Smith, S. A. (2015). Smartphone Applications for Promoting Healthy Diet and Nutrition: A Literature Review. *Jacobs Journal of Food and Nutrition*, *2*(3), 021.

Credé, M., Bashshur, M., & Niehorster, S. (2010). Reference Group Effects in the Measurement of Personality and Attitudes. *Journal of Personality Assessment*, *92*(5), 390–399. https://doi.org/10.1080/00223891.2010.497393

De Fruyt, F., Bartels, M., Van Leeuwen, K. G., De Clercq, B., Decuyper, M., & Mervielde, I. (2006). Five types of personality continuity in childhood and adolescence. *Journal of Personality and Social Psychology*, *91*(3), 538–552. https://doi.org/10.1037/0022-3514.91.3.538

de Guinea, A. O., Titah, R., & Léger, P.-M. (2013). Measure for Measure: A two study multi-trait multi-method investigation of construct validity in IS research. *Computers in Human Behavior*, *29*(3), 833–844. https://doi.org/10.1016/j.chb.2012.12.009

de Haan, A. D., Deković, M., van den Akker, A. L., Stoltz, S. E. M. J., & Prinzie, P. (2013). Developmental Personality Types From Childhood to Adolescence: Associations With Parenting and Adjustment. *Child Development*, *84*(6), 2015–2030. https://doi.org/10.1111/cdev.12092

Denissen, J. J. A., Luhmann, M., Chung, J. M., & Bleidorn, W. (2019). Transactions between life events and personality traits across the adult lifespan. *Journal of Personality and Social Psychology*, *116*(4), 612–633. https://doi.org/10.1037/pspp0000196

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

DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology*, *91*(6), 1138–1151. https://doi.org/10.1037/0022-3514.91.6.1138

Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review*, *108*(2), 346–369. https://doi.org/10.1037/0033-295X.108.2.346

Donaldson, S. I., & Grant-Vallone, E. J. (2002). Understanding Self-Report Bias in Organizational Behavior Research. *JOURNAL OF BUSINESS AND PSYCHOLOGY*, 16.

Donnellan, M. B., & Lucas, R. E. (2008). Age differences in the big five across the life span: Evidence from two national samples. *Psychology and Aging*, *23*(3), 558–566. https://doi.org/10.1037/a0012897

Edwards, A. L. (1953). The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of Applied Psychology*, *37*(2), 90–93. https://doi.org/10.1037/h0058073

Elleman, L. G., McDougald, S. K., Condon, D. M., & Revelle, W. (2020). *That takes the BISCUIT: A comparative study of predictive accuracy and parsimony of four statistical learning techniques in personality data, with data missingness conditions* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/tuqap

Ellis, B. J., Del Giudice, M., Dishion, T. J., Figueredo, A. J., Gray, P., Griskevicius, V., Hawley, P. H., Jacobs, W. J., James, J., Volk, A. A., & Wilson, D. S. (2012). The evolutionary basis of risky adolescent behavior: Implications for science, policy, and practice. *Developmental Psychology*, *48*(3), 598–623. https://doi.org/10.1037/a0026220

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). **qgraph**: Network Visualizations of Relationships in Psychometric Data. *Journal of Statistical Software*, *48*(4). https://doi.org/10.18637/jss.v048.i04

Epstein, S. (1979). The Stability of Behavior: I. On Predicting Most of the People Much of the Time. *Journal of Personality and Social Psychology*, *37(7)*, 30.

Espie, C. A., Kyle, S. D., Hames, P., Gardani, M., Fleming, L., & Cape, J. (2014). The Sleep Condition Indicator: A clinical screening tool to evaluate insomnia disorder: *BMJ Open*, *4*(3), e004183. https://doi.org/10.1136/bmjopen-2013-004183

Eysenck, H. J., & Eysenck, S. B. G. (1967). On the unitary nature of extraversion. *Acta Psychologica*, *26*, 383–390. https://doi.org/10.1016/0001-6918(67)90034-0

Ferdosi, M., Nikkar Isfahani, B., & Kolahdozan, M. S. (2020). The Relationship between Personality Factors, Social Support, and Regulation with Lifestyle. *Evidence Based Health Policy, Management and Economics*. https://doi.org/10.18502/jebhpme.v4i3.4164

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, *33*(1). https://doi.org/10.18637/jss.v033.i01

Friedrich Götz, Sam Gosling, & Rentfrow, J. (2021). *Small effects: The indispensable foundation for a cumulative psychological science.*

Gold, A. K., Otto, M. W., Deckersbach, T., Sylvia, L. G., Nierenberg, A. A., & Kinrys, G. (2018). Substance use comorbidity in bipolar disorder: A qualitative review of treatment strategies and outcomes: Substance Use Disorders in Bipolar Disorder. *The American Journal on Addictions*, *27*(3), 188–201. https://doi.org/10.1111/ajad.12713

Goldberg, L. R. (1990). *An Alternative 'Description of Personality': The Big-Five Factor Structure*. 14.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, *40*(1), 84–96. https://doi.org/10.1016/j.jrp.2005.08.007

Goldberg, W. A. (1999). *Marital Quality, Parental Personality, and Spousal Agreement About Perceptions and Expectations for Children*. 27.

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLOS ONE*, *12*(6), e0174035. https://doi.org/10.1371/journal.pone.0174035

Golino, H., Moulder, R. G., Shi, D., Christensen, A. P., Garrido, L. E., Nieto, M. D., Nesselroade, J. R., Sadana, R., Thiyagarajan, J. A., & Boker, S. M. (2019). *Entropy Fit Indices: New Fit Measures for Assessing the Structure and Dimensionality of Multiple Latent Variables.* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/mtka2

Graham, E. K., Rutsohn, J. P., Turiano, N. A., Bendayan, R., Batterham, P. J., Gerstorf, D., Katz, M. J., Reynolds, C. A., Sharp, E. S., Yoneda, T. B., Bastarache, E. D., Elleman, L. G., Zelinski, E. M., Johansson, B., Kuh, D., Barnes, L. L., Bennett, D. A., Deeg, D. J. H., Lipton, R. B., … Mroczek, D. K. (2017). Personality predicts mortality risk: An integrative data analysis of 15 international longitudinal studies. *Journal of Research in Personality*, *70*, 174–186. https://doi.org/10.1016/j.jrp.2017.07.005

Greenfield, D. (2011). *The addictive properties of ınternet usage. Internet addiction: A handbook and guide to evaluation and treatment*. 135–153.

Ha, S. E., & Kim, S. (2013). Personality and Subjective Well-Being: Evidence from South Korea. *Social Indicators Research*, *111*(1), 341–359. https://doi.org/10.1007/s11205-012-0009-9

Hampson, S. E., Goldberg, L. R., Vogt, T. M., & Dubanoski, J. P. (2007). Mechanisms by which childhood personality traits influence adult health status: Educational attainment and healthy behaviors. *Health Psychology*, *26*(1), 121–125. https://doi.org/10.1037/0278-6133.26.1.121

Hang, Y., Soto, C. J., Lee, B., & Mõttus, R. (2021). *Social Expectations and Abilities to Meet Them as Possible Mechanisms of Youth Personality Development* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/8yp6x

Hang, Y., Soto, C. J., Speyer, L. G., Haring, L., Lee, B., Ostendorf, F., & Mõttus, R. (2021). *Age Differences in the Big Five Personality Domains, Facets and Nuances: A Replication across the life span* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/bzkwd

Hartshorne, H., May, M. A., & Shuttleworth, F. K. (1930). *Studies in the nature of character, Vol. 3: Studies in the organization of character.* MacMillan Co. https://doi.org/10.1037/13357-000

Hathaway, S. R., & McKinley, J. C. (1940). A multiphasic personality schedule (Min- nesota) III. The measurement of symptomatic depression. *Journal of Psychology*, *14*, 73–84.

Hathaway, S. R., & McKinley, J. C. (1942). *The Minnesota Multiphasic Personality Schedule manual*. MN: University of Minnesota Press.

Hayes, N., & Joseph, S. (2003). Big 5 correlates of three measures of subjective well-being. *Personality and Individual Differences*, *34*(4), 723–727. https://doi.org/10.1016/S0191-8869(02)00057-0

Heilmayr, D., & Friedman, H. S. (2020). Personality and Health Outcomes. In R. H. Paul, L. E. Salminen, J. Heaps, & L. M. Cohen (Eds.), *The Wiley Encyclopedia of Health Psychology* (1st ed., pp. 429–438). Wiley. https://doi.org/10.1002/9781119057840.ch93

Hengartner, M. P., Kawohl, W., Haker, H., Rössler, W., & Ajdacic-Gross, V. (2016). Big Five personality traits may inform public health policy and preventive medicine: Evidence from a cross-sectional and a prospective longitudinal epidemiologic study in a Swiss community. *Journal of Psychosomatic Research*, *84*, 44–51. https://doi.org/10.1016/j.jpsychores.2016.03.012

Hennecke, M., Bleidorn, W., Denissen, J. J. A., & Wood, D. (2014). A Three–Part Framework for Self–Regulated Personality Development across Adulthood. *European Journal of Personality*, *28*(3), 289–299. https://doi.org/10.1002/per.1945

Henry, S., & Mõttus, R. (2020). Traits and Adaptations: A Theoretical Examination and New Empirical Evidence. *European Journal of Personality*, *34*(3), 265–284. https://doi.org/10.1002/per.2248

Hill, P. L., & Edmonds, G. W. (2017). Personality development in adolescence. In *Personality Development Across the Lifespan* (pp. 25–38). Elsevier. https://doi.org/10.1016/B978-0-12-804674-6.00003-X

Hoffmann, A. O. I., & Risse, L. (2020). *Do good things come in pairs? How personality traits help explain individuals' simultaneous pursuit of a healthy lifestyle and financially responsible behavior*. 39.

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Hopwood, C. J., Donnellan, M. B., Blonigen, D. M., Krueger, R. F., McGue, M., Iacono, W. G., & Burt, S. A. (2011). Genetic and environmental influences on personality trait stability and growth during the transition to adulthood: A three-wave longitudinal study. *Journal* *of Personality and Social Psychology*, *100*(3), 545–556. https://doi.org/10.1037/a0022409

Horstmann, K. T., & Ziegler, M. (2020). Assessing Personality States: What to Consider when Constructing Personality State Measures. *European Journal of Personality*, *34*(6), 1037–1059. https://doi.org/10.1002/per.2266

Hudson, N. W., Roberts, B. W., & Lodi-Smith, J. (2012). Personality trait development and social investment in work. *Journal of Research in Personality*, *46*(3), 334–344. https://doi.org/10.1016/j.jrp.2012.03.002

Humeniuk, R., Ali, R., Babor, T. F., Farrell, M., Formigoni, M. L., Jittiwutikarn, J., de Lacerda, R. B., Ling, W., Marsden, J., Monteiro, M., Nhiwatiwa, S., Pal, H., Poznyak, V., & Simon, S. (2008). Validation of the alcohol, smoking and substance involvement screening test (ASSIST). *Addiction*, *103*(6), 1039–1047. https://doi.org/10.1111/j.1360-0443.2007.02114.x

Hunt, J. M. (1965). TRADITIONAL PERSONALITY THEORY IN THE LIGHT OF RECENT EVIDENCE. *American Scientist*, *53(1)*, 80–96.

Hurlock, E. B. (1994). Developmentalpsychology:A Iife-span approach. *New Delhi: Tata McGraw-Hill.*

Israel, A., Lüdtke, O., & Wagner, J. (2019). The longitudinal association between personality and achievement in adolescence: Differential effects across all Big Five traits and four achievement indicators. *Learning and Individual Differences*, *72*, 80–91. https://doi.org/10.1016/j.lindif.2019.03.001

Jackson, J. J., Bogg, T., Walton, K. E., Wood, D., Harms, P. D., Lodi-Smith, J., Edmonds, G. W., & Roberts, B. W. (2009). Not all conscientiousness scales change alike: A multimethod, multisample study of age differences in the facets of conscientiousness. *Journal of Personality and Social Psychology*, *96*(2), 446–459. https://doi.org/10.1037/a0014156

Jang, K. L., McCrae, R. R., Angleitner, A., Riemann, R., & Livesley, W. J. (1998). *Heritability of Facet-Level Traits in a Cross-Cultural Twin Sample: Support for a Hierarchical Model of Personality*. 10.

Jokela, M., Hintsanen, M., Hakulinen, C., Batty, G. D., Nabi, H., Singh-Manoux, A., & Kivimäki, M. (2013). Association of personality with the development and persistence of obesity: A meta-analysis based on individual–participant data. *Obesity Reviews*, *14*(4), 315–323. https://doi.org/10.1111/obr.12007

Jokela, M., Kivimäki, M., Elovainio, M., & Keltikangas-Järvinen, L. (2009). Personality and having children: A two-way relationship. *Journal of Personality and Social Psychology*, *96*(1), 218–230. https://doi.org/10.1037/a0014058

Kandler, C., Bleidorn, W., Riemann, R., Spinath, F. M., Thiel, W., & Angleitner, A. (2010). Sources of cumulative continuity in personality: A longitudinal multiple-rater twin study. *Journal of Personality and Social Psychology*, *98*(6), 995–1008. https://doi.org/10.1037/a0019558

Kandler, C., & Zapko-Willmes, A. (2017). Theoretical perspectives on the interplay of nature and nurture in personality development. In *Personality Development Across the Lifespan* (pp. 101–115). Elsevier. https://doi.org/10.1016/B978-0-12-804674-6.00008-9

Kapure, R. (2017). *A relationship among personality factors (Neuroticism, Openness to experience and Agreeableness) and subjective well being of female post graduate students*. *4*(3), 4.

Kelley, T. L. (1927). *Interpretation of Educational Measurements*. NY: World Book.

Konstabel, K., Aavik, T., & Allik, J. (2006). Social desirability and consensual validity of personality traits. *European Journal of Personality*, *20*(7), 549–566. https://doi.org/10.1002/per.593

Kööts-Ausmees, L., Kandler, C., Realo, A., Allik, J., Borkenau, P., Hrebickova, M., & Mõttus, R. (2020). *Age Differences in Personality Traits and Social Desirability: A Multi-Rater Multi-Sample Study* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/bmv9r

Lamers, S. M. A., Westerhof, G. J., Kovács, V., & Bohlmeijer, E. T. (2012). Differential relationships in the association of the Big Five personality traits with positive mental health and psychopathology. *Journal of Research in Personality*, *46*(5), 517–524. https://doi.org/10.1016/j.jrp.2012.05.012

Lee, K., & Ashton, M. C. (2004). Psychometric Properties of the HEXACO Personality Inventory. *Multivariate Behavioral Research*, *39*(2), 329–358. https://doi.org/10.1207/s15327906mbr3902_8

Lee, K., & Ashton, M. C. (2020). Sex differences in HEXACO personality characteristics across countries and ethnicities. *Journal of Personality*, jopy.12551. https://doi.org/10.1111/jopy.12551

Lehnart, J., Neyer, F. J., & Eccles, J. (2010). Long-Term Effects of Social Investment: The Case of Partnering in Young Adulthood. *Journal of Personality*, *78*(2), 639–670. https://doi.org/10.1111/j.1467-6494.2010.00629.x

Lodi-Smith, J., & Roberts, B. W. (2007). Social Investment and Personality: A Meta-Analysis of the Relationship of Personality Traits to Investment in Work, Family, Religion, and

Volunteerism. *Personality and Social Psychology Review*, *11*(1), 68–86. https://doi.org/10.1177/1088868306294590

Lucas, R. E., & Donnellan, M. B. (2009). Age differences in personality: Evidence from a nationally representative Australian sample. *Developmental Psychology*, *45*(5), 1353–1363. https://doi.org/10.1037/a0013914

Luciano, E. C., & Orth, U. (2017). Transitions in romantic relationships and development of self-esteem. *Journal of Personality and Social Psychology*, *112*(2), 307–328. https://doi.org/10.1037/pspp0000109

Markon, K. E. (2009). Hierarchies in the Structure of Personality Traits: Hierarchies in Structure of Personality Traits. *Social and Personality Psychology Compass*, *3*(5), 812–826. https://doi.org/10.1111/j.1751-9004.2009.00213.x

Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the Structure of Normal and Abnormal Personality: An Integrative Hierarchical Approach. *Journal of Personality and Social Psychology*, *88*(1), 139–157. https://doi.org/10.1037/0022-3514.88.1.139

Marsh, H. W., & Grayson, D. (1995). *Latent variable models of multitrait-multimethod data. In R. H. Hoyle (Ed.), Structural equation modeling: Concepts, issues, and applications*. CA: SagE.

Mazza, C., Ricci, E., Biondi, S., Colasanti, M., Ferracuti, S., Napoli, C., & Roma, P. (2020). A Nationwide Survey of Psychological Distress among Italian People during the COVID-19 Pandemic: Immediate Psychological Responses and Associated Factors. *International Journal of Environmental Research and Public Health*, *17*(9), 3165. https://doi.org/10.3390/ijerph17093165

McCrae, R. R. (2015). A More Nuanced View of Reliability: Specificity in the Trait Hierarchy. *Personality and Social Psychology Review*, *19*(2), 97–112. https://doi.org/10.1177/1088868314541857

McCrae, R. R. (2018). Method biases in single-source personality assessments. *Psychological Assessment*, *30*(9), 1160–1173. https://doi.org/10.1037/pas0000566

McCrae, R. R., & Costa, P. T. (1997). Personality Trait Structure as a Human Universal. *American Psychologist*, 8.

McCrae, R. R., & Costa, P. T. (2008). *The Five--Factor Theory of Personality*. 23.

McCrae, R. R., Costa, P. T., de Lima, M. P., Simões, A., Ostendorf, F., Angleitner, A., Marušić, I., Bratko, D., Caprara, G. V., Barbaranelli, C., Chae, J.-H., & Piedmont, R. L. (1999). Age differences in personality across the adult life span: Parallels in five cultures. *Developmental Psychology*, *35*(2), 466–477. https://doi.org/10.1037/0012-1649.35.2.466

McCrae, R. R., Costa, P. T., Martin, T. A., Oryol, V. E., Rukavishnikov, A. A., Senin, I. G., Hřebíčková, M., & Urbánek, T. (2004). Consensual validation of personality traits across cultures. *Journal of Research in Personality*, *38*(2), 179–201. https://doi.org/10.1016/S0092-6566(03)00056-4

McCrae, R. R., Costa, P. T., Ostendorf, F., Angleitner, A., Hřebíčková, M., Avia, M. D., Sanz, J., Sánchez-Bernardos, M. L., Kusdil, M. E., Woodfield, R., Saunders, P. R., & Smith, P. B. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology*, *78*(1), 173–186. https://doi.org/10.1037/0022-3514.78.1.173

McCrae, R. R., De Bolle, M., Löckenhoff, C. E., & Terracciano, A. (2021). Lifespan trait development: Toward an adequate theory of personality. In *The Handbook of Personality*

Dynamics and Processes (pp. 621–641). Elsevier. https://doi.org/10.1016/B978-0-12-813995-0.00023-6

McCrae, R. R., Gaines, J. F., & Wellington, M. A. (2013). *The Five-Factor Model in fact and fiction.*

McCrae, R. R., & John, O. P. (1992). An Introduction to the Five-Factor Model and Its Applications. *Journal of Personality*, *60*(2), 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

McCrae, R. R., & Mõttus, R. (2019). What Personality Scales Measure: A New Psychometrics and Its Implications for Theory and Assessment. *Current Directions in Psychological Science*, *28*(4), 415–420. https://doi.org/10.1177/0963721419849559

McKinley, J. C., Hathaway, S. R., & Meehl, P. E. (1948). The Minnesota Multiphasic Personality Inventory: VI. The K Scale. *J Consul Psychol*, *12*, 20–31.

McNeely, J., Strauss, S. M., Wright, S., Rotrosen, J., Khan, R., Lee, J. D., & Gourevitch, M. N. (2014). Test–retest reliability of a self-administered Alcohol, Smoking and Substance Involvement Screening Test (ASSIST) in primary care patients. *Journal of Substance Abuse Treatment*, *47*(1), 93–101. https://doi.org/10.1016/j.jsat.2014.01.007

Meehl, P. E., & Hathaway, S. R. (1946). The K factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. *Journal of Applied Psychology*, *30(5)*, 525–564.

Millsap, R. E. (1995). *The statistical analysis of method effects in multitrait-multimethod data: A review. In P. E. Shrout & S. T. Fiske (Eds.), Personality method, methods, and theory: A festschrift honoring Donald W. Fiske*. NJ: Erlbaum.

Mischel, W. (1968). *Personality and assessment*. New York: Wiley.

Mischel, W. (2009). From Personality and Assessment (1968) to Personality Science, 2009. *Journal of Research in Personality*, *43*(2), 282–290. https://doi.org/10.1016/j.jrp.2008.12.037

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., Poulton, R., Roberts, B. W., Ross, S., Sears, M. R., Thomson, W. M., & Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, *108*(7), 2693–2698. https://doi.org/10.1073/pnas.1010076108

Moffitt, T. E., & Caspi, A. (2001). Childhood predictors differentiate life-course persistent and adolescence-limited antisocialpathways among males and females. *Development and Psychopathology*, *13*(2), 355–375. https://doi.org/10.1017/S0954579401002097

Moffitt, T. E., Caspi, A., Dickson, N., Silva, P., & Stanton, W. (1996). Childhood-onset versus adolescent-onset antisocial conduct problems in males: Natural history from ages 3 to 18 years. *Development and Psychopathology*, *8*(2), 399–424. https://doi.org/10.1017/S0954579400007161

Moreira, P. A. S., Cloninger, C. R., Dinis, L., Sã¡, L., Oliveira, J. T., Dias, A., & Oliveira, J. (2015). Personality and well-being in adolescents. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01494

Morey, L. C., Hopwood, C. J., Gunderson, J. G., Skodol, A. E., Shea, M. T., Yen, S., Stout, R. L., Zanarini, M. C., Grilo, C. M., Sanislow, C. A., & McGLASHAN, T. H. (2007). Comparison of alternative models for personality disorders. *Psychological Medicine*, *37*(7), 983–994. https://doi.org/10.1017/S0033291706009482

Morgan, S. L. (2007). Expectations and Aspirations. In G. Ritzer (Ed.), *The Blackwell Encyclopedia of Sociology* (p. wbeose085). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781405165518.wbeose085

Mõttus, R. (2006). Towards More Rigorous Personality Trait-Outcome Research: Personality traits and outcomes. *European Journal of Personality*, *30*(4), 292–303. https://doi.org/10.1002/per.2041

Mõttus, R., & Allerhand, M. (2018). *Why Do Traits Come Together? The Underlying Trait and Network Approaches*. SAGE Publications Ltd. https://doi.org/10.4135/9781526451163

Mõttus, R., Allik, J., Hřebíčková, M., Kööts–Ausmees, L., & Realo, A. (2016). Age Differences in the Variance of Personality Characteristics. *European Journal of Personality*, *30*(1), 4–11. https://doi.org/10.1002/per.2036

Mõttus, R., Allik, J., & Realo, A. (2019). Do self-reports and informant-ratings measure the same personality constructs? *European Journal of Psychological Assessment*.

Mõttus, R., Bates, T. C., Condon, D. M., Mroczek, D., & Revelle, W. (2017). *Leveraging a more nuanced view of personality: Narrow characteristics predict and explain variance in life outcomes* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/4q9gv

Mõttus, R., Briley, D. A., Zheng, A., Mann, F. D., Engelhardt, L. E., Tackett, J. L., Harden, K. P., & Tucker-Drob, E. M. (2019). Kids becoming less alike: A behavioral genetic analysis of developmental increases in personality variance from childhood to adolescence. *Journal of Personality and Social Psychology*, *117*(3), 635–658. https://doi.org/10.1037/pspp0000194

Mõttus, R., Kandler, C., Bleidorn, W., Riemann, R., & McCrae, R. R. (2017). Personality traits below facets: The consensual validity, longitudinal stability, heritability, and utility of personality nuances. *Journal of Personality and Social Psychology*, *112*(3), 474–490. https://doi.org/10.1037/pspp0000100

Mõttus, R., McCrae, R. R., Allik, J., & Realo, A. (2014). Cross-rater agreement on common and specific variance of personality scales and items. *Journal of Research in Personality*, *52*, 47–54. https://doi.org/10.1016/j.jrp.2014.07.005

Mõttus, R., Pullmann, H., & Allik, J. (2006). Toward More Readable Big Five Personality Inventories. *European Journal of Psychological Assessment*, *22*(3), 149–157. https://doi.org/10.1027/1015-5759.22.3.149

Mõttus, R., Realo, A., Allik, J., Esko, T., Metspalu, A., & Johnson, W. (2015). Within-Trait Heterogeneity in Age Group Differences in Personality Domains and Facets: Implications for the Development and Coherence of Personality Traits. *PLOS ONE*, *10*(3), e0119667. https://doi.org/10.1371/journal.pone.0119667

Mõttus, R., & Rozgonjuk, D. (2019). *Development is in the details: Age differences in the Big Five domains, facets and nuances*. 22.

Mõttus, R., Sinick, J., Terracciano, A., Hřebíčková, M., Kandler, C., Ando, J., Mortensen, E. L., Colodro-Conde, L., & Jang, K. L. (2019). Personality characteristics below facets: A replication and meta-analysis of cross-rater agreement, rank-order stability, heritability, and utility of personality nuances. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspp0000202

Mõttus, R., Soto, C. J., & Slobodskaya, H. R. (2017). Are All Kids Alike? The Magnitude of Individual Differences in Personality Characteristics Tends to Increase from Early Childhood to Early Adolescence: Personality trait variance in children. *European Journal of Personality*, *31*(4), 313–328. https://doi.org/10.1002/per.2107

Mõttus, R., Wood, D., Condon, D. M., Back, M., Baumert, A., Costantini, G., Epskamp, S., Greiff, S., Johnson, W., Lukaszewski, A., Murray, A. L., Revelle, W., Wright, A. G. C., Yarkoni, T., Ziegler, M., & Zimmermann, J. (2020). *Descriptive, predictive and explanatory personality research: Different goals, different approaches, but a shared need to move beyond the Big Few traits* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/hvk5p

Murray, A. L., Eisner, M., & Ribeaud, D. (2016). The Development of the General Factor of Psychopathology 'p Factor' Through Childhood and Adolescence. *Journal of Abnormal Child Psychology*, *44*(8), 1573–1586. https://doi.org/10.1007/s10802-016-0132-1

Musek, J. (2007). A general factor of personality: Evidence for the Big One in the five-factor model. *Journal of Research in Personality*, *41*(6), 1213–1233. https://doi.org/10.1016/j.jrp.2007.02.003

Nelson, L. J., & Barry, C. M. (2005). Distinguishing Features of Emerging Adulthood: The Role of Self-Classification as an Adult. *Journal of Adolescent Research*, *20*(2), 242–262. https://doi.org/10.1177/0743558404273074

Nelson, M. C., Story, M., Larson, N. I., Neumark-Sztainer, D., & Lytle, L. A. (2008). Emerging Adulthood and College-aged Youth: An Overlooked Age for Weight-related Behavior Change. *Obesity*, *16*(10), 2205–2211. https://doi.org/10.1038/oby.2008.365

Neyer, F. J., & Asendorpf, J. B. (2001). Personality–relationship transaction in young adulthood. *Journal of Personality and Social Psychology*, *81*(6), 1190–1204. https://doi.org/10.1037/0022-3514.81.6.1190

Olaru, G., Schroeders, U., Wilhelm, O., & Ostendorf, F. (2018). A confirmatory examination of age-associated personality differences: Deriving age-related measurement-invariant solutions using ant colony optimization. *Journal of Personality*, *86*(6), 1037–1049. https://doi.org/10.1111/jopy.12373

Olino, T. M., & Klein, D. N. (2015). Psychometric Comparison of Self- and Informant-Reports of Personality. *Assessment*, *22*(6), 655–664. https://doi.org/10.1177/1073191114567942

Ones, D. S., & Viswesvaran, C. (1996). Bandwidth–fidelity dilemma in personality measurement for personnel selection. *Journal of Organizational Behavior*, *17*(6), 609–626. https://doi.org/10.1002/(SICI)1099-1379(199611)17:6<609::AID-JOB1828>3.0.CO;2-K

Onifade, P. O. (2014). Psychometric Properties of Alcohol Smoking and Substance Involvement Screening Test (Assist V3.0) Among University Students. *Journal of Addictive Behaviors Therapy & Rehabilitation*, *03*(03). https://doi.org/10.4172/2324-9005.1000126

Ortiz de Guinea, A., Titah, R., & Léger, P.-M. (2013). Measure for Measure: A two study multi-trait multi-method investigation of construct validity in IS research. *Computers in Human Behavior*, *29*(3), 833–844. https://doi.org/10.1016/j.chb.2012.12.009

Ostendorf, F., & Angleitner., A. (2004). *Neo-Persönlichkeitsinventar nach Costa und McCrae: Neo-PI-R.* Göttingen: Hogrefe.

Paulhus, D. L., & Vazire, S. (2007). *The self-report method.* (Vol. 1). Handbook of research methods in personality psychology.

Paunonen, S. V. (1998). Hierarchical organization of personality and prediction of behavior. *Journal of Personality and Social Psychology*, *74*(2), 538–556. https://doi.org/10.1037/0022-3514.74.2.538

Paunonen, S. V., & Ashton, M. C. (2001). Big Five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology*, *81*(3), 524–539. https://doi.org/10.1037/0022-3514.81.3.524

Pei, R., Kranzler, E. C., Suleiman, A. B., & Falk, E. B. (2019). Promoting adolescent health: Insights from developmental and communication neuroscience. *Behavioural Public Policy*, *3*(1), 47–71. https://doi.org/10.1017/bpp.2018.30

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Pollard, C. M., Pulker, C. E., Meng, X., Kerr, D. A., & Scott, J. A. (2015). Who Uses the Internet as a Source of Nutrition and Dietary Information? An Australian Population Perspective. *Journal of Medical Internet Research*, *17*(8), e209. https://doi.org/10.2196/jmir.4548

Pringle, J., Mills, K., McAteer, J., Jepson, R., Hogg, E., Anand, N., & Blakemore, S.-J. (2016). A systematic review of adolescent physiological development and its relationship with health-related behaviour: A protocol. *Systematic Reviews*, *5*(1), 3. https://doi.org/10.1186/s13643-015-0173-5

Pusch, S., Mund, M., Hagemeyer, B., & Finn, C. (2019). Personality Development in Emerging and Young Adulthood: A Study of Age Differences. *European Journal of Personality*, *33*(3), 245–263. https://doi.org/10.1002/per.2181

Raag, T., & Rackliff, C. L. (1998). Preschoolers' Awareness of Social Expectations of Gender: Relationships to Toy Choices. *Sex Roles*, *38*(9/10), 685–700. https://doi.org/10.1023/A:1018890728636

Revelle, W. (2019). *An introduction to the psych package: Part II Scale construction and psychometrics*. 119.

Revelle, W., Dworak, E. M., & Condon, D. M. (2020). Exploring the persome: The power of the item in understanding personality structure. *Personality and Individual Differences*, 109905. https://doi.org/10.1016/j.paid.2020.109905

Reynolds, S. K., & Clark, L. Anna. (2001). Predicting Dimensions of Personality Disorder From Domains and Facets of the Five Factor Model. *Journal of Personality*, *69*(2), 199–222. https://doi.org/10.1111/1467-6494.00142

Roberts, B. W. (2018). A Revised Sociogenomic Model of Personality Traits: Revised Sociogenomic Model of Personality. *Journal of Personality*, *86*(1), 23–35. https://doi.org/10.1111/jopy.12323

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., Walton, K., Bogg, T., & Caspi, A. (2006). De-investment in work and non-normative personality trait change in young adulthood. *European Journal of Personality*, *20*(6), 461–474. https://doi.org/10.1002/per.607

Roberts, B. W., Walton, K. E., & Bogg, T. (2005). Conscientiousness and Health across the Life Course. *Review of General Psychology*, *9*(2), 156–168. https://doi.org/10.1037/1089-2680.9.2.156

Roberts, B. W., & Wood, D. (2006). Personality Development in the Context of the Neo-Socioanalytic Model of Personality. In *Handbook of personality development.* (pp. 11–39). Lawrence Erlbaum Associates Publishers.

Roberts, B. W., Wood, D., & Caspi, A. (2008). The development of personality traits in adulthood. In *Handbook of personality: Theory and research, 3rd ed.* (pp. 375–398). The Guilford Press.

Roberts, B. W., Wood, D., & Smith, J. L. (2005). Evaluating Five Factor Theory and social investment perspectives on personality trait development. *Journal of Research in Personality*, *39*(1), 166–184. https://doi.org/10.1016/j.jrp.2004.08.002

Roese, N. J., & Sherman, J. W. (2007). *Expectancy* (In A. W. Kruglanski&E. T. Higgins). The Guilford Press.

Rohrer, J. M. (2018). Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data. *Advances in Methods and Practices in Psychological Science*, *1*(1), 27–42. https://doi.org/10.1177/2515245917745629

Rohrer, J. M., Egloff, B., Kosinski, M., Stillwell, D., & Schmukle, S. C. (2018). In your eyes only? Discrepancies and agreement between self- and other-reports of personality from age 14 to 29. *Journal of Personality and Social Psychology*, *115*(2), 304–320. https://doi.org/10.1037/pspp0000142

Rowe, D. C., & Rodgers, J. L. (2002). Expanding variance and the case of historical changes in IQ means: A critique of Dickens and Flynn (2001). *Psychological Review*, *109*(4), 759–763. https://doi.org/10.1037/0033-295X.109.4.759

Rushton, J. P., Bons, T. A., & Hur, Y.-M. (2008). The genetics and evolution of the general factor of personality. *Journal of Research in Personality*, *42*(5), 1173–1185. https://doi.org/10.1016/j.jrp.2008.03.002

Samuel, D., & Widiger, T. (2008). A meta-analytic review of the relationships between the five-factor model and DSM-IV-TR personality disorders: A facet level analysis☆. *Clinical Psychology Review*, *28*(8), 1326–1342. https://doi.org/10.1016/j.cpr.2008.07.002

Santee, R. T., & Maslach, C. (1982). To agree or not to agree: Personal dissent amid social pressure to conform. *Journal of Personality and Social Psychology*, *42*(4), 690–700. https://doi.org/10.1037/0022-3514.42.4.690

Schwaba, T., & Bleidorn, W. (2018). Individual differences in personality change across the adult life span. *Journal of Personality*, *86*(3), 450–464. https://doi.org/10.1111/jopy.12327

Schwartz, S. J., Zamboanga, B. L., Luyckx, K., Meca, A., & Ritchie, R. A. (2013). Identity in Emerging Adulthood: Reviewing the Field and Looking Forward. *Emerging Adulthood*, *1*(2), 96–113. https://doi.org/10.1177/2167696813479781

Seeboth, A., & Mõttus, R. (2018). Successful Explanations Start with Accurate Descriptions: Questionnaire Items as Personality Markers for More Accurate Predictions: Items as personality markers. *European Journal of Personality*, *32*(3), 186–201. https://doi.org/10.1002/per.2147

Shah, S., McRae, A. F., Marioni, R. E., Harris, S. E., Gibson, J., Henders, A. K., Redmond, P., Cox, S. R., Pattie, A., Corley, J., Murphy, L., Martin, N. G., Montgomery, G. W., Starr, J. M., Wray, N. R., Deary, I. J., & Visscher, P. M. (2014). Genetic and environmental exposures constrain epigenetic drift over the human life course. *Genome Research*, *24*(11), 1725–1733. https://doi.org/10.1101/gr.176933.114

Shek, D. T. L., & Chan, L. K. (1999). Hong Kong Chinese Parents’ Perceptions of the Ideal Child. *The Journal of Psychology*, *133*(3), 291–302. https://doi.org/10.1080/00223989909599742

Shiner, R., & DeYoung, C. G. (2013). *Running Head: PERSONALITY STRUCTURE*. 58.

Shrier, I., & Platt, R. W. (2008). Reducing bias through directed acyclic graphs. *BMC Medical Research Methodology*, *8*(1), 70. https://doi.org/10.1186/1471-2288-8-70

Shweder, R. A. (1975). How relevant is an individual difference theory of personality?1. *Journal of Personality*, *43*(3), 455–484. https://doi.org/10.1111/j.1467-6494.1975.tb00716.x

Soto, C. J. (2016). The Little Six Personality Dimensions From Early Childhood to Early Adulthood: Mean-Level Age and Gender Differences in Parents’ Reports: Mean-Level Development of the Little Six. *Journal of Personality*, *84*(4), 409–422. https://doi.org/10.1111/jopy.12168

Soto, C. J. (2019). How Replicable Are Links Between Personality Traits and Consequential Life Outcomes? The Life Outcomes of Personality Replication Project. *Psychological Science*, *30*(5), 711–727. https://doi.org/10.1177/0956797619831612

Soto, C. J., & John, O. P. (2012). Development of Big Five Domains and Facets in Adulthood: Mean-Level Age Trends and Broadly Versus Narrowly Acting Mechanisms: Big Five Domains and Facets in Adulthood. *Journal of Personality*, *80*(4), 881–914. https://doi.org/10.1111/j.1467-6494.2011.00752.x

Soto, C. J., & John, O. P. (2014). Traits in Transition: The Structure of Parent-Reported Personality Traits from Early Childhood to Early Adulthood: Structure of Youths’ Personality Traits. *Journal of Personality*, *82*(3), 182–199. https://doi.org/10.1111/jopy.12044

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*(1), 117–143. https://doi.org/10.1037/pspp0000096

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

Specht, J. (2017). Personality development research. In *Personality Development Across the Lifespan* (pp. 3–6). Elsevier. https://doi.org/10.1016/B978-0-12-804674-6.00001-6

Specht, J., Bleidorn, W., Denissen, J. J. A., Hennecke, M., Hutteman, R., Kandler, C., Luhmann, M., Orth, U., Reitz, A. K., & Zimmermann, J. (2014). What Drives Adult Personality Development? A Comparison of Theoretical Perspectives and Empirical Evidence: Personality development in adulthood. *European Journal of Personality*, *28*(3), 216–230. https://doi.org/10.1002/per.1966

Specht, J., Egloff, B., & Schmukle, S. C. (2013). Examining Mechanisms of Personality Maturation: The Impact of Life Satisfaction on the Development of the Big Five Personality Traits. *Social Psychological and Personality Science*, *4*(2), 181–189. https://doi.org/10.1177/1948550612448197

Spector, P. E. (2006). Method Variance in Organizational Research: Truth or Urban Legend? *Organizational Research Methods*, *9*(2), 221–232. https://doi.org/10.1177/1094428105284955

Straub, D., & Gefen, D. (2004). Validation Guidelines for IS Positivist Research. *Communications of the Association for Information Systems*, *13*. https://doi.org/10.17705/1CAIS.01324

Strickhouser, J. E., Zell, E., & Krizan, Z. (2017). Does personality predict health and well-being? A metasynthesis. *Health Psychology*, *36*(8), 797–810. https://doi.org/10.1037/hea0000475

Sultson, H., Kukk, K., & Akkermann, K. (2017). Positive and negative emotional eating have different associations with overeating and binge eating: Construction and validation of the Positive-Negative Emotional Eating Scale. *Appetite*, *116*, 423–430. https://doi.org/10.1016/j.appet.2017.05.035

Swann, A. C. (2010). The strong relationship between bipolar and substance-use disorder: Mechanisms and treatment implications. *Annals of the New York Academy of Sciences*, *1187*(1), 276–293. https://doi.org/10.1111/j.1749-6632.2009.05146.x

Szumski, G., & Karwowski, M. (2019). Exploring the Pygmalion effect: The role of teacher expectations, academic self-concept, and class context in students' math achievement. *Contemporary Educational Psychology*, *59*, 101787. https://doi.org/10.1016/j.cedpsych.2019.101787

Terracciano, A., McCrae, R. R., Brant, L. J., & Costa, P. T. (2005). Hierarchical linear modeling analyses of the NEO-PI-R Scales in the Baltimore Longitudinal Study of Aging. *Psychology and Aging*, *20*(3), 493–506. https://doi.org/10.1037/0882-7974.20.3.493

Terracciano, A., McCrae, R. R., & Costa, P. T. (2006). Longitudinal Trajectories in Guilford-Zimmerman Temperament Survey Data: Results From the Baltimore Longitudinal Study of Aging. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *61*(2), P108–P116. https://doi.org/10.1093/geronb/61.2.P108

Thorndike, E. L. (1904). *An Introduction to the Theory of Mental and Social Measurements*. Columbia University.

Tsiplakides, I., & Keramida, A. (2010). The Relationship between Teacher Expectations and Student Achievement in the Teaching of English as A Foreign Language. *English Language Teaching*, *3*(2), p22. https://doi.org/10.5539/elt.v3n2p22

Tzelgov, J., & Henik, A. (1991). Suppression Situations in Psychological Research: Definitions, Implications, and Applications. *Psychological Bulletin*, *109*(3), 13.

Vainik, U., Dagher, A., Realo, A., Colodro-Conde, L., Mortensen, E. L., Jang, K., Juko, A., Kandler, C., Sørensen, T. I. A., & Mõttus, R. (2019). Personality-obesity associations are driven by narrow traits: A meta-analysis. *Obesity Reviews*, obr.12856. https://doi.org/10.1111/obr.12856

Vainik, U., Eun Han, J., Epel, E. S., Janet Tomiyama, A., Dagher, A., & Mason, A. E. (2019). Rapid Assessment of Reward-Related Eating: The RED-X5. *Obesity*, *27*(2), 325–331. https://doi.org/10.1002/oby.22374

Vainik, U., García-García, I., & Dagher, A. (2019). Uncontrolled eating: A unifying heritable trait linked with obesity, overeating, personality and the brain. *European Journal of Neuroscience*, *50*(3), 2430–2445. https://doi.org/10.1111/ejn.14352

Vainik, U., Misic, B., Zeighami, Y., Michaud, A., Mõttus, R., & Dagher, A. (2020). Obesity has limited behavioural overlap with addiction and psychiatric phenotypes. *Nature Human Behaviour*, *4*(1), 27–35. https://doi.org/10.1038/s41562-019-0752-x

Vainik, U., Mõttus, R., Allik, J., Esko, T., & Realo, A. (2015). Are Trait–Outcome Associations Caused by Scales Or Particular Items? Example Analysis of Personality Facets and Bmi. *European Journal of Personality*, *29*(6), 622–634. https://doi.org/10.1002/per.2009

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

Wagner, J., Becker, M., Lüdtke, O., & Trautwein, U. (2015). The First Partnership Experience and Personality Development: A Propensity Score Matching Study in Young Adulthood. *Social Psychological and Personality Science*, *6*(4), 455–463. https://doi.org/10.1177/1948550614566092

Wagner, J., Orth, U., Bleidorn, W., Hopwood, C. J., & Kandler, C. (2020). Toward an Integrative Model of Sources of Personality Stability and Change. *Current Directions in Psychological Science*, *29*(5), 438–444. https://doi.org/10.1177/0963721420924751

Wang, C.-S., Shyu, M.-L., & Lee, Y.-L. (2020). Personality traits and subjective well-being among fathers of preterm infants in Taiwan: A cross-sectional study. *Journal of Reproductive and Infant Psychology*, *38*(4), 408–420. https://doi.org/10.1080/02646838.2020.1748875

Weiss, A., Gale, C. R., Batty, G. D., & Deary, I. J. (2013). A questionnaire-wide association study of personality and mortality: The Vietnam Experience Study. *Journal of*

Psychosomatic Research, *74*(6), 523–529. https://doi.org/10.1016/j.jpsychores.2013.02.010

Widaman, K. F. (1985). Hierarchically Nested Covariance Structure Models for Multitrait-Multimethod Data. *Applied Psychological Measurement*, *9*(1), 1–26. https://doi.org/10.1177/014662168500900101

Williams, T. C., Bach, C. C., Matthiesen, N. B., Henriksen, T. B., & Gagliardi, L. (2018). Directed acyclic graphs: A tool for causal studies in paediatrics. *Pediatric Research*, *84*(4), 487–493. https://doi.org/10.1038/s41390-018-0071-3

Wilt, J., & Revelle, W. (2015). Affect, Behaviour, Cognition and Desire in the Big Five: An Analysis of Item Content and Structure. *European Journal of Personality*, *29*(4), 478–497. https://doi.org/10.1002/per.2002

Wood, A. M., Brown, G. D. A., Maltby, J., & Watkinson, P. (2012). How Are Personality Judgments Made? A Cognitive Model of Reference Group Effects, Personality Scale Responses, and Behavioral Reactions: Relative Personality Judgments. *Journal of Personality*, *80*(5), 1275–1311. https://doi.org/10.1111/j.1467-6494.2012.00763.x

Wood, D., & Roberts, B. W. (2006). The Effect of Age and Role Information on Expectations for Big Five Personality Traits. *Personality and Social Psychology Bulletin*, *32*(11), 1482–1496. https://doi.org/10.1177/0146167206291008

World Health Organization. (2021). *Adolescent health*. World Health Organization. https://www.who.int/health-topics/adolescent-health#tab=tab_1

Wortman, J., Lucas, R. E., & Donnellan, M. B. (2012). Stability and Change in the Big Five Personality Domains: Evidence From a Longitudinal Study of Australians. *Psychology and Aging*, a0029322.supp. https://doi.org/10.1037/a0029322.supp

Yarkoni, T. (2020). *Implicit realism impedes progress in psychology: Comment on Fried (2020)*. 12.

Yarkoni, T., & Westfall, J. (2017). Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. *Perspectives on Psychological Science*, *12*(6), 1100–1122. https://doi.org/10.1177/1745691617693393

Zhou, A., Ma, X., & Hajime, A. (2007). Parental expectation of early childhood education: Comparison between China, Japan, and Korea. *Frontiers of Education in China*, *2*(1), 140–147. https://doi.org/10.1007/s11516-007-0011-1

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *67*(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

Appendix

Cross prediction

In order to assess the robustness of the correlations among personality profiles (personality correlations), we re-created the personality correlations between all outcomes using cross-prediction of each outcome pair member (x and y) from the prediction models trained for the other pair member (average of the cross-predictions in both x → y and y → x direction were taken), divided by the average of the predictions accuracies of both pair member from models trained for themselves (because models' accuracies in predicting the outcomes for which they were trained were the upper boundary for cross-prediction). This procedure yielded a symmetrical matrix of personality correlations with ones on the diagonal. Each cross-prediction analysis was repeated 200 times with 2/3 – 1/3 training-validation split; the averages across the repetitions were taken as the final cross-predictions (see Figure S1). The matrix obtained this way was highly similar to the correlation matrix of nuance-level personality profiles described in the manuscript (r = .96; correlation between the off-diagonal elements of the two correlation matrices), thus indicating high robustness of the former.

Table S1 *Mean Absolute Errors between Observed Lifestyle Aspects and Mental Health Outcomes and Their Predicted Values from Personality Traits*

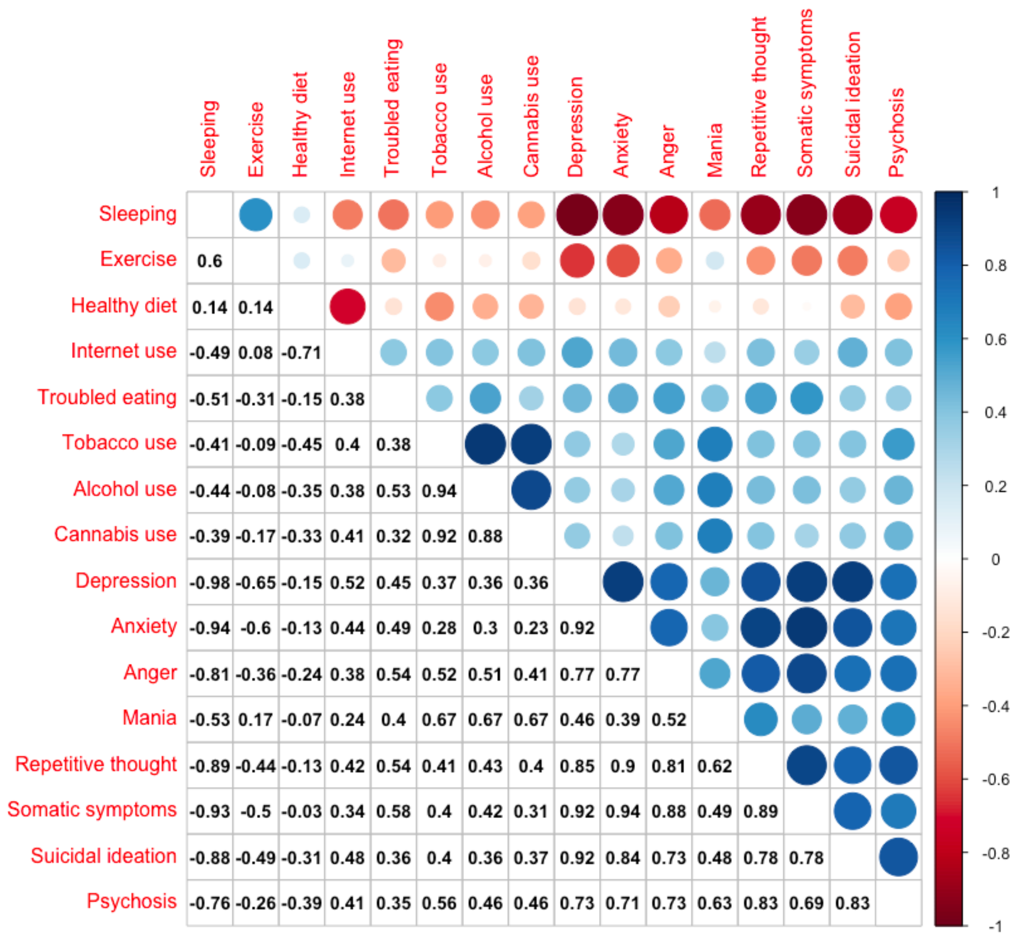
		Parameter	**Domains**	**Facets**	**Nuances**
Health promoting lifestyle	Sleeping	Mean	.57	.54	.54
		SD	.01	.01	.01
	Exercise	Mean	.64	.63	.63
		SD	.02	.02	.02
	Healthy diet	Mean	.79	.78	.77
		SD	.02	.02	.02
Health harming lifestyle	Internet use	Mean	.77	.76	.74
		SD	.02	.02	.02
	Troubled eating	Mean	.68	.63	.56
		SD	.01	.01	.01
	Tobacco	Mean	.98	.94	.90
		SD	.03	.03	.03
	Alcohol	Mean	.87	.85	.80
		SD	.02	.02	.02
	Cannabis	Mean	.56	.55	.54
		SD	.02	.02	.02
Mental health outcomes	Depression	Mean	.68	.58	.57
		SD	.02	.01	.01
	Anxiety	Mean	.64	.61	.58
		SD	.02	.02	.02
	Anger	Mean	.60	.57	.56
		SD	.02	.02	.01
	Mania	Mean	.73	.71	.70

	SD	.02	.02	.02
Repetitive thoughts	Mean	.64	.63	.61
	SD	.02	.02	.02
Somatic symptoms	Mean	.68	.66	.66
	SD	.02	.02	.02
Suicidal ideations	Mean	.61	.57	.55
	SD	.02	.02	.02
Psychosis	Mean	.29	.29	.29
	SD	.01	.01	.01

Note. The prediction was trained across 100 replications with the training sample of 75% of the total sample.

Figure S1

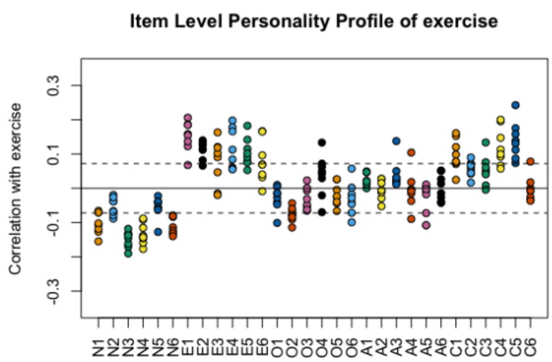
Cross predictions of 16 nuance-level personality profiles.

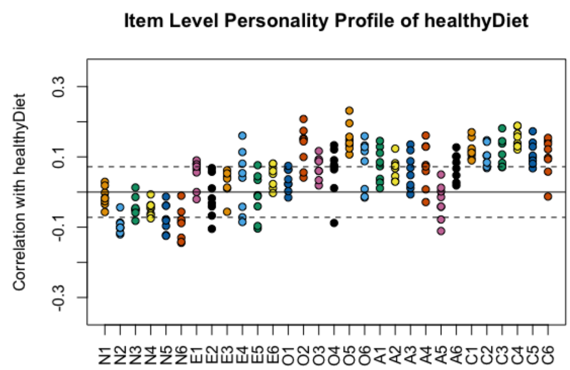


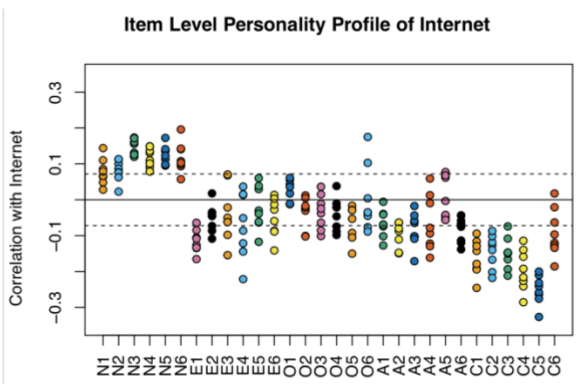
Note. Darker colours indicate higher correlations between two personality profiles which further suggest more similarities between these two profiles.

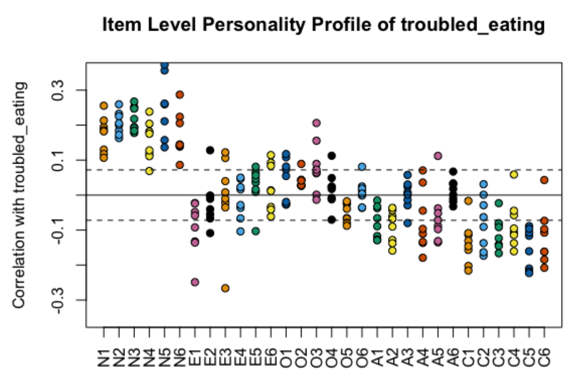
Figure S2 to S15

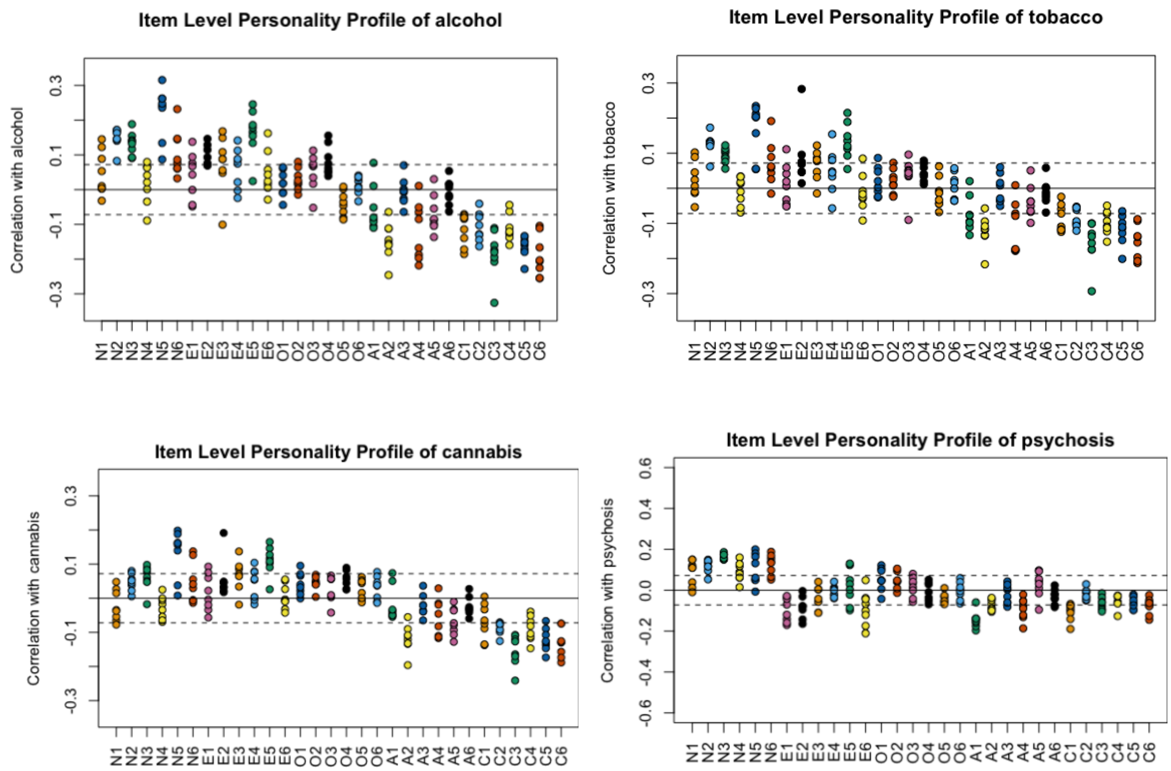
Plots for the correlations of the 240 items of the EE.PIP-NEO and lifestyle and mental health outcomes (exercise, healthy diet, internet use, troubled eating, tobacco use, alcohol use and cannabis use, psychosis, depression, anxiety, anger, mania, repetitive thought, somatic symptoms, suicidal ideation).

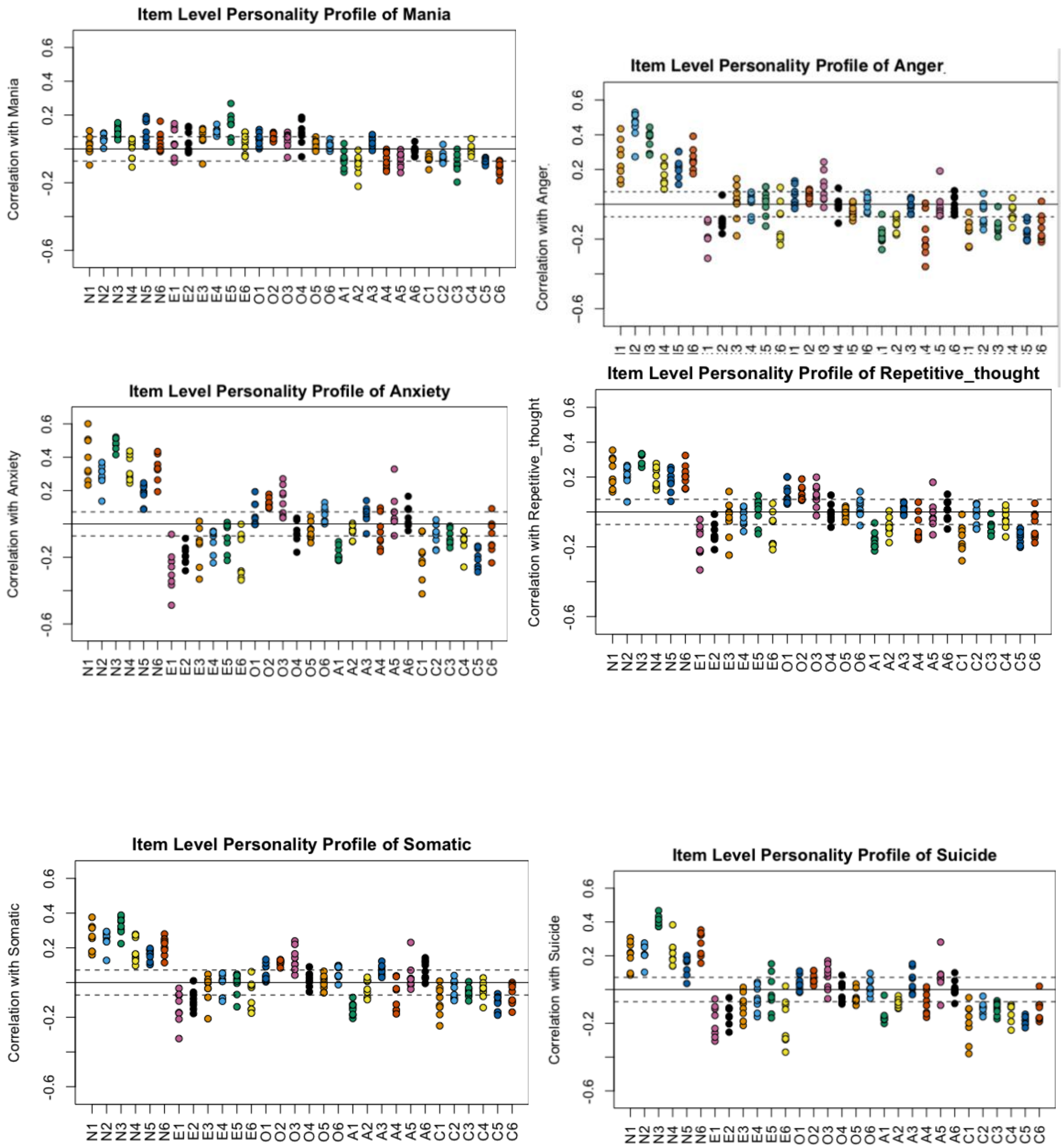









Item Level Personality Profile of alcohol
Correlation with alcohol
Item Level Personality Profile of tobacco
Correlation with tobacco
Item Level Personality Profile of cannabis
Correlation with cannabis
Item Level Personality Profile of psychosis
Correlation with psychosis
N1 N2 N3 N4 N5 N6 E1 E2 E3 E4 E5 E6 O1 O2 O3 O4 O5 O6 A1 A2 A3 A4 A5 A6 C1 C2 C3 C4 C5 C6



Note. The correlations are grouped according to the Big Five domains (indicated by letter) and their facets (indicated by number). The dashed line indicates the threshold for statistical significance after correcting for multiple comparisons using Holm's correction (Holm, 1979). N1, Anxiety; N2, Anger; N3, Depression; N4, Self-Consciousness; N5, Immoderation; N6 Vulnerability; E1, Friendliness; E2, Gregariousness; E3, Assertiveness; E4, Activity Level; E5, Excitement-Seeking; E6, Cheerfulness; O1, Imagination; O2, Artistic Interests; O3, Emotionality; O4, Adventurousness; O5, Intellect; O6, Liberalism; A1, Trust; A2, Morality; A3 Altruism; A4, Cooperation; A5, Modesty; A6, Sympathy; C1, Self-Efficacy; C2, Orderliness; C3, Dutifulness; C4, Achievement Striving; C5, Self-Discipline; C6, Cautiousness.