Personality Development in Adulthood:

Studies of stability and change in a nationally representative sample of adult New Zealanders

Petar Milojev

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy in Psychology

University of Auckland, 2016

Abstract

The development of personality in adulthood remains a topic of debate with direct implications for the conceptualisation of personality traits (McCrae & Costa, 1999; Roberts et al., 2008). However, with notable exceptions, comprehensive longitudinal investigations of the development of personality traits across a broad range of the adult life span are surprisingly rare. Through four systematic longitudinal investigations, the present thesis investigates the change and stability in the Big Six personality traits – Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility – in a nationally representative panel study of adult New Zealanders. Study 1 demonstrated the very high stability of the six traits, thus confirming the prediction of high temporal stability of personality traits. This study also provided the crucial test of the test-re-test reliability of the personality scales used throughout this thesis. Study 2 demonstrated the systematic variability in the stability of personality traits across the adult life span. That is, personality traits tend to stabilise across younger adulthood, reach peak stability in middle age (i.e., late 40's and early 50's), and systematically destabilise thereafter. The findings of Study 3 indicated remarkable consistency in personality traits associated with a major natural disaster (namely, the 2010/2011 Earthquakes in Canterbury, New Zealand). The only mean-level change from before to after the earthquakes was a small decrease in Emotional Stability (or an increase in Neuroticism) among those affected by the event. Finally, Study 4 investigated normative patterns of mean level change in personality traits and identified distinct developmental trajectories for the different traits with changes occurring across the entire adult life span. Collectively, these studies paint an interesting picture of both stability and change in personality traits, highlighting the need for an integrated perspective on personality development – one that incorporates both the intrinsic stability of the construct and early development, as well the continuing process of change.

Acknowledgements

Thanks Everyone!

Co-authorship Form

THE UNIVERSITY OF AUCKLAND
SCHOOL OF GRADUATE STUDIES

Co-Authorship Form

Graduate Centre
ClockTower – East Wing
22 Princes Street, Auckland
Phone: +64 9 373 7599 ext 81321
Fax: +64 9 373 7610
Email: postgraduate@auckland.ac.nz
www.postgrad.auckland.ac.nz

This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

Chapter 7 Milojev, P., Osborne, D., Greaves, L.M., Barlow, F.K., & Sibley, C.G. (2013). The Mini-IPIP: Tiny yet highly stable markers of the Big Six personality. Journal of Research in Personality, 47, 936-944. http://dx.doi.org/10.1016/j.jrp.2013.09.004.

Chapter 8 Milojev, P. & Sibley, C.G. (2014). The stability of adult personality varies across age: Evidence from a two-year longitudinal sample of adult New Zealanders. Journal of Research in Personality, 51, 29-37. http://dx.doi.org/10.1016/j.jrp.2014.04.005.

Chapter 9 Milojev, P., Osborne, D., & Sibley, C.G. (2014). Personality Resilience Following a Natural Disaster. Social Psychological and Personality Science, 5 (7), 760-768. DOI: 10.1177/1948550614528545.

Chapter 10 Milojev, P., & Sibley, C.G. (2016). Normative personality trait development in adulthood: A six-year cohort-sequential growth model. Journal of Personality and Social Psychology, manuscript conditionally accepted.

Nature of contribution by PhD candidate	I was the lead author for all of these papers. I conceived the initial ideas for the studies, conducted all analyses, formulated the models, and wrote the manuscripts for publication.
Extent of contribution by PhD candidate (%)	95%

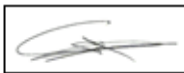
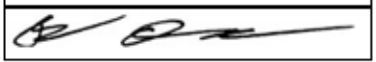
CO-AUTHORS

Name	Nature of Contribution
Dr. Chris Sibley	Thesis Supervisor. Provided feedback on manuscript drafts, offered conceptual suggestions in the initial stages of data analysis and model formulation. Offered suggestions about theoretical framing of manuscripts.
Dr. Danny Osborne	Secondary Thesis supervisor. Provided feedback on a draft of the manuscript. Offered suggestions about the theoretical framing of the manuscript.
Dr. Fiona Barlow	Provided feedback on a draft of the manuscript. Offered suggestions about the theoretica framing of the manuscript.
Ms. Lara Greaves	Helped collect data for the broader study, the New Zealand Attitudes and Values Study, which I analyzed for my thesis. Provided feedback on a draft of the manuscript.

Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- in cases where the PhD candidate was the lead author of the work that the candidate wrote the text.

Name	Signature	Date
Assoc. Prof. Chris Sibley		26/07/2016
Dr. Danny Osborne		22/07/2016

Last updated: 25 March 2013

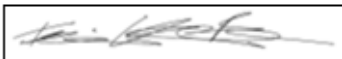
Dr. Fiona Barlow		Click here
Ms. Lara Greaves		22/07/2016
		Click here
		Click here

Table of Contents

List of Tables

List of Figures

graph.). The estimations are based on the mean-levels of the given trait (y axis) across age and assessments (x-axis). The 95% confidence intervals are presented as error bars around each point estimate. Within each cohort, i = intercept for that parameter, s = the fixed effect for the slope.

Figure 5.3. Estimated cohort effects between each of the cohorts as seen in Figure 1. Notes: Within each panel the bars represent the difference between the mean-level of the given trait at the participants' age as estimated by the latent growth trajectory from the preceding cohort and the intercept of the adjacent cohort. For example, the first bar in the panel showing cohort effects for Extraversion indicates the difference between the level of Extraversion at age 24 as estimated by the latent growth model from the youngest cohort - representing ages 19 through to 24 - and the level of Extraversion at age 24 based on the intercept of the subsequent cohort – representing ages 24 through to 29. Error bars represent 95% confidence intervals.

Figure A.2.1. Estimated power for detecting the men of the latent slope across a range of sample sizes. The estimated power (*y axis*) refers to the proportion of Monte Carlo replications for which the null hypothesis that a parameter is equal to zero is rejected at the .05 level (Muthén & Muthén, 2002). The effect size for the focal parameter – the mean of the slope – was systematically varied between .005 and .10 (*x-axis*). The simulations were conducted at two different sample sizes based on the sizes of the birth cohorts presented in Table 1: 1) small cohort size (n = 300), 2) large cohort size (n = 500); as well as 3) a large sample size (n = 700), and 4) a very large sample size (n = 5000). These are represented as separate functions in the figure.

Comment on publications

This thesis is based on the four published papers listed below. The papers are re-printed here as they appear in press, with formatting changes. Throughout the text to follow, these are referred to as study 1, study 2, study 3, and study 4, respectively.[1]

Milojev, P., Osborne, D., Greaves, L.M., Barlow, F.K., & Sibley, C.G. (2013). The Mini-IPIP: Tiny yet highly stable markers of the Big Six personality. *Journal of Research in Personality, 47,* 936-944. http://dx.doi.org/10.1016/j.jrp.2013.09.004.

Milojev, P. & Sibley, C.G. (2014). The stability of adult personality varies across age: Evidence from a two-year longitudinal sample of adult New Zealanders. *Journal of Research in Personality, 51,* 29-37. http://dx.doi.org/10.1016/j.jrp.2014.04.005.

Milojev, P., Osborne, D., & Sibley, C.G. (2014). Personality Resilience Following a Natural Disaster. *Social Psychological and Personality Science, 5 (7),* 760-768. DOI: 10.1177/1948550614528545.

Milojev, P., & Sibley, C.G. (2016). Normative personality trait development in adulthood: A six-year cohort-sequential growth model. *Journal of Personality and Social Psychology, Manuscript in press.*

[1] *MPlus* model syntax for studies 1 and 2 are presented in Appendix B. The sample *MPlus* model syntax for study 4 is presented in Appendix A.

Introduction

Preface

"Man's character is his fate" (Heraclitus).

"Personality is less a finished product than a transitive process. While it has some stable features, it is at the same time continually undergoing change" (Allport, 1955).

For many, the notion of 'personality' is associated with stability, consistency, and even predetermination of behaviour, cognitive patterns, and emotional reactions. This attractive proposition suggests that if we could accurately gauge a person's personality, we would be able to predict what they do and how they live. To a large extent, this represents the dominant position in the personality literature. However, as suggested in Gordon Allport's (1955) classic work, some argue that personality is a continuing process and that behaviour is a function of both personal dispositions and environmental processes. These positions highlight the central debate regarding the development of human personality and, indeed, the nature of personality traits. Brought forward by the contextual arguments of Mischel (1968), the question of whether personality is a matter of genetic influences and early temperament (e.g., McCrae & Costa, 1999) or a process of person-environment interactions (e.g., Roberts et al., 2008) remains unanswered.

This thesis contributes novel insights to this ongoing debate, and provides important contributions to the cumulative process of achieving a consensus regarding the developmental patterns of stability and change in personality across the adult life span. Through four systematic longitudinal studies, this thesis presents a comprehensive investigation of stability and change in the Big Six personality traits – Extraversion,

Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility – in a nationally-representative sample of adult New Zealanders. Study 1 (Milojev, Osborne, Greaves, Barlow, & Sibley, 2013) investigated the rank-order stability of personality traits over a 1-year test-re-test period, showing that personality traits are highly stable in adulthood. Study 2 (Milojev & Sibley, 2014) investigated the patterns of rank-order stability of personality across the adult life span. The findings of study 2 indicate that personality stability varies systematically across the entire adulthood, showing a process of stabilisation in younger adulthood and a period of de-stabilisation in older adulthood. Study 3 (Milojev, Osborne, & Sibley, 2014) investigated the changes in mean-levels of personality associated with experiencing a large-scale natural disaster, showing remarkable consistency in personality traits, with small domain-specific changes. Finally, study 4 (Milojev & Sibley, 2016) investigated patterns of mean-level changes in personality traits across the adult life span, showing evidence for both systematic change and consistency.

The chapters that follow present a detailed account of the present state of developmentally oriented personality research pertinent to the debate in question. The later chapters present the four studies as they appear in publication, followed by a discussion of the findings and the still uncertain landscape of personality theory and the nature of normative personality development.

Personality Defined

Understanding the person, or what people are like is arguably the ultimate goal of personality psychology as a discipline. This leaves the construct of "personality" with the complex job of allowing the formulation of "…an adequate psychological profile of *the whole person*." (Emmons, 1995, p.341). Given the complexity of human experience such a construct must be multi-faceted and provide multiple levels of understanding and analysis

(e.g., McAdams, 1995). Accordingly, organising the framework of individual differences in personality has been a major concern in personality psychology. Early theorists such as Allport (1937) and Cattell (1957) emphasised different levels of personality traits, while the contextual perspective put forward by Mischel (1968) inspired a focus on more domain-specific constructs. However, the focus of personality psychology in recent decades has centred on different conceptualisations and structure of broad bandwidth personality traits – such as the five-factor model of personality (McCrae & Costa, 1987, 1999) – as the key taxonomy of personality. While the debate about the conceptualisation of personality as a broad construct is ongoing, and while it cannot be assumed that personality is limited only to established personality traits, traits remain the central construct in personality research and inquiry. The work here presented is thus concerned with the developmental processes of broad bandwidth personality traits and the conceptualisation of 'personality' within the trait framework.

The systematic study of personality traits or "*relatively enduring styles of thinking, feeling, and acting*" (McCrae & Costa, 1997, p.509) dates as far back as the early work of Galton (1884) and the efforts by Thurnstone (1934), Klages (1926) and Baumgarten (1933). These are the first personality-oriented efforts employing the lexical hypothesis – that is, the working hypothesis that those individual differences that are most important to human transactions will be represented in the languages across the world (Goldberg, 1981). This tradition in the personality literature involves efforts to systematically organise the language of personality by identifying the number of relevant descriptive terms and, more importantly, the extent to which these terms share a common meaning. Indeed, these early insights about the relationships between the descriptive terms inspired questions about the structure of those associations which dominated most of personality research to come and formed what we commonly refer to as personality traits (Goldberg, 1981, 1990). With the advancement of

factor analytic methods the identification of the number and the structure of the different factors, or meaningful clusters of personality-descriptive terms, has been the major focus of personality research for almost a century (Ashton & Lee, 2007). This has led to the establishment of a number of models of personality over the years (e.g., Cattell, 1947, 1968; Eysenck, 1970; Gilford, 1975), including the Five Factor model which consists of the familiar Big Five personality traits - Extraversion (or Surgency), Agreeableness, Conscientiousness, Neuroticism (vs Emotional Stability),and Openness to Experience (or Intellect; Goldberg, 1990, 1992; McCrae & Costa, 1987, 1999). Indeed, the debate about the structure of personality in terms of the number and conceptualisation of the personality traits is ongoing with the recent developments of the six factor HEXACO model of personality (Ashton & Lee, 2007, 2009).

However, this evolution of our theoretical models has been far from smooth. The construct of a personality trait has been contested in the past, leading to the temporary abandonment of personality trait research (see Roberts, 2009). The driving force in this contention is represented in the work of Walter Mischel (1968). Mischel argued that the definition of personality must, at its core, necessitate invariance of behavioural tendencies across time and across different situations. In the same swoop, he presented a meta-analytic review of extant research to show that people's behaviour is indeed very variable and context-dependent (Mischel, 1968; Mischel & Shoda, 1995). This strict definition and the arguments put forward had a tremendous impact, threatening the validity of a personality trait, but also inspiring a response that led to the rejuvenation of personality trait research. Indeed, most of the highly influential work behind the Five Factor model of personality (McCrae & Costa, 1999; Terraciano, McCrae, & Costa, 2006) was aimed at showing the utility of personality traits, their consistency over time, and the cross-cultural universality of the personality trait structure. Within the revival of personality trait research in the past two

decades (Goldberg, 1981, 1990, 1992), scholars argue that it is *temporal consistency*, rather than cross-sectional context invariance, that is at the core of the construct of personality and personality traits (Roberts, 2009). Thus, the spotlight of the personality debate has moved towards the processes of personality development and its change and stability over time, with competing perspectives arguing for consistency (McCrae & Costa, 1999, 2008) and those arguing for both consistency and meaningful change (Roberts et al., 2008). It is this debate with which the present work is concerned.

Early Models

In the early 1900s, guided by the lexical hypothesis, a number of researchers set about cataloguing the terms used to describe individual characteristics and differences (Baumgarten, 1933; Galton, 1884; Klages, 1926; Thurnstone, 1934). Perhaps the most influential of these early analyses was Allport and Oldbert's (1936) analysis of the Webster's Underbridged Dictionary which influenced the development of Cattell's (1943) personality models. Cattell's early models used the list of terms developed by Allport and Oldbert (1936) to derive a set of 35 variables used in factor analyses in later studies to develop a personality model with 16 primary factors and 8 second-order factors (Cattell, 1947, 1948; Cattell et al., 1970). While Cattell's 16PF is in one form or another still used today, most of the researchers who attempted to replicate the findings failed. Early attempts at replication by Fiske (1949), and later ones by Tupes and Christal (1961), were unable to replicate the structural complexity proposed by Cattell. Instead, Tupes and Christal found evidence supporting a much more parsimonious five factor structure – Surgency, Agreeableness, Dependability, Emotional Stability, and Culture. Further work by Norman (1963), Borgotta (1964), and Smith (1967) replicated the five factors identified by Fiske, Tupes and Christal, offering further support for this taxonomy of personality traits.

Throughout the personality literature, other models of personality structure have been proposed such as Eysenck's (1970) three personality dimensions – Neuroticism, Extraversion (vs. Introversion), and the later addition of Psychoticism. A similar three-factor model was proposed by Tellegen (1985), which consisted of Positive Emotionality, Negative Emotionality, and Constraint. Others have proposed slightly more complex models with Thurnstone's (1953) seven-factor Temperament Schedule developed from Guilford scales, or Hogan's (1983) or Brand's (1984) six factor models. However, commonalities with the five factor solutions run through all of these models. For instance, Hogan's six factor model achieves the sixth factor by dividing the dimension of extraversion into Ambition and Sociability, arguably the two facets of Extraversion (see DeYoung, 2010). Indeed, the five factor solution has been the dominant model of personality structure throughout the history of personality research (Goldberg, 1981, 1990, 1993; McCrae & Costa, 1985, 1999).

The Big Five

The suggestion that a five factor structure provides the most adequate taxonomy of personality traits is evident from the early work by Thurnstone (1934) onwards. As mentioned above, analyses inspired by Cattell's (Cattell, 1947, 1948; Cattell et al., 1970) work consistently replicated a five-factor solution of personality (e.g., Borgotta, 1964; Fiske, 1949; Goldberg, 1990, 1992; McCrae & Costa, 1985, 1987; Norman, 1963; Smith, 1967; Tupes & Christal, 1961). Accordingly, the Big Five, a term coined by Goldberg (1981), has been the most dominant model of personality structure in the past three decades of personality research, with broad consensus as to the validity of the model and the utility of its dimensions (Goldberg, 1993; McCrae & Costa, 2008). The Big Five factors have generally been labelled Extraversion (or Surgency; describing people as talkative, assertive, or energetic), Agreeableness (describing people as good natured, cooperative, or trustful), Conscientiousness (describing people as orderly, responsible, and dependable), Emotional

Stability (versus Neuroticism; describing people as calm and not easily upset), and Intellect (or Openness to Experience; describing people as intellectual, imaginative, and open-minded). Despite some disagreement regarding the fifth factor, with some favouring the interpretation as Intellect (e.g., Peabody & Goldberg, 1989) and others favouring Openness to Experience (McCrae & Costa, 1987), there is generally strong consensus about the interpretations of the Big Five personality dimensions.

With the development and validation of questionnaire measures designed to assess the Big Five (e.g. NEO PI-R; Costa & McCrae, 1992) the validity of the trait constructs and their predictive utility were extensively tested (Costa & McCrae, 1992; McCrae & Costa, 1990, 1999). Over the years, the Big Five personality traits have been shown to predict important life outcomes such as marital satisfaction, occupational achievement, and even mortality (e.g., Caspi et al, 2005; Roberts et al., 2007; Ozer & Benet-Martinez, 2006). Similarly, an extensive amount of attention was directed toward demonstrating the cross-cultural consistency of the Five Factor model (McRae & Costa, 1994; McCrae et al., 1999, 2000), as well as the temporal consistency of the five dimensions of personality across time and across the life span (Costa & McCrae, 1997; McCrae & Costa, 1999; Terraciano, McCrae & Costa, 2006). Indeed, as discussed above, the question of temporal consistency is of direct importance for the validation of the trait as a construct in the face of the situationalist arguments such as those of Mischel (1968). Moreover, as discussed below, the question of the temporal consistency of the Big Five and the patterns of change and stability that characterise their development across the life span is at the centre of the current debate over the conceptualisation of personality traits and their development (e.g., McCrae & Costa, 1999; 2008; Roberts et al, 2008).

The Big Six

The universality of the Big Five personality structure, however, was brought into question in recent years, driven by the work of Ashton and Lee (2007). Through a series of lexical studies across a number of cultures and languages, these researchers found that a six factor, rather than the expected five factor solution, emerges in at least 12 different cultures, including lexical analyses of English (Ashton, Lee, & Goldberg, 2004; Ashton, Lee, Perugini et al., 2004). Three out of the six of these traits correspond to the Five Factor Model's dimensions of Extraversion, Conscientiousness and Intellect (or Openness to Experience). However, Ashton and Lee proposed a reconceptualization of Agreeableness that excludes the terms relating to sentimentality and includes the terms relating to the lack of anger. Similarly, the dimension labelled Emotionality is similar to the dimension of Emotional Stability (versus Neuroticism), however, it is conceptualised as including the terms related to sentimentality and excluding those terms relating to anger (Ashton & Lee, 2007). Finally, the additional sixth factor is conceptualised as including those terms relating to fairness (suggesting honesty) and lack of self-entitlement and greed (suggesting humility). This sixth factor is thus labelled Honesty-Humility. Accordingly, the six factor model presented by Ashton and Lee is coined the HEXACO model and presents a personality structure with six broad traits – Honesty-Humility (H), Emotionality (E), Extraversion (X), Agreeableness (A), Conscientiousness (C), and Openness to Experience (O). The six factor model is argued to better account for cross-cultural nuances and variations that the five-factor model cannot explain, and to allow for the subtle distinctions between factors better than the Big Five model (Ashton & Lee, 2007).

Further support for the six-factor model including the trait of Honesty-Humility comes from evidence of improvement in practical utility over the Big Five model in relation to those variables conceptually related to this additional trait. For instance, prediction of

workplace delinquency and likelihood of engaging in sexual harassment is significantly improved when items assessing Honesty-Humility are added to measures of the Big Five (Lee, Ashton, & De Vries, 2005; Lee, Gizzarone, & Ashton, 2003). Indeed, there is considerable evidence supporting the benefit of the six factor solution. This has inspired recent efforts to develop questionnaire based measures that allow for the assessment of the Big Six personality traits – Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility (e.g., Sibley et al., 2011).

Personality Change and Development

With the common conceptualisation of personality traits as "*relatively enduring styles of thinking, feeling and acting*" (McCrae & Costa, 1997, p.509), trait consistency and change across time are central concerns (Asendorpf, 1992; McCrae & Costa, 1999; Roberts, 2009). Research concerned with processes of normative personality development is thus pertinent to the understanding of the core construct of personality. This research largely suggests that personality is indeed a highly stable construct, however, substantive changes do occur across different periods of the life span (e.g., Anusic, Lucas, & Donnellan, 2012; Ardlet, 2000; Ferguson, 2010; Lucas & Donnellan, 2011; Roberts & DelVecchio, 2000; Roberts et al., 2006; Specht, Egloff, & Schmukle, 2011). Despite the evidence of systematic variation in personality across the life span, important questions remain both in terms of the patterns of change, as well as the processes that govern the development of personality.

At the centre of this debate are two broad perspectives or theoretical approaches to personality. On the one hand, the trait perspective, or the Five Factor Theory, emphasises intrinsic maturation processes and the general consistency of traits (McCrae & Costa, 1999; McCrae et al., 1999, 2000). On the other hand, the transactional life span perspective, or the Neo-Socioanalytic perspective, highlights the dynamics of person-environment transactions

as explaining the normative change in personality across the life span (Roberts & Caspi, 2003; Roberts et al., 2008). The different approaches have different implications for the expected patterns of personality development. For example, the theorists from the trait perspective tend to argue that most personality development occurs before the age of 30 and that personality traits remain largely unchanged thereafter (McCrae & Costa, 1997; Terraciano, Costa, & McCrae, 2006). The Neo-Socioanalytic perspective would, however, suggest specific patterns of changes in personality that may extend throughout the adult life span (Lucas & Donnellan, 2011; Roberts et al., 2008; Roberts et al., 2006).

These differing perspectives clearly have important implications for how we see personality – as static and predetermined or more dynamic in terms of continual person-environment interactions (Roberts et al., 2006). As such, personality development has inspired a considerable amount of research over the decades. This research overall aims to systematically investigate how the different types of personality change and stability are associated with age. The two types of personality change most commonly investigated are changes in rank-order stability – referring to the longitudinal stability in individual differences in personality or rank ordering of a group of people – and mean-level change – referring to changes in the absolute levels of personality traits over time (Caspi et al., 2005; Roberts et al., 2008). Extant research is considerable, however, studies utilising extensive longitudinal samples necessary to comprehensively address questions of rank-order stability and mean-level change over the life span are still relatively scarce, hindering the achievement of consensus on the question of trait development.

Theories of Personality Change

In considering the reasons for or causes of normative personality change and development, at least three broad perspectives can be outlined (Kogan, 1990; Neyer &

Asendorf, 2001). Focusing on the influence of genetic factors, the essentialist perspective suggests continuity of traits and universality of trait structures (McCrae et al., 1999, 2000). On the other hand, the contextualist perspective focuses on the influence of environmental factors (Kogan, 1990). More recently, however, developments in of transactional models – focusing on the transactions between person and the environment across the life span as explaining patterns of normative personality development – have gathered considerable support (Roberts, Wood, & Caspi, 2008). The debate as to why personality changes and how it develops is, however, ongoing.

The trait perspective of personality emphasises the continuity of personality traits and trait heritability. The Five Factor theory of personality has been, and continues to be, the dominant trait perspective on personality development (McCrae & Costa, 1999, 2008). The Five Factor theory suggests that the normative changes in personality are attributed to intrinsic maturation emphasising biological processes in personality development (McCrae et al., 2000; McCrae & Costa, 1999). In other words, personality trait development is determined by temperament and genetic factors. This development occurs through childhood and adolescence, stabilising in adulthood. Indeed, very little changes in personality, both in terms of stability or absolute changes, are expected across the adult life span (McCrae & Costa, 1999, 2008). Supporting this perspective are numerous studies showing a considerable genetic factor in terms of heritability of personality (Janng, Livesley, & Vemon, 1996; Loehlin, McCrae, Costa, & John, 1998; McCartney, Harris, & Bernieri, 1990), as well as cross-cultural studies showing the consistency of the five factor structure of personality (McRae & Costa, 1994; McCrae et al., 1999, 2000). Indeed, Five Factor theorists argue, based on longitudinal samples, that personality development occurs up to age 30, and is essentially unchanged, and unaffected by environmental factors thereafter (Costa & McCrae, 1997). However, as discussed below, a considerable amount of more recent empirical studies

and meta-analyses suggest substantive developmental changes in personality across adulthood indicating a continual process of personality development (e.g., Lucas & Donnellan, 2011; Roberts & DelVecchio, 2000; Roberts et al., 2006; Specht et al., 2011). Furthermore, longitudinal twin studies suggest that while there is a considerable component of heritability in personality, there is also a significant environmental component accounting for variability in personality (McGue, Bacon, & Lykken, 1993).

Contrary to the trait based perspectives, the contextual models of personality development emphasise the influence of environmental contingencies, particularly social roles (Kogan, 1990; Mischel, 1968). Social-cognitive approaches (e.g., Bandura, 1999; Zelli & Dodge, 1999) are the typical contextual approaches to explaining personality development. Such models emphasise the contextual factors such as social roles and personal goals. Consistency in personality is thought to be entirely due to the transactions with the social and environmental factors. As such, these models by and large deemphasise the notion of a trait with inherent consistency, but rather, explain personality development through environmental contingencies and changes therein (Zelli & Dodge, 1999). The issue with such models in considering normative changes in personality across the life span is that they do not actually inform changes in personality traits and, in the extreme case, render personality change relatively unpredictable as it depends on environmental contingencies which are themselves relatively unpredictable (Lewis, 1999).

Conceptually, somewhere between the essentialist trait theories and the contextual approaches lays a third set of developmental perspectives – the transactional models (Kogan, 1990). The transactional models emphasise the process of person-environment transactions, incorporating both the presupposed continuity of a trait within a person and the variation in environmental factors (Baltes, 1997; Roberts et al., 2008). While transactional models have a long history in Levinson's (1978) model focusing on building of life structures at different

period of the life span, or Erickson's (1950) stage theory emphasising specific life-tasks at different ages, more recent models can be formalised into a life span perspective or the Neo-Socioanalytic perspective of personality development (Baltes, 1997; Roberts et al., 2008). The perspective put forward by Baltes and colleagues (Baltes, 1997; Baltes, Lindenberger, & Staudinger, 1998) emphasises adaptation as central to personality development and proposes that both stability and change in personality occur across the life span. The person-environment interactions suggest a cumulative continuity whereby the effects of social, environmental or psychological influences decrease over the life span because people will tend to select the social roles and normative environmental circumstances based on their traits or dispositions.

Recent developments of the Neo-Socioanalytic perspective by Roberts, Caspi and colleagues (Roberts & Caspi, 2003; Roberts et al., 2008) are largely consistent with the model put forward by Baltes (1997). Specifically, this approach proposes identity achievement across the life span as associated with increased continuity or consistency in personality. Moreover, this model emphasises the importance of normative social institutions and life events – such as school, work, family, and normative events associated with these institutions – in promoting systematic personality change (Roberts et al., 2008). That is, as people engage in and commit to normative social roles at similar points in the life span (at least for most people), there should be a normative increase in those traits that facilitate performance of tasks associated with those social roles – or a decrease in those traits that hinder it. Personality traits generally associated with the norms, expectations and prescriptions of social roles and institutions associated with increased maturity (e.g., Conscientiousness, Agreeableness, or Neuroticism) should, therefore, show corresponding mean-level changes across the life span. This is formalised as the Neo-Socioanalytic perspective's *maturity principle* (Caspi et al., 2005). An interesting implication of this

perspective is that, to the extent that normative life events and commitments to normative social institutions occur at similar points of the life span for most people, substantive patterns of change and malleability in personality traits should be observed across the entire life span. This is particularly important when examining personality stability in younger and older age – periods of life associated with higher amounts of relative changes in identifying social institutions and the associated social roles (e.g. Lucas & Donnellan, 2011).

The debate regarding the different perspectives on personality development is ongoing. Clearly, the different perspectives have distinct implications regarding personality trait change and stability, both in terms of mean-level change and rank-order stability. Focusing specifically on the debate between the Five Factor Theory and the Neo-Socioanalytic perspective, different developmental patterns of mean-level change would be expected. If personality traits were largely determined by genetic factors and development dependent on intrinsic maturation processes, very little in the way of systematic changes in personality traits would be observed across the life span, particularly beyond early adulthood (Costa & McCrae, 1997). On the other hand, the Neo-Socioanalytic perspective would suggest systematic normative changes in personality traits associated with the interaction between personal and environmental factors (Roberts et al., 2008). For example, a systematic increase in traits associated with normative expectations of maturity and social roles therein, such as Conscientiousness, and Agreeableness, should be observed. Furthermore, to the extent that notable changes in commitment to social institutions occur in other periods of the life span, such as older adulthood, systematic changes to relevant personality traits should again be observed. This would suggest the possibility of curvilinear patterns of personality change across the life span such as a decrease in Conscientiousness in older age – perhaps associated with retirement and breakdown of task oriented social institutions, as well as intrinsic, biological changes (Lucas & Donnellan, 2011; Specht et al., 2011).

The Five Factor model would also predict a very specific pattern of rank-order stability across the life span. Specifically, given the argument that personality development is dictated by temperament and genes and that most of the development occurs in childhood and adolescence, a fairly uniform pattern of stability should be observed across adulthood (McCrae & Costa, 1999, 2000). Indeed, trait theorists have argued that most of personality development is complete by age 30 (Costa & McCrae, 1997). Thus, a pattern of increasing stability towards 30 years of age should emerge, followed be high (and unchanged) levels of stability thereafter. If, on the other hand, personality development is a function of person-environment transactions (Roberts et al., 2008), a different pattern of stability across the adult life span should be observed. Rank-order stability of personality traits should systematically increase through early adulthood due to selection and compensation processes whereby individuals should selectively commit to social institutions and roles that fit their personalities, which would in turn increase the identity formation regarding those personality traits. Indeed, this process has been formalised as the *cumulative continuity principle* (Caspi et al., 2005; Roberts et al., 2008). Furthermore, as mentioned above, the Neo-Socioanalytic perspective would predict further variation in rank-order stability across the adult life span. Specifically, with the normative changes in social institutions and social roles associated with older adulthood – such as retirement, bereavement, changes in social networks, as well as intrinsic changes in terms of deteriorating health and cognitive capacities – a systematic destabilisation of personality traits should be expected in older age (Lucas & Donnellan, 2011; Specht et al., 2011). These different perspectives, as discussed below, are differentially supported by extant research on personality development. However, a definitive conclusion has not been achieved, in part due to the relative scarcity of longitudinal studies adequate for addressing this debate.

Mean-level Change across the Life Span

Mean-level changes in personality reflect the magnitude and direction of absolute change in levels of a trait dimension over time (Caspi & Roberts, 1999; Roberts et al., 2006). Sometimes referred to as normative change, mean-level change - particularly change across the life span - is central to the considerations of generalizable patterns of personality development. In other words, patterns of normative change across the life span reflect the patterns of personality development that can be expected for most people (Roberts et al., 2006).

While the conceptualisation of personality traits emphasises temporal consistency in the patterns of individual differences in terms of thinking, feeling and behaving, a building body of literature indicates that notable changes do occur across the lifespan. A number of meta-analyses (Ardlet, 2000; Ferguson, 2010; Roberts & DelVecchio, 2000; Roberts *et al.*, 2006) and empirical investigations (e.g., Anusic, Lucas, & Donnellan, 2012; Lucas & Donnellan, 2011; Roberts, Caspi, & Moffit, 2001; Specht et al., 2011; Wortman, Lucas, & Donnellan, 2012) have reported systematic mean-level changes in personality traits at different stages of the life span, suggesting continuous personality development across adulthood.

As discussed in the previous section, while there is considerable evidence of personality change, there is ongoing debate as to the mechanisms of personality development and change across the life span, and the patterns of change that should be expected (McCrae & Costa, 1999l; Roberts et al, 2008). For example, it has been argued that developmental changes in personality should largely occur before the age of 30, with personality remaining stable thereafter (Costa & McCrae, 1988). Others, however, have observed systematic changes in personality beyond that point, with increases in stability as far as middle age, and

mean-level changes in older age (Scollon & Diener, 2006; Lucas & Donnellan, 2011; Roberts & DelVecchio, 2000; Roberts et al., 2006; Specht et al., 2011). Some of these discrepancies in the literature can be attributed to different perspectives on personality development, such as intrinsic maturation processes from the Five Factor theory and the trait perspectives (McCrae & Costa, 1999) or the transactional processes between person and the environment suggested by the life span perspectives, or the Neo-Socioanalytic perspective (Roberts *et al.*, 2008). However, consensus on the processes is complicated by the paucity of longitudinal studies with samples containing enough heterogeneity in terms of age and the periods of the life span represented (Specht et al., 2011). Furthermore, the discrepancies between extant longitudinal studies in terms of measures of personality used, as well as the specific periods of the life span investigated, hinder achievement of consensus on this issue.

Due to the relative paucity of longitudinal samples, most of the suggestions about the developmental patterns of change in personality traits across the life span have been extrapolated from the findings of cross-sectional studies of mean-level differences in personality across age (e.g., Anusic, Lucas, & Donnellan, 2012; Donnellan & Lucas, 2008; Lucas & Donnellan, 2009; McCrae et al., 2000; Sristave et al., 2003; Terraciano et al., 2005). Interestingly, the cross-sectional patterns of difference suggest developmental patterns that are broadly consistent with those suggested by the relatively few longitudinal samples (Anusic, Lucas, & Donnellan, 2012; Roberts et al., 2006; Specht et al, 2011; Wortman et al., 2012).

Overall, extant literature suggests that mean-levels of Extraversion tend to decrease over the life span. However, the different facets of Extraversion may show different developmental patterns (Lucas & Donnellan, 2011; Roberts et al., 2006). That is, while the social vitality facet may decrease with age (particularly across younger to middle adulthood), the social dominance facet may show an increase with increasing age (Roberts et al., 2006).

In terms of the trait of Agreeableness, there exists relative consensus that the levels of this trait tend to increase with age (Lucas & Donnellan, 2009; Lucas & Donnellan, 2011; Terraciano et al., 2005). Further, the literature seems to suggest a possibility of a curvilinear developmental pattern of change in Conscientiousness. While there are robust findings of an increase in mean-levels of Conscientiousness with increasing age, particularly in younger adulthood (Lucas & Donnellan, 2009; Lucas & Donnellan, 2011; Roberts et al., 2006), there is evidence to suggest a subsequent decrease in older age (Lucas & Donnellan, 2011; Terraciano et al., 2005). Neuroticism has tended to show a decrease with increasing age (Lucas & Donnellan, 2009; Roberts et al., 2006). However, the recent studies failed to replicate this trend showing instead relatively unchanged levels across the life span (Lucas & Donnellan, 2011). Finally, Openness to Experience tends to show a decrease in absolute levels with increasing age (Lucas & Donnellan, 2011; Roberts et al., 2006).

These patterns of normative change in the Big Five identified in the literature support the notion that personality changes across the life span. The general consistency of findings has been argued to lend support to the theoretical perspectives suggesting transactional processes between the person's dispositions and environmental influences that continue throughout adulthood (Caspi et al., 2005; Roberts & Caspi, 2003; Roberts et al., 2008). As discussed above, one of the aspects of this perspective is that personality change may in part be due to the different contingencies and prescriptions associated with different environmental circumstances and social roles into which a person enters (Roberts et al., 2008). For example, the findings of developmental patters of age differences in Conscientiousness, Agreeableness, and even Neuroticism, have been incorporated into the *maturity principle* (Caspi et al., 2005). That is, mean-level changes should be expected for those traits that are associated with performance of the roles of adulthood. The observed patterns of increases in Conscientiousness and Agreeableness, and the decreases in

Neuroticism, have largely been explained in these terms (Caspi et al., 2005). Indeed, recent findings have shown domain-specific changes in personality traits associated with normative life events that signify entering a normative social role. For example, Specht and colleagues (2011) observed systematic increases in Conscientiousness among those who entered the job market and commenced a new job – a social role and a circumstance that requires higher levels of Conscientiousness to enhance performance, achievement, and market competitiveness.

Overall, extant literature suggests that personality change is a continuous process that is at least in part due to social demands (Roberts et al., 2008) rather than intrinsic maturation (McCrae & Costa, 1999). However, while the literature suggests normative variability in personality across the whole range of the life span (e.g. Lucas & Donnellan, 2011; Specht et al., 2011; Wortman et al., 2012), the longitudinal studies needed to establish a coherent picture of normative life-span development are still few and far between. Moreover, just as the extrapolation of longitudinal patterns from cross-sectional findings is problematic in terms of the possibility of cohort effects, so do the few longitudinal studies suffer from this problem (Lucas & Donnellan, 2012; Specht et al., 2011). The entanglement of cohort effects and the effects of time in investigating normative changes in personality traits across the life span continues to be a significant obstacle. Consequently, the literature has not yet seen a closely approximated full life-span trend in personality change trajectories.

Rank-order Stability across the Life Span

Differential or rank-order stability reflects whether—and to what degree—relative individual differences - or placement of a group of people relative to each other - in a trait dimension are maintained over time (Ferguson, 2010; Roberts & DelVecchio, 2000). This temporal stability is of crucial importance to the conceptualisation of personality traits which

should exhibit high levels of stability over time (Asendorpf, 1992; McCrae & Costa, 1999; Roberts, 2009). Furthermore, as discussed above, the developmental patterns of differential stability of personality traits across the life span pertain to the understanding of personality and the causes of developmental processes therein (McCrae & Costa, 2008; Roberts et al., 2008).

Several meta-analytic reviews (Ardlet, 2000; Ferguson, 2010; Roberts & DelVecchio, 2000) and recent empirical studies (Lucas & Donnellan, 2011; Specht, Egloff, & Scukle, 2011; Wortman, Lucas, & Donnellan, 2012) have investigated the developmental patterns of differential stability of the Big Five personality across the life span. Consistent with the conceptualisation of personality as stable patterns of individual difference, these analyses indicate high stability estimates for the Big Five personality traits across test re-test periods of several years (Ferguson, 2010; Roberts & DelVecchio, 2000). More importantly, a consistent finding among these investigations is that differential stability of personality traits increases with increasing age. However, while the finding that individual differences in personality tend to stabilise with age is robust, there are inconsistencies in terms of the point in the life span at which differential stability peaks, as well as the pattern of stability that emerges thereafter. Ferguson (2010) provided evidence that is broadly consistent with the idea that personality stability peaks after around the age of 30 (e.g. Costa & McCrae, 1988), while data reported by Roberts and Del Vecchio (2000) indicated a somewhat later peak (around the age of 50). This general pattern of increasing stability with increasing age is consistent with the arguments from the Neo-Socioanalytic perspective, particularly the *cumulative continuity principle* (Caspi et al., 2005; Roberts et al., 2008) whereby personality is expected to stabilise with increasing age, largely through selection and compensation processes. Indeed, the evidence of inconsistencies in terms of the point of the life span that represents peak stability seems to be better accounted for by this perspective than the more static Five Factor trait

theory (McCrae & Costa, 1999; 2008). Interestingly, a recent meta-analysis of behavioural genetic studies aiming to decompose personality trait stability into genetic and environmental sources lends further support to this suggestion (Briley & Tucker-Drob, 2014). That is, while the genetic effects on stability tend to remain constant with age, the environmental effects increase substantially from childhood into adulthood. Thus, the increasing stability of personality across the life span seems to be more due to environmental rather than genetic effects, lending support to the *cumulative continuity principle* and the Neo-Socioanalytic perspective (Caspi et al., 2005; Roberts et al., 2008).

Ardlet's (2000) meta-analysis suggests that the stability of personality peaks around the age of 50, and declines afterward. Such findings might indicate a curvilinear pattern characterised by a decrease in stability towards the late adulthood and old age, such as that observed for other characteristics (e.g., self-esteem; Trzesniewski, Donnellan, & Robins, 2003). As discussed above, this pattern might be expected from the Neo-Socioanalytic perspective as older age may represent a period of normative change in social institutions and social norms. This in turn may be associated with a systematic destabilisation of individual differences in personality traits (Rioberts et al., 2008). However, a major barrier in personality development research is the relative scarcity of studies utilising longitudinal samples, particularly those that include older populations.

Several recent studies sought to remedy this issue by investigating the stability of Big Five personality traits across a broader range of the adult life span extending into old-age (up to 80 years old; Lucas & Donnellan, 2011; Specht, Egloff, & Scukle, 2011; Wortman, Lucas, & Donnellan, 2012). These studies corroborate the suggestions raised in the meta-analytic results of Ardlet (2000). Specifically, these studies observed a curvilinear – quadratic inverted U - pattern with regards to the distribution of differential stability of the broad personality dimensions across the adult life span. For most personality dimensions, the four-

year rank-order stability increased with age – towards high estimates between .70 and .90 – up to middle age, and then decreased towards old age, almost to the levels observed at the youngest age. Indeed, with only slight inconsistencies, these studies show remarkably similar results with regards to the general pattern of differential stability across the life span in large samples from Germany (Lucas & Donnellan, 2011; Specht, Egloff, & Scukle, 2011) and Australia (Wortman, Lucas, & Donnellan, 2012).

These curvilinear patterns are broadly consistent with the Neo-Socioanalytic perspective on personality development, suggesting that differential stability is likely to be lower in life-periods associated with instability and change (Caspi & Roberts, 2003; Roberts et al., 2008). As mentioned, it can be argued that, much like younger adulthood, older age is a period of substantive instability and changes pertinent to identity formation and stability or changeability in personality. That is, normative changes in older age such as retirement, bereavement, changes in social networks and roles therein, as well as intrinsic changes relating to health factors or cognitive functioning may be associated with substantive changes in personality and the consequent decrease in differential stability due to individual differences (Lucas & Donnellan, 2011; Roberts et al., 2008; Specht et al., 2011). While these recent findings show remarkable consistency across the three studies, longitudinal studies investigating these patterns on diverse samples are needed to examine their replicability and thus to build consensus as to the expected developmental patterns of trait stability across the life span.

Methodological Challenges and Gaps in Extant Literature

Developmentally oriented personality research has a long tradition with implications for understanding the construct of a personality trait. However, this research continues to be faced with glaring methodological challenges that hinder our ability to arrive at a definite

consensus regarding the issues of stability and change in personality. Most obviously, investigating the patterns of rank-order stability and mean-level change (rather than mean-level differences) requires longitudinal samples, with adequate heterogeneity in terms of the participant characteristics, most notably their age (e.g., Schmukle et al., 2011). Furthermore, given the subtlety of the effects associated with personality change (e.g., Roberts et al., 2006; Schmukle et al., 2011), these samples need to be large enough to allow for necessary statistical power.

The issue of longitudinal samples is particularly pertinent to the literature relating to mean-level changes in personality. As discussed above, most of the studies looking at normative changes in personality across the life span based their inferences on studies of cross-sectional differences across age (e.g., Anusic, Lucas, & Donnellan, 2012; Donnellan & Lucas, 2008; Lucas & Donnellan, 2009; McCrae et al., 2000; Sristave et al., 2003; Terraciano et al., 2005). Comparatively few studies have investigated developmental patterns of normative change in longitudinal samples (see Roberts et al., 2006). While the evidence of cross-sectional differences is informative and suggest patterns of change similar to those identified in the longitudinal studies, we are presented with the problem of teasing apart of the effects of age or time and birth cohorts in personality development. In other words, the observed mean-level differences in personality traits across age may represent developmental trends associated with aging. However, they may also represent cohort differences in personality. A similar problem is presented in studies of rank-order stability – it may be the case that there are systematic cohort differences in the stability of personality that account for some of the patterns of variability across the life span observed in the literature (e.g., Ardlet, 2000; Ferguson, 2010; Roberts & DelVecchio, 2000).

Moreover, longitudinal studies on personality development have tended to be limited in the scope of the population they represent. Many of the longitudinal studies limit their

scope to adolescence and early adulthood (e.g., McCrae et al., 2002) with very few studies representing later periods of the life span (Roberts et al., 2006). If we are interested in assessing the predictions from competing perspectives on personality and personality development – namely the Five Factor theory (McCrae & Costa, 1999, 2000) and the Neo-Socioanalytic perspective (Roberts et al.¸2008), a representation of the entire normative adult life span, or as close as it can be approximated, is crucial. As discussed in previous sections, the two perspectives have drastically different predictions regarding the patterns of rank-order stability and mean-level change across the adult life span, particularly beyond early adulthood (or beyond the age of 30 as suggested by McCrae & Costa, 1997). Indeed, the Neo-Socioanalytic perspective may imply systematic variation that continues across the life span and into old age, both in terms of normative change and rank-order stability. The narrow scope of representation in the extant research presents an additional issue. Namely, if we are interested in normative patterns of personality development across the adult life span (i.e., patterns of development as they apply to most people in the population) we must work with samples that allow for such generalisations. That is, the samples we work with must be representative of our population of interest.

Overcoming these methodological challenges and thus addressing the gaps in the literature regarding longitudinal patterns of change, beginning to address cohort effects, and the representation of the normative life span and the population of interest is a difficult task. Recent studies by Lucas ad Donnellan (2011), Specht and colleagues (2011) and Wortman and colleagues (2011) have made great strides in addressing some of these issues. By studying large, longitudinal, representative samples from Germany and Australia, and through the use of sophisticated analytic methods, these three studies have propelled the research on personality development across the life span. With samples that are representative of a broad range of the life span (through to 80 years old), these studies have

consistently shown continued variation in personality both in terms of mean-level change and rank-order consistency.

However, while these studies represent a considerable advance in developmental personality research, they also highlight the rarity of such investigations. The quadratic – inverted U – developmental patterns of rank-order stability observed across these three studies are the first three instances where such patterns have been identified (aside from suggestion by Ardlet, 2000). Similarly, the patterns of normative change, though largely consistent with suggestions from the extant literature (Lucas & Donnellan, 2011; Wortman et al., 2012), mark only the beginning of systematic analyses of personality change in representative longitudinal samples. Replication is necessary to arrive at the consensus of what might be the 'typical' developmental patterns of personality traits across the adult life span. Furthermore, these studies still face the problem of disentangling the effects of time or aging and cohort differences. More extensive samples and analytic methods are necessary to address these issues; nevertheless, they must be addressed. As such, even with the meta-analytic reviews such as those by Roberts and colleagues (2006), personality literature is still far from seeing a close approximation of a developmental pattern of normative change in traits across the life span.

Specht and colleagues (2011) in particular made notable advances in studying the person-environment transactions by investigating the relationships between personality change and life events. Given the arguments of the life span perspective (Roberts et al., 2008), investing the relationship between experiences of life events and personality change are pertinent to the proposed processes. While investigation of the effects of life events on personality change have been undertaken in the literature (e.g., Löckenhoff et al., 2008; Specht et al., 2011), there is still little evidence to confidently infer the types of life events that are associated with systematic personality change, and thus personality development.

Evidently, despite the rich history and ongoing efforts in developmentally oriented personality research, there are still glaring gaps in our knowledge about the pertinent processes. More importantly, with the methodological advances in terms of longitudinal studies with representative samples and as analytic methods, very exciting prospects and opportunities are now tantalisingly close, making the research on personality development a vibrant field of study.

Measuring Personality

In terms of questionnaire-based assessments of the Big Five, a broad range of personality measures have been developed over the years. Scales such as the NEO-FFI and the NEO PI-R (Costa & McCrae, 1992), the BFI (Big Five Inventory; John & Sristava, 1999), and the publically-available IPIP-FFM (International Personality Item Pool – Five Factor Model; Goldberg, 1999) are familiar examples of scales used to assess the broad personality traits of Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience. Such scales are used throughout the personality literature, with different studies employing different measurement approaches (e.g., Lucas & Donnellan, 2011, used a short-form BFI). Given that the different scales are developed from slightly different conceptualisations of the Big Five (e.g., the NEO PI-R and the CPI; California Psychological Inventory; Conn & Reike, 1994), integrating the findings pertinent to personality development becomes challenging. Indeed, influential meta-analyses have developed taxonomies that allow for categorisation of this variety within the Big Five framework (Roberts & DelcVecchio, 2000; Roberts et al., 2006). However, even this organisation has been controversial (Costa & McCrae, 2006), suggesting the issue remains a problem in a diverse field arguably studying the ‘same’ concept.

An additional note about personality measurement worth mentioning relates to the practicality of the majority of personality scales assessing the Big Five. Specifically, scales such as the Neo-PI-R or the IPIP-FFM are, quite simply, too long (e.g., two versions of the IPP have 50 and 100 items). This poses practical problems in terms of the demands on the participants (e.g., the time it takes to complete), and the questionnaire space these items occupy. This may force researchers to compromise by excluding measures of other variables they would otherwise include. Given the need for, and the advantages associated with, longitudinal studies with large representative samples (e.g., Lucas & Donnellan, Specht et al., 2011), the practicality of using such scales is a notable issue. In response to this problem, short form scales such as the Ten Item Personality Inventory (TIPI; Gosling *et al*., 2003) have been developed providing an alternative. However, a number of studies advise caution when weighing up the need for very short-form scales and the drawbacks of their implementation (e.g., Credé et al., 2012; Kruyen, Emons, & Sijtsma, 2013).

Credé and colleagues (2012) conducted a comprehensive investigation in which they compared the psychometric performance of various short-form personality scales, including the two-item TIPI, and the four-item Mini-IPIP (Donnellan et al., 2006). The authors found that one and two item scales tended to produce measures of personality that had attenuated correlations with criterion variables, though longer (i.e., four, six or eight item) scales performed comparably better. Because assessing the effects of personality on relevant variables is one of the main applications of short-form personality scales, the possibility that they yield relatively unreliable estimates is a widespread problem that must be considered when researchers are deciding to use such scales. Moreover, Credé and colleagues (2012) indicated that short-form scales are also associated with inflated Type 1 error. In other words, short-form measures of personality - particularly those that only include one or two items per personality dimension - tend to overestimate the effects of additional constructs included in a

regression model and the variance in the outcome that they explain. However, this research showed that there are appreciable differences in psychometric performance across short-form scales. Specifically, scales with four or more items per personality dimension produced considerably fewer Type 1 and Type 2 errors relative to the shorter (i.e., one or two item) measures. For instance, the Mini-IPIP - a short-form, 20-item version of the IPIP assessing the Big Five by four items each – was associated with fewer problems than the shorter scales. Furthermore, Donnellan and colleagues (2006) have reported good criterion related validity of the 20-item Mini-IPIP relative to longer measures of the Big-Five, as well as strong support for the critical five-factor structure. Such a scale can thus provide a practical balance that combines the practicality of the short-form, while minimising the associated drawbacks.

Given the recent developments of the six-factor HEXACO model discussed in previous chapters (Ashton & Lee, 2007), additional issues are presented. Firstly, researchers interested in investigating the Big Six face the same problem of long and potentially impractical scales – the original measurement scale of this model consists of the 60-item HEXACO-60 scale (Ashton & Lee, 2009). However, given the additional sixth factor of Honesty-Humility within this model, existing short-form scales such as the Mini-IPIP are not adequate. To address this issue, Sibley and colleagues (2011) proposed an extension of the Mini-IPIP by developing the Mini-IPIP6. The Mini-IPIP6 is a 24-item scale that extends the original Mini-IPIP by adding four items assessing Honesty-Humility. As with its five factor counterpart, the Mini-IPIP6 shows promising psychometric properties. Confirmatory Factor Analysis (CFA) supports the six factor structure. The scale correlates with various criteria, and demonstrates good construct and discriminant validity (Sibley et al., 2011). Likewise, the Mini-IPIP6 has acceptable Item Response Properties (Sibley, 2012), and scale norms in a national probability sample (Sibley & Pirie, 2013). This development provides a promising

option for personality research that is consistent with the recent developments in terms of the structure of personality traits.

Aims of the Present Research

This thesis is guided by the central aim of investigating the patterns of personality development in adulthood. Through a series of four longitudinal investigations, this thesis provides a systematic investigation of stability and change in the Big Six personality dimensions across the adult life span within a nationally representative sample of adult New Zealanders. Because the specific aims and hypotheses associated with each of the four investigations are discussed in the subsequent chapters, this section will address the overall aims of the thesis and provide an overview of the studies within it.

This thesis is comprised of four longitudinal; studies published in the *Journal of Research in Personality*, *Social Psychological and Personality Science*, and the *Journal of Personality and Social Psychology*[2]. The studies are based on different aspects of the representative panel sample of the New Zealand Attitudes and Values Study (NZAVS). The details about the NZAVS sampling procedure are presented in the following chapter. Thus, the four studies represent a systematic investigation of stability and change in the Big Six personality dimensions within a large, longitudinal, nationally representative sample of adult New Zealanders.

The first study of this thesis had the specific aim of establishing the test re-test reliability of the Mini-IPIP6, thus establishing the temporal reliability of the scales that were used in the subsequent studies. As discussed in the related chapter below, given that the definition of personality necessitates that personality traits be "*relatively enduring*" (Roberts, 2009; p. 140), high levels of stability are expected over moderate time-periods. The first

[2] Please note that study 4 is currently *in press*.

study of this thesis (Milojev, Osborne, Greaves, Barlow, & Sibley, 2013) sought to establish the stability of the Big Six personality traits over a one-year re-test period, thus establishing the test re-test reliability of the Mini-IPP6 scales.

The second study was specifically designed to assess the developmental patterns in the rank-order stability of the Big Six personality dimensions across the life span. The study sought to test whether the patterns predicted by the Neo-Socioanalytic perspective on personality development (Roberts et al., 2008) would be supported in the representative New Zealand sample. Specifically, as discussed above, the Neo-Socioanalytic perspective would predict systematic variation in rank-order stability of personality across the entire life span, rather than a levelling off after early adulthood as suggested by the Five Factor theory (McCrae & Costa, 1999). Support for the Neo-Socioanalytic perspective has, however, been seen in the findings of Lucas ad Donnellan (2011), Specht and colleagues (2011), and Wortman and colleagues (2011). The second study of this thesis (Milojev & Sibley, 2014) sought to provide a necessary test and replication of these developmental patterns.

The first and second studies addressed the questions related to rank-order stability of personality traits and the developmental patterns of stability across the life span. The specific aim of the third study, however, focused on examining the mean-level change in personality in the context of adverse life-events. Specifically, the third study (Milojev, Osborne, & Sibley, 2014) examined changes in the Big Six personality dimensions before and after a major natural disaster in New Zealand – the 2010/2011 Christchurch earthquakes. This study thus aimed to contribute to the building literature directly assessing the person-environment transactions that, according to the Neo-Socioanalytic perspective, in part govern the developmental changes in personality (Löckenhoff et al., 2008; Specht et al., 2011). Specifically, this study examined the extent to which personality change is associated with the non-normative, rare, but dramatic negative life-events, such as a large-scale natural

disaster. Moreover, these effects are assessed in the nationally representative sample of the NZAVS, thus providing a practical contribution in terms of the effects of such natural disasters on people's personality that may be generalizable to other similar events.

Finally, the aim of the fourth study was to provide a comprehensive investigation of the patterns of normative change in personality over the adult life span. More importantly, this study (Milojev & Sibley, 2016) sought to present a modelling approach that allows an investigation into personality change due to aging and change due to cohort differences. This study aimed to begin addressing the pertinent issues of the confluence of the effects of time and cohort differences in personality development. As discussed above, this is a particularly notable issue in extant research, and is especially relevant to examinations of normative change. Therefore, the models presented in the fourth study of this thesis –Cohort Sequential Latent Growth Models (e.g., Pronzie & Onghena, 2005) - are an important contribution to the literature on personality development as they provide the closest approximation of the pattern of normative change over the adult life span presently available.

Overall, this thesis aims to systematically investigate the developmental patterns of stability and change in the Big Six personality traits across the adult life span in a nationally representative sample of New Zealanders. The four studies provide an assessment of both the rank-order stability and the normative change in personality and seek to contribute important evidence to the debate regarding the nature of personality development. The four studies are presented in order in the later chapters, as they appear in publication (with minimal formatting changes).

Overview of Methodology

The methodological details, including the analytic strategies, as they pertain to each of the four investigations that comprise this thesis are presented in the chapters that follow.

Furthermore, the power analyses and Monte Carlo simulations as pertaining to each of the four studies are presented in the *"Methodological Notes"* of the general discussions section – note that the power simulations for study 4 are presented within the supplementary materials accompanying the publication (presented in Appendix A). All four studies are based on the NZAVS sample. Similarly, all four studies are concerned with the Big Six personality dimensions as assessed by the Mini-IPIP6 short-form scales (Sibley et al., 2011). This chapter discusses the NZAVS sampling procedure, and the Mini-IPIP6 scales in more detail than they appear in the chapters to follow.

Sampling and Participants – The NZAVS

The New Zealand Attitudes and Values Study (NZAVS) is an ongoing effort that has been conducting yearly surveys of adult New Zealanders since 2009 (Time 1) through 2010 (Time 2), 2011 (Time 3), 2012 (Time 4), 2013 (Time 5) and 2014 (Time 6). The Time 1 (2009) NZAVS contained responses from 6518 participants sampled from the 2009 New Zealand electoral roll. The electoral roll is publicly available for scientific research and, in 2009, contained 2,986,546 registered voters. This represented all citizens over 18 years of age who were eligible to vote (regardless of whether they chose to vote), barring people who had their contact details removed due to specific case-by-case concerns about privacy. The sample frame was spilt into three parts. Sample Frame 1 constituted a random sample of 25,000 people from the electoral roll (4,060 respondents). Sample Frame 2 constituted a second random sample of a further 10,000 people from the electoral roll (1,609 respondents).

Sample Frame 3 constituted a booster sample of 5,500 people randomly selected from meshblock area units of the country with a high proportion of Māori, Pacific Nations and Asian peoples (671 respondents). Statistics New Zealand (2013) define the meshblock as "the smallest geographic unit for which statistical data is collected and processed by Statistics

New Zealand. A meshblock is a defined geographic area, varying in size from part of a city block to large areas of rural land. Each meshblock abuts against another to form a network covering all of New Zealand including coasts and inlets, and extending out to the two hundred mile economic zone. Meshblocks are added together to 'build up' larger geographic areas such as area units and urban areas. They are also the principal unit used to draw-up and define electoral district and local authority boundaries." Meshblocks were selected using ethnic group proportions based on 2006 national census data. A further 178 people responded but did not provide contact details and so could not be matched to a sample frame.

In sum, postal questionnaires were sent to 40,500 registered voters or roughly 1.36% of all registered voters in New Zealand. The overall response rate (adjusting for the address accuracy of the electoral roll and including anonymous responses) was 16.6%.

The Time 2 (2010) NZAVS contained responses from 4442 participants. The Time 2 (2010) NZAVS retained 4423 from the initial Time 1 (2009) NZAVS sample of 6518 participants, and included an additional 20 respondents who could not be matched to the Time 1 participant database (a retention rate of 67.9% over one year). Participants were posted a copy of the questionnaire, with a second postal follow-up two months later. Participants who provided an email address were also emailed and invited to complete an online questionnaire if they preferred.

The Time 3 (2011) NZAVS contained responses from 6884 participants (3918 retained, 2965 new participants). The Time 3 (2011) NZAVS retained 3916 from the initial Time 1 national probability sample (a 60.1% retention rate over two years). A further three participants who joined at Time 2 were also retained.

To boost sample size at Time 3 and compensate for sample attrition, a booster sample was recruited through an unrelated survey posted on the website of a major New Zealand

newspaper in 2011. A total of 3208 participants registered an initial expression of interest in being contacted to participate in the NZAVS via this survey. Participants in this non-random booster sample were emailed an invitation to participate in an online version of the NZAVS, and those who did not respond to the email were also sent a postal version of the questionnaire. A total of 2961 participants completed the questionnaire when subsequently contacted (92.4%). This yielded a total sample size for the Time 3 (2012) NZAVS of 6884 (3918 retained from Time 1, 3 additions retained from Time 2, and 2961 recruited from the newspaper website at Time 3).

The Time 4 (2012) NZAVS contained responses from 12,182 participants (6805 retained from one or more previous wave, 5377 new additions from booster sampling). The sample retained 4051 participants from the initial Time 1 (2009) NZAVS of 6518 participants (a retention rate of 62.2% over three years). The sample retained 5762 participants from the full Time 3 (2011) sample (a retention rate of 83.7% from the previous year).

Non-respondents were emailed a follow-up reminder email approximately two months later. Three attempts were then made using each provided phone number (typically home and cell phone) to contact non-respondents to encourage participation. These attempts were made on separate days, approximately one week apart. When possible, a phone message was left for each phone number after the third attempt. Participants were also posted a pamphlet outlining recent findings from the study mid-way through the year. Finally, participants were posted a Season's Greetings from the NZAVS research team, and informed that they had been automatically entered into a bonus seasonal grocery voucher prize draw for a total pool of $NZ 1000). Participants were informed that the draw would happen automatically and winners contacted. The Season's Greetings card also asked participants to contact us (online, email or phone) to let us know if any of their contact details had changed

before the prize draw was conducted. These additional materials are presented by Huang, Greaves, and Sibley (2014) in an online NZAVS technical report.

To boost sample size at Time 4 and increase sample diversity for subsequent waves, five independent booster samples using different sample frames were also conducted. The first sample frame consisted of a randomly selected sample of 20,000 people from the 2012 New Zealand Electoral Roll (the electoral roll list all eligible voters, barring those removed on a case-by-case basis due to privacy concerns). A total of 2431 participants responded to this booster sample (response rate = 12.34% when adjusting for the 98.5% accuracy of the 2012 electoral roll). The second sample frame consisted of a regional booster of 10,000 people randomly selected from people listed in the 2012 Electoral Roll who lived in the Auckland region. A total of 890 participants responded to this booster sample (adjusted response rate = 9.04). The Auckland region was oversampled because it is the fastest growing and most ethnically diverse region of the country with an increasing number of Asian and Pacific peoples in particular. The questionnaire used for this Auckland sample was longer than the standard NZAVS questionnaire, and contained additional unrelated questions that are not included in the NZAVS dataset (these related to the use of community facilities). Exit interviews conducted during Time 5 indicated that the longer length of this questionnaire may have contributed to the low response rate in this case.

The third sample frame consisted of 3,000 people randomly selected from the 2012 Electoral Roll who lived in the Christchurch region. A total of 333 participants responded to this booster sample (adjusted response rate = 13.52%). The Christchurch region was oversampled because it has experienced significant hardship and change due to the Christchurch earthquakes of 2010 and 2011 with many people moving out of the region (Statistics New Zealand, 2013) and problems with mail delivery with some city zones being

placed under restricted entry due to safety concerns and considerable infrastructure destroyed).

The fourth sample frame consisted of 9000 respondents selected from meshblock area units across the country that were moderate-to-high in deprivation according to the index developed by Salmond, Crampton and Atkinson (2007). Regions with levels of deprivation were selected using scores on the decile-ranked NZ Deprivation index from 6-10, with 10 being the most deprived). This sample frame used scaled weighting so that people in increasingly derived regions were increasingly more likely to be selected (with random sampling of people within regions that had a given level of deprivation). The scaling factor was as follows: $n_i = n_{base} * weight_i$, where $n_{base} = 600$, and $weight_i$, ranged from 1 to 5 and increased by 1 for each one-unit increased in deprivation score. Thus, 600 people were randomly selected from regions with a deprivation score of 6, 1200 people were randomly selected from regions with a deprivation score of 7, and so on. This sampling strategy was designed to increase the representativeness of the sample across regions with different levels of deprivation, as the NZAVS showed increased an attrition rate in increasingly more deprived regions over the first three years of the study. A total of 767 participants responded to this booster sample (adjusted response rate = 9.73%). The fifth sample frame consisted of 9,000 people randomly selected from those who indicated on the 2012 Electoral Roll that they were of Maori ethnicity (ethnic affiliation as Maori is listed on the role, but other ethnic affiliations are not). A total of 690 participants responded to this booster sample (adjusted response rate = 7.78%). The questionnaire administered to the Maori booster sample included questions specifically designed for Maori.

The Time 5 (2013) NZAVS contained responses from 18,264 participants (10,502 retained from one or more previous wave, 7,518 new additions from a booster sample, and 181 unmatched participants or unsolicited opt-ins). The sample retained 3,934 participants

from the initial Time 1 (2009) NZAVS of 6518 participants (a retention rate of 60.4% over four years). The sample retained 9,844 participants from the Time 4 (2012) sample (a retention rate of 80.8% from the previous year). As described in the Time 4 procedure, we offered a prize draw for participation, non-respondents were emailed and phoned multiple times, and all participants were posted a Season's Greetings card from the NZAVS research team and informed that they had been automatically entered into a bonus seasonal grocery voucher prize draw. We also posted our yearly pamphlet summarizing key research findings published during the current wave of the study.

To boost sample size and increase sample diversity for subsequent waves, two booster samples were also conducted. As with previous booster samples, sampling was conducted without replacement (i.e. all people included in previous sample frames were identified and removed from the 2014 roll). The first sample frame consisted of 70,000 people aged from 18-60 randomly selected from the 2014 New Zealand Electoral Roll. The New Zealand electoral roll contains participants' date of birth (within a one-year window), and we limited our frame to people who were 60 years or younger due to our aim of retaining participants for the following 15 years. A total of 7,489 participants responded to this booster sample (response rate = 10.9% when adjusting for the 98.6% accuracy of the 2014 electoral roll). The second sample frame consisted of 1,500 people who were listed as having Maori ancestry aged from 18-60 randomly selected from the 2014 New Zealand Electoral Roll. A total of 92 participants responded to this booster sample (response rate = 6.2% when adjusting for electoral roll accuracy).

The Time 6 (2014) NZAVS contained responses from 15,822 participants (15,740 retained from one or more previous wave, and 82 unmatched participants or unsolicited opt-ins). The sample retained 3,727 participants from the initial Time 1 (2009) NZAVS of 6,518 participants (a retention rate of 57.2% over five years). The sample retained 14,875

participants from the full Time 5 (2012) sample (a retention rate of 81.5% from the previous year). Participants were posted a copy of the questionnaire, with a second postal follow-up two months later. Participants who provided an email address were also emailed and invited to complete an online version if they preferred. As described in the Time 5 procedure, we offered a prize draw for participation, non-respondents were emailed and phoned multiple times, and all participants were posted a Season's Greetings card from the NZAVS research team and informed that they had been automatically entered into a bonus seasonal grocery voucher prize draw. We also emailed participants an online pamphlet containing a series of video interviews with the researchers summarizing different research findings.

Personality Measures – The Mini-IPIP6

Across the four studies, the Big Six personality dimensions are assessed by the Mini-IPIP6 (Sibley et al., 2011). The Mini-IPIP6 is a short-form, 24-item scale assessing the six personality dimensions – Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility – by 4 items each. Specifically, the scale extends the 20-item Mini-IPIP scales developed by Donnellan, Oswald, Baird, and Lucas (2006) to assess the Big Six personality dimensions. As such, the Mini-IPIP6 includes the 20 items developed by Goldberg (1999). An additional four items were used to assess Honesty-Humility – two items from the Narcissism scale developed by Campbell *et al*. (2004), and two items adapted from Ashton and Lee's (2009) HEXACO measure of Honesty-Humility.

Within the NZAVS questionnaire, the items were administered with the following instructions: "This part of the questionnaire measures your personality. Please circle the number that best represents how accurately each statement describes you." Items were rated on a scale ranging from 1 (very inaccurate) to 7 (very accurate). Extraversion was assessed by the following four items: "Am the life of the party."; "Don't talk a lot." (reverse scored);

"Keep in the background." (reverse scored); "Talk to a lot of different people at parties." ($\alpha_{Time\ 1}$ = .71; $\alpha_{Time\ 2}$ = .75; $\alpha_{Time\ 3}$ = .75; $\alpha_{Time\ 4}$ = .73; $\alpha_{Time\ 5}$ = .75; $\alpha_{Time\ 6}$ = .75). Agreeableness was assessed by the following four items: "Sympathize with others' feelings."; "Am not interested in other people's problems." (reverse scored); "Feel others' emotions."; "Am not really interested in others." (reverse scored) ($\alpha_{Time\ 1}$ = .66; $\alpha_{Time\ 2}$ = .69; $\alpha_{Time\ 3}$ = .69; $\alpha_{Time\ 4}$ = .68; $\alpha_{Time\ 5}$ = .70; $\alpha_{Time\ 6}$ = ..72). Conscientiousness was assessed by the following four items: "Get chores done right away."; "Like order."; "Make a mess of things." (reverse scored); "Often forget to put things back in their proper place." (reverse scored) ($\alpha_{Time\ 1}$ = .65; $\alpha_{Time\ 2}$ α = .65; $\alpha_{Time\ 3}$ = .65; $\alpha_{Time\ 4}$ = .64; $\alpha_{Time\ 5}$ = .66; $\alpha_{Time\ 6}$ = ..67). The Neuroticism scale included the following four items: "Have frequent mood swings."; "Am relaxed most of the time." (reverse scored); "Get upset easily."; "Seldom feel blue." (reverse scored) ($\alpha_{Time\ 1}$ = .64; $\alpha_{Time\ 2}$ = .72; $\alpha_{Time\ 3}$ = .72.; $\alpha_{Time\ 4}$ = .68; $\alpha_{Time\ 5}$ = .70; $\alpha_{Time\ 6}$ = ..71). Openness to Experience was assessed by the following four items: "Have a vivid imagination."; "Have difficulty understanding abstract ideas." (reverse scored); "Do not have a good imagination." (reverse scored); "Am not interested in abstract ideas." (reverse scored) ($\alpha_{Time\ 1}$ = .67; $\alpha_{Time\ 2}$ = .70; $\alpha_{Time\ 3}$ = .70; $\alpha_{Time\ 4}$ = .68; $\alpha_{Time\ 5}$ = .69; $\alpha_{Time\ 6}$ = ..71). Finally, Honesty-Humility was assessed using the four items as per Sibley and colleagues (2001). These items (all of which were reverse scored) included: "Would like to be seen driving around in a very expensive car."; "Would get a lot of pleasure from owning expensive luxury goods."; "Feel entitled to more of everything."; "Deserve more things in life." ($\alpha_{Time\ 1}$ = .78; $\alpha_{Time\ 2}$ = .77; $\alpha_{Time\ 3}$ = .77; $\alpha_{Time\ 4}$ = .77; $\alpha_{Time\ 5}$ = .77; $\alpha_{Time\ 6}$ = ..79).

Within this framework, and the framework of the HEXACO model (Ashton & Lee, 2007, 2009; Sibley et al., 2011), Extraversion reflects friendliness and involvement in social activities. Agreeableness covers tolerance and cooperation and relates to the facet of compassion rather than politeness. Conscientiousness includes diligence, organisation and

motivation to carry out tasks. Neuroticism relates to the monitoring of inclusionary status through insecurity and anxiety. Openness to Experience relates to engagement in abstract ideas and curiosity. Finally, Honesty-Humility, consists of reciprocal altruism, sincerity and (the absence of) entitlement. As discussed above, the Mini-IPIP6 has been comprehensively validated for use in the New Zealand population demonstrating good construct and discriminant validity (Sibley et al., 2011). The scale has also been shown to have acceptable Item Response Properties (Sibley, 2012), and scale norms on a national probability sample (Sibley & Pirie, 2013).

Chapter II

Study 1: "The Mini-IPIP6: Tinny yet highly stable markers of Big Six personality"

Reference:

Milojev, P., Osborne, D., Greaves, L.M., Barlow, F.K., & Sibley, C.G. (2013). The Mini-IPIP: Tiny yet highly stable markers of the Big Six personality. *Journal of Research in Personality, 47,* 936-944. http://dx.doi.org/10.1016/j.jrp.2013.09.004.

Abstract

We assessed the stability of a short-form six-factor personality measure over a one-year period in a large national probability sample (N = 4,289). Personality was assessed using the Mini-IPIP6—a short-form measure assessing Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility. Standardized estimates calculated using Bayesian Structural Equation Modelling (BSEM) indicated that all six personality dimensions were extremely stable. An alternative model using Maximum Likelihood estimation, in which residual item variances were associated over repeated assessments, yielded similar findings. These results highlight the stability of personality in the general population, even when assessed using short-form scales. The use of Bayesian models to examine the stability of personality and their application for study of change in specific developmental periods is discussed.

Introduction

The reliable assessment of personality has been a major concern of psychological research for over a century (see Goldberg, 1981, 1990). To these ends, various personality measures have been developed, a process which has been aided by the relatively recent emergence of unifying frameworks such as the "Big-Five" model of personality (Goldberg,

1981, 1990). Familiar scales - which include the NEO-FFI (NEO-Five Factor Inventory; Costa & McCrae, 1992), the BFI (Big Five Inventory; John & Srivastave, 1999), and the publically available IPIP-FFM (International Personality Item Pool – Five Factor Model; Goldberg, 1999) - are examples of measures used to assess the five dimensions of this model: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience (Goldberg, 1999). In addition, recent short-form scales have been developed to address concerns relating to the practical drawbacks of traditional long-form personality scales. Arguably one of the most notable of these brief measures is the 20 item Mini-IPIP developed by Donnellan, Oswald, Baird, and Lucas (2006).

Because personality is largely defined as a temporally stable pattern of responding to the environment, the validity of such scales – and, in fact, of any measure of personality - is particularly dependent on its test-retest reliability (i.e., the temporal stability of the scales). This is especially relevant to short-form scales such as the Mini-IPIP, as they have only been recently developed and are associated with various psychometric problems (Donnellan et al., 2006; Credé, Harms, Niehorster, & Gaye-Valentine, 2012; Kruyen, Emmos, & Sijtsma, 2013). Moreover, the fewer items on short-form (relative to long-form) scales inherently limit the reliability of these brief, albeit sometimes necessary, measures (Crocker & Algina, 1986).

With this in mind, we investigate the one year test-retest stability of the Mini-IPIP6 (Sibley et al., 2011). The Mini-IPIP6 is a 24-item measure of six broad bandwidth personality dimensions and is a direct extension of the original five-factor Mini-IPIP (Donnellan et al., 2006). Inspired by the recent proposition of the HEXACO personality model (Ashton & Lee, 2007, 2009), this scale includes additional items measuring the sixth factor of Honesty-Humility. Therefore, the Mini-IPIP6 extends the Mini-IPIP to assess the six broad-bandwidth personality dimensions of Extraversion, Agreeableness, Conscientiousness, Openness to Experience, Neuroticism and Honesty-Humility (Sibley et al., 2011). Using a large nationally

representative sample in New Zealand (N = 4,289), we employ Structural Equation Models using both Bayesian estimation and Maximum Likelihood estimation methods to test the temporal stability of the Mini-IPIP6. In doing so, we examine a particularly crucial psychometric property for any "good" measure of personality (namely, the temporal stability of the scale).

Stability of Personality

Test-retest reliability is an important assessment of the psychometric properties of a scale that speaks directly to the stability of the constructs measured. This is particularly important when these constructs relate to personality, which is defined as "relatively enduring styles of thinking, feeling, and acting" (McCrae & Costa, 1997, p. 509). Because personality is conceptualized as a consistent pattern of responses, one would expect to observe high levels of stability across *time* and place (Mischel, 1968). More specifically, if a scale, long or short, is to be considered an acceptable measure of personality, it should show particularly high levels of stability.

Research looking at the stability of personality traits largely suggests that personality is stable over moderate and long-term test-retest periods, particularly in adult populations (e.g., Costa & McCrae, 1988, 1994; Soldz & Vaillant, 1999). Indeed, test-retest correlations of around .90 have been reported using long-form measures of the Big Five over a six year period (Costa & McCrae, 1988a, 1994). Studies investigating life-span development in personality find some change in youth or the adolescent years, however, they generally report largely stable scores on personality measures in adulthood. Likewise, meta-analytic studies of personality have shown test-retest correlation coefficients of above .64 and .74, indicating high stability of personality in adults as assessed by various measurement scales (e.g., Roberts & DelVecchio, 2000; Soldz & Vaillant, 1999). Finally, recent investigations suggest

that the five personality dimensions captured by the Big-Five show moderate to high stability over 8 years in a large longitudinal national probability sample (Lökenhoff et al., 2008). Therefore, in line with theorizing, research demonstrates that adult personality is largely stable, particularly in relation to rank-order stability.

Although past research shows remarkably high levels of stability across time, recent meta-analytic findings suggest that standard estimates of test-retest reliability are actually attenuated due to measurement error. Indeed, non-corrected stability coefficients tend to be significantly lower than stability coefficients that account for measurement error (Ferguson, 2010). Thus, it is possible that the stability of personality is routinely *under*estimated. As such, it is important to account for measurement error in any examination of the stability of personality (Ferguson, 2010).

Related to this concern, recent studies conducted in Europe have shown very high levels of stability in the Big Five over a 9 year period when advanced Structural Equation Modelling (SEM) techniques are employed, such as those used in the present paper (Rantanen et al., 2007). Using SEM techniques to account for measurement error through the estimation of latent personality variables, the researchers showed notably high reliability estimates and addressed the issue of attenuated stability coefficients as discussed by Ferguson (2010). The stability coefficients reported differed somewhat between men and women, with men showing greater stability of Neuroticism and Extraversion, but were generally in the .65 to .97 range, higher than the test-retest correlations generally reported (Rantanen et al., 2007).

The above research suggests that personality is largely stable over time, although there is some variability across the lifespan. These suggestions are partially supported by research on the genetic influences on personality and development. Such research shows that the variability in the five personality factors is, to a notable degree, heritable (Jang, Livesley,

& Vemon, 1996; Loehlin, McCrae, Costa, & John, 1998; McCartney, Harris, & Bernieri, 1990). However, these investigations also suggest that the genetic influences on personality decrease with age (McCartney, Harris, & Bernieri, 1990). Furthermore, longitudinal twin studies identified significant effects of the environment in accounting for the variability, or instability in personality (McGue, Bacon, & Lykken, 1993). As such, while correlational studies suggest high levels of stability in personality, heritability studies also suggest instability on account of the interaction between genetic influences and the environment. Such findings have obvious implications for arguments in relation to intrinsic stability of personality in the long term. However, in terms of moderate length test-retest investigations one can expect results consistent with the high personality stability findings in the literature.

Although there is good evidence for the stability of long-form personality scales, the stability of short-form measures remains relatively unexplored by comparison. A notable exception to this general trend is a study by Donnellan and Lucas (2011). Their research examined the effects of age on the stability of the Big Five as assessed by a 15-item version of the Big Five Inventory. While the goals of these analyses were different to the aims of our investigation - we do not examine the effects of age on stability - the researchers showed generally high test-retest reliabilities of the short-form personality scale over a four year period. High test-retest reliabilities on short-form scales have also been reported for a number of other scales including the BFI-10, as well as the TIPI and the Mini-IPIP (Rammstedt & John, 2007; Gosling, Rentfrow, & Swann, 2003; Donnellan et al., 2006).

Thus, there is good evidence that short-form personality measures can attain high test-retest reliability. Nevertheless, relative to the literature on the stability of long-form personality scales, investigations of the stability of short-form personality scales are scarce. This is particularly problematic considering the rising popularity of various short-form personality scales developed over the last 15 years, as well as the potential psychometric

problems associated with reducing the number of items included in a given scale (e.g., low levels of reliability). Moreover, there is a relative absence of research on the stability of the Big-Six model of personality.

The current study addresses these oversights by assessing the stability of the Mini-IPIP6, a short measure of the six personality factors. To do this, we use a large national probability sample of New Zealand adults. We argue that if the Mini-IPIP6 is an acceptable short-form measure of the Big Six personality, then the stability estimates over a one-year period should be very high (with estimates above .90). Such evidence would help validate one of the only short-form measures of the Big Six, thereby providing researchers with a critical alternative to long-form measures of the six factor model of personality.

Personality Assessment and Short-Form Measures

In recent years, attention has shifted towards the numerous practical drawbacks of traditional long-form measurement scales. One of the common critiques of these scales is that the time needed to complete a long-form scale necessarily leads scholars to exclude other constructs from their questionnaire (e.g., the IPIP-FFM is a 50-item scale). That is, researchers often must choose between including a quality measure of personality at the cost of assessing other constructs, or assess more attitudes at the expense of a potentially poor-performing short-form scale of personality. In response to this problem, short-form scales such as the Ten Item Personality Inventory (TIPI; Gosling, Rentfrow, & Swann, 2003) have been developed. While these trade-offs are sometimes inevitable in large-scale surveys, a number of studies suggest that researchers should exercise considerable caution when balancing the need for very short-form measures of personality with the drawbacks of their implementation (e.g., Credé et al., 2012; Kruyen, Emmos, & Sijtsma, 2013).

In particular, recent research indicates that the use of very short-form personality scales is associated with increased Type 1 and Type 2 errors (Credé et al., 2012; Kruyen, Emons, & Sijtsma, 2013). This is especially problematic for very short scales such as the TIPI (Gosling, Rentfrow, & Swann, 2003), a measurement instrument that assesses the Big Five with one or two items per personality dimension. Indeed, Credé and colleagues (2012) conducted a comprehensive investigation in which they compared the psychometric performance of various short-form personality scales, including the two-item TIPI, and the four-item Mini-IPIP (Donnellan et al., 2006). The authors found that one and two item scales tended to produce measures of personality that had attenuated correlations with criterion variables, though longer (i.e., four, six or eight item) scales performed comparably better.

The results of Credé and colleagues' (2012) study indicate that the use of short-form personality scales can lead to the underestimation of the effects of personality on other constructs of interest. Because assessing the effects of personality on relevant variables is one of the main applications of short-form personality scales, the possibility that they yield relatively unreliable measures is a widespread problem that must be considered when researchers are deciding to use such scales. Moreover, Credé and colleagues (2012) indicated that short-form scales are also associated with inflated Type 1 error. In other words, short-form measures of personality - particularly those that only include one or two items per personality dimension - tend to overestimate the effects of additional constructs included in the regression model and the variance in the outcome that they explain.

Credé and colleagues' (2012) research also showed that there are appreciable differences in psychometric performance across short-form scales. Specifically, scales with four or more items per personality dimension produced considerably fewer Type 1 and Type 2 errors relative to the smaller (i.e., one or two item) measures (Credé et al., 2012). While caution must be taken when deciding which short-form scale to use, slightly longer scales

with four or eight items for each personality dimension will considerably improve the reliability of one's findings. Nevertheless, some research contexts will require the use of short-form scales. Indeed, the current study was conducted within a large national probability study in which we assessed a variety of values and attitudes. Short-form personality measures are particularly useful in this context, as space for additional items in large-scale population surveys is limited.

As mentioned above, the Mini-IPIP is a short-form, 20-item version of the IPIP. In their investigations, Credé and colleagues (2012) found that this particular scale produced fewer Type 1 and Type 2 errors than the shorter scales using one or two items. Furthermore, Donnellan and colleagues (2006) have demonstrated that this scale has favourable psychometric properties. Specifically, they reported good criterion related validity of the 20-item Mini-IPIP relative to longer measures of the Big-Five, as well as strong support for the critical five-factor structure. The scale thus provides a practical approach to measuring the Big-Five that combines the benefits of the long-form measures of personality (e.g., low error rates) with the practicality of the short-form personality scales.

However, as suggested above, a critical component of any measure of personality is the scale's stability over time (i.e., the test-retest reliability of the measure). As such, Donnellan and colleagues (2006) conducted typical short-term (3 weeks and 9 months) tests of test-retest reliability of the Mini-IPIP, albeit with relatively small samples. Their results showed high correlations between the scales over the two time frames, indicating high short-term stability of their measures (9 months test-retest correlations r = .86, .68, .77, .82, .75, for Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience respectively). Comparatively, the test-retest correlations of the five dimensions of personality, as measured by the TIPI, have been found to be somewhat lower over the short term (two week test-retest correlations r = .77, .71, .76, .70, .62, for Extraversion,

Agreeableness, Conscientiousness, Emotional Stability, and Openness to Experience, respectively; Gosling, Rentfrow, & Swann, 2003). Thus, the Mini-IPIP shows considerable promise as a brief and reliable measure of personality

Six-Factor Model of Personality

While the Big-Five has been the dominant approach in personality research for over two decades, recent investigations have identified a sixth *independent* personality dimension referred to as Honesty-Humility (Ashton & Lee, 2007, 2009). The six independent factors then are: Honesty-Humility (H), Emotionality (E), Extraversion (X), Agreeableness (A), Conscientiousness (C), and Openness to Experience (O). This six dimension model, referred to as the HEXACO model, is argued to account for cross-cultural lexical findings, as well the subtle distinctions between personality traits better than the Big-Five (Ashton & Lee, 2007). The original measurement scale of this model consists of the 60-item HEXACO-60 scale (Ashton & Lee, 2009).

Given the recent development of the HEXACO model, neither the original Mini-IPIP nor the TIPI is equipped to assess Honesty-Humility. To address this issue Sibley and colleagues (2011) proposed an extension of the Mini-IPIP by developing the Mini-IPIP6. The Mini-IPIP6 directly builds on the original Mini-IPIP of Donnellan and colleagues (2006) by adding four items to the existing scale in order to assess Honesty-Humility (Sibley et al., 2011; Ashton & Lee, 2007, 2009). The Mini-IPIP6 is thus a 24-item scale measuring six broad bandwidth personality dimensions.

As with its five factor counterpart, the Mini-IPIP6 has shown promising psychometric properties. Specifically, Confirmatory Factor Analysis (CFA) supports the six factor structure; the scale correlates with various criteria, and demonstrates good construct and discriminant validity (Sibley et al., 2011). Likewise, the Mini-IPIP6 has acceptable Item

Response Properties (Sibley, 2012), and scale norms (Sibley & Pirie, 2013). Nevertheless, to date there have been no tests of the temporal stability of this scale, nor its test-retest reliability. Our research thus examines both the stability and test-retest reliability of the Mini-IPIP6 to establish the psychometric properties and assess its suitability for use in future research.

Although the estimates of test-retest reliability for the Mini-IPIP reported by Donnellan and colleagues (2006) are indeed high, these estimates may be considerably higher when assessed using advanced modelling procedures in a large nationally representative sample. In fact, in light of the literature discussed above (Ferguson, 2010), in our models of test-retest stability of the Mini-IPIP6, we argue that very high levels of stability are to be expected, given that both the Mini-IPIP and the Mini-IPIP6 show such favourable psychometric properties. We further propose the use of Bayesian estimation methods and Structural Equation Models as a general framework for testing the test-retest stability of such scales over traditional correlational analyses.

Overview of the Present Study

In this paper, we build on the work of Donnellan and colleagues (2006) by assessing the stability of the Mini-IPIP6 scale across two waves of a national probability sample conducted over a one year period (2009 to 2010). We do this using a large (N = 4,289), nationally representative sample in New Zealand derived from the New Zealand Attitudes and Values Study (NZAVS).

In our investigation of the psychometric properties of the Mini-IPIP6 scale, we conducted two Confirmatory Factor Analyses, enforcing a 6 factor solution on the 24 items of the scale at Time 1 (2009) and at Time 2 (2010). Using Bayesian estimation methods (see Kruschke, Aquinis & Joo, 2012), we then assessed the stability of the 6 latent variables

representing the 6 broad-bandwidth dimensions of personality: Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Honest-Humility. This approach allowed us to adjust for measurement error and estimate the posterior distribution of each of the six personality dimensions assessed via the Mini-IPIP6.

We also provide a further model investigating the short-term stability of the Mini-IPIP6 using the same data, albeit via a different estimation method. Specifically, we present the results of a model corresponding to the main Bayesian model above, using traditional Maximum Likelihood estimates. As an alternative (and arguably more conservative) test, this model controls for the correlations between the residual variances of each of the 24 item pairs over the one year period. Therefore, as well as adjusting for measurement error, this model also adjusts the stability estimates for residual error associated with multiple assessments of each of the 24 separate items in the scales. Because this approach includes additional parameters, model fit was likely to be arbitrarily improved. However, because the variance that was misattributed to the latent factors of the personality dimensions was correctly partitioned to the shared error variances unique to each item, we expected this model to show somewhat lower reliability estimates than the Bayesian model for each of the six Mini-IPIP6 scales.

Predictions

We expected each of the dimensions of the Mini-IPIP6 to show high stability over the one year period, which we defined as standardized stability coefficients greater than .90. We suggest this as a standard for stability of an “excellent” personality measure over the short-term, when taking into account past estimates of the stability of the Mini-IPIP, and analytic strategies that allow us to identify and account for measurement error. We further expected

all six dimensions of personality to demonstrate relatively similar levels of stability, with any differences in the stability of the dimensions being trivial in magnitude.

Method

Sampling Procedure

The New Zealand Attitudes and Values Study (NZAVS) Time 1 questionnaire was posted to 40,500 participants randomly-selected from the publicly available version of the 2009 NZ electoral roll. The NZAVS-2009 sampled a total of 6,518 participants. The overall response rate (adjusting for address accuracy of the electoral roll and including anonymous responses) was 16.6%. Roughly two-thirds of the sample (N = 4,442) responded to a follow-up survey at Time 2 one year later (a sample retention rate of 68.1%).

Participant Details

We limited our analyses to the 4,289 participants who provided responses to the Time 1 and 2 data analysed here (2633 women and 1656 men). Roughly 85% of the sample analysed were members of the NZ European ethnic group (n = 3,658), 15% were members of the Māori ethnic group (n = 654), about 4% were Asian (n = 171), around 4% were of Pacific descent (n = 153), and roughly 3.5% did not report their ethnicity or identified with another minority group. The mean age of the participants was 49.69 (SD = 15.15).

With regard to other demographics, 43.7% of the sample identified as religious (n = 1,874). In terms of parental status, 77% had at least one child (n = 3,304). Most participants were in some form of romantic relationship (71.4%, n = 3,062). As for educational status, 21.7% said they had no education or failed to report their highest level of education (n = 930), 28.3% reported at least some high school (n = 1,213), 15.9% reported having studied towards a diploma or certificate (n = 682), 24% reported having studied at undergraduate

level (*n* = 1,028) and 10.8% reported having pursued post-graduate study (*n* = 436). The majority (79.5%) of participants were born in New Zealand (*n* = 3,408), and roughly 72% were in some kind of employment (*n* = 3,093).

Questionnaire Measures

The Mini-IPIP6 is a 24-item scale containing the 20 items developed by Goldberg (1999) and included by Donnellan et al. (2006) in the original Mini-IPIP. In addition, four items were used to index Honesty-Humility. Items H01 and H02 were adapted from the Narcissism scale developed by Campbell et al. (2004), whereas items H03 and H04 were adapted from Ashton and Lee's (2009) HEXACO measure of Honesty-Humility.

As in the Mini-IPIP6 validation study by Sibley and colleagues (2011), the items were administered with the following instructions: "This part of the questionnaire measures your personality. Please circle the number that best represents how accurately each statement describes you." Items were rated on a scale ranging from 1 (very inaccurate) to 7 (very accurate). Extraversion was assessed by the following four items: "Am the life of the party."; "Don't talk a lot." (reverse scored); "Keep in the background." (reverse scored); "Talk to a lot of different people at parties." ($\alpha_{\text{Time 1}} = .709$; $\alpha_{\text{Time 2}} = .709$). Agreeableness was assessed by the following four items: "Sympathize with others' feelings."; "Am not interested in other people's problems." (reverse scored); "Feel others' emotions."; "Am not really interested in others." (reverse scored) ($\alpha_{\text{Time 1}} = .661$; $\alpha_{\text{Time 2}} = .681$). Conscientiousness was assessed by the following four items: "Get chores done right away."; "Like order."; "Make a mess of things." (reverse scored); "Often forget to put things back in their proper place." (reverse scored; $\alpha_{\text{Time 1}} = .649$; $\alpha_{\text{Time 2}}$ $\alpha = .646$). The Neuroticism scale included the following four items: "Have frequent mood swings."; "Am relaxed most of the time." (reverse scored); "Get upset easily."; "Seldom feel blue." (reverse scored; $\alpha_{\text{Time 1}} = .639$; $\alpha_{\text{Time 2}} = .669$). Openness

to Experience was assessed by the following four items: “Have a vivid imagination.”; “Have difficulty understanding abstract ideas.” (reverse scored); “Do not have a good imagination.” (reverse scored); “Am not interested in abstract ideas.” (reverse scored; $\alpha_{\text{Time 1}} = .669$; $\alpha_{\text{Time 2}} = .668$). Finally, Honesty-Humility was assessed using the four items as per Sibley and colleagues (2011). These items (all of which were reverse scored) included: “Would like to be seen driving around in a very expensive car.”; “Would get a lot of pleasure from owning expensive luxury goods.”; “Feel entitled to more of everything.”; “Deserve more things in life.” ($\alpha_{\text{Time 1}} = .776$; $\alpha_{\text{Time 2}} = .789$).

Bayesian Estimation

Bayesian analysis is less well known than the standard Maximum Likelihood estimation method. Our purpose here is not to review Bayesian methods, but a few brief comments are warranted before introducing our model (for a review see Kruschke et al., 2012). Bayesian probability estimates and associated credible intervals tend to reflect what people naturally think of when interpreting p-values (see Gelman, Carlin, Stern & Rubin, 2003, for more extensive discussion). In short, Bayesian probability estimates provide an indication of the extent to which the observed patterns are significant and are not due to chance.

Yuan and MacKinnon (2009) put it nicely when discussing the difference between Bayesian credible intervals relative to more well-known Confidence Intervals (CIs) when they commented that “Bayesian credible intervals have more natural probability interpretations than CIs. A 95% credible interval means that there is a 95% chance that the credible interval contains the true value of the parameter on the basis of the observed data” (p. 304). Because of this, credible intervals in Bayesian analysis do not have to be symmetric. Rather, intervals are determined based on the specific percentile values around the

distribution of each parameter (known as the posterior distribution). For example, the 95% credible interval for the posterior distribution of a regression slope, such as those that we estimate here, would take the 2.5 and 97.5 percentile values of the posterior distribution, thus accommodating for any potential skew in the distribution of coefficient estimates.

The p-values we report in our analysis thus reflect the proportion of the posterior distribution for a given parameter (or regression slope) that is above or below zero. Muthén (2010, p. 7) summarized the nature of p-values in Bayesian path analysis as follows: "For a positive estimate, the p-value is the proportion of the posterior distribution that is below zero. For a negative estimate, the p-value is the proportion of the posterior distribution that is above zero." An advantage of Bayesian methods of estimation is the specification of prior information in terms of the level of uncertainty about a parameter, which allows for direct replication of, or building on, previous work. One can specify two aspects of the prior information: the point estimate and the distribution of the parameter in question. Alternatively, diffuse priors can be used such that no specifications are made in terms of the prior point estimate or the distribution of the parameter. That is, a diffuse or "flat" prior distribution is used in estimation. Our model used diffuse priors and was conducted in Mplus 7.0.

In order to overcome the problems associated with measurement error, we constructed a Bayesian Structural Equation Model estimating the test re-test stability of each latent Mini-IPIP6 factor over a one-year period. This allowed us to estimate the stability of each of the six Mini-IPIP6 personality subscales while adjusting for measurement error (MPlus syntax specifying this model is available online or on request; see Author's note for details).

Results

Descriptive statistics and frequentist bivariate correlations for the mean personality variables at Time 1 and at Time 2 are presented in Table 2.1. Test-retest correlations (represented in bold in Table 1) indicated reasonable levels of stability the six personality dimensions (*r*s = .753, .618, .691, .686, .658, .714, for Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility, respectively).

Table 2.1. Descriptive statistics and bivariate correlations between the mean personality variables at Time 1 and Time 2

	1	2	3	4	5	6	7	8	9	10	11	12
1 Extraversion T1												
2 Extraversion T2	**.753***											
3 Agreeableness T1	.213*	.163*										
4 Agreeableness T2	.173*	.187*	**.618***									
5 Conscientiousness T1	.016	.011	.140*	.106*								
6 Conscientiousness T2	.033*	.045*	.113*	.162*	**.691***							
7 Neuroticism T1	-.088*	-.058*	-.033*	-.025	-.112*	-.115*						
8 Neuroticism T2	-.075*	-.095*	-.027	-.053*	-.105*	-.115*	**.686***					
9 Openness to Experience T1	.264*	.212*	-.233*	.190*	.0021	.014	-.010	-.006				
10 Openness to Experience T2	.214*	.234*	.199*	.252*	-.009	.050*	.007	-.010	**.658***			
11 Honesty-Humility T1	-.106*	-.104*	.156*	.147*	.072*	.074*	-.192*	-.183	.018	.011		
12. Honesty-Humility T2	-.125*	-.148*	.153*	.185*	.072*	.100*	-.144*	-.178*	.013	.019	**.714***	
M	4.00	3.97	5.27	5.26	5.17	5.12	3.39	3.45	4.76	4.73	5.15	5.11
SD	1.16	1.12	.98	.96	1.05	1.02	1.10	1.08	1.12	1.09	1.32	1.30

N = 4,289; * $p < .05$.

Preliminary CFA measurement analysis using Maximum Likelihood estimation was conducted to assess the fit of the base factorial solution using the Time 1 and Time 2 data. A twelve factor solution was conducted representing each of the six latent personality variables at Time 1 and at Time 2. Each latent factor was estimated by its four corresponding indicators such that Extraversion at Time 1 was estimated by the four Time 1 Extraversion items, and Extraversion at Time 2 was estimated by the corresponding Time 2 items, and so forth. Items were allowed to relate only to the hypothesised latent factor at that point in time. Factors

were allowed to correlate. This solution provided an acceptable fit to the data, with a significant chi square test of absolute fit ($X^2(1014) = 19568.911$, $p < .001$). Tests of relative model fit provided results in acceptable ranges (RMSEA = .065, 95%CI = .065, .066; sRMR = .060). These results indicate that the Mini-IPIP6 reasonably captures the six distinct personality dimensions at both time points.

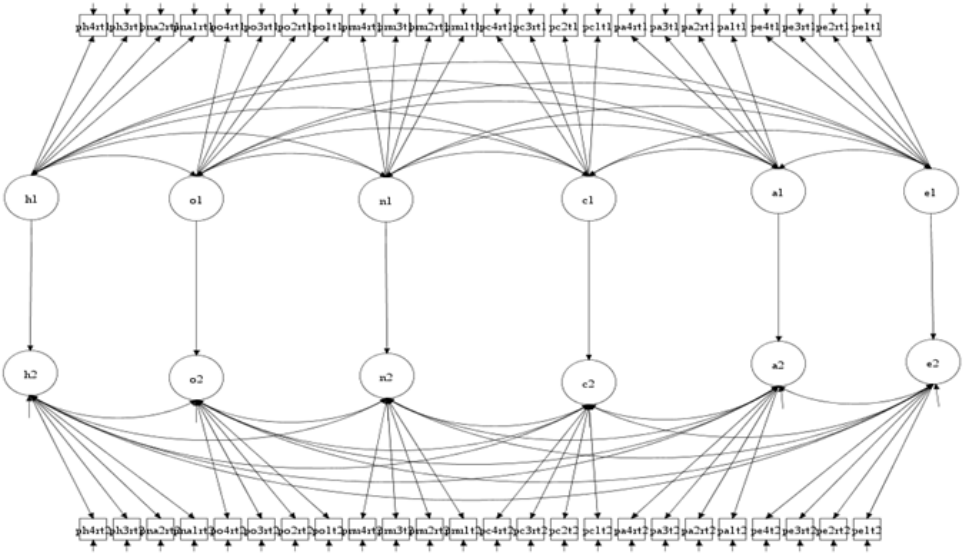


Figure 2.1. Overview of the Bayesian stability model.

Main Bayesian Model

As with our prior CFA, to test the stability of the Mini-IPIP6 in our focal model, the four marker items for each scale were allowed to relate only to the hypothesized latent variable at that same point in time. The six Time 1 latent personality factors were allowed to correlate. We also allowed the six Time 2 latent personality factors to relate to one another, thus adjusting for their residual association. Finally, we allowed each Time 1 latent factor to predict its Time 2 counterpart. These regression slopes for each Time 1 factor predicting its Time 2 counterpart thus represent stability estimates for each Mini-IPIP6 subscale after adjusting for measurement error. This Structural Equation Model is presented in Figure 2.1. The standardised factor loadings for each latent factor estimated are presented in Table 2.2.

Table 2.2. Standardized factor loadings for the Bayesian and the ML model at Time 1 and Time 2.

	Bayesian Model		ML Model	
	Time 1	Time2	Time1	Time 2
Extraversion				
Am the life of the party.	.738	.718	.688	.660
Don't talk a lot. (reverse scored)	.550	.549	.570	.572
Keep in the background. (reverse scored)	.604	.595	.635	.627
Talk to a lot of different people at parties.	.668	.652	.633	.612
Agreeableness				
Sympathize with others' feelings.	.718	.705	.612	.605
Am not interested in other people's problems. (reverse scored)	.475	.481	.552	.560
Feel others' emotions.	.616	.634	.525	.547
Am not really interested in others.	.568	.580	.656	.663
Conscientiousness				
Get chores done right away.	.690	.690	.584	.590
Like order.	.548	.554	.518	.524
Make a mess of things. (reverse scored)	.511	.491	.559	.538
Often forget to put things back in their proper place. (reverse scored)	.557	.558	.588	.585
Neuroticism				
Have frequent mood swings.	.708	.714	.705	.706
Am relaxed most of the time. (reverse scored)	.541	.563	.519	.552
Get upset easily.	.623	.611	.626	.614
Seldom feel blue. (reverse scored)	.452	.462	.438	.450
Openness to Experience				
Have a vivid imagination.	.458	.455	.346	.342
Have difficulty understanding abstract ideas. (reverse scored)	.644	.655	.665	.695
Do not have a good imagination. (reverse scored)	.550	.542	.497	.481
Am not interested in abstract ideas. (reverse scored)	.675	.681	.730	.743
Honesty-Humility				
Would like to be seen driving around in a very expensive car.	.571	.563	.589	.572
Would get a lot of pleasure from owning expensive luxury goods.	.676	.697	.690	.705
Feel entitled to more of everything.	.747	.779	.734	.771
Deserve more things in life.	.718	.743	.701	.732

N = 4,289. All loadings are significant at the .05 level

Personality Stability Estimates

The regression coefficients representing the one-year stability estimates for each personality factor are presented in Table 2.3. The standardised point estimates of the regression parameters for each pair of the Time 1 and Time 2 latent factors were all high. Extraversion showed a standardised stability estimate of $\beta = .982$, with a posterior $SD = .004$, $p < .001$, and 95% C.I. = .974, .988. Likewise, the standardised stability estimate for Agreeableness was $\beta = .908$, with a posterior $SD = .010$, $p < .001$, and 95% C.I. = .888, .927.

Conscientiousness showed a standardised stability estimate of β = .978, with a posterior *SD* = .004, p < .001, and 95% C.I. = .968, .985. For Neuroticism, the standardised stability estimate was β = .979, with a posterior *SD* = .006, p < .001, and 95% C.I. = .966, .988. The standardised stability estimate for Openness to Experience was β = .926, with a posterior *SD* =.011, p < .001, and 95% C.I. = .902, .946. Finally, Honesty-Humility showed a standardised stability estimate of β = .903, with a posterior *SD* = .008, p < .001, and 95% C.I. = .887, .918. These results indicate extremely high levels of stability across a 1-year period for each subscale of the Mini-IPIP6, after adjusting for measurement error and using Bayesian estimation with non-informative priors.

Table 2.3. Standardized Bayesian regression coefficients representing the stability of each latent Mini-IPIP6 personality factor over a one year period (see Figure 1 for model overview).

Stability coefficients	*β*	*Posterior SD*	*95% C.I. (lower 2.5%, upper 2.5%)*	*p*
Extraversion	.982	.004	.974, .988	.000
Agreeableness	.908	.010	.888, .927	.000
Conscientiousness	.978	.004	.968, .985	.000
Neuroticism	.979	.006	.966, .988	.000
Openness to Experience	.926	.011	.902, .946	.000
Honesty-Humility	.903	.008	.887, .918	.000

N = 4,289. Model used diffuse priors; *p*-values were one-tailed and give the proportion of the posterior distribution below or above 0 (for positive and negative estimates, respectively). 95% Confidence Interval for difference between observed and the replicated χ^2 values = 18300.752, 18570.900]. Bayesian Information Criterion = 698896.840.

In terms of the differences in relative stability between the six personality dimensions, we conducted difference tests using the unstandardized stability estimates and Bayesian posterior distributions. Save for a few exceptions, these analyses largely supported our predictions that the different Mini-IPIP6 subscales would demonstrate comparable levels of stability across time. Small differences in stability estimates were observed between the following three contrast pairs: (a) Extraversion and Agreeableness (*b* = .083, posterior SD =

.033, 95% C.I. = .019, .148, p = .006), (b) Agreeableness and Conscientiousness (b = -.091, posterior SD = .036, 95% C.I. = -.159, -.021, p = .004), and (c) Agreeableness and Neuroticism (b = -.102, posterior SD = .034, 95% C.I. = -.171, -.037, p = .001). These findings suggest that Agreeableness is slightly less stable over a one year period than Extraversion, Conscientiousness, and Neuroticism. However, while still significant, the magnitudes of these differences were relatively small and arguably trivial.

Frequentist Model with Correlated Residuals

We examined a more restrictive stability model in which we (a) allowed the residuals of each item to correlate across our two time points, (b) estimated using Maximum Likelihood. The model thus adjusted for shared residual error unique to each item, or item pair, which was consistent across the one year test-retest period. As such, variance in the prior models that was attributed to the latent factor of each personality dimension was re-partitioned to the shared residual variance unique to each item. In this model, standardised regression parameters thus represented (Frequentist) stability estimates for each Mini-IPIP6 subscale after adjusting for both measurement error and shared residual item-level error variance over the test-retest period.

The model was estimated using Full Information Maximum Likelihood, which allowed us to account for missing data on the individual responses to the Mini-IPIP6 items at each time point. Not surprisingly given our sample size, a chi-square test of absolute model fit was significant ($\chi^2(1020) = 9750.017$, $p < .001$). This indicates that the model fit the data significantly worse than a fully-saturated baseline model. However, the model fit fairly well according to the relative fit indices. The RMSEA was .045, and the SRMR was .053. The standardised factor loadings for each latent factor estimated (Personality variables at Time 1 and Time 2) are presented in Table 2.2. The residual variance correlations between each item

pair across the retest period were all significant, as would be expected given the sample size. The ranges of the parameters for each set of item pairs were as follows: for Extraversion, the parameters ranged from b = .242 to .461 (t = 13.179 to 29.125); for Agreeableness, the parameters ranged from b = .120 to .390 (t = 6.360 to 21.556); for Conscientiousness, the parameters ranged from b = .280 to .539 (t = 16.191 to 38.344); for Neuroticism, the parameters ranged from b = .283 to .397 (t = 16.058 to 27.145); for Openness to Experience, the parameters ranged from b = .085 to .580 (t = 3.361 to 54.397); and for Honesty-Humility, the parameters ranged from b = .282 to .429 (t = 17.043 to 26.570).

Considering the imposed correlations of the item pair residual variances, this model presented a more stringent (i.e., conservative) test of the one-year stability of the six Mini-IPIP6 personality subscales. As such, it was expected that the standardised regression parameters for each latent variable pair at Time 1 predicting Time 2 would be slightly lower in comparison to the previously-presented model based on Bayesian estimates.

In this model, Extraversion showed a standardised regression estimate of β = .905 (*SE* = .008). The standardised regression estimate for Agreeableness was β = .814 (*SE* = .012). Conscientiousness had a standardised regression estimate of β = .880 (*SE* = .011). Neuroticism had a standardised regression estimate of β = .874 (*SE* = .010). The standardised regression estimate for Openness to Experience was β = .814 (*SE* = .012). Finally, the standardised regression estimate for Honesty-Humility was β = .805 (*SE* = .009; $p < .001$ for all estimates).

These stability estimates, while lower than those reported in our Bayesian model, still demonstrate remarkably high levels of test-retest reliability over a one year period for each of the six Mini-IPIP6 subscales. Critically, this model adjusted for both measurement error and shared residual item-level error over the test-retest period. Most of the estimates were

somewhat lower than we would have initially expected of a good personality measure over a short-term period (most of the regression coefficients were $\beta < .90$). Considering that the one-year time frame is of a moderate length in terms of test-retest assessments, as well as the conditions imposed on this model by correlating the residual variances of individual item pairs, these parameters indicate impressively high levels of stability for each of these personality scales.

Discussion

Arguably one of the most central components of the operationalization of personality is that it represents "relatively *enduring* styles of thinking, feeling, and acting" (McCrae & Costa, 1997, p. 509; emphasis added). Indeed, personality is - by definition - a temporally stabile construct. As such, test-retest reliability should be one of the most important psychometric properties for a personality measure, whether it is long-form (e.g., NEO-FFI, IPIP-FFM), or the recently-developed short-form scales (e.g., Mini-IPIP). This study investigated the stability of personality as measured by a recently developed short-form scale of the Big-Six (the Mini-IPIP6; see Sibley et al., 2011).

Our aim in this study was to establish the test-retest reliability of the Mini-IPIP6 over a one year period. The Mini-IPIP6 (Sibley et al., 2011) is an extension of the earlier Mini-IPIP developed by Donnellan and colleagues (2006). Whereas the Mini-IPIP is a 20-item measure of the Big-Five, the Mini-IPIP6 includes an additional four items to assess a sixth personality dimension (namely, Honesty-Humility; see Ashton & Lee, 2007, 2009). Therefore, by assessing the temporal stability of the Mini-IPIP6, we were able to assess a critical component of the operationalization of both Donnellan and colleagues' (2006) Mini-IPIP five factor measure and the Mini-IPIP6 (Sibley et al., 2011).

We assessed the test-retest reliability of these scales using three differently-estimated models in a large nationally representative sample of New Zealand adults. The results of our analyses show that the Mini-IPIP6 factorial structure holds in the New Zealand context and that the different dimensions of personality indexed by the Mini-IPIP6 (and, by extension, the Mini-IPIP) show high levels of stability over a one year period that replicate across estimation methods. Our results also demonstrate the utility of the Bayesian SEM approach to assessing the test-retest reliability of measurement scales, as evidenced by the difference between the traditional test-retest correlation estimates and the stability estimates produced by our Bayesian-based SEM models (namely, the former approach yielded slightly higher stability estimates than the latter approach).

Reliability of Short-form Scales

The current study extends the validation of the Mini-IPIP6 by providing an assessment of the test-retest reliability of the scale across a one year time frame. Moreover, we use a large nationally representative sample and advanced model estimation methods in order to ensure that our models are both (a) representative of the population and (b) based on sound methodological grounds. Crucially, our results complement those of Sibley and colleagues (2011) – as well as those of Donnellan and colleagues (2006) - by showing that each of the six personality dimensions assessed by our scale was remarkably stable over a one year period. Indeed, the standardised stability coefficients for each dimension of the Big-Six were around $\beta = .95$ or above (depending on the estimation method used). Furthermore, there were relatively few differences in terms of the size of these stability estimates observed across these six dimensions. That said, our results show that Agreeableness is less stable than (a) Extraversion, (b) Conscientiousness, and (c) Neuroticism. Nevertheless, the differences in these stability estimates are trivial in magnitude. Therefore, the five dimensions of Extraversion, Agreeableness, Conscientiousness, Openness to Experience, and Neuroticism

assessed by the Mini-IPIP scales, as well as the sixth dimension of Honesty-Humility added by the Mini-IPIP6, had remarkably high levels of temporal stability.

High Standard for Short-term Stability

The original Mini-IPIP and the Mini-IPIP6 have both shown favourable psychometric properties in terms of reliability, as well as construct and discriminant validity (Donnellan et al., 2006). Our analyses contribute to this literature by replicating these results using a nationally representative sample of the adult population in New Zealand. In addition, our analyses were done using multiple estimation methods including Bayesian and Maximum Likelihood-based SEM to control for both measurement error and shared residual item-level error in our third estimation model. As predicted, our results demonstrated high levels of test-retest reliability within a nationally representative sample, although we did find that the stability estimates were slightly lower in the most conservative test of our model. Therefore, we argue that such stringent expectations constitute a standard that should be expected from every "good" personality scale over a short to moderate-term test-retest period.

Caveats and Future Directions

There are several strengths to our approach to assessing the test-retest reliability of the short-form personality scales. The models were tested within a large nationally representative sample of New Zealand adults. More importantly, we utilised Structural Equation Modelling techniques to address some of the issues raised in the literature associated with measurement error in tests of personality stability (e.g. Ferguson, 2010). Specifically, we advance the use of Bayesian estimation methods for SEM and demonstrate the utility of such models in testing the stability of personality. Investigations across age cohorts or over specific developmental periods would benefit from utilising the stability models implemented in the current study in order to examine differences in the stability of personality across critical

predictor variables (e.g., age). The use of SEM techniques within these contexts is rising in popularity, as the increased availability of large samples has rendered such techniques possible (e.g., Donnellan & Lucas, 2011; Rantanen et al., 2007). We suggest that these estimations could be further improved through Bayesian methods which allow for the estimation of posterior distributions as well as the specification of prior point estimates and variable distributions. Such an approach allows for research to directly build upon past studies. It is hoped that the methods for estimating the stability of personality will help in the development of future research in the area.

Mini-IPIP6 Psychometrics and Concluding Comments

The current study builds upon the work of Sibley and colleagues (2011) by assessing the test-retest reliability of the Mini-IPIP6. In doing so, we help further establish the psychometric properties of the scale, as well as the original Mini-IPIP (Donnellan et al., 2006) upon which it is based. The Mini-IPIP6 has thus far been shown to be both a valid and reliable measure of the Big-Six. In terms of stability, the Mini-IPIP6 performed exceptionally well under the stringent standards imposed in the current study. This is particularly impressive, considering that this measure, like its five factor predecessor, is a short-form scale measuring broad bandwidth personality dimensions with only four items per trait.

The Mini-IPIP6 (Sibley et al., 2011), and the Mini-IPIP (Donnellan et al., 2006), address a fundamental conceptual issue presented by traditional personality measures. These short-form scales provide a practical tool for personality assessment in situations where long-form scales may be too costly and time consuming, while at the same time retaining the validity and the reliability of their long-form counterparts. The Mini-IPIP6 is, therefore, a viable alternative to long-form measures of the six factor personality model. This is particularly important, as the Big-Six is currently lacking short form measures that have

proliferated in recent work on the Big-Five (e.g., Donnellan et al., 2006; Gosling, Rentfrow & Swann, 2003). Importantly, the Mini-IPIP6 assesses the traditional Big-Five personality dimensions as well as the Honesty-Humility personality dimension proposed by the six-factor HEXACO model (Ashton & Lee, 2007, 2009). Thus, we present a practical alternative (with only 24 items) to long-form measures that is openly available to researchers. It is hoped that the availability and psychometric strengths of measures such as the Mini-IPIP6 and the Mini-IPIP will increase their use in personality research.

Authors' Note. Collection of the New Zealand Attitudes and Values Study data analyzed in this paper was funded by University of Auckland FRDF (#3624435/9853) and ECREA (#3626075) grants awarded to Chris G. Sibley. Mplus syntax for the models reported here (including syntax documenting the model constraint commands we implemented in our analyses) will be posted on the NZAVS website upon acceptance of this article. Syntax and data are also available upon request for reviewing purposes.

http://www.psych.auckland.ac.nz/uoa/NZAVS

Bridging Comments

Study 1 was an important foundational study of the present thesis serving two broad purposes. Firstly, the findings of the models employed demonstrated the expected temporal stability of personality tarts across a moderate – a 1-year – test-re-test period. Indeed, temporal stability is a crucial property of a personality trait. Study 1 thus supported the predicted very high stability of the Big Six personality traits – Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility – in a nationally representative sample of adult New Zealanders. Moreover, the modelling strategy employed – the BSEM – provided accurate estimates of rank-orders stability, correcting for the possible attenuation due to measurement error. Given this modelling approach, the stability estimates for all of the personality traits were around $\beta = .90$ or higher indicating extremely high sample-level stability over a moderate test-re-test period.

Secondly, this study provided an important test of the test-re-test reliability of the Mini-IPIP6 – the short form personality scales used throughout all of the subsequent studies. Given that the conceptualisation of the construct of a personality trait necessitates high temporal stability, high test-re-test reliability is a necessary requirement of any reliable personality measure. While the test-re-test reliability of the original Mini-IPIP had been previously demonstrated (Donnellan et al., 2006), Study 1 demonstrated the test-re-test reliability of the extended Mini-IPIP6 in the NZAVS sample over a longer time period. This study thus established the reliability if the Mini-IPIP6, paving the way of the subsequent three studies.

Having established the temporal stability of the Big Six personality traits as assessed by the short form Mini-IPIP6, study 2 sought to investigate the developmental pattern of stability in these personality traits across the cohorts of the adult life span – from 20 to 80

years old. Such developmental patterns provide an important test of the different perspectives on personality development. On the one hand, the Five Factor Theory (McCrae & Costa, 1999) would predict that the stability of personality should increase across young adulthood, reaching peak stability around the age of 30, and remaining unchanged and unchangeable thereafter. On the other hand, the Neo-Socioanalytic perspective (Roberts et al., 2008) argues for a continuing process of person-environment interactions, suggesting that the rank-order stability of personality traits should vary systematically across the life span. Specifically, in the periods of life that are normatively associated with more life-changes – such as younger and older adulthood – personality traits should be less stable.

Indeed, extant research has documented that while highly stable at the sample level, as supported by study 1, the rank-order stability of personality tends to vary across the different cohorts of the life span (e.g., Ardelt, 2000; Ferguson, 2010). While most of such research agrees that personality traits increase in stability through younger adulthood towards middle age, recent studies have demonstrated a systematic destabilisation beyond middle age and towards older age (Lucas ad Donnellan, 2011; Specht et al., 2011; and Wortman et al., 2011). Specifically, these studies suggest that personality traits tend to stabilise through younger adulthood, being at their most stable in middle age. Following this, however, this stability tends to decrease into older age. These findings support the suggestions of continued personality change and development showing lower stability of personality traits in periods of life that are associated with more normative changes in terms of both social roles and institutions, as well as cognitive and biological factors.

Study 2 sought to test these developmental patterns in a nationally representative sample of adult New Zealanders. Furthermore, this study sought to provide a much needed replication of the recent findings of systematic variation of personality stability across adulthood (Lucas ad Donnellan, 2011; Specht et al., 2011; and Wortman et al., 2011).

Chapter III

Study 2: "The Stability of Adult Personality Varies across Age: Evidence from a two-year longitudinal sample of adult New Zealanders"

Reference:

Milojev, P. & Sibley, C.G. (2014). The stability of adult personality varies across age: Evidence from a two-year longitudinal sample of adult New Zealanders. *Journal of Research in Personality, 51,* 29-37. http://dx.doi.org/10.1016/j.jrp.2014.04.005.

Abstract

We model developmental patterns in the stability of Big-Six personality markers over a two-year period in a multi-cohort panel study of adult New Zealanders (n = 3910). Structural Equation Models testing patterns of linear, quadratic and cubic effects identified considerable variation in the stability of most aspects of personality across the adult age range (20-80 years). Agreeableness showed a slight linear decrease in consistency with increasing age. Extraversion, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility exhibited quadratic 'inverted U' patterns of rank-order stability. Unlike Agreeableness, the two-year consistency of these other personality dimensions increased from young adulthood towards middle age, then decreased in stability as people progressed toward old age. Implications for personality research and recently emerging trends are discussed.

Introduction

The stability of personality traits across the life-span is a central consideration in personality psychology. Indeed, the extent to which personality is stable has implications for the very definition of personality as a construct (Caspi, Roberts, & Shiner, 2005; Mischel,

1968; Roberts, Wood, & Caspi, 2008). Although personality traits, such as those assessed by the Big Five or the Big Six frameworks, show considerable temporal consistency over moderate time periods, systematic variation in stability has been observed over different portions of the life-span (Ardelt, 2000; Ferguson, 2010; Lucas & Donnellan, 2011; Milojev, Osborne, Greaves, Barlow, & Sibley, 2013; Mischel & Shoda, 1995; Roberts & DelVecchio, 2000; Specht, Egloff, & Schmukle, 2011; Wortman, Lucas, & Donnellan, 2012). Recent reports have documented emerging trends in personality development that suggest curvilinear developmental patterns characterised by increases in stability towards middle age. This then tends to be followed by a decrease in personality stability as people progress toward old age (Lucas & Donnellan, 2011; Specht, Egloff, & Schmukle, 2011; Wortman, Lucas, & Donnellan, 2012).

In this study we assess the distribution of the two-year differential or rank-order stability estimates of the Big Six personality dimensions: Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility. We do so across a broad section of the adult life span (from those aged 20 to 80 years). To these ends we construct Structural Equation Models with latent interaction variables to analyse data from the New Zealand Attitudes and Values Study (NZAVS). This is a large nationally representative sample of New Zealanders from a multi-cohort panel study. We thus contribute to the growing but still scarce corpus of longitudinal data on personality by providing both a replication and an extension of previous research on the stability of personality in different developmental periods in adulthood. By using the latent variable approach our models provide estimates of differential stability controlling for the effects of measurement error that has been found to affect and attenuate personality stability estimates under certain conditions, particularly over short re-test periods (e.g., Ferguson, 2010).

Finally, we also extend previous research by investigating the stability of the Big Six personality dimensions.

The Six Factor Model of Personality and Study Overview

In order to avoid splitting the sample by cohorts (e.g., Lucas & Donnellan, 2011; Wortman, Lucas, & Donnellan, 2012), and to account for the issues that might arise from measurement error (e.g., Ferguson, 2010), we implement a Structural Equation Modelling approach to estimate stability coefficients for the latent personality traits over the two-year longitudinal period dependent upon participants' age. Our analytic strategy is thus similar to the models implemented by Specht, Egloff, and Schmukle (2011). As we outline in detail below, the models control for gender differences and test the linear, quadratic (Age^2) and cubic (Age^3) effects of age. The conditional stability for each participants' age is calculated to include both the linear and the quadratic (and the cubic where appropriate) effect of age on personality stability. Given the reasonably broad age-range the analyses test the hypothesised curvilinear patterns of stability, and thus provide a replication of extant research.

We assess marker items for the Big Six broad bandwidth personality dimensions over a two-year re-test period from 2009 to 2011 using the recently developed Mini-IPIP6 scale (Sibley et al., 2011). Recent developments in the research on personality structure have extended the Big Five model of personality to include a sixth independent personality dimension – Honesty-Humility (Ashton & Lee, 2007, 2009). This six-factor model – referred to as the HEXACO model – is argued to account for subtle distinctions between personality traits better than does the Big Five (Ashton & Lee, 2009). The Mini-IPIP6 is an extension of the original five-factor Mini-IPIP (Donnellan, Oswald, & Lucas, 2006) that includes items marker items also assessing Honesty-Humility. The factor structure and internal reliability of the Mini-IPIP6 scale has been extensively validated (Sibley et al, 2011; Sibley, 2012; Sibley

& Pirie, 2013). Previous analyses of NZAVS data further indicate that the Mini-IPIP6 personality markers have excellent rank-order stability over a one-year period (Milojev et al, 2013).

In line with extant research we expected to find support for the life-span perspective of personality development and to observe the 'inverted U' pattern of distribution of the personality stability estimates across the life -span (Lucas & Donnellan, 2011; Specht, Egloff, & Schmukle, 2011; Wortman, Lucas, & Donnellan, 2012). We expected this pattern to emerge for all six of the Big Six dimensions – Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility.

Methods

Participants and Sampling Procedure

This report is based on data from the NZAVS collected in 2009 (Time 1) and 2011 (Time 3). In 2009, invitations to participate in a mail-based survey were sent to a total of 40,500 people listed on the electoral roll (i.e., a compulsory list of eligible voters). After adjusting for inaccuracy on the electoral roll, the response rate to the Time 1 survey was 16.6% (n = 6,518). %). Of these, 3,914 also participated at Time 3 (retention = 60.0%). Due to the longitudinal focus of our investigation, our analyses were limited to the retained sample of 3901 (2409 women, 1492 men) who responded to the relevant measures[3].

Table 3.1 presents the sample size across the 20 to 80 age range. Note that we based our models using the full longitudinal sample including all ages, not just the 95.6% aged

[3] The Time 1 and the retained longitudinal sample at Time 3 had a gender bias, with 2414 (61.7%) women and 1496 (38.3%) men responding to both time points. In contrast, 52% of adult New Zealand residents are women and 48% are men (Statistics New Zealand, 2006). The longitudinal sample was also more likely to include Europeans (n = 3395, 86.7%) relative to 2006 census estimates (68.7%). In most other respects, the NZAVS was reasonably consistent with national population estimates. Please see the NZAVS website for additional technical information (http://www.psych.auckland.ac.nz/en/about/our-research/research-groups/new-zealand-attitudes-and-values-study/nzavs-tech-docs.html).

between 20-80 years. However, we focused on estimating stability coefficients for those aged between 20 and 80 because we were wary about estimating parameters based on diminishing sample sizes beyond this age range.

As of 2009, 53.6% of the sample (n = 2090) identified as non-religious (2.1% not reporting their religious status). 77.9% (*n* = 3040) of the sample were parents, 73% (*n* = 2848, 2.3% not responding) were in some kind of employment, and 20.1% (*n* = 783) were born overseas. Mean age of the sample was 50.29 (SD = 15.06). Finally, socioeconomic status was assessed using the NZDep 2006 (see Salmond, Crampton, & Atkinson, 2007) index, ranging from 1 to 10 with higher numbers indicating more deprived residential areas. As such, mean neighbourhood deprivation was 4.82 (SD = 2.79).[4]

Table 3.1. Sample size and the proportion of women across the age-range split into 5-year cohorts for clarity.

Age Cohort	*N*	Proportion of women
21-25 years	135	.711
25-30 years	173	.769
31-35 years	275	.709
36-40 years	353	.680
41-45 years	434	.631
46-50 years	488	.600
51-55 years	478	.615
56-60 years	425	.560
61-65 years	399	.591
66-70 years	271	.557
71-75 years	194	.521
76-80 years	104	.625

Note: People in the age-range from 20 to 80 comprise 95.6% (*n* = 3834) of the full (*n* = 3901) sample. The models were conducted using the full sample. The conditional simple slopes were calculated for people aged from 20 to 80.

[4] Participants who responded in both 2009 and 2011 were more likely to be female (61.7% vs. 56.1% for those who did not follow-up; $\chi^2(1, 6518) = 20.946$, p<.001), more likely to be European (86.8% vs. 74.7%; $\chi^2(1, 6518) = 153.642$, p<.001), more likely to have been born in New Zealand (79.9% vs. 75.7%; $\chi^2(1, 6517) = 16.321$, p<.001),more likely to have children (77.9% vs. 70.3%; $\chi^2(1, 6518) = 26.134$, p<.001), and to be in a relationship (72.8% vs. 66.9%; $\chi^2(1, 6424) = 26.134$, p<.005) compared to respondents who did not follow up in 2011. They also lived in more affluent or less deprived neighbourhoods (NZDep = 4.82 (2.794) vs. 5.42 (2.885), F (1.6362) = 68.310, p<.001), and were older (50.32 (15.037) vs. 44.22 (16.149) F (1, 6362) = 235.268, p<.001).

Questionnaire Measures

The Mini-IPIP6 is a 24-item scale containing the 20 items developed by Goldberg (1999) and included by Donnellan, Oswald, Baird, and Lucas (2006) in the original Mini-IPIP. In addition, four items were used to index Honesty-Humility. Items H01 and H02 were adapted from the Narcissism scale developed by Campbell et al. (2004), whereas items H03 and H04 were adapted from Ashton and Lee's (2009) HEXACO measure of Honesty-Humility.

The items were administered with the following instructions: "This part of the questionnaire measures your personality. Please circle the number that best represents how accurately each statement describes you." Items were rated on a scale ranging from 1 (very inaccurate) to 7 (very accurate). Extraversion was assessed by the following four items: "Am the life of the party."; "Don't talk a lot." (reverse-scored); "Keep in the background." (reverse-scored); "Talk to a lot of different people at parties." ($\alpha_{\text{Time 1}} = .71$; $\alpha_{\text{Time 2}} = .75$). Agreeableness was assessed by the following four items: "Sympathize with others' feelings."; "Am not interested in other people's problems." (reverse-scored); "Feel others' emotions."; "Am not really interested in others." (reverse-scored) ($\alpha_{\text{Time 1}} = .66$; $\alpha_{\text{Time 2}} = .69$). Conscientiousness was assessed by the following four items: "Get chores done right away."; "Like order."; "Make a mess of things." (reverse-scored); "Often forget to put things back in their proper place." (reverse-scored) ($\alpha_{\text{Time 1}} = .65$; $\alpha_{\text{Time 2}}\ \alpha = .65$). The Neuroticism scale included the following four items: "Have frequent mood swings."; "Am relaxed most of the time." (reverse-scored); "Get upset easily."; "Seldom feel blue." (reverse-scored) ($\alpha_{\text{Time 1}} = .64$; $\alpha_{\text{Time 2}} = .72$). Openness to Experience was assessed by the following four items: "Have a vivid imagination."; "Have difficulty understanding abstract ideas." (reverse-scored); "Do not have a good imagination." (reverse-scored); "Am not interested in abstract ideas." (reverse-scored) ($\alpha_{\text{Time 1}} = .67$; $\alpha_{\text{Time 2}} = .70$). Finally, Honesty-Humility was assessed using

the four items as per Sibley and colleagues (2001). These items (all of which were reverse-scored) included: “Would like to be seen driving around in a very expensive car.”; “Would get a lot of pleasure from owning expensive luxury goods.”; “Feel entitled to more of everything.”; “Deserve more things in life.” ($\alpha_{\text{Time 1}} = .78$; $\alpha_{\text{Time 2}} = .77$).

Model Estimation Methods

We advance a model of the distribution of two-year rank-order stability estimates of the Big Six personality traits from age 20 to 80 years old. Using MPlus 7.1 modelling software (Muthén & Muthén, 1998-2012) we used SEM methods to estimate the moderation of the latent two-year (2009 – 2011) rank-order stability coefficients by age including both the linear and the curvilinear (quadratic and cubic) effects. All of our models control for the effects of gender. The models were estimated using Maximum Likelihood estimation with numeric integration algorithms, using random effects modelling to estimate the multiple latent interaction variables. The measurement component imposed strict measurement invariance for each personality dimension over the re-test period. The four marker items at Time 1 were specified to estimate the exogenous factor of the personality variable at Time 1. The four corresponding marker items at Time 2 were specified to estimate the endogenous latent personality variable at Time 2. The factor loading were constrained to equality across time, as were the residual variances of the items and the item intercepts. The residual item-variances were allowed to correlate over the repeated assessments. All of the parameters were standardised using the variances of the latent personality variables at Time 1 and Time 2. The standardised effects of the Time 1 latent personality variables on the corresponding Time 2 latent personality variables represent the estimates of rank-order stability over the 2-years.

As mentioned, our models included the latent interaction effects, testing the effects of gender, age, age^2, and age^3. Age was mean-centered to allow for easy interpretation of all

interaction effects and the simple slopes. As such, significant effects of interaction terms represent the effect of that demographic covariate on the differential stability of the given latent personality dimension. All of the models tested the full set of interactions; however, age^3 and age^2 were removed from the models where no significant effect was observed. Age and gender were included in all models even when not significant. Implementing constraints to these structural models we estimate constrained stability coefficients for the joint interactions at every age point in the available range – from 20 to 80 years old. We thus estimate a continuous effect of age on the two year rank-order stability of each of the Big Six personality dimensions. This approach avoids having to split the available sample into age cohorts as has been done in previous research (e.g., Lucas & Donnellan, 2011; Wortman, Lucas, & Donnellan, 2012), and also controls for errors of measurement through estimating the personality traits as latent variables (Ferguson, 2010; see Specht, Egloff, & Schmukle, 2011, for a similar approach). The model here described was applied to each of the six personality dimensions as assessed by the Mini-IPIP6 (Sibley et al., 2011) – Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility.

We introduced additional parameters estimating the stability coefficients for those aged 20 to 80 years. To estimate constrained stability coefficients including both the linear and the quadratic interaction terms we specified the simple slope (s_i) computation as:

$$s_i = za + (zab * A_i) + (zabb * A_i^2) + (zabbb * A_i^3)$$

Where *za* refers to the standardised effect of the personality variable at Time 1 on the corresponding personality variable at Time 2; *zab* refers to the standardised linear interaction effect; *zabb* refers to the standardised quadratic interaction effect; zabbb refers to the standardised cubic interaction effect; with *Ai* reflecting sample mean-centred age values from

20 to 80 years old. The cubic and the quadratic effects of age were only included in the equation if significant effects were observed. Otherwise they were excluded from the model. Thus, in addition to specifying the stability coefficients and the linear, quadratic and cubic moderation by age, we also estimated the corresponding 60 standardised simple slopes reflecting the distribution of the two-year stability coefficients across the adult life-span integrating the multiple effects of age on rank-order stability. These stability coefficients are presented in a figure with their estimated 95% Confidence Intervals.[5]

[5] Mplus syntax for the models here described is available through the NZAVS website, or by contacting the corresponding author.

Table 3.2. Descriptive statistics and bivariate correlations for the mean personality variables and age.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. T1.EXTRAVERSION														
2. T3.EXTRAVERSION	**.738****													
3. T1.AGREEABLENESS	.215**	.164**												
4. T3.AGREEABLENESS	.164**	.197**	**.587****											
5. T1.CONSCIENTIOUSNESS	.012	.017	.144**	.127**										
6. T3.CONSCIENTIOUSNESS	.027	.075**	.114**	.215**	**.683****									
7. T1.NEUROTICISM	-.095**	-.069**	-.036*	-.027	-.120**	-.114**								
8. T3.NEUROTICISM	-.098**	-.135**	-.023	-.086**	-.132**	-.193**	**.663****							
9. T1.OPENNESS	.244**	.191**	.222**	.159**	.015	-.010	-.022	-.034*						
10. T3.OPENNESS	.210**	.229**	.166**	.250**	-.017	.047**	-.018	-.054**	**.649****					
11. T1.HONESTY_HUMILITY	-.087**	-.081**	.156**	.153**	.091**	.083**	-.202**	-.171**	.021	.003				
12. T3.HONESTY_HUMILITY	-.097**	-.106**	.159**	.205**	.070**	.098**	-.146**	-.184**	.022	.041*	**.707****			
13. Gender T1	-.064**	-.057**	-.295**	-.298**	-.102**	-.095**	-.135**	-.132**	.012	.030	-.128**	-.135**		
14. Age.T1	-.080**	-.090**	-.002	-.021	.077**	.057**	-.181**	-.144**	-.164**	-.159**	.223**	.218**	.121**	
M	3.998	3.914	5.297	5.280	5.160	5.089	3.382	3.400	4.755	4.722	5.209	5.172	.38	50.29
SD	1.160	1.123	.971	.930	1.055	1.003	1.096	1.075	1.112	1.086	1.296	1.250	.486	15.056

**. Correlation is significant at the 0.01 level (2-tailed); *. Correlation is significant at the 0.05 level (2-tailed).

Results

Bivariate correlations and descriptive statistics of the averaged personality scores at both time points, as well as the relevant demographics are presented in Table 3.2. Simple re-test correlations are presented in bold font. As can be seen in the table, the two-year re-test correlations for the mean personality variables (*r* = .738, .587, .683, .663, .649, .707, for Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility, respectively) are comparable, if only marginally smaller than the one-year re-test correlations for the mean personality variables reported by Milojev et al. (2013; *r* = .753, .618, .691, .686, .658, .714, for Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility, respectively).[6]

With respect to the measurement component of our SEMs, we conducted tests of measurement invariance, over time and across age for each of the six personality dimensions. The respective model fit indices are presented in Table 3.3. Specifically, we tested the baseline configural model over the two time points, subsequently testing more restrictive metric and scalar models. We then tested the measurement model specifying strict scalar temporal invariance across age by splitting the participant sample into 5-year cohorts and running tests of multi-group comparisons. As can be seen in the table, our measurement specifications show reasonable fit for the strict invariance over time and between age-groups for Extraversion, Conscientiousness, and Neuroticism, and less than optimal fit for Agreeableness, Openness to Experience, and Honesty-humility. However, as the strict factorial invariance solution provided the best fit across the six dimensions, we ran our SEMs as detailed above, specifying a very strict and conservative factorial solution. Specifically, our measurement specifications constrained the factor loadings over time to equality. The

[6] While mean-level change is not a focus of this investigation, simple analyses indicated a slight decrease in mean-level Extraversion ($t(3274) = 5.806$, $p < .001$) and a slight decrease in mean-level Conscientiousness ($t(3275) = 3.092$, $p = .002$) over the re-test period.

same constraints were applied to the item intercepts, and item residual variances. The item level residual variances were allowed to correlate over repeated assessments.

Table 3.4 presents the standardised stability estimates over the two-year re-test period as well as the effects of gender and age derived from the SEMs as specified above. We also present the estimates of model fit - χ^2, RMSEA, CFI, and SRMR – derived form comparable models excluding the latent interaction terms. The latent two-year stability estimates estimated here are slightly smaller in magnitude than the one-year Bayesian stability coefficients reported by Milojev et al. (2013; $\beta > .09$) using the same scale and a similar sample. This is to be expected given the increased re-test period and the statistical controlling for the effects of age and gender. However, our models indicate high stability of personality across all six broad bandwidth dimensions, as wold be expected - $\beta = .922, .728, .910, .853, .841, .824$; for Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility, respectively.

Table 3.3. Fit indices for tests of Measurement Invariance.

Model	χ^2	*df*	RMSEA	CFI	TLI	SRMR
Extraversion						
1. Configural over time	363.199	*15*	.077	.955	.916	.039
2. Metric over time	384.197	*19*	.070	.953	.931	.041
3. Scalar over time	403.585	*22*	.067	.951	.937	.041
4. Between age-groups configural	552.960	*176*	.078	.953	.918	.050
5. Between age-groups metric	624.080	*229*	.070	.951	.934	.058
6. Between age-groups scalar	816.829	*303*	.069	.936	.935	.067
Agreeableness						
1. Configural over time	640.153	*15*	.103	.865	.749	.062
2. Metric over time	682.584	*19*	.095	.857	.789	.085
3. Scalar over time	710.789	*22*	.089	.852	.811	.086
4. Between age-groups configural	874.065	*176*	.106	.861	.757	.088
5. Between age-groups metric	947.220	*229*	.094	.857	.808	.086
6. Between age-groups scalar	1080.609	*303*	.085	.845	.843	.088
Conscientiousness						
1. Configural over time	111.349	*15*	.041	.981	.964	.024
2. Metric over time	128.459	*19*	.038	.978	.968	.032
3. Scalar over time	139.433	*22*	.037	.977	.970	.033
4. Between age-groups configural	303.292	*176*	.045	.976	.958	.046
5. Between age-groups metric	373.199	*229*	.042	.973	.964	.056
6. Between age-groups scalar	567.501	*303*	.050	.951	.950	.072
Neuroticism						
1. Configural over time	130.404	*15*	.044	.980	.962	.024
2. Metric over time	613.473	*19*	.089	.895	.845	.106
3. Scalar over time	151.767	*22*	.039	.977	.971	.028
4. Between age-groups configural	336.989	*176*	.051	.972	.951	.040
5. Between age-groups metric	397.505	*229*	.046	.971	.961	.049
6. Between age-groups scalar	708.482	*303*	.061	.930	.928	.077
Openness to Experience						
1. Configural over time	655.189	*15*	.104	.890	.794	.074
2. Metric over time	909.924	*19*	.110	.846	.774	.119
3. Scalar over time	945.991	*22*	.104	.841	.797	.119
4. Between age-groups configural	1111.915	*176*	.123	.842	.724	.112
5. Between age-groups metric	1066.856	*229*	.102	.859	.810	.133
6. Between age-groups scalar	1277.135	*303*	.095	.836	.833	.158
Honesty-Humility						
1. Configural over time	648.869	*15*	.104	.921	.853	.056
2. Metric over time	763.561	*19*	.100	.907	.863	.099
3. Scalar over time	1058.524	*22*	.110	.871	.836	.105
4. Between age-groups configural	962.218	*176*	.112	.902	.829	.106
5. Between age-groups metric	949.256	*229*	.094	.911	.880	.077
6. Between age-groups scalar	1521.491	*303*	.107	.849	.846	.113

Notes. Strict Scalar Factorial invariance over time tested between age groups.

No cubic effects of age (Age^3) were identified in our models and the respective latent interaction terms were removed from the analyses. As can be seen in Table 8.4, we observed significant quadratic effects of age (Age^2) for Extraversion (β = -.029, p = .002), Conscientiousness (β = -.055, p = .002), Neuroticism (β = -.033, p = .050), and Openness to Experience (β = -.069, p = .021). We also observed a marginally significant quadratic effect

of age on Honesty Humility (β = -.029, p = .052) which we decided to retain in the model on balance with the consideration of the confidence intervals around the parameter (CI = -.051, -.004). The only personality dimension that showed no curvilinear effect of age – cubic or quadratic – was Agreeableness, showing only a linear effect of age (β = -.036, p = .033). No significant effects of gender were observed; however, the effect was retained and controlled for in all of the models. The findings presented in Table 8.4 thus suggest that age is a significant moderator of the two-year rank order stability estimates for Extraversion, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility, with only linear effects on Agreeableness.

Using the approach described above, we estimated the simple slopes for including both the linear and quadratic age-interactions for Extraversion, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility, and the linear age interaction for Agreeableness, for the entire age-range of 20 to 80 years old. That is, we estimated constrained standardised stability coefficients, point estimates and the 95% Confidence Intervals, for each of the six personality dimensions at every age. These stability coefficients are presented in Figure 3.1. Thus, each panel of Figure 3.1 presents the distribution of the two-year rank order stability estimates across the adult life course for the corresponding personality dimension, as assessed in a large national sample of New Zealand adults.

As shown in Figure 3.1, the overall pattern of results is relatively similar across the personality dimensions, with the exception of Agreeableness. The estimated rank-order stability was reasonably high across the board, with the latent stability estimates ranging from β = .534 at the lowest point (stability of Openness to Experience at age 80), to β = .943 at the highest point (stability of Extraversion at age 36).

For the traits of Extraversion, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility there was a pattern of increasing rank-order stability in younger adulthood up to 40's and 50's, with a subsequent decrease in older age towards the age of 80. Examining the distributions of the differential stability estimates of these latent traits might suggest an earlier peak in stability for Extraversion and Neuroticism - with highest estimates in the late 30's – compared to Conscientiousness, Openness to Experience and Honesty-Humility – all of which show highest estimates in late 40's and 50;s. In the case of Extraversion, Conscientiousness, Neuroticism, and Openness to Experience, declines in stability in later adulthood and older age were such that the parameter estimates decreased to the level or below the parameter estimates for early adulthood. The difference calculations between the estimated simple slopes at age 20 and at age 80 indicated significantly lower scores at age 80 for Extraversion (difference = .194, p = .001, 95% CI = .098, .290), Agreeableness (difference = .145, p = .033, 695%.033, .257), and Neuroticism (difference = .287, p < .001, 95% CI = .174, .400). In contrast, no significant differences between the differential stability at age 20 and at age 80 were found for Conscientiousness (95% CI = -.085, .145 p = .668), Openness to Experience (95% CI = -.049, .182 p = .342), and Honesty-Humility (95% CI = -.122, .070 p = .657).

Table 3.4. Standardised SEM stability estimates and the effects of gender and age (linear, quadratic and cubic where significant) for the six latent personality dimensions. Models estimated using Maximum Likelihood estimation, and numeric integration, computing the latent interaction as random effects.

				95% CT	
	Point Estimate	*Posterior S.D*	*p-value*	*Lower 2.5%*	*Upper 2.5%*
Predicting Extraversion (T2)					
Extraversion (T1)	..922	.023	.000	.884	.959
Gender	-.024	.014	.096	-.048	-.000
Age	-.050	.015	.001	-.074	-.026
Age Squared	-.029	.015	.002	-.053	-.005
N = 3901; Effect of Age^3 not significant at p<.05 and excluded from the model;					
Model Fit [a]: χ^2 (*df*) = 822.018(54), RMSEA (95% C.I.) = .060 (.057, .064), CFI = .932, SRMR = .049.					
Predicting Agreeableness (T2)					
Agreeableness (T1)	.728	.025	.000	.687	.769
Gender	.021	.017	.213	-.007	.050
Age	-.036	.017	.033	-.064	-.008
N = 3901; Effect of Age ^{2} and Age^3 not significant at p<.05 and excluded from the model;					
Model Fit [a]: χ^2 (*df*) = 1349.111(54), RMSEA (95% C.I.) = .078 (.075, .082), CFI = .840, SRMR = .075.					
Predicting Conscientiousness (T2)					
Conscientiousness (T1)	.910	.028	.000	.864	.955
Gender	-.007	.017	.696	-.035	.022
Age	-.009	.018	.594	-.038	.020
Age Squared	-.055	.018	.002	-.085	-.026
N = 3814; Effect of Age^3 not significant at p<.05 and excluded from the model;					
Model Fit [a]: χ^2 (*df*) = 435.701(54), RMSEA (95% C.I.) = .043 (.039, .046), CFI = .953, SRMR = .045.					
Predicting Neuroticism (T2)					
Neuroticism (T1)	.853	.027	.000	.809	.898
Gender	-.004	.017	.803	-.032	.024
Age	-.073	.017	.000	-.102	-.045
Age Squared	-.033	.017	.050	-.061	-.005
N = 3814; Effect of Age^3 not significant at p<.05 and excluded from the model;					
Model Fit [a]: χ^2 (*df*) = 639.426(54), RMSEA (95% C.I.) = .053 (.049, .056), CFI = .930, SRMR = .058.					
Predicting Openness to Experience (T2)					
Openness to Experience (T1)	.841	.028	.000	.796	.887
Gender	-.013	.017	.455	-.040	.015
Age	-.019	.018	.279	-.048	.010
Age Squared	-.069	.030	.021	-.118	-.020
N = 3813; Effect of Age^3 not significant at p<.05 and excluded from the model;					
Model Fit [a]: χ^2 (*df*) = 1166.560(54), RMSEA (95% C.I.) = .073 (.069, .076), CFI = .876, SRMR = .068.					
Predicting Honesty-Humility (T2)					
Honesty-Humility (T1)	.824	.022	.000	.787	.861
Gender	-.014	.014	.334	-.037	.010
Age	.006	.015	.706	-.019	.030
Age Squared*	-.028	.014	.052	-.051	-.004
N = 3811; Effect of Age^3 not significant at p<.05 and excluded from the model;					
* While the effect of Age ^{2} is only marginally significant, given parameter confidence intervals it is here retained in the model;					
Model Fit [a]: χ^2 (*df*) = 1782.867(54), RMSEA (95% C.I.) = .091 (.087, .094), CFI = .870, SRMR = .086.					

RMSEA = root-mean-square error of approximation; CFI = comparative fit index; SRMR = standardised root-mean-square residual;

[a] Model fit estimates based on models not including the latent interaction.

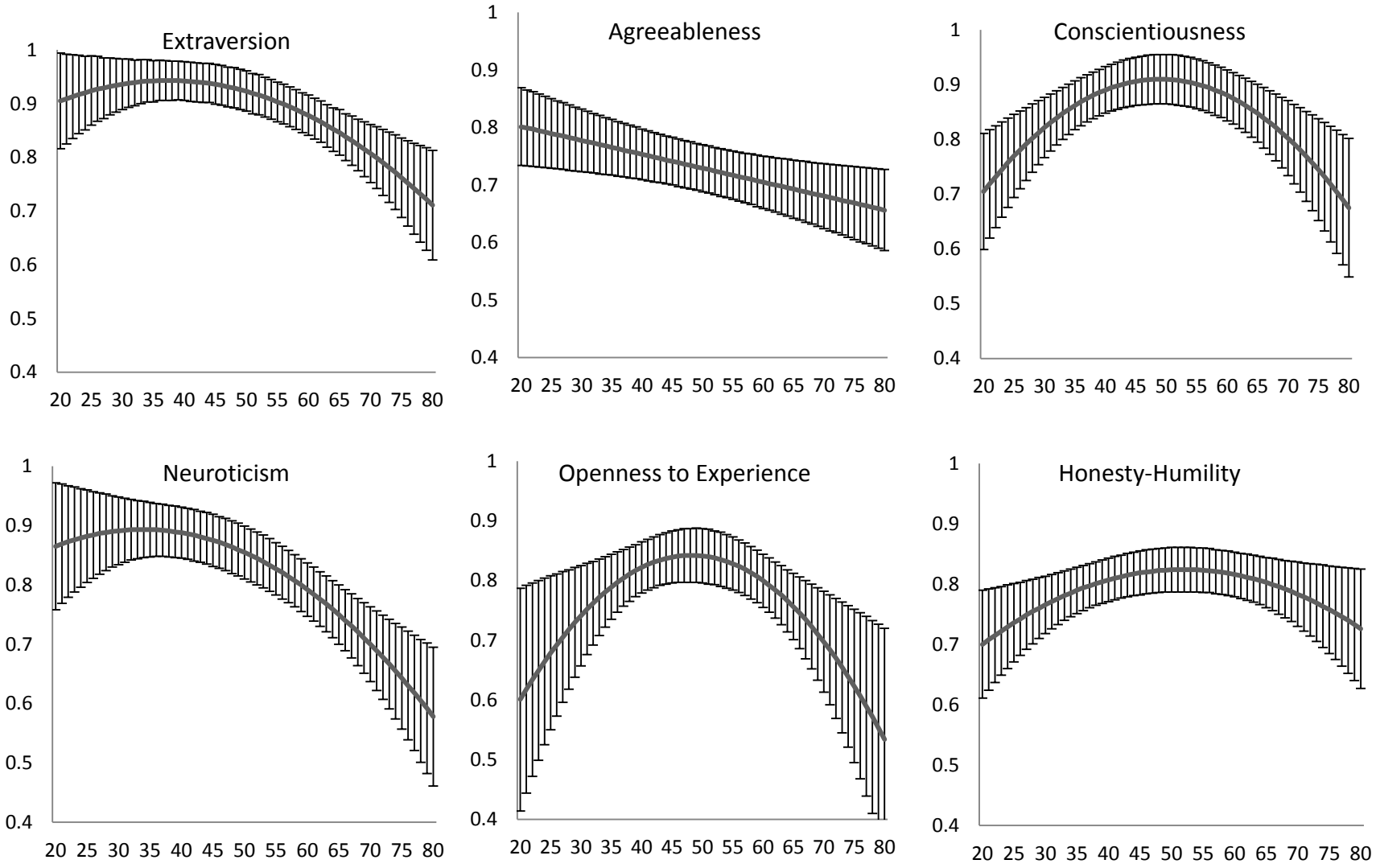


Figure 3.1. The distributions of the latent differential stability coefficients for personality over the 2-year re-test period from age 20 to 80. 95 % Confidence Intervals (Upper 2.5%, Lower 2.5%) plotted as error bars around stability point estimates.

Consistent with the marginally significant quadratic effect of age on the differential stability of Honesty-Humility, Figure 3.1 indicates that the rate of decline of the stability coefficient magnitudes for this latent trait in older age was lower than that for the other dimensions. While these findings highlight the slight differences in the life-course patterns of stability for trait-level Honesty-Humility, they might also highlight the limitation in our study with regards to the shortened age range and the relatively short re-test period (i.e. 2 years). Should the analysis be conducted with an extended range into both younger and older adulthood we might see a more prominent inverted 'U' pattern as associated with Extraversion, Conscientiousness, Neuroticism, and Openness to Experience.

As shown in Figure 3.1 and in the SEM estimates in Table 3.4, the pattern of results for Agreeableness differed notably from those observed for the other five personality dimensions. Specifically, Agreeableness showed a slight linear decrease in the magnitudes of rank-order stability estimates from the highest point in early adulthood ($\beta = .801$ at age 20) to the lowest point in older age ($\beta = .656$ at age 80). These linear effects across age in the differential stability of Agreeableness are relatively small in magnitude. This pattern thus suggests that the rank-order stability of Agreeableness decreases very slightly with increasing age.

Discussion

This study examined changing patterns of stability in latent markers of Big-Six personality markers over a two year period in a multi-cohort panel study of adult New Zealanders. Our analyses revealed considerable variation in the two-year stability for markers of most personality dimensions across considerable range of the adult life-course - 20 to 80 years of age. Moreover, our use of SEM controlled for measurement error while assessing the possible linear, quadratic and cubic interaction effects. Given prior inconsistencies in the

reports of differential stability of personality traits across the lifespan (Ardelt, 2000; Ferguson, 2010; Roberts & DelVecchio, 2000), and the more recent comprehensive findings of curvilinear patterns (e.g., Lucas & Donnellan, 2011; Specht, Egloff, & Schmukle, 2011, Wortman, Lucas, & Donnellan, 2012), replications of such analyses using alternative methods of estimation and different participant samples, as is provided by this study, are critical for the field. We hope that our work, using a national probability sample in New Zealand, can contribute to this endeavour.

The six personality traits showed impressive rank-order stability over the two-year re-test period. The latent stability estimates were largely in the .80 range, with the lowest estimate for Agreeableness ($\beta = .728$) and highest for Extraversion ($\beta = .922$). The magnitude of these estimates is consistent with those previously observed using short form personality scales across slightly different re-test periods (Lucas & Donnellan, 2011; Milojev et al., 2013; Rantanen, Metsäpelto, & Kokko, 2007; Specht, Egloff, & Schmukle, 2011, Wortman, Lucas, & Donnellan, 2012). Thus, our findings are consistent with the large body of research suggesting that personality in adults is highly stable over varying re-test periods (Costa & McCrae, 1988; Solds & Valiant, 1999).

The distribution patterns of differential stability of personality observed in our models are overall consistent with the evidence suggesting curvilinear patterns of stability across the lifespan (Ardelt, 2000; Lucas & Donnellan, 2011; Specht, Egloff, & Schmukle, 2011, Wortman, Lucas, & Donnellan, 2012). With the exception of Agreeableness, our results indicate the 'inverted – U' pattern of rank-order stability across the life-course, exhibiting increases in stability coefficient magnitudes from young to middle adulthood (i.e., from 20's to 40's –50's), followed by decreases past middle age towards late adulthood or old age (i.e., towards 80 years of age). Furthermore, in the case of Conscientiousness, Openness to Experience, and Honesty-Humility the relative stability levels at the oldest age were

comparable to the stability estimates at the youngest age in the analysis. In the case of Extraversion and Neuroticism, the stability at the oldest age was actually lower than the estimated stability at the youngest.

These findings highlight a number of important points. Firstly, as consistent with the cumulative continuity principle (Caspi, Roberts, & Shiner, 2005), the differential stability of these traits increases with progressing adulthood (20 years old onwards), which might reflect the stabilising of social and environmental pressures and demands, and the consequent stabilising of individual differences in personality traits. On the other hand, our findings corroborate the recent evidence of curvilinear developmental patterns (Lucas & Donnellan, 2011; Specht, Egloff, & Schmukle, 2011, Wortman, Lucas, & Donnellan, 2012), such that differential stability of personality traits peaks at a certain point in adulthood and then decrease towards older age (80 years old in this study). These findings are broadly consistent with the life-span argument which would predict such curvilinear patterns across the personality traits. Specifically, periods in life characterised by notable social, cognitive and biological changes, such as young and adulthood, but also older age due to notable changes in the professional and family lives, should be associated with declining stability of personality (Roberts & DelVecchio, 2000; Trzesniewski, Donnellan, & Robins, 2003).

Considering Honesty-Humility, the characteristic curvilinear pattern was somewhat less pronounced, with the SEM estimates for the latent quadratic interaction effect of age being only marginally significant. This finding might reflect a distinct developmental pattern of Honesty-Humility. On a related note, our analyses revealed different distributional patterns between Extraversion and Neuroticism on the one hand, and Conscientiousness, Openness to Experience, and Honesty-Humility on the other. Specifically, the findings seem to suggest different timing of peak stability between these traits, with Extraversion and Neuroticism showing the highest levels of estimated rank-order stability in the mid to late 30’s,

considerably earlier than the peak for the other three traits in late 40's – early 50's. The results of our models thus seem to indicate that systematically different developmental patterns might be inferred for different personality traits, in terms of rank-order stability. Following the argument that systematic stabilisation or change in environmental and social demands across the life-span should be associated with stabilisation or destabilisation of personality differences, these findings would imply that different kinds of developmental pressures differentially affect the stability of the six personality dimensions. Specifically, those events and life changes most strongly associated with Extraversion and Neuroticism might be occurring earlier in life compared to those associated with Conscientiousness, Openness to Experience, and Honesty-Humility.

In addition, the results here observed serve to address an inconsistency among the recent studies investigating differential stability of the Big Five across the life span. While Lucas and Donnellan (2011) and Wortman et al. (2012) reliably identified the characteristic curvilinear pattern of the life span distribution of stability of Conscientiousness, Specht et al. (2011) found no evidence for a curvilinear pattern beyond a steady linear increase with advancing age. Our findings, however, serve to replicate and support the characteristic 'inverted – U' distribution of stability coefficients across the adult life span using modelling strategies similar to those employed by Specht et al. (2011).

Our findings are, however, somewhat inconsistent with regard to the pattern of results associated with Agreeableness. Our results indicate that differential stability of Agreeableness over the two-year re-test period shows a slight linear decrease with increasing age, but that Agreeableness nevertheless remains reasonably stable across the adult life-span. The stability estimates across the ages of the participants might indicate real decreases in the stability of Agreeableness with increasing age, such that trait Agreeableness might change more in response to changing environmental pressures over the adult life span, or due to

systematic maturation processes (e.g., McCrae et al., 2000; Roberts, Wood, & Smith, 2005). On the other hand, this might reflect effects of transient error that our models, despite the precautions taken, could not control for (Chmielewki & Watson, 2009). If this was a pattern largely due to real change in rank-order stability of Agreeableness, our finding then might indicate cross-cultural differences, or life-span development patterns characterised by pressures and cultural norms specific to New Zealand. This inconsistency and the existing differences in the literature discussed above, highlight the importance of further investigation of the distribution of personality stability estimates across the life-span.

Limitations and Future Directions

As previously discussed, the analytic strategy put forward in this report presents considerable advantages. Specifically, the modelling techniques allow simultaneous investigation of the linear and quadratic effects of age on personality stability estimates. Furthermore, the SEM approach provides an estimation of rank-order stability controlling for attenuation due to measurement error through use of latent variable estimation, and does not require splitting the sample by specified cohorts as has been previously done (e.g., Lucas & Donnellan, 2011; Wortman, Lucas, & Donnellan, 2012). However, the current analyses are presented with an important issue evident in the tests of measurement invariance. As outlined above, the model indices of these tests show sub-optimal results with respect to latent variable estimation of Agreeableness, Openness to Experience, and Honesty-Humility, suggesting notable variation with respect to measurement error. While the implementation of SEM procedures and the strict measurement component for all latent personality dimensions controls for this measurement error, the violations of measurement equivalence leave open the possibility that some of the result are influenced by transient error not controlled for (e.g., Chmielewki & Watson, 2009). However, given the recent suggestion that reliability estimates are of limited utility with respect to longer-term re-test reliabilities, and the previous findings

of good re-test reliability of the Minin-IPIP6 over a shorter time-frame (Miojev et al., 2013), the use of SEM procedures with strict measurement estimates affords considerable confidence in interpretations of these results.

Finally, only a moderate re-test period of two years was used in this study, notably shorter than the periods investigated in some of the previous research (Ardelt, 2000; Ferguson, 2010; Roberts & DelVecchio, 2000). By contrast, the recent studies observing the curvilinear distributional patterns of personality stability across the life span and the declines in stability associated with older age, investigated four-year re-test periods (Lucas & Donnellan, 2011; Specht, Egloff, & Schmukle, 2011, Wortman, Lucas, & Donnellan, 2012). This limitation is clearly an issue that can and will be addressed by future research and longitudinal follow-ups. However, we assert that while a moderate time-frame in the context of previous research, a two year period involves a considerable amount of life-events that could affect the stability of personality traits. Furthermore, being able to reliably identify the characteristic curvilinear patterns as we have here done over a different and shorter time-frame than previously investigated, serves to further support the overall findings and the life-span perspective of personality development.

Along with testing the life span changes in the rank-order stability over longer re-test periods, our findings highlight the importance of considering the specific life-changes and events that might be responsible for the observed results. This report further highlights the need to test, as Specht et al. (2011) have done, the effects of events that might cause the lower stability in younger and older adulthood. In addition, our finding of systematically different peaks in stability between different personality dimensions suggest the need to further investigate age-specific changes in environmental and social pressures that are associated with such domain-specific effects. Perhaps, for example, the social pressures associated with the stability of Extraversion stabilise in mid 20s to early 30’s, but then begin

to destabilise with life-changes such as starting one's own family, having one's first child, buying one's first family home and so on, mid-30's onwards. These changes might, therefore, produce effect on personality earlier than the changes that might be associated with the stability of Openness to Experience in later adulthood. Studies examining the life-events associated with the stability and change in personality would be a promising avenue for future investigations into the processes behind the variation in rank-order stability of personality over the life-span.

Concluding Comments

A number of recent studies have reported variability in personality stability across the life span. These findings general converge upon a general developmental pattern characterised by the curvilinear increase in stability estimates towards middle age and a decrease towards old age. However, work in this area remains relatively scarce (longitudinal panel studies are fairly hard to come by) and there are also some inconsistencies across studies. Furthermore, the recent developments in the matter of personality structure and the emergence of the Big Six model of personality brought a need for the extension of the traditional five factor framework. This study provides a further test of the life span development with regards to the differential stability of personality and presents important corroborating evidence for the changing stability across the adult life span in a national panel sample of New Zealanders. Employing recently developed short scales assessing the Big Six personality dimensions this investigation further serves to address the lack of comprehensive stability models of the six factor framework. We advance a Structural Equation Model of the distribution of the two-year differential stability estimates of the six personality factors at every age from 20 to 80 years old. The results obtained through this analytic approach highlight the importance of further investigations into the stability and change of personality traits through cross-cultural investigations and various methods of analysis.

Authors' Note. This manuscript is based on part of Petar Milojev's PhD thesis. Collection of the New Zealand Attitudes and Values Study data analyzed in this paper was funded by University of Auckland FRDF (#3624435/9853) and ECREA (#3626075) grants awarded to Chris G. Sibley.

Mplus syntax for the models reported here (including syntax documenting the model constraint commands we implemented in our analyses) will be posted on the NZAVS website upon acceptance of this article. Syntax and data are also available upon request for reviewing purposes.

Bridging Comments

Studies 1 and 2 provided an investigation of the rank-order stability of the Big Six personality traits in adulthood. While study 1 demonstrated the high temporal stability of personality traits, study 2 showed that the stability of personality varies systematically across the cohorts of the adult life span. Specifically, personality traits, with the possible exception of Agreeableness in study 4, tend to increase in stability across young adulthood, reach peak stability in late 40's and early 50's, and systematically decrease in stability from middle age into older age. These findings thus support the suggestions that personality development is an ongoing process of person-environment interactions, with personality being less stable in periods of life that are normatively associated with more life changes. These findings also provide an important replication of the recent findings of similar patterns (Lucas ad Donnellan, 2011; Specht et al., 2011; and Wortman et al., 2011), thus aiding in the achievement of a general consensus on the developmental patterns of stability of personality traits across adulthood.

Studies 3 and 4 focused on the investigations of mean-level change in the Big Six personality traits. One of the suggestions of the Neo-Socioanalytic perspective (Roberts et al., 2008) is that personality change and development is a function of person-environment interactions throughout the life span. That is, systematic changes in people's lives, in terms of life events and social roles invested in – are systematically associated with changes in personality traits. Indeed, some recent evidence of such processes was reported by Specht and colleagues (2011). For example, these researchers found an increase in Conscientiousness among those who started a new job – a change in personality associated with a life event that signals change in social roles and demands associated with it. Study 3 sought to contribute to this literature by investigating the changes in the Big Six personality traits associated with a major natural disaster in New Zealand – the 2010/2011 Christchurch earthquakes.

Investigation of personality change associated with an event such as this provides interesting opportunities. Firstly, the Neo-Socaioanalytic perspective suggests that personality change can be expected only if the given life event represents substantive changes in social roles and institutions invested in. While a major natural disaster is indeed a major and impactful life event, it is a non-normative life event which may not necessarily be associated with systematic long lasting changes in social roles and identity. On the other hand, experience of such a negative life event may be associated with domain-specific changes in personality traits. For instance, extant research (e.g., Löckenhoff et al., 2009) has suggested increases in Neuroticism among those who reported experience of negative events. It is, therefore, possible that being affected by a large scale natural disaster such as the Christchurch earthquakes may be associated with specific changes in personality characterised by an increase in Neuroticism (or decrease in Emotional Stability) from before to after the event. Study 3 was one of the first studies to investigate such effects in the context of a large-scale natural disaster. Secondly, given the rarity of such investigations, for obvious methodological reasons, this study allows an opportunity to provide important and useful information on the impacts of large scale natural disasters in a nationally-representative sample. Indeed, study 3 asks the important question: Do people change in terms of what they are like when faced with a disastrous event such as the Christchurch earthquakes.

Chapter IV

Study 3: "Personality Resilience Following a Natural Disaster"

Reference:

Milojev, P., Osborne, D., & Sibley, C.G. (2014). Personality Resilience Following a Natural Disaster. *Social Psychological and Personality Science, 5 (7),* 760-768. DOI: 10.1177/1948550614528545.

Abstract

We examine changes in the Big-Six personality markers (Extraversion, Agreeableness, Conscientiousness, Emotional Stability, Openness to Experience, and Honesty-Humility) before and after the 2010/2011 Christchurch earthquakes in a longitudinal study of New Zealand residents (N = 3914). Results show remarkable stability in personality, save for one exception: those who were affected by the earthquakes evidenced a slight decrease in Emotional Stability over the two-year test-retest period relative to those unaffected by the earthquakes. These findings indicate that most aspects of personality are resilient following a major natural disaster. The slight decrease in Emotional Stability, however, points to a possible increase in vulnerability to depression and anxiety for those affected by the earthquakes. Our study provides important insights into a central question about stability and change in personality following major life events.

Introduction

The question of systematic personality change has captured considerable research interest over the years (Roberts, Wood, & Caspi, 2008). Indeed, the scope and nature of personality change has been the subject of many reviews and meta-analyses (Ardelt, 2000; Ferguson, 2010; Mischel & Shoda, 1995; Roberts & DelVecchio, 2000; Roberts, Walton, &

Veichtbauer, 2006), as well as numerous empirical examinations (e.g., Anusic, Lucas, & Donnellan, 2012; Lucas & Donnellan, 2011; Specht, Egloff, & Schmukle, 2011; Wortman, Lucas, & Donnellan, 2012). This literature indicates that, while personality – typically examined from the framework of the Big Five or the Big Six (i.e., the HEXACO model; Ashton & Lee 2007, 2009) – is remarkably stable across time, notable changes do occur over the life-course (Anusic, Lucas, & Donnellan, 2012; Lucas & Donnellan, 2011; Milojev, Osborne, Greaves, Barlow, & Sibley, 2013; Roberts, Walton, & Veichtbauer, 2006; Specht, Egloff, & Scukle, 2011; Wortman, Lucas, & Donnellan, 2012). A critical question arising from this research, then, is under what conditions does change occur? Relatedly, to what extent is personality change due to developmental processes/maturation (McCrae et al., 2000) versus environmental pressures such as changing social roles and/or experiencing major life events (e.g., Löckenhoff, Terraciano, Bienvenu, Patriciu, Nestadt, McCrae, Eaton, & Costa, 2009; Roberts, Wood, & Smith, 2005)?

A common finding across the literature is that the effects of major life events on changes in personality vary with respect to the types of events and personality dimensions in question. Specifically, research shows that domain-specific changes in personality correspond with various major life events including increases in Conscientiousness upon getting a first job (Specht, Egloff, & Schmukle, 2011), as well as increases in Neuroticism following self-reported experiences with adverse life events (Löckenhoff et al., 2009). In other words, though personality is typically stable across time, major life events can substantively alter people's personality.

We present a novel contribution to the personality literature by examining changes in personality before and after experiencing a major natural disaster amongst a national sample. For obvious reasons, such investigations are rare. Indeed, extant studies of personality change versus resilience among the survivors of natural disasters often entail small-scale clinical

investigations or rely on relatively small convenience samples (e.g., Nolen-Hoeksema & Morrow, 1991; Phifer, 1990). As such, our ability to assess stability and change in personality within a national sample following a natural disaster offers an important contribution to the literature.

We report analyses from a unique longitudinal dataset comprised of a national sample of New Zealand adults (N = 3914). Critically, our sample includes a group of participants who were affected by a large-scale natural disaster that occurred in the middle of a two-year test-retest period. Specifically, we assess changes in the Big-Six (Extraversion, Agreeableness, Conscientiousness, Emotional Stability/Neuroticism, Openness to Experience, and Honesty-Humility) among people who were affected by the 2010/2011 Christchurch earthquakes and use those unaffected by the earthquakes as baseline of comparison.

The Christchurch Earthquake and the Present Study

The city of Christchurch and the wider Canterbury region of New Zealand was shook by an M_w 6.3 earthquake on 22 February 2011. Tragically, 185 people were killed—some of whom remain unidentified (Source: New Zealand Police, website accessed: http://www.police.govt.nz/major-events/previous-major-events/christchurch-earthquake/list-deceased; Accessed December 11, 2013). In addition to the lives lost, a number of multi-story buildings collapsed and much of the city's power, water and phone lines were destroyed. Indeed, the New Zealand Treasury indicates that about a third of the buildings in Christchurch were damaged beyond repair, resulting in an estimated rebuild cost of 20 billion NZD (Source: New Zealand Treasury, website accessed: http://www.teara.govt.nz/en/historic-earthquakes/page-13; http://www.treasury.govt.nz/budget/2012/speech/07.htm; Accessed December 11, 2013). In

total, over 300,000 people were affected by the earthquake (Source: International Disaster Database, website accessed: http://www.emdat.be/result-country-profile; Accessed December 11, 2013). Moreover, the 22 February event came on the heels of an earlier—and stronger (M_w 7.1), albeit non-fatal—earthquake that produced thousands of aftershocks that continue to shake the Canterbury region. In short, the 2010/2011 earthquakes placed an immensely stressful toll on the Christchurch community.

The current study is uniquely positioned to assess the effects of this disaster on people's personality. Specifically, the data analysed here are part of a larger national probability panel study - the New Zealand Attitudes and Values Study (NZAVS) - that began in 2009 and re-interviewed participants at the end of 2011. As such, participants completed measures of personality (namely, the Mini-IPIP6; Sibley et al, 2011) both *before and after* the Christchurch earthquakes. This provides us with the rare opportunity to examine stability and change in personality following exposure to a major life event. Specifically, we assess the interaction between time and participants' self-report of being affected by the earthquakes on their personality. Because self-report data can be problematic, we replicate our analyses with an objective measure of exposure to the earthquakes (i.e., whether participants lived in the Canterbury region before the earthquakes[7]).

In accordance with the literature discussed above, we expected that, to the extent that any changes in personality occur over a two year period, they will be small in magnitude. Given that exposure to adverse life events is associated with increases in Neuroticism (e.g., Löckenhoff et al., 2009) coupled with the conceptualisation of Emotional Stability/Neuroticism as one's responsiveness to (and resilience in the face of) situational stressors (DeYoung, 2010), we expected a slight decrease in Emotional Stability among those

[7] Within the retained longitudinal sample, about 86% of the people living in the Canterbury region reported having been personally affected by the earthquake.

affected by the earthquakes. No other changes in personality were expected to occur across our two time points.

Method

Participants and Sampling Procedure

This report is based on data from 3914 participants who completed both the 2009 (Time 1) and 2011 (Time 3) waves of the NZAVS.[8] The Time 1 NZAVS was completed by 6,518 participants randomly selected from the New Zealand Electoral Roll (response rate = 16.6%). Of these, 3,914 also participated at Time 3(retention rate 60.0%). Variations in sample size reported below are due to missing data.

The Time 1 and the retained longitudinal sample at Time 3 had a gender bias, with 2414 (61.7%) women and 1496 (38.3%) men responding to both time points. In contrast, 52% of adult New Zealand residents are women and 48% are men (Statistics New Zealand, 2006). The longitudinal sample was also more likely to include Europeans (n = 3395, 86.7%) relative to 2006 census estimates (68.7%). In most other respects, the NZAVS was reasonably consistent with national population estimates. The mean age of participants in the retained sample was 50.31 in 2009 (SD = 15.045). 79.9% (n = 3124) reported having been born in New Zealand, 44.3% (n = 1733) identified as religious, 77.9% (n = 3047) of the sample were parents, 72% (n = 2814) were in a relationship, and 73% (n = 2855) were employed. 20.9% reported no formal education (n = 816), 28.3 % having some high-school education (n = 1106), 16% reported working towards a diploma or a certificate (n = 624), 24.4% reported an undergraduate degree (n = 955), and 10.5% reported a postgraduate degree

[8] The collection of Time 3 (2011) responses was staggered from September 2011. Most responses were collected in October and November of 2011.

(n = 409).[9] Detailed information about the NZAVS sample statistics for each wave and comparisons with census data are available upon request.

Of the retained sample, 21.4% (n = 838) reported being personally affected by the 2011 Christchurch Earthquakes, whereas 77.6% (n = 3034) indicated they were unaffected (1%, n = 38 did not respond to this question). As for an objective indicator of exposure to the earthquakes, 13.5% (n = 529) of the retained sample resided in the Canterbury region. Descriptive summaries for the retained sample are presented in Table 1. Additional descriptive summaries comparing Cantabrians with non-Cantabrians are available upon request. With the exception of ethnicity (i.e., 86% of the non-Cantabrians were European, whereas 93% of the Cantabrians were European), the two sub-samples had remarkably similar demographic characteristics. Nevertheless, our analyses adjusted for demographic covariates.

Questionnaire Measures

The Mini-IPIP6 (Sibley et al., 2011) is a 24-item personality measure containing the 20 items developed by Goldberg (1999) and included in Donnellan et al.'s (2006) original Mini-IPIP. Four items were added to this scale to index Honesty-Humility. Two of these items were adapted from Campbell et al.'s (2004) Narcissism scale, whereas the remaining two items were adapted from Ashton and Lee's (2009) HEXACO measure of Honesty-

[9] Participants who responded in both 2009 and 2011 were more likely to be female (61.7% vs. 56.1% for those who did not follow-up; $\chi^2(1, 6518) = 20.946$, $p<.001$), more likely to be European (86.8% vs. 74.7%; $\chi^2(1, 6518) = 153.642$, $p<.001$), more likely to have been born in New Zealand (79.9% vs. 75.7%; $\chi^2(1, 6517) = 16.321$, $p<.001$),more likely to have children (77.9% vs. 70.3%; $\chi^2(1, 6518) = 26.134$, $p<.001$), and to be in a relationship (72.8% vs. 66.9%; $\chi^2(1, 6424) = 26.134$, $p<.005$) compared to respondents who did not follow up in 2011. They also lived in more affluent or less deprived neighbourhoods (NZDep = 4.82 (2.794) vs. 5.42 (2.885), $F(1.6362) = 68.310$, $p<.001$), and were older (50.32 (15.037) vs. 44.22 (16.149) $F(1, 6362) = 235.268$, $p<.001$). Participants who responded to both wave I and III had lower Extraversion scores ($F(1,6462) = 15.090$, $p<.001$, $\eta^2_p = .002$), higher Agreeableness scores ($F(1, 6462) = 30.263$, $p<.001$, $\eta^2_p = .005$), higher Conscientiousness scores ($F(1, 6462) = 34.533$, $p<.001$, $\eta^2_p = .005$), higher Emotional Stability or lower Neuroticism ($F(1, 6462) = 31.597$, $p<.001$, $\eta^2_p = .005$) and higher Honesty-Humility scores ($F(1, 6462) = 163.587$, $p<.001$, $\eta^2_p = .0025$). While these personality differences are significant the effects are small with most of them accounting for less than 1% of the variance.

Humility. The items assessing Neuroticism were reverse-coded for this study to capture levels of Emotional Stability. Items from each scale were averaged to reflect each of the six personality dimensions in the Big Six.

TheMini-IPIP6 was administered with the following instructions: “This part of the questionnaire measures your personality. Please circle the number that best represents how accurately each statement describes you.” Items were rated on a scale ranging from 1 (very inaccurate) to 7 (very accurate). Extraversion items included “Am the life of the party.”; “Don't talk a lot.” (reverse-scored); “Keep in the background.” (reverse-scored); “Talk to a lot of different people at parties” ($\alpha_{\text{Time 1}} = .71$; $\alpha_{\text{Time 3}} = .75$). Agreeableness items included “Sympathize with others' feelings.”; “Am not interested in other people's problems.” (reverse-scored); “Feel others' emotions.”; “Am not really interested in others.” (reverse-scored; $\alpha_{\text{Time 1}} = .66$; $\alpha_{\text{Time 3}} = .69$). Conscientiousness items included “Get chores done right away.”; “Like order.”; “Make a mess of things.” (reverse-scored); “Often forget to put things back in their proper place” (reverse-scored; $\alpha_{\text{Time 1}} = .65$; $\alpha_{\text{Time 3}}$ $\alpha = .65$). Emotional Stability items included “Am relaxed most of the time”; “Have frequent mood swings” (reverse-scored); “Seldom feel blue.”; “Get upset easily” (reverse-scored; $\alpha_{\text{Time 1}} = .64$; $\alpha_{\text{Time 3}} = .72$). Openness to Experience items included “Have a vivid imagination.”; “Have difficulty understanding abstract ideas.” (reverse-scored); “Do not have a good imagination.” (reverse-scored); “Am not interested in abstract ideas” (reverse-scored; $\alpha_{\text{Time 1}} = .67$; $\alpha_{\text{Time 3}} = .70$). Honesty-Humility items included “Would like to be seen driving around in a very expensive car.”; “Would get a lot of pleasure from owning expensive luxury goods.”; “Feel entitled to more of everything.”; “Deserve more things in life.” (all reverse-scored; $\alpha_{\text{Time 1}} = .78$; $\alpha_{\text{Time 3}} = .77$).

Participants were also asked if they were “…personally affected by the Christchurch earthquakes” (Yes/No). Finally, the questionnaire assessed relevant demographics including:

gender (0 female, 1 male), age, socioeconomic status, ethnicity (0 minority group, 1 majority group), immigration status (0 born overseas, 1 born in New Zealand), religiosity (0 not religious, 1 religious), parental status (0 no children, 1 at least one child), relationship status (0 single, 1 in relationship), employment (0 unemployed, 1 employed) and education (ordinal ranging from -2 none reported to 2 postgraduate study). Socioeconomic status was assessed using a measure of neighbourhood deprivation of the participants residential area (NZDep 2006; see Salmond, Crampton, & Atkinson, 2007), with scores ranging from 1 (most affluent) to 10 (most deprived).

Results

Descriptive statistics and bivariate correlations for all of the Big Six personality dimensions, as well as our demographic covariates, are presented in Table 4.1. As shown here, the rank-order stability estimates for the two-year period range from .58 and above for all sub-samples. While these are somewhat lower than the stability estimates reported elsewhere, these estimates indicate that personality is highly stable over a two year period[10].

To test our hypotheses, we conducted multiple mixed-design repeated-measures ANCOVAs assessing change in each of the Mini-IPIP6 subscale scores. In doing so, we controlled for the effects of our demographic covariates including gender, age, socioeconomic status, ethnicity, immigration status, religiosity, parental status, relationship status, employment, and education. Consistent with the view that personality is stable over time, these analyses revealed no significant main effects of Time on changes in personality across the six personality dimensions ($p>.05$). Likewise, no significant main effects of being

[10] While not a primary aim of our investigation, we also examined the potential effects of having experienced (or been affected by) the Christchurch Earthquake on rank-order stability of the six personality dimensions. Using Mplus 7.1 (Muthén & Muthén, 1998-2012) and specifying Maximum Likelihood (ML) estimation methods with 1000 Bootstrap resampling, we conducted tests of moderation of the rank-order stability coefficients by whether or not one was affected by the earthquakes. The analyses indicated no significant interactions with all of the estimated Bootstrapped CI's (90%, 95%, 99%) including 0. Please contact the corresponding author for further details.

affected by the earthquakes (versus not having been affected) were observed for mean personality levels when averaging across the two assessment periods ($p>.05$).

Table 4.1. Descriptive Statistics and Bivariate correlations for the full sample (*N* = 3,914). Test-retest correlations for the personality variables are printed in bold.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.
1.Extraversion T1																						
2.Extraversion T2	**.738***																					
3.Agreeableness T1	.215*	.165*																				
4.Agreeableness T2	.166*	.199*	**.588***																			
5.Conscientiousness T1	.011	.018	.144*	.126*																		
6.Conscientiousness T2	.026	.075*	.113*	.213*	**.682***																	
7.Emotional Stability T1	.096*	.070*	.035*	.027	.120*	.114*																
8.Emotional Stability T2	.086*	.113*	.031	.086*	.133*	.183*	**.627***															
9.Openness T1	.245*	.193*	.221*	.158*	.016	-.009	.022	.007														
10.Openness T2	.210*	.229*	.166*	.249*	-.017	.047*	.018	.019	**.648***													
11.Honesty-Humility T1	-.091*	-.082*	.157*	.152*	.091*	.082*	.199*	.154*	.022	.003												
12.Honesty-Humility T2	-.096*	-.106*	.159*	.205*	.069*	.096*	.144*	.159*	.021	.041*	**.705***											
13.Gender	-.064*	-.058*	-.293*	-.296*	-.102*	-.095*	.136*	.103*	.012	.030	-.127*	-.133*										
14.Age	-.081*	-.091*	-.003	-.022	.076*	.056*	.180*	.154*	-.163*	-.160*	.221*	.217*	.121*									
15.Socioeconomic status	-.026	-.038*	-.037*	-.046*	-.054*	-.069*	-.060*	-.071*	-.009	-.038*	-.052*	-.057*	-.022	-.018								
16.Ethnicity	.012	.013	.043*	.058*	-.024	-.027	.011	.025	.015	.054*	.138*	.142*	-.032*	.084*	-.175*							
17.Immigration status	-.004	-.024	-.003	.000	-.026	-.041*	-.006	-.014	-.043*	-.048*	.010	-.002	-.036*	-.061*	.039*	.133*						
18.Religiosity	.008	.023	.121*	.090*	.061*	.064*	.028	.031	-.051*	-.057*	.067*	.049*	-.066*	.196*	.020	-.123*	-.056*					
19.Parent	.003	.034*	.001	-.007	.078*	.066*	.100*	.075*	-.119*	-.109*	.084*	.078*	.023	.419*	-.011	.019	-.012	.095*				
20.Relationship	.001	.036*	-.046*	-.021	.055*	.063*	.032*	.034*	-.051*	-.031	.048*	.044*	.112*	.060*	-.174*	.050*	-.038*	-.040*	.297*			
21.Employement	.070*	.069*	.001	.034*	-.019	.006	-.018	-.022	.102*	.115*	-.098*	-.098*	.030	-.399*	-.094*	-.020	.010	-.098*	-.130*	.085*		
22.Education	.029	.023	.120*	.142*	.005	.024	-.001	-.001	.215*	.223*	.077*	.092*	-.093*	-.179*	-.174*	.069*	-.081*	.000	-.089*	.066*	.206*	
Mean	4.00	3.91	5.30	5.28	5.16	5.09	4.62	4.62	4.75	4.72	5.21	5.17	.38	50.31	4.82	.87	.80	.45	.78	.73	.75	-.25
SD	1.16	1.12	.97	.93	1.05	1.00	1.10	1.13	1.11	1.09	1.30	1.25	.49	15.05	2.79	.34	.40	.50	.42	.45	.44	1.31

Note: The test-retest correlations over the two year period are highlighted in bold.
*. Correlation is significant at the 0.05 level (2-tailed).

The hypothesised interaction between Time and self-report of being affected by the earthquakes was, however, observed for Emotional Stability (F (1, 3552) = 9.422, p = .002, η^2_p = .003). Specifically, those who were affected by the earthquake showed a slight decrease in trait-level Emotional Stability from Time 1 (M = 4.596, SE = .039) to Time 3 (M = 4.514, SE = .040; Estimated 95% CI's for the mean difference = -.027, .192; d = -.075). Conversely, those who were unaffected by the earthquake showed a slight increase in Emotional Stability from Time 1 (M = 4.611, SE = .020) to Time 3 (M = 4.648, SE = .021; estimated 95% CI's for the mean difference = -.093, .019; d = -.035). Although the former effect is small in magnitude, the decrease in Emotional Stability associated with being affected by the Christchurch earthquake demonstrates that environmental events can substantively alter domain-specific personality dimensions. Figure 4.1 highlights changes (or lack thereof) in the six personality variables under examination.

In order to replicate our analyses using an objective measure of participants' exposure to the Christchurch earthquakes, we compared responses from people living in the Canterbury region in 2009 with those who were living in other regions of New Zealand. These analyses revealed results comparable to those reported above. Specifically, there were no significant main effects of Time on each of the six personality dimensions ($p>.05$). Likewise, living in the Canterbury region (versus other regions of New Zealand) was unassociated with mean levels of personality averaging across the two assessment periods ($p>.05$). There was, however, the predicted interaction between Time and participants' residence on Emotional Stability (F(1, 3587) = 4.465, p = .035, η^2_p = .001). Cantabrians experienced a mean-level decrease in Emotional Stability from 2009 (M = 4.624, SE =.049) to 2011 (M = 4.548, SE = .051; Estimated 95% CI's for the mean difference = -.063, .251; d = -.069), whereas those living in other areas of the country maintained comparable levels of Emotional Stability across time points (M = 4.608, SE = .019 and M = 4.632, SE = .020, for

2009 and 2011, respectively; Estimated 95% CI's for the mean difference = -.078, .030; d = -.022). Though the interaction between Time and area of residence on Emotional Stability was slightly smaller than the interaction between Time and self-reported measure of being affected by the earthquakes (η^2_p = .001, vs. η^2_p = .003, respectively), both of the predicted interactions were reliable. [11]

[11] We assessed the possibility that the effects observed were associated with a sampling bias, particularly with personality differences between Cantabrians who responded to the post-earthquake survey and those who did not follow-up. Cantabrians who followed-up in 2011 had slightly higher Agreeableness (F(1,856) = 6.780, p = .009, η^2_p = .008, d = .-.181), Conscientiousness (F(1, 856) = 9.569, p = .002, η^2_p = .011, d = -.214), and Honesty-Humility (F(1,856) = 11.124, p = .01, η^2_p = .013, d = -.230). No significant differences were observed for Extraversion (p = .205, η^2_p = .002, d = .089), Emotional Stability (p = .161, η^2_p = .002, d = -.098), or Openness to Experience (p = .088, η^2_p = .003, d = .118).

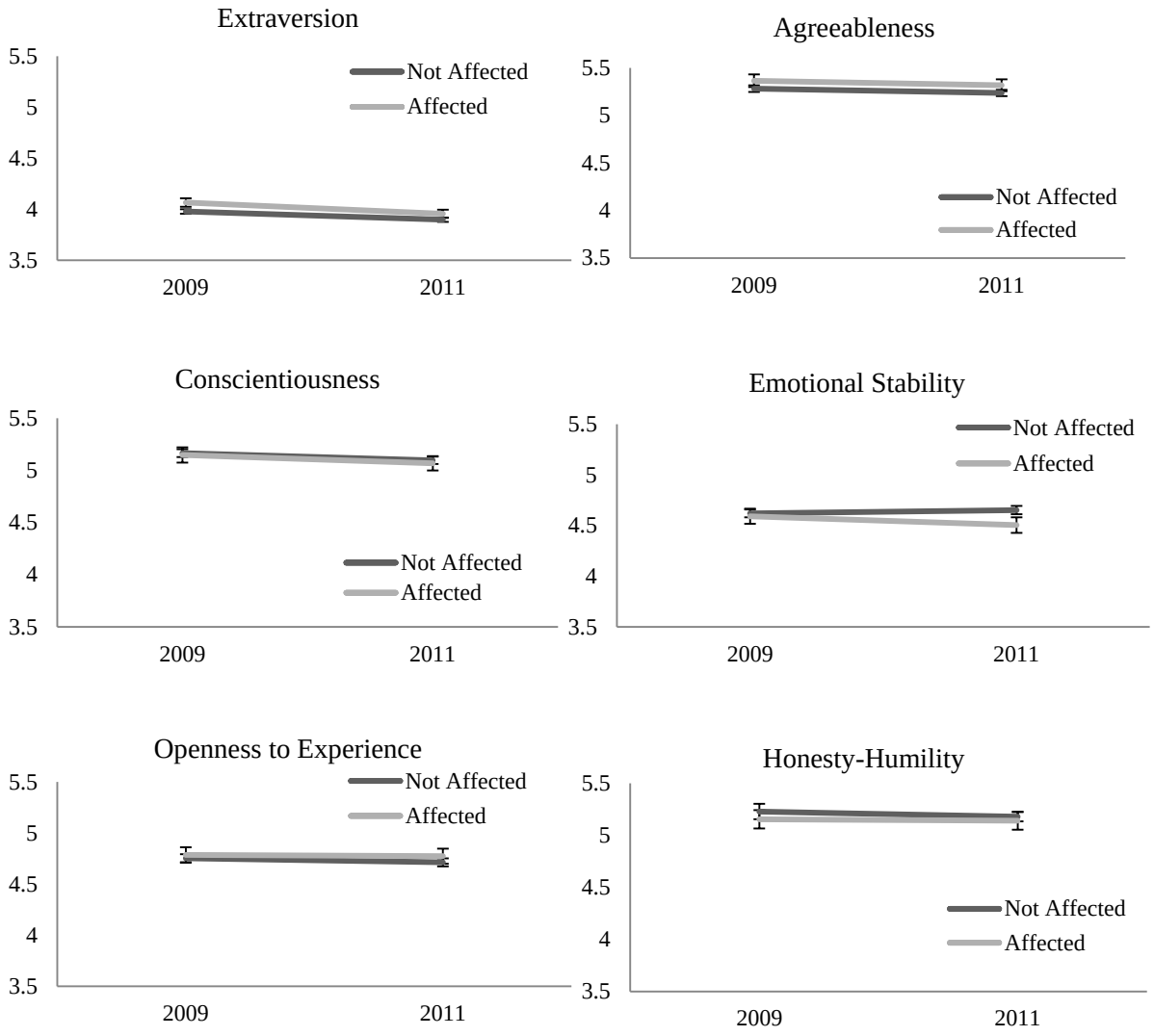


Figure 4.1. Mean-scores (with 95% confidence intervals) showing the small magnitudes of the intercation effects observed for the main analysis.

Exploratory cross-lag models

We conducted additional exploratory analyses predicting Time 3 Emotional Stability from participants' gender, age, socioeconomic status, ethnicity, immigration status, religiosity, parental status, relationship status, employment, and education after adjusting for Time 1 Emotional Stability. This model therefore tested the extent to which each demographic factor predicted unique variance in residualized change in Emotional Stability. As with our prior analyses, we tested two versions of this model, each of which included one of our two indicators of subjective or objective exposure to the earthquakes. As shown in

Table 4.2, the only demographic factors with a significant effect (beyond that of being affected by the earthquake or living in the affected area) was the positive effect of age (β = .054 and β = .051 for the two models) and the negative effect of neighbourhood deprivation (β = -.032 for both models). That is, age was associated with a slight increase in Emotional Stability over the two years, whereas living in more deprived neighbourhoods was associated with a decrease in Emotional Stability.

We also explored the possible effects participants' initial levels of personality had on Emotional Stability at Time 3, after controlling for Emotional Stability at Time 1 and earthquake exposure. The separate regression models using our two indicators of earthquake exposure are presented in Table 4.3. Both models show that initial levels of Extraversion (β = .039 for the subjective assessment, and β = .037 for the objective assessment), Conscientiousness (β = .058 for both models), and Honesty-Humility (β = .028 for the subjective assessment, and β = .028 for the objective assessment) were associated with increases in Emotional Stability after the earthquakes. These findings indicate that these three personality dimensions may constitute a protective personality profile that is resilient to natural disasters. These analyses are, however, exploratory and beyond the scope of the present paper. As such, future research should aim to replicate these findings before any firm conclusions can be drawn.

Table 4.2. Cross-lagged demographic models predicting residual change in Emotional Stability in 2011.

	b	*se*	*β*	*t*
Subjective assessment of earthquake exposure				
2009 Emotional Stability	..640*	.014	.621	46.692
Affected by the Earthquakes (0 no 1 yes)	-.125*	.036	-.045	-3.501
Gender	.012	.031	.005	.383
Age	.004*	.001	.054	3.291
Socioeconomic status	-.013*	.005	-.032	-2.393
Ethnic Group	.027	.046	.008	.595
Immigration Status	-.011	.037	-.004	-.297
Religious status	.017	.030	.008	.565
Parental status	-.034	.041	-.012	-.811
Relationship	.016	.036	.006	.430
Employment	.020	.038	.008	.537
Education	.009	.012	.010	.750
Objective assessment of earthquake exposure				
2009 Emotional Stability	.640*	.014	.621	46.916
Resided in Canterbury region 2009 (0 no 1 yes)	-.093	.043	-.028	-2.175
Gender	.010	.031	.004	.334
Age	.004*	.001	.051	3.120
Socioeconomic status	-.013*	.005	-.032	-2.343
Ethnic Group	.034	.046	.010	.754
Immigration Status	-.013	.037	-.004	-.338
Religious status	.019	.030	.008	.631
Parental status	-.034	.041	-.013	-.834
Relationship	.020	.036	.008	.543
Employment	.021	.038	.008	.559
Education	.009	.012	.010	.734

* $p < .05$.

Table 4.3. Cross-lagged personality models predicting residual change in Emotional Stability in 2011.

	b	*se*	*β*	*t*
Subjective assessment of earthquake exposure				
2009 Emotional Stability	.631*	.013	.612	47.100
Affected by the Earthquakes (0 no 1 yes)	-.129*	.035	-.047	-3.729
2009 Extraversion	.038*	.013	.039	2.881
2009 Agreeableness	-.008	.016	-.007	-.517
2009 Conscientiousness	.062*	.014	.058	4.517
2009 Openness to Experience	-.015	.013	-.015	-1.155
2009 Honesty-Humility	.025*	.012	.028	2.153
Objective assessment of earthquake exposure				
2009 Emotional Stability	.630*	.013	.611	47.160
Resided in Canterbury region 2009 (0 no 1 yes)	-.098*	041	-.030	-2.356
2009 Extraversion	.036*	.013	.037	2.747
2009 Agreeableness	-.006	.016	-.005	-.364
2009 Conscientiousness	.063*	.014	.058	4.578
2009 Openness to Experience	-.017	.013	-.017	-1.287
2009 Honesty-Humility	.027*	.011	.031	2.385

* $p < .05$.

Discussion

The earthquakes that shook Christchurch and the wider Canterbury region in late 2010 and early 2011 left a mark of devastation in the area and affected the whole country. Undeniably, the disaster had great effects on every New Zealander, particularly Cantabrians. Indeed, previous research has shown that the earthquakes produced an increase in people's tendency to identify as religious (Sibley & Bulbulia, 2012), while also suggesting that baseline levels of Emotional Stability provided a protective barrier against mental health decrements among the survivors (Osborne & Sibley, 2013). The current study adds to this literature by examining the fundamental effect that such a large-scale natural disaster has on survivors' personalities. This is an important investigation that provides critical insight into the impetus behind personality change (Löckenhoff et al., 2009; McCrae et al., 2000; Roberts, Walton, & Veichtbauer, 2006; Roberts, Wood, & Smith, 2005).

Our findings indicate that personality is (by and large) unaffected by exposure to catastrophic events such as the Christchurch earthquakes. Indeed, the only change found over the two-year test-retest period was a slight decrease in Emotional Stability among those affected (vs. unaffected) by the disaster. It is important to note that the estimated effect sizes were very small, which demonstrates that personality is remarkable stable—even among those affected by the earthquakes. In fact, the test-retest correlations for all of the personality dimensions were .58 or higher. These findings indicate reasonably high levels of stability, though they are slightly lower than previous reports that use advanced modelling techniques (e.g., Milojev et al., 2013). The somewhat lower stability estimates generated here are likely due to measurement error, as we used manifest personality variables (see Ferguson, 2010) derived from a short-form scale.

To the extent that personality changed, it was localised to the specific dimension most relevant to the given event, namely, Emotional Stability/Neuroticism (i.e., the personality dimension associated with stress and reactivity to threat). This finding is in line with previous research showing domain-specific effects of major life events on mean-level personality change including increased levels of Neuroticism in response to aversive life events (e.g., Löckenhoff et al., 2009). Our findings thus contribute to a long-standing discussion about the malleability of personality associated with maturation, environmental and social influences, and transactional processes (e.g., Roberts, Walton, & Veichtbauer, 2006).

Future Directions and Limitations

The time interval between assessment periods in this investigation allowed us to examine change in personality before and after the earthquakes. Nevertheless, we were unable to investigate the potential long-term consequences of the earthquakes on mean-levels of personality. For instance, previous research has indicated that individual-difference variables such as life satisfaction can show dynamic change patterns whereby, after an initial change, they return to baseline levels after several years (e.g., Lucas, 2007). Furthermore, recent analyses by Ogle, Rubin, and Siegler (2013) indicate that Neuroticism does not exhibit meaningful longitudinal change in association with trauma exposure. In considering such evidence, the most relevant discrepancy between our study and prior research is the much longer test-retest period used by these researchers – about 8 years - compared to our 2-year test-retest period (Ogle, Rubin, & Siegler, 2013). It might, therefore, be the case that the changes we observed in Emotional Stability represent a temporary reactive effect. Specifically, longitudinal follow-ups might find that, after the salience of the earthquakes has decreased, mean-levels of Emotional Stability among survivors return to their baseline levels. Similarly, the personality dimensions that exhibited no change may show delayed effects of the earthquakes in the following years. Differential development trajectories of personality

might also be exhibited between those affected and those unaffected by the earthquakes. Such investigations are crucial for our understanding of personality. Thus, our study points to the need for longitudinal follow-ups that allow for the estimation of long-term development trajectories.

A related consideration arising from our findings regards the issues of personality styles associated with vulnerability to anxiety and depression – particularly (a) Dependency and Self-Criticism and (b) Sociotropy and Autonomy (Beck, 1983; Blatt & Zuroff, 1992; Zuroff, Mongrian, & Santor, 2004). These personality variables are associated with the Big Five personality dimensions, particularly Neuroticism/Emotional Stability (Zuroff, 1994). Therefore, our finding that being affected by the earthquakes was associated with a decrease in Emotional Stability suggests that survivors may be at an increased risk of developing a personality-based vulnerability to depression. On a positive note, our exploratory analyses indicated that Extraversion, Conscientiousness, and Honesty-Humility protected participants from decreases in Emotional Stability following exposure to a natural disaster. These possibilities further highlight the need to pursue longitudinal follow-ups that investigate the developmental trajectories of Emotional Stability as they relate to the mental health of survivors (e.g., Osborne & Sibley, 2013).

Concluding Comments

This report makes an important contribution to the study of personality processes and change, as well as providing useful information about possible long-term changes in the well-being of Cantabrians (and survivors of other natural disasters). We hope that our findings will provide a message to the survivors of the devastating Christchurch earthquakes. The message is this: many aspects of people's core personality—at least those aspects that personality psychologists use to describe people remained unchanged in the wake of the Christchurch

earthquakes. While the devastation and turmoil caused by the earthquakes undoubtedly affected many aspects of people's lives and many other aspects of their psychology (such as stress and anxiety, and also likely for many, potential concerns about the timeframe for decisions provided by the Earthquake commission), people's core personality remained largely unchanged. Indeed, our data indicate that the broadest and most central aspects of people's personality remained—and we hope will continue to remain—resilient amongst those affected by the Christchurch earthquakes.

Authors' Note. This research was supported by a Templeton World Charity Foundation Grant (ID: 0077). Mplus syntax for the models reported here (including syntax documenting the model constraint commands we implemented in our analyses) will be posted on the NZAVS website upon acceptance of this article. Syntax and data are also available upon request for reviewing purposes.

Bridging Comments

Study 3 demonstrated the remarkable consistency in the Big Six personality traits in the face of a large-scale natural disaster – the 2010/2011 earthquakes in Canterbury, New Zealand. The only systematic change observed was a slight decrease in Emotional Stability (or an increase in Neuroticism) among those affected by the event. This finding of domain specific changes in personality traits is in line with the predictions of the Neo-Socioanalytic perspective on personality change and development (Roberts *et al*., 2008) and with extant literature showing increases in Neuroticism associated with experiences of negative life events (e.g., Löckenhoff et al., 2009). Study 3 demonstrated that while personality is highly stable, some specific, systematic change is possible and does occur in adulthood and in relation to life events. The finding of a decrease in Emotional Stability also highlights the potential for an increase in personality based vulnerability of depressive symptoms among those affected by such natural disasters (e.g., Zuroff, 1994).

Study 4 continued the investigation into mean-level change in the Big Six personality traits by investigating the normative developmental patterns of change across the cohorts of the adult life span. Specifically, study 4 put forward a series of 6-year Cohort-Sequential Latent Growth Models (e.g., Pronzie & Onghena, 2005) to identify linear and curvilinear change trajectories in the six personality traits across the adult life span from 19 to 74 years old. As discussed above, normative developmental change in personality traits is a central issue in the debate about the very nature of the personality trait as a construct. The two broad perspectives here discussed put forward different predictions of the expected change trajectories. On the one hand, the Five Factor Theory (McCrae & Costa, 1999) would predict that most, if not all of the normative development in personality will occur in younger adulthood, with very little changes expected thereafter. This perspective argues that personality development is mainly a matter of genetic influences and temperament and that

personality is relatively unchangeable in adulthood. On the other hand, the Neo-Socioanalytic perspective (Roberts et al., 2008) would predict a process of change in personality traits that continues throughout the entire adult life span, into older age. As people age the social roles and institutions they normatively engage in, as well as life events they experience will change – such as finishing formal education, starting careers, or starting a family in younger age; or retirement, bereavement, and cognitive and biological changes in older age. The Neo-Socioanalyic perspective would argue that systematic personality changes should thus be expected through the adult life span. For example, traits such as Agreeableness and Conscientiousness tend to increase, while Neuroticism tends to decrease with increasing age. This finding, formalised as the *maturity principle* (Caspi et al., 2005), suggests that those traits that facilitate the success in social roles normally associated with maturity – such as career achievement or starting a family – namely Agreeableness and Conscientiousness, should increase as people age, or mature. On the other hand, those traits that would hinder such achievement – namely Neuroticism – should decrease with increasing age.

To this effect, the aim of study 4 was to estimate the patterns of normative mean-level change in personality traits across the adult life span, thus effectively modelling the development of personality in adulthood. Therefore, this study sought to provide an important test of the different predictions regarding personality development in adulthood. Moreover, study 4 employed a novel modelling approach – complementing the Cohort-Sequential Latent Growth Models with cohort-specific analyses – in order to begin to address a major problem in most developmental personality research – that is the unavoidable entanglement of cohort differences and systematic effects of time or ageing in the estimation of developmental trajectories of change.

Chapter V

Study 4: "Normative personality trait development in adulthood: A 6-year cohort-sequential growth model."

Reference:

Milojev, P., & Sibley, C.G. (2016). Normative personality trait development in adulthood: A six-year cohort-sequential growth model. *Journal of Personality and Social Psychology, manuscript in press.*

Please note that the supplementary files accompanying this chapter are presented in Appendix A.

Abstract

The present study investigated patterns of normative change in personality traits across the adult life span (19 through 74 years of age). We examined change in Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience and Honesty-Humility using data from the first six annual waves of the New Zealand Attitudes and Values Study (N = 10,416; 61.1% female, average age = 49.46). We present a Cohort-Sequential Latent Growth Model assessing patterns of mean-level change due to both aging and cohort effects. Extraversion decreased as people aged, with the most pronounced declines occurring in young adulthood, and then again in old age. Agreeableness, indexed with a measure focusing on empathy, decreased in young adulthood and remained relatively unchanged thereafter. Conscientiousness increased among young adults then levelled off and remained fairly consistent for the rest of the adult life span. Neuroticism and Openness to Experience decreased as people aged. However, the models suggest that these latter effects may also be partially due to cohort differences, as older people showed lower levels of Neuroticism and

Openness to Experience more generally. Honesty-Humility showed a pronounced and consistent increase across the adult life span. These analyses of large-scale longitudinal national probability panel data indicate that different dimensions of personality follow distinct developmental processes throughout adulthood. Our findings also highlight the importance of young adulthood (up to about the age of 30) in personality trait development, as well as continuing change throughout the adult life span.

Introduction

Understanding and mapping the development of traits and other aspects of personality is a central endeavour of personality research, and also a topic of much debate (McCrae & Costa, 2008; Roberts Wood, & Caspi, 2008; Specht et al., 2014). Developmental patterns of change in personality have direct implications for the understanding of personality as a construct and the processes that govern individuals' development (Roberts, Walton, & Viechtbauer, 2006). Accordingly, the nature of personality change and stability has been the focus of numerous distinct theories (see e.g., Specht et al., 2014), meta-analyses (Ardlet, 2000; Ferguson, 2010; Roberts & DelVecchio, 2000; Roberts e al., 2006), and empirical studies (e.g., Anusic, Lucas, & Donnellan, 2012; Lucas & Donnellan, 2011; Milojev & Sibley, 2014; Milojev, Osborne, & Sibley, 2014; Soto & John, 2012; Specht, Egloff, & Schmukle, 2011).

There is general consensus in the research literature that personality traits such as those estimated within a Big Five framework (McCrae & Costa, 1999, 2008) show both consistency and change across the life span. In this context, mean-level change refers to changes in the average levels of a given personality trait in the population of interest, over time and across different periods of the life span. Mean-level change is sometimes referred to as normative change, as the patterns of change across the life span are assumed to be

generalizable to most people most of the time (Roberts et al., 2006). Mean-level change, in other words, represents the average trajectory of individual change in personality traits across the life span. Comprehensively testing patterns of individual development over the life span (or within-person change) requires longitudinal samples that are heterogeneous in terms of participant characteristics such as age or the life span duration that is represented (Specht et al., 2011). As many researchers are well aware, inferring developmental trends from largely cross-sectional data raises some serious issues with regard to the differentiation between developmental versus cohort effects. Such analyses offer hints of the developmental trends, but cannot rule out cohort effects. Yet studies employing representative longitudinal samples spanning long periods of time remain few and far between (Lucas & Donnellan, 2011; Specht et al., 2011).

The paucity of data available to investigate normative change in personality – that is longitudinal, heterogeneous, multi-cohort designs – highlights one of the key challenges in developmental research. Specifically, we do not yet know the extent to which suggested mean-level changes in personality (e.g., Anusic, Lucas, & Donnellan, 2012; Donnellan & Lucas, 2008; Lucas & Donnellan, 2009; McCrae et al., 2000; Srivastava et al., 2003; Terraciano et al., 2005) are due to normative change (the developmental process of ageing, or change over time) versus cohort differences (systematic differences in the level of personality traits across different age cohorts). This entanglement of the effects of time and cohort is an ongoing issue in personality research, and one which we hope the methodological approach presented here may help address.

In the present investigation we leverage data from a large national longitudinal panel study of New Zealand adults, the New Zealand Attitudes and Values Study (NZAVS).[12] The NZAVS provides repeated measurements of personality traits at six consecutive annual

[12]For details about the NZAVS see: www.psych.auckland.ac.nz/uoa/NZAVS

assessments from 2009-2014. The NZAVS data have the sample characteristics desired for investigations of normative change in personality across the life span. That is, the NZAVS is a large and heterogeneous multi-cohort longitudinal sample of adult New Zealanders. Here, we employ Cohort-Sequential Latent Growth Models (Preacher, Wichman, MacCallum, & Briggs, 2008; Prinzie & Onghena, 2005) to investigate normative development in the Big Six personality traits (Sibley et al., 2011) over six years of assessment. The Big Six framework is based on the recent (H) honesty-humility, (E) emotionality, (X) estraversion, (A) agreeableness, (C) conscientiousness, (O) openness to experience model (HEXACO; Ashton & Lee, 2007, 2009; Sibley et al., 2011; see also Milojev & Sibley, 2014; Milojev et al, 2013; Milojev, Osborne & Sibley, 2015, for investigations of personality stability and change using the Big Six framework).[13]

Theoretical Perspectives

There are two broadly different theoretical approaches in the field of personality development (Kogan, 1990; Neyer & Asendorf, 2001, Specht et al., 2014). These two approaches differ primarily in terms of the weight they place on biological versus environmental processes in personality change and development. On the one hand, McCrae and Costa's (2008) the Five Factor theory emphasises the biological roots of personality. On the other hand, Roberts and Wood's (2006) Neo-Socioanalytic theory of personality tends to emphasise environmental effects on personality development (Roberts & Wood, 2006). Of course, these theories are not in opposition, as both processes are almost certainly at play.

The Five Factor theory emphasises the continuity and heritability of personality traits. This perspective suggests that the normative changes in personality are produced by intrinsic

[13] As discussed in a latter section, it is important to note that the short-form personality scales used in the present study are not a formal short-form measure of the HEXACO model, or a short-form of the HEXACO-60 (Ashton & Lee, 2009). The present study is based on the Mini-IPIP6; an extension of the Five Factor Mini-IPIP (Donnellan *et al*, 2006) designed to assess Honesty-Humility along with the Big Five using short-form four-item markers. We refer the reader to Sibley et al. (2011) for details on scale development and validation.

maturation and biological processes (McCrae et al., 2000; McCrae & Costa, 1999). In other words, personality development is determined primarily by temperament and genetic factors. This development occurs through childhood and adolescence, stabilising in adulthood. Very little change in personality, in terms of both rank-order stability or absolute change, is expected across the adult life span (assuming no non-normative biological changes; McCrae & Costa, 1999, 2008).

Numerous studies have documented considerable genetic variance indicative of heritable aspects of personality (e.g., Janng, Livesley, & Vemon, 1996; Loehlin, McCrae, Costa, & John, 1998; McCartney, Harris, & Bernieri, 1990). Indeed, a recent meta-analysis indicates that up to 40% of variance in personality may be due to genetic factors (Vukasovic & Bratko, 2015). In addition, cross-cultural studies indicate that the five factor model of personality is relatively consistent across cultures (McRae & Costa, 1994; McCrae et al., 1999, 2000). Early analysis of longitudinal data tended to suggest that personality development occurs up to 30 years of age, and was essentially unchanged and unaffected by environmental factors thereafter (Costa & McCrae, 1997). However, as discussed below, more recent empirical studies and meta-analyses have identified substantive developmental changes in personality across adulthood indicating a continuous process of personality development (e.g., Lucas & Donnellan, 2011; Roberts & DelVecchio, 2000; Roberts et al., 2006; Specht et al., 2011). Furthermore, longitudinal twin studies suggest that, while there is a considerable component of heritability, there is also a significant environmental component to personality (see McGue, Bacon, & Lykken, 1993; and meta-analysis by Vuksanovic & Bratko, 2015). Indeed, there is even some evidence to suggest that the effects of heritability or genetics, versus the effects of the environment, vary across different periods of the life span (Kandler, 2012).

Neo-Socioanalytic theory, in contrast, emphasises the process of person-environment transactions, incorporating both the continuity of a trait and the variation in environmental factors (Roberts & Caspi, 2003; Roberts et al., 2006; Roberts & Wood, 2006). This perspective emphasises the contribution that normative social institutions and life events—such as school, work, family, and normative events associated with these institutions—have in shaping personality (Roberts et al., 2006). According to this perspective, change in social roles and role expectations should lead to change in personality. As people's roles and associated identities become more stable – such as across middle age – their personalities should also become more stable in terms of both rank-order stability and mean-level change. Conversely, in life periods associated with changes in social roles – such as in younger and older age –personality should be less stable. An interesting implication of this perspective is that to the extent to which normative life events and commitments to normative social institutions occur at similar points across the life span for most people, we should be able to detect this signal and identify substantive patterns of change and stability in personality traits across the life span.

Normative Change across the Life-Span

Developmental research assessing normative change in personality traits derives many of its inferences from cross-sectional data assessing difference across age groups (e.g., Anusic, Lucas, & Donnellan, 2012; Donnellan & Lucas, 2008; Lucas & Donnellan, 2009; McCrae et al., 2000; Srivastava et al., 2003; Terraciano et al., 2005). The developmental patterns inferred from such research tend to be broadly consistent with those observed in the available longitudinal samples (Anusic, Lucas, & Donnellan, 2012; Roberts et al., 2006; Specht et al, 2011; van Der Akker et al., 2014; Wortman et al., 2012).

Perhaps the most comprehensive investigation of the longitudinal study of personality change is the meta-analysis by Roberts et al. (2006). This meta-analysis suggests that mean-levels of Extraversion tend to decrease over the life span; however, different facets may show different developmental patterns (see also, Lucas & Donnellan, 2011). That is, while the social vitality facet may decrease with age, particularly across younger to middle adulthood, the social dominance facet may increase with age. Although the meta-analysis by Roberts et al. (2006) suggested that increases in Agreeableness are limited to old age, more recent findings in the literature suggest that the levels of Agreeableness may tend to increase fairly consistently across the life span as people age (Lucas & Donnellan, 2009; Lucas & Donnellan, 2011; Specht et al., 2011). Available data further indicate that the development of Conscientiousness follows a curvilinear pattern. That is, although the finding that mean levels of Conscientiousness increase as people age, and particularly in younger adulthood, are robust, (Lucas & Donnellan, 2009; Roberts et al., 2006), there is also evidence to suggest a subsequent decrease in older age (Lucas & Donnellan, 2011; Specht et al., 2011; Terraciano et al., 2005). Neuroticism tends to show a decrease with age (Roberts et al., 2006). However, recent studies have failed to replicate this trend showing instead relatively unchanged levels across the adult life span (Lucas & Donnellan, 2011). Finally, evidence from meta-analyses (Roberts et al., 2006) and more recent studies (Lucas & Donnellan, 2011, Specht et al., 2011) indicates that Openness to Experience tends to decrease in absolute level across adulthood, and particularly in old age (see Roberts. et al., 2006).

To sum up thus far, the patterns of normative change in the Big Five traits reported in the literature lend support to the notion of continued change in personality across the life span. The general consistency of these findings suggests that transactional processes between a person's dispositions and environmental influences continue throughout adulthood (Caspi et al., 2005; Roberts & Caspi, 2003; Roberts et al., 2006).

One feature of this perspective is that personality change may in part be due to the different contingencies and prescriptions associated with different environmental circumstances and social roles that people engage with (Roberts et al., 2006). For example, the developmental patterns of Conscientiousness, Agreeableness, and even Neuroticism, are indicative of a *maturity principle* (Caspi et al., 2005). That is, mean-level changes should be expected for those traits that are associated with performance of the roles of adulthood – such as family or career. Successfully adapting to such roles and role expectations may require adaptations associated with higher levels of Conscientiousness (e.g., industriousness), and Agreeableness (e.g. working with others), and lower levels of neuroticism (e.g., higher emotional stability). Intriguingly, recent findings suggest the possibility of a temporary deviation from this latter pattern and also potential gender differences in it (see Van den Akker, Dekovic, & Prinzie, 2014). Van den Akker and colleagues found evidence of a temporary decrease in these maturity-related personality traits in adolescence, followed by an increase in later adolescence and early adulthood. Indeed, recent findings have also detected domain-specific changes in personality traits associated with life events that signify engagement in normative social roles. For example, Specht and colleagues (2011) documented systematic increases in Conscientiousness among those people who entered the job market and commenced a new job – a social role and a circumstance that require higher levels of Conscientiousness to enhance performance, achievement, and market competitiveness (for further evidence of domain-specific changes in personality associated with life-events see Löckenhoff *et al.*, 2009; Milojev, Osborne, & Sibley, 2014).

The observation of mean-level change in older age, particularly the well-documented decrease in Conscientiousness (Lucas & Donnellan, 2011; Specht et al., 2011; Terraciano et al., 2005), has been referred to as the *la dolce vita* effect (Marsh, Nagengast, & Morin, 2012). This decrease in Conscientiousness in old age may be associated with retirement and

the breakdown of task-oriented social institutions, as well as intrinsic, biological changes. However, the *la dolce vita* effect suggests that as people enter into old age they become more content with themselves resulting in being less preoccupied with productivity (decreased Conscientiousness), as well as in generally becoming happier (decreased Neuroticism and increased Agreeableness), less sociable and more satisfied with smaller groups of relations (lower Extraversion), and less open (lower Openness to Experience; Marsh, Nagengast, & Morin, 2012; also see Roberts et al., 2006).

The aforementioned literature generally suggests that personality development is an integrative process characterised by genetic factors or heritability (e.g., Janng, Livesley, & Vemon, 1996; Loehlin, McCrae, Costa, & John, 1998; McCartney, Harris, & Bernieri, 1990; Vukasovic & Bratko, 2015), development in early life (Costa & McCrae, 1997; Soto, *in press*; Soto et al., 2011), as well as a continuous process that is at least in part due to social demands (Roberts et al., 2006). However, while the research evidence tends to support the notion of normative variability in personality across the whole range of the adult life span (e.g. Lucas & Donnellan, 2011; Specht et al., 2011; Wortman et al., 2012), longitudinal studies remain necessary to establish a coherent picture of normative life span development.

Furthermore, extant longitudinal studies (e.g., Lucas & Donnellan, 2011; Specht et al., 2011) are not yet able to account for the possibility of cohort effects in explaining the observed patterns of normative personality change. With some notable exceptions, the majority of the literature in this area is based on investigations of cross-sectional differences (i.e. cohort differences), or longitudinal within-person change over a time frame that is too short, or focused on too narrow a range of age cohorts (Roberts et al., 2006). While it would take a lifetime of repeated assessments with a sample that is heterogeneous in terms of the age-cohorts to estimate the patterns of personality change across the full life span, with the availability of longitudinal panel samples and advanced modelling techniques, we can

provide a better approximation of such processes than ever before. This is the goal of the current study.

Overview of the Present Study and Guiding Hypotheses

First, we contribute to the corpus of literature investigating developmental patterns of normative change in personality traits firstly through the use of a large, heterogeneous, longitudinal sample. Until recently, methodologically adequate longitudinal studies of the developmental patterns of mean-level change have been rare (Lucas & Donnellan, 2011; Roberts et al., 2006; Specht, et al., 2011; Wortman et al., 2012). Furthermore, in agreement with many of the researchers who have paved the way for our work, we believe that to adequately investigate normative change in personality it is necessary to employ samples that represent the population of interest as closely as possible. The most notable samples in this tradition have been conducted in Germany and Australia (Lucas & Donnellan, 2011; Specht, et al., 2011; Wortman et al., 2012). We contribute to this research corpus and expand upon it by investigating normative change in personality in a longitudinal panel sample of adult New Zealand residents, the New Zealand Attitudes and Values Study.

Second, we investigate change across the adult life span by considering both developmental and cohort effects. In doing so, we employ a modelling approach that may allow researchers to disentangle the change due to passage of time from the differences attributable to cohort effects. To these ends, we estimate a series of Cohort-Sequential Latent Growth Models (Preacher, Wichman, Mac, Callum, & Briggs, 2008; Prinzie & Onghena, 2005) through two complementary approaches – a single-group approach and a multi-group framework. When considered together these two approaches allow us to observe both cohort differences and change over time. The details of the modelling approach are presented in the

sections below.[14] This framework builds on the recent work of Lucas and Donnellan (2011) and Wortman and colleagues (2012), as well as the latent growth models presented by Specht and colleagues (2011; also see Van der Akker et al., 2014, for an application of cohort sequential models to investigating personality change in childhood and adolescence).

We derived our hypotheses of expected patterns of mean-levels change in the Big Five personality factors from the propositions of the Neo-Socioanalytic perspective (Roberts & Woods, 2006; Roberts et al., 2006). Following the *maturity principle*, we expected to observe increases in Agreeableness and decreases in Neuroticism with increasing age (Caspi et al., 2005; Lucas & Donnellan, 2009; Lucas & Donnellan, 2011; Terraciano et al., 2005; Roberts et al., 2006; Soto et al., 2011). Similarly, we expected to detect increases in Conscientiousness with increasing age, particularly across the cohorts of early adulthood. However, we also expected to observe a general curvilinear pattern of change in Conscientiousness. Indeed, as suggested by the *la dolce vita* effect (Marsh, Nagengast, & Morin, 2012) we expected to detect a decrease in Conscientiousness among the oldest cohorts in our model. We also expected to observe a decrease in Extraversion with increasing age (Lucas & Donnellan, 2011; Roberts et al., 2006). Similarly, we expected to observe a decrease in absolute levels of Openness to Experience, particularly in older age (Lucas & Donnellan, 2011; Roberts et al., 2006).

In terms of the extent to which these patterns might be associated with cohort differences versus change over time, we made no concrete predictions. In simple terms, more data are needed to ascertain the extent to which the normative changes in personality discussed above may be due to aging versus cohort differences. However, our multi-group cohort-sequential latent growth model offers a method that advances our interpretation of

[14] The full syntax for our models is available as a supplementary document, on the NZAVS website, or by contacting the corresponding author.

such effects. Our reasoning is as follows: if normative change in a given trait is largely due to change over time, we should expect to see significant and comparable latent trajectories in relevant age cohorts with seamless continuation from one cohort to the subsequent one. If, however, there is a discrepancy in trajectories between two adjacent cohorts and if the end values estimated for one cohort—*values estimated for the final year of assessment*—are different from the starting values of the subsequent cohort—*values at the initial point of measurement*—we can assume cohort effects. In addition to this, we utilised a multi-group Structural Equation Modelling framework to formally investigate cohort differences in change trajectories.

Finally, rather than investigating the developmental patterns of only the Big Five personality traits, we also investigated normative change in a marker of Honesty-Humility within a Big Six framework (Ashton & Lee, 2007, 2009; Sibley et al., 2011; for an empirical study using a different six-factor model of personality see Soto, *in press*). Honesty-Humility is a trait that reflects, for instance, one's sense of self-entitlement or lack thereof (Ashton & Lee, 2007, 2009).

There are very few studies to our knowledge investigating developmental patterns of the Big-Six personality traits (e.g., Milojev & Sibley, 2014) and no studies at all specifically pertaining to developmental patterns of normative change in Honesty-Humility (but see Twenge, Konrath, Foster, Campbell, & Bushman, 2008 for work focusing specifically on Narcissism). Based on the conceptualisation of Honesty-Humility as being related to low Narcissism (Ashton & Lee, 2007; Sibley et al., 2011) and also some facets of Agreeableness (DeYoung, 2010), we expected to observe a systematic increase in this trait with increasing age.

There are two primary reasons for investigating personality change within the broad HEXACO framework and including the sixth personality marker of Honesty-Humility. First, the six-factor HEXACO-based model represents a viable alternative to the traditional Big Five, and it may better account for cross-cultural nuances (Ashton & Lee, 2007). Second, given that Honesty-Humility is conceptualised as relating to fairness (honesty) and the lack of self-entitlement (humility), and is thus distinct from Agreeableness, we were interested in the extent to which this trait would show a distinct, or similar, pattern of normative change in our Cohort-Sequential Latent Growth Models. Tracking the developmental trends in Honesty-Humility may thus further inform the pattern of personality change formalised as the *maturity principle*. Indeed, it can be argued that increased levels of Honesty-Humility would indicate an adaptive process of change, or maturation. An increase in Honesty-Humility paired with cohort differences where higher levels of the trait were observed among older cohorts would also be consistent with research suggesting that Narcissism (which bears some overlap with low Honesty-Humility) may be increasing in younger generations (Twenge et al., 2008).

Method

Sampling procedure and sample Participants

Participants

The models presented here were based on 10,416 (61.1% women) participants who responded to at least three out of the six waves of the New Zealand Attitudes and Values Study (NZAVS).[15] The NZAVS is an ongoing study that has been conducting an annual

[15] Differences on key variables were identified between the sample of participants meeting the selection criterion and the sample of participants not meeting this criterion. Specifically looking at differences at the first point of measurement (wave 1), those who met the selection criterion tended to be older (M = 49.46, SD = 14.98) than those not selected (M = 44.62, SD = 17.08, *d* = .30). The selected sample was also more gender biased with women comprising 61.1% in the selected sample, versus 55.5% of those not selected. They also tended to live in more affluent neighbourhoods – i.e. those with lower scores on the NZ Deprivation Index (M = 4.87, SD = 2.79, and M = 5.57, SD = 2.93, for those selected and not selected respectively, *d* = .24). In terms of personality traits, those who met the selection criterion of having completed at least three waves had lower

longitudinal panel survey of adult New Zealanders since 2009 (Time 1) through 2010 (Time 2), 2011 (Time 3), 2012 (Time 4), 2013 (Time 5), and 2014 (Time 6).[16]

The average age of the sample in 2009 (Time 1) was 49.46 (*SD* = 14.98). For the purposes of the present analyses (in estimating the multi-group cohort-sequential latent growth models), participants were grouped into 5-year cohorts based on the year of their birth. These birth cohorts and their respective sample sizes are presented in Table 5.1.

The majority of participants (85.7%) identified as NZ European (the majority ethnic group in New Zealand). In terms of socioeconomic status, the New Zealand Deprivation index is a decile based measure of deprivation in neighbourhood units across the country with 1 representing the most affluent neighbourhoods and 10 representing the most deprived (for more detail see Salmond, Crampton & Atkinson, 2007, Salmond & Crampton, 2012, White, Gunston, Salmond, Atkinson, & Crampton, 2008). Within the NZAVS, the NZ Deprivation index is calculated based on the participants' residential address. The average score on the NZ Deprivation index in the present sample was 4.87 (*SD* = 2.79). Recent work by Satherley and colleagues (2015) provides an extensive investigation of sample attrition over the first four years of the NZAVS. The analyses highlight the relatively low levels of attrition with over 60% of participants retained from the first wave (2009) and over 80% wave to wave retention in the later years.[17]

Extraversion scores (M = 4.02, SD = 1.15, versus M = 4.11, SD = 1.17, *d* = .07), higher scores on Agreeableness (M = 5.29, SD = .97, versus M = 5.12, SD = 1.04, *d* = .16), higher scores on Conscientiousness (M = 5.15, SD = 1.06, versus M = 4.98, SD = 1.12, *d* = .15), lower scores on Neuroticism (M = 3.38, SD = 1.10, versus M = 3.89, SD = 1.07, *d* = .18) and higher scores on Honesty-Humility (M = 5.16, SD = 1.30, versus M = 4.71, SD = 1.43, *d* = .33). For a more detailed analysis of sample retention and bias associated with the longitudinal nature of the NZAVS please refer to Satherley *et al.* (2015).

[16] Full information on the sampling procedure, retention strategies, attrition rates and sample sizes across the five years is provided in the Supplementary file 1.

[17] Interestingly, the analyses by Satherley *et al* (2015) do indicate some systematic attrition related to personality traits. For example, Conscientiousness and Honesty-Humility tend to be associated with higher sample retention (i.e. lower likelihood of a participant dropping out or abandoning the study). The implications of these effects are covered in the discussion section.

Table 5.1. Sample sizes and the represented age range for the 5-year birth cohorts.

Birth Cohort	N	Age in T1 (approx. 2009)	Age in T5 (approx.. 2013)
1940 and older	396	69	74
1941 to 1945	696	64	69
1946 to 1950	1159	59	64
1951 to 1955	1246	54	59
1956 to 1960	1306	49	54
1961 to 1965	1287	44	49
1966 to 1970	1059	39	44
1971 to 1975	934	34	39
1976 to 1980	697	29	34
1981 to 1985	461	24	29
1986 to 1990	549	19	24
1991 and younger	251	18	23

Measures

Markers of the Big Six personality were assessed using the Mini-IPIP6 (Sibley et al., 2011). The Mini-IPIP6 is a 24-item scale containing the 20 items developed by Goldberg (1999) and included by Donnellan et al. (2006) in the original Mini-IPIP. In addition, four items were used to index Honesty-Humility. Two of the items were adapted from the Narcissism scale developed by Campbell et al. (2004), whereas the remaining two were adapted from Ashton and Lee's (2009) HEXACO measure of Honesty-Humility. The factor structure and internal reliability of the Mini-IPIP6 scale has been extensively validated (Sibley et al., 2011; Sibley, 2012; Sibley & Pirie, 2013). Previous analyses of NZAVS data further indicate that the Mini-IPIP6 personality markers have high rank-order stability over a one-year and a three-year period (Milojev & Sibley, 2014; Milojev et al., 2013).

The Mini-IPIP6 was administered with the following instructions: "This part of the questionnaire measures your personality. Please circle the number that best represents how

accurately each statement describes you." Items were rated on a scale ranging from 1 (very inaccurate) to 7 (very accurate). Extraversion was assessed by the following four items: "Am the life of the party."; "Don't talk a lot." (reverse scored); "Keep in the background." (reverse scored); "Talk to a lot of different people at parties.". Agreeableness was assessed by the following four items: "Sympathize with others' feelings."; "Am not interested in other people's problems." (reverse scored); "Feel others' emotions."; "Am not really interested in others." (reverse scored). Conscientiousness was assessed by the following four items: "Get chores done right away."; "Like order."; "Make a mess of things." (reverse scored); "Often forget to put things back in their proper place." (reverse scored). The Neuroticism scale included the following four items: "Have frequent mood swings."; "Am relaxed most of the time." (reverse scored); "Get upset easily."; "Seldom feel blue." (reverse scored). Openness to Experience was assessed by the following four items: "Have a vivid imagination."; "Have difficulty understanding abstract ideas." (reverse scored); "Do not have a good imagination." (reverse scored); "Am not interested in abstract ideas." (reverse scored). Finally, Honesty-Humility was assessed using the four items as per Sibley and colleagues (2011). These items (all of which were reverse scored) included: "Would like to be seen driving around in a very expensive car."; "Would get a lot of pleasure from owning expensive luxury goods."; "Feel entitled to more of everything."; "Deserve more things in life."

Scale reliability estimates (Cronbach's α) for these six sub-scales across the five assessment occasions ranged between α = .64 to α = .78 (see Table 5.2). For tests of measurement invariance across the repeated assessments and age cohorts see Supplementary File 4. Mean scale scores were constructed by reverse-coding and taking the mean of each four item scale. These means were used in analyses.

Table 5.2. Internal reliability estimates (Cronbach's alpha) for the six sub scales of the Mini-IPIP6 across the six assessment occasions (T1-T6).

	Cronbach's α					
	T1	T2	T3	T4	T5	T6
Extraversion	.71	.75	.75	.73	.75	.75
Agreeableness	.66	.69	.69	.68	.70	.72
Conscientiousness	.65	.65	.65	.64	.66	.67
Neuroticism	.64	.72	.72	.68	.70	.71
Openness to Experience	.67	.70	.70	.68	.69	.71
Honesty-Humility	.78	.77	.77	.77	.77	.79

Analytic Strategy

Scale means for each of the Big Six personality were assessed in a separate Cohort-Sequential Latent Growth Model. The two types of models, detailed below, were estimated in *MPlus 7.4* (Muthén & Muthén, 1998-2015) using maximum likelihood with robust estimation of standard errors (MLR).

To assess the pattern of normative change in personality across the adult life span we estimated Cohort-Sequential Latent Growth Models (Preacher, Wichman, MacCallum, & Briggs, 2008; Prinzie & Onghena, 2005) using two different and, we argue, complementary approaches. Specifically, we estimated a single-group cohort-sequential latent growth model based on the participants' age at the times of assessment, which allowed us to estimate an overall change trajectory of personality traits from ages 19 to 74.[18] We complemented this approach by estimating a multi-group cohort-sequential latent growth model across five-year birth cohorts representing the same age range, based on individually varying time indicators. This approach in turn allowed us to assess the presence of cohort effects in the estimated change trajectories of personality traits across the adult life span.

[18] Note that the analyses included all participants, including those older than 74 in the analysis. However, we limit our estimation of model-implied growth trajectories to the 19-74 age range as the slopes may be less reliable for more extreme values of age due to decreasing sample size.

Single-group cohort-sequential latent growth models were estimated based on all of the six waves of annual assessments. Due to the sampling strategy of the NZAVS the repeated assessments correspond to the years 2009 (Time 1), 2010 (Time 2), 2011 (Time 3), 2012 (Time 4), 2013 (Time 5), and 2014 (Time 6). [19] The models were based on scale means at each of the six time points. In order to estimate a change trajectory over the available age-range of the adult life span, participants' exact ages (to two decimal places) at each assessment point were used as individually varying time indicators. For instance, a participant who was 19.50 years of age at their first response at Time 1, would generally be about 20.50 at the follow-up response at Time 2, and 21.50 at the second follow-up response at Time 3, assuming that they completed each wave at exactly one-year intervals. By estimating exact ages at each time point, we allowed for variation in age due to the varying intervals of assessment (i.e., 350 days for some people, 400 days for others, and so on). Continuing with the example of someone who was 19 at Time 1, that person's responses would thus inform the estimation of the growth curve in that particular age range. Similarly, a person who was 41 at their first response time would then inform a later portion of the growth curve. This means that, due to the diverse range of age-cohorts represented and the large number of participants of overlapping ages, a growth curve could be estimated representing change from ages 19 to 74, with the different participants' data informing different portions of the curve.

Using the TSCORE function in *MPlus* a latent intercept (i) and a latent slope (s) were estimated based on the participants' ages estimated as individually varying time indicators. For the purposes of the models, age was centred on the sample mean. Quadratic (q) and cubic (c) slopes were also estimated. However, where the higher order polynomial did not reach significance it was removed from that particular model. The latent intercept was estimated by

[19] For details on NZAVS sampling strategy pertinent to the survey and participant response timeline see the NZAVS website (http://www.psych.auckland.ac.nz/en/about/our-research/research-groups/new-zealand-attitudes-and-values-study/nzavs-tech-docs.html)

fixing the six factor loadings (T1 to T6) to 1. The latent intercept thus estimated the mean level of the given personality variable at the sample mean age. The latent slope was estimated based on individually-varying indicators of participants' age over time (T1 to T6). Similarly the quadratic and the cubic latent slope were estimated based on the quadratic and cubic functions of the individually varying indicators of participants' age over time, respectively. Thus, the latent slope(s) represent the linear or curvilinear change trajectory for a given personality trait across the available age range (19 through to 74).

To estimate the multi-group cohort-sequential latent growth model the sample was organised into 12 sequential 5-year birth cohorts as presented in Table 5.1.[20] Note that the 12th birth cohort (born in 1991 and later) was removed from the multi-group analyses due to small sample size. In congruence with the 5-year birth-cohorts, the multi-group growth model was estimated based on the first five waves of annual assessment (Time1 through to Time 5). Again, because of the sampling strategy of the NZAVS, exact times of collection were approximately annual and individually varying. The multi-group cohort-sequential latent growth models were thus estimated based on the individually varying time indicators (i.e. date of response rather than age at time of response) using the TSCORE function in the *MPlus* framework. Within these models a latent intercept (i) was estimated as above, along with the latent slope (s), for each of the 5-year age cohorts seen in Table 5.1. The variances of the latent intercept and the latent slope, and the covariance between the intercept and slope were constrained to equality across the birth-cohorts. Only the linear slope was estimated in these models.[21]

[20] Use of 5-year birth cohorts (rather than the available 6) was decided on based on the common practice in the available literature on aging and normative change (e.g. Lucas & Donnellan, 2011).
[21] Given the varying sample sizes across the birth cohorts, we conducted Monte Carlo Simulations to estimate statistical power to detect effects (latent slope) of varying magnitude at the different sample sizes available. Please see Supplementary file 2.

Because the multi-group models spanned consecutive 5-year periods, the organisation of birth cohorts into 5 year bands allowed us to sequentially organise the multiple LGMs. That is, each cohort LGM estimated change in the given personality trait over 5 years of assessment for that cohort – i.e. the 1986 to 1990 birth-cohort represented change from 19 years of age to 24 years of age; the 1981 to 1985 cohort represented change from 24 years of age to 29 years of age, and so on. Note that for each 5-year age cohort, the youngest age represented by that cohort was taken as an indicator of age in this framework. By employing this approach, the estimated levels of personality traits, the intercepts, and the latent change trajectories (i.e., the slopes) could be plotted across the broad range of the adult life. This allows for simultaneous investigation of estimated cross-sectional cohort differences in the latent intercepts (i.e. the cohort differences in levels of a personality trait at time 1), the change trajectories in each cohort and the cohort differences in the rate of change (i.e. the latent slopes for each 5-year birth cohort), as well as the overall pattern of change in a given personality trait that may be observed across the adult life span. Most importantly, this approach allows one to appreciate the age differences that are due to cohort differences, and those that are due to change over time. [22]

Results

Descriptive statistics and bivariate correlations for the six personality dimensions, Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility, at the six assessment occasions are presented in Supplementary File 5. The test-retest correlations indicate high stability for all of the Big Six personality traits across the six re-test periods. One-year re-test correlations (Time1 to Time 2) ranged from .61 to .75 - Extraversion r = .75, Agreeableness r = .61, Conscientiousness r = .69, Neuroticism r

[22] Sample *Mplus* syntax for the Single-Group and the Multi-Group Cohort-Sequential Growth Models can be found in the Supplementary file 3.

= .89, Openness to Experience *r* = .66, and Honesty-Humility *r* = .72. Five-year re-test correlations (Time 1 to Time 6) ranged from .59 to .71 - Extraversion *r* = .71, Agreeableness *r* = .59, conscientiousness *r* = .64, Neuroticism *r* = .62, Openness to Experience *r* = .64, Honesty-Humility *r* = .68.These correlations are comparable to those reported in previous research (Milojev et al., 2013; Milojev & Sibley, 2014).

Single-Group Cohort-Sequential Latent Growth Models

Table 5.3 presents the parameter estimates for the six single-group cohort-sequential latent growth models estimating mean-level change in the six personality dimensions over the six annual assessments across the ages of 19 to 74. Figure 5.1 shows the estimated values of Extraversion, Agreeableness and Conscientiousness from ages 19 through to 74. Figure 5.2 shows the estimated values of Neuroticism, Openness to Experience and Honesty-Humility. The Single-Group Cohort Sequential LGM is represented as the dark line, while the estimates based on the Multi-Group LGMs (see later section) are presented as the lighter lines in respective age-cohorts. Looking at the model estimating change in Extraversion, the estimated mean-level of Extraversion at the sample mean age (about 50 years of age; or 49.46 specifically) was 3.96, 95% Confidence Interval [3.92, 3.99]. Furthermore, the LGM indicated a decreasing cubic change trajectory from age 19. This trajectory is characterised by an initial decrease, decelerating by about the age of 40, followed by a further decrease in older age (around the age 70). This change trajectory is shown in the first panel of Figure 5.1.

As can be seen in Table 5.3, the estimated level of Agreeableness at 50 years of age was 5.30, 95% CI [5.27, 5.32]. The change trajectory of Agreeableness can be seen in panel 2 of Figure 5.1. This change trajectory followed a cubic functions, decelerating in younger age (up to about age 35), and remaining largely stable, with only small variations thereafter.

The estimated latent intercept for Conscientiousness indicated that the average level of Conscientiousness at age 50 was 5.10, 95% CI [5.07, 5.12]. The model further indicated that the change trajectory of Conscientiousness followed a cubic function, characterised by a decelerating increase through young adulthood through to middle age (up to about the age of 45). This trajectory then remained largely stable through the rest of the represented life span. This change trajectory can be seen in the third panel of Figure 5.1.

The LGM estimating change in Neuroticism indicated that the average level of Neuroticism at age 50 was 3.50, 95% CI [3.46, 3.52]. As can be seen in panel 1 of Figure 5.2, the change trajectory for Neuroticism followed a cubic function characterised by an accelerating decrease which then decelerated from around the age 60 and over.

Looking at Openness to Experience, Table 5.3 shows that the estimated level of Openness to Experience at age 50 was 4.94 [4.91, 4.96]. As can be seen in panel 2 of Figure 5.2, the change trajectory of Openness to Experience followed a cubic function characterised by a decrease in younger adulthood through to middle age, followed by relative stability, and an accelerating decrease in older age.

Finally, the LGM estimating change in Honesty-Humility indicated that the average level of Honesty-Humility at age 50 was 5.10 [5.07, 5.12]. As can be seen in Figure 5.2, the change trajectory of Honesty-Humility followed a linear increasing function indicating a consistent increase in Honesty-Humility from age 19 to 74.

Table 5.3. Parameter coefficients for the single-group age-based Latent Growth Model for the six personality traits estimating the change trajectory from age 19 to 74.

				95% CIs		
	Estimate	***se***	***p -value***	***Low***	***High***	***Variances***
Extraversion						
Intercept	3.96	.02	<.001	3.92	3.99	1.09*
Linear Slope	-.04	.01	<.001	-.017	-.02	.42*
Quadratic Slope	.01	.01	.042	-.00	.02	.05
Cubic Slope	-.00	.00	.048	-.01	.00	.00*
AIC = 98651.417; aBIC = 98712.516.						
Agreeableness						
Intercept	5.30	.01	<.001	5.27	5.32	.63*
Linear Slope	.01	.01	.13	-.00	.03	.19*
Quadratic Slope	.02	.00	<.001	.01	.03	.02
Cubic Slope	-.01	.00	<.001	-.01	-.01	.00*
AIC = 95825.928; aBIC = 95887.026.						
Conscientiousness						
Intercept	5.10	.01	<.001	5.07	5.12	.76*
Linear Slope	.05	.01	<.001	.03	.06	.29*
Quadratic Slope	-.03	.00	<.001	-.04	-.02	.03
Cubic Slope	.00	.00	.001	.00	.01	.00*
AIC = 95589.504; aBIC = 95650.603.						
Neuroticism						
Intercept	3.50	.01	<.001	3.46	3.52	.87*
Linear Slope	-.15	.01	<.001	-.17	-.13	.31*
Quadratic Slope	-.01	.01	.019	-.02	-.00	.03
Cubic Slope	.01	.00	.002	.00	.01	.00
AIC = 103766.735; aBIC = 103827.834.						
Openness to Experience						
Intercept	4.94	.01	<.001	4.91	4.96	.82*
Linear Slope	-.09	.01	<.001	-.11	-.08	.28*
Quadratic Slope	.01	.00	.110	-.00	.02	.02
Cubic Slope	-.00	.00	.010	-.01	-.00	.00*
AIC = 103196.507; aBIC = 103257.605; * p< .05.						
Honesty-Humility						
Intercept	5.10	.01	<.001	5.07	5.12	1.07*
Linear Slope	.18	.01	<.001	.17	.19	.01
AIC = 110119.783; aBIC = 110144.222.						

N = 10416; * p< .05.
AIC – Akaike Information Criterion; aBIC = Sample-size adjusted Bayesian Information Criterion

Multi-Group Cohort-Sequential Growth Models

The Multi-Group Cohort-Sequential LGMs for the six personality dimensions are also presented in Figures 1 and 2. These models estimate mean-level change over the five annual assessments – from October 2009 to July 2013 – in each of the 11 five-year birth cohorts.

As can be seen in the first panel of Figure 5.1, the defining characteristic of the pattern of change in Extraversion, as indicated by this modelling approach, is the systematic decrease in younger cohorts. Specifically, between the ages of 19 and 24 and between the ages of 24 and 29 the LGMs indicate a significant decrease in Extraversion (s = -.05 and s = -.03 for the two cohorts respectively). A levelling off in terms of change over time can be observed across middle aged cohorts, with a further decrease in older cohorts. Specifically, a negative slope can be seen between the ages of 59 and 64 (s = -.02) and between the ages of 69 and 74 (s = -.03). In terms of the cohort differences the model suggests that older age is associated with lower levels of Extraversion. The Multi-Group Cohort Sequential Model indicates an overall change trajectory that is subjectively comparable to that estimated by the Single-Group Cohort-Sequential modelling framework (represented as the darker line in Figure 5.1, and described in the previous section).

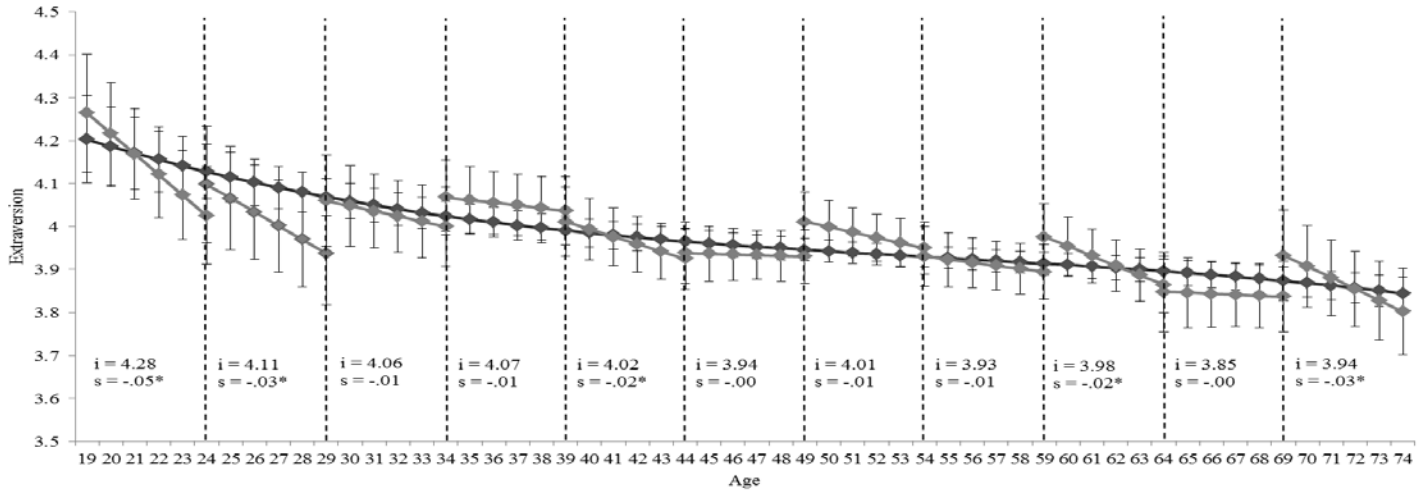
The estimated mean-level change in Agreeableness across the cohorts is presented in the second panel of Figure 5.1. An initial decrease can be seen between the ages of 19 and 24 (s = -.04), with no systematic changes in the later cohorts. The initial decrease in Agreeableness followed by relative stability is consistent with the change trajectory estimated by the Single-Group Cohort-Sequential Model described above.

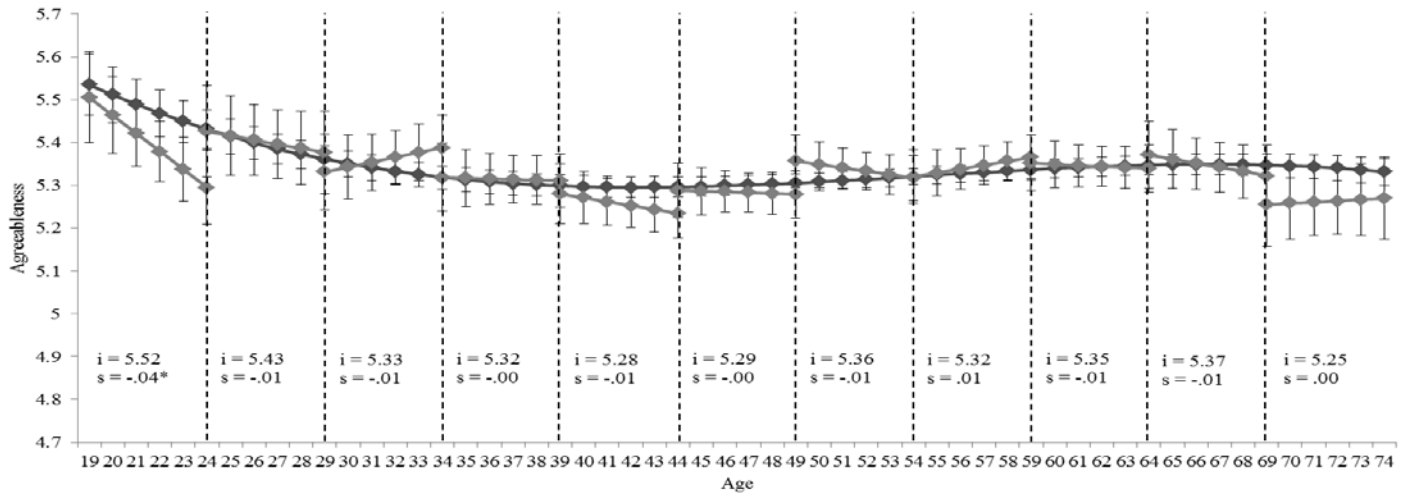
Consistent with the findings of the Single-Group LGM above, the Multi-Group Cohort-Sequential model indicated an initial increase in Conscientiousness, particularly between 19 and 24 years of age (s = .03). Further, cohort specific effects can also be observed, suggesting a decrease over the 5 years between 34 and 39 (s = .02).

The first panel of Figure 5.2 shows the developmental patterns of normative change in Neuroticism as estimated by the Multi-Group Cohort-Sequential LGM. A cohort based decrease can be observed with older cohorts showing systematically lower levels of Neuroticism. However, there was no apparent effect of aging, or within-cohort change over time, suggesting that the change pattern indicated by the Single-Group LGM may be largely due to cohort differences.

The Multi-Group LGM indicates a decrease in Openness to Experience across the cohorts of the adult life span as can be seen in Figure 5.2. However, an effect of aging can only be observed in the youngest cohort, with the only significant latent slope suggesting a decrease between the ages of 19 and 24 (s = .03). The decrease in Openness to Experience across the rest of the adult life span suggested by the Single-Group LGM seems to be largely due to cohort-differences.

Finally, the third panel of Figure 5.2 shows the developmental trend of normative change in Honesty-Humility. As suggested by the Single-Group Cohort-Sequential LGM, a distinct systematic increase across the life span can be observed. Significant latent slopes within cohorts can be observed between the ages of 19 and 64. This undisrupted increase across age cohorts indicates a developmental pattern whereby mean-levels of Honesty-Humility increase with increasing age, until the age of about 64. The change trajectories between the ages of 64 and 69 and 69 and 74 are not significant, suggesting a levelling off of Honesty Humility at about this age.





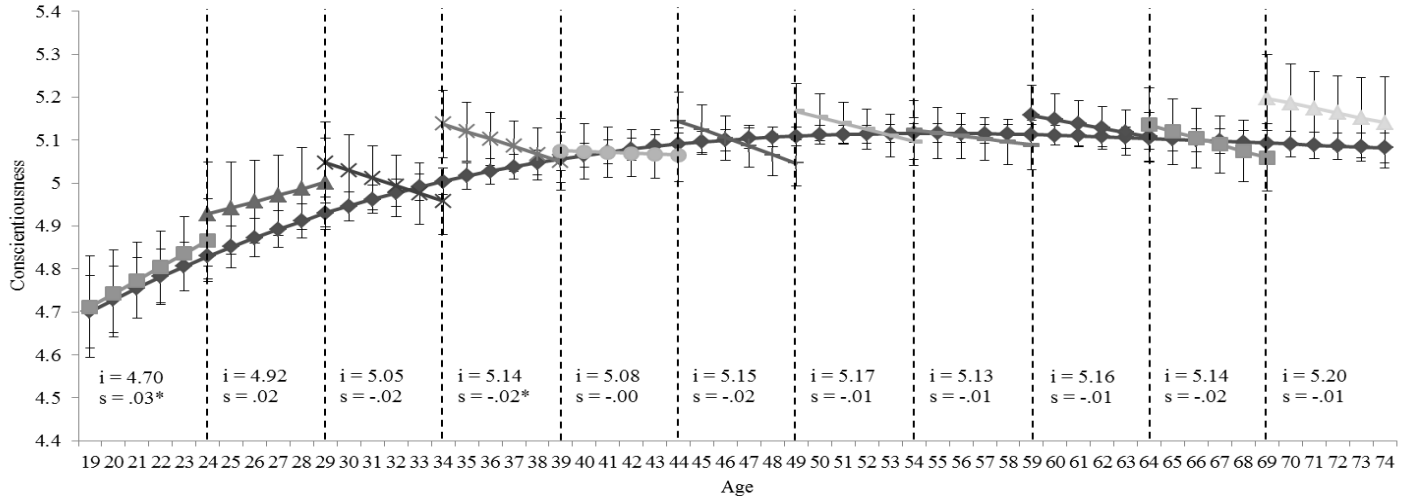


Figure 5.1. Developmental patterns of normative (mean-level) change in Extraversion, Agreeableness and Conscientiousness. Notes: Each panel shows the latent change trajectories for the given trait based on the (a) six-year single-group cohort-sequential latent growth model (dark line), and (b) five-year multi-group cohort sequential latent growth models across the 5-year birth cohorts presented in separate sections (light lines). The estimates of the latent intercept and the latent slope for each cohort are presented in the graph. The estimations are based on the mean-levels of the given trait (y axis) across age and assessments (x-axis). The 95% confidence intervals are presented as error bars around each point estimate. Within each cohort, i = intercept for that parameter, s = the fixed effect for the slope (* $p < .05$).

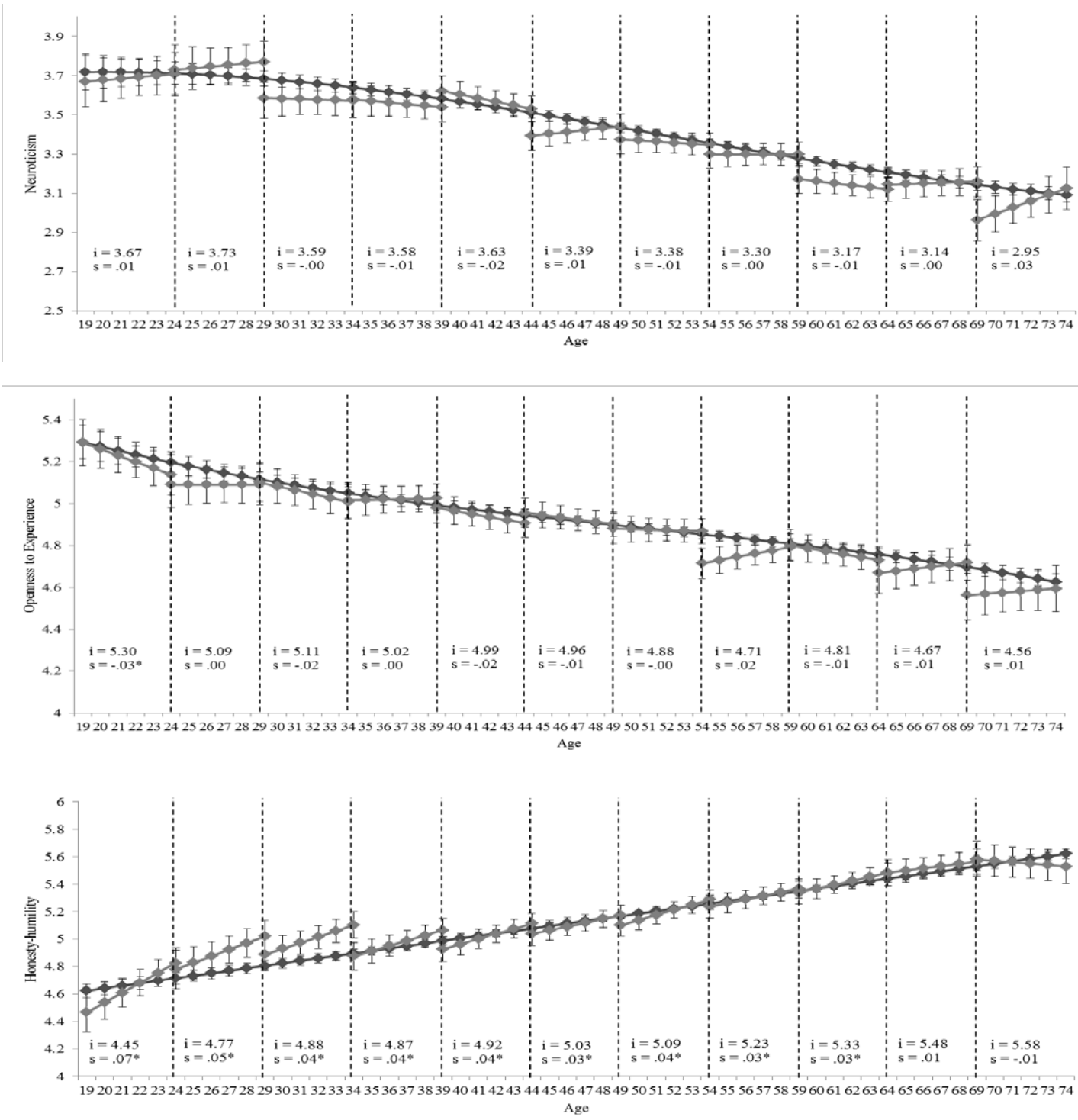


Figure 5.2. Developmental patterns of normative (mean-level) change in Neuroticism, Openness to Experience, and Honesty-Humility. Notes: Each panel shows the latent change trajectories for the given trait based on the (a) six-year single-group cohort-sequential latent growth model (dark line), and (b) five-year multi-group cohort sequential latent growth models across the 5-year birth cohorts presented in separate sections (light lines). The estimates of the latent intercept and the latent slope for each cohort are presented in the graph.). The estimations are based on the mean-levels of the given trait (y axis) across age and assessments (x-axis). The 95% confidence intervals are presented as error bars around each point estimate. Within each cohort, i = intercept for that parameter, s = the fixed effect for the slope (* $p < .05$).

A Note on Cohort Effects

Our multi-group models provide a potential avenue for quantifying the cohort effects in mean-level change across the adult life span. This is because the specification of 5-year age cohorts and the 5 yearly assessments provides for an overlap between the estimated mean level of a trait at the last assessment and the mean-level of the trait for the same age at the first assessment point of the next sequential cohort. To clarify, the LGM for the youngest cohort used in our model estimates change from age 19 at the first assessment to age 24 at the final – fifth – assessment. The LGM for the next cohort estimates change over the five yearly assessments starting at age 24 through to age 29. The models thus provide two estimates of mean-levels of a trait at age 24 – one based on the five-year latent change trajectory from age 19 to 24, and the other based on the initial level of the trait at the first assessment occasion. The discrepancy between these two estimates thus provides an indicator of the difference between the model-implied value due to change over time, and the model-implied value based on cohort differences.

Looking at the estimated trajectories for Extraversion in the first panel of Figure 5.1, we can see that all of the differences between estimated values at time 5 (2013) and the estimated values for the same age at time 1 (2009) are non-significant with overlapping confidence intervals. This lack of difference, plotted in the first panel of Figure 5.3, suggests that cohort effects do not play a major role in the developmental patterns observed for Extraversion. A similar lack of cohort differences can be observed for Agreeableness. On the other hand, some cohort differences can be seen for Conscientiousness (at age 34), Neuroticism (at ages 29, 44, and 59), Openness to Experience (at age 54), and Honesty-Humility (at age 34).

In the case of Conscientiousness the cohort effect indicates that the 29 to 34 year old cohort has lower estimated levels of the trait than the 34 to 39 year old cohort. The opposite pattern is observed for Neuroticism whereby the 24 to 29, 39 to 44, and the 54 to 59 year old cohorts have higher estimated levels of the trait than their respective subsequent age cohorts. Similarly, the cohort effect observed for Openness to Experience indicates that the 49 to 54 year old cohort has higher estimated levels of the trait than the subsequent, 54 to 59 year old cohort. Finally, the estimated level of Honesty-Humility for the 29 to 34 year old cohort is higher than for the subsequent 34 to 39 year old cohort, an interesting effect given the consistent increase in the trait across all of the age cohorts.

The above differences are not ideal indicators of cohort effects in these developmental patterns. However they do provide a novel method for approximating such effects, and we hope will pave the way for future research extending this approach into a more general framework. Refinement of this modelling framework may allow for a more accurate estimation of these cohort effects in the years to come and, moreover, may allow quantification of the amount of normative change that is due to these cohort effects versus the effects of age (i.e. change over time).

To complement our tentative analyses of cohort effects, we also conducted a formal test of cohort differences in the change trajectories in our Cohort-Sequential LGMs. Given the multi-group framework of the Multi-Group Cohort-Sequential LGMs we ran a subsequent set of models where the latent intercepts and the latent slopes were constrained to be equal across the age cohorts. The fit of these constrained models was then compared to the model fit of the free, or baseline, models – that is, the models presented in Figures 5.1 and 5.2 where the intercepts and the slopes were allowed to freely vary across the age-cohorts. The estimates for the free versus constrained models for the six personality traits are presented in Table 10.4. As can be seen in the table, in the case of Neuroticism, Openness to Experience,

and arguably Honesty-Humility, the unconstrained model provided a better fit. This is indicated by decreased values of the information criteria in the models with free parameters relative to the constrained model. This in turn implies that cohort differences may play a role in the estimation of the developmental pattern over the adult life span, in the case of Neuroticism, Openness to Experience, and Honesty-Humility.

Table 5.4. Model Fit indices comparing the Multi-Group Cohort Sequential Growth models allowing for cohort differences (Free Model) versus constraining the latent intercepts and latent slopes to equality across age-cohorts (Constrained Model).

	Free Model	Constrained Model
Extraversion		
AIC	77909.294	77923.636
aBIC	78260.788	78186.246
Agreeableness		
AIC	74767.170	74784.243
aBIC	75118.654	75046.846
Conscientiousness		
AIC	74963.814	75064.572
aBIC	75315.122	75327.182
Neuroticism		
AIC	80880.742	81269.034
aBIC	81232.235	81531.644
Openness to Experience		
AIC	80557.599	80805.007
aBIC	80909.092	81067.617
Honesty-humility		
AIC	86207.659	86611.144
aBIC	86559.153	86873.754

AIC – Akaike Information Criterion; aBIC = Sample-size adjusted Bayesian Information Criterion

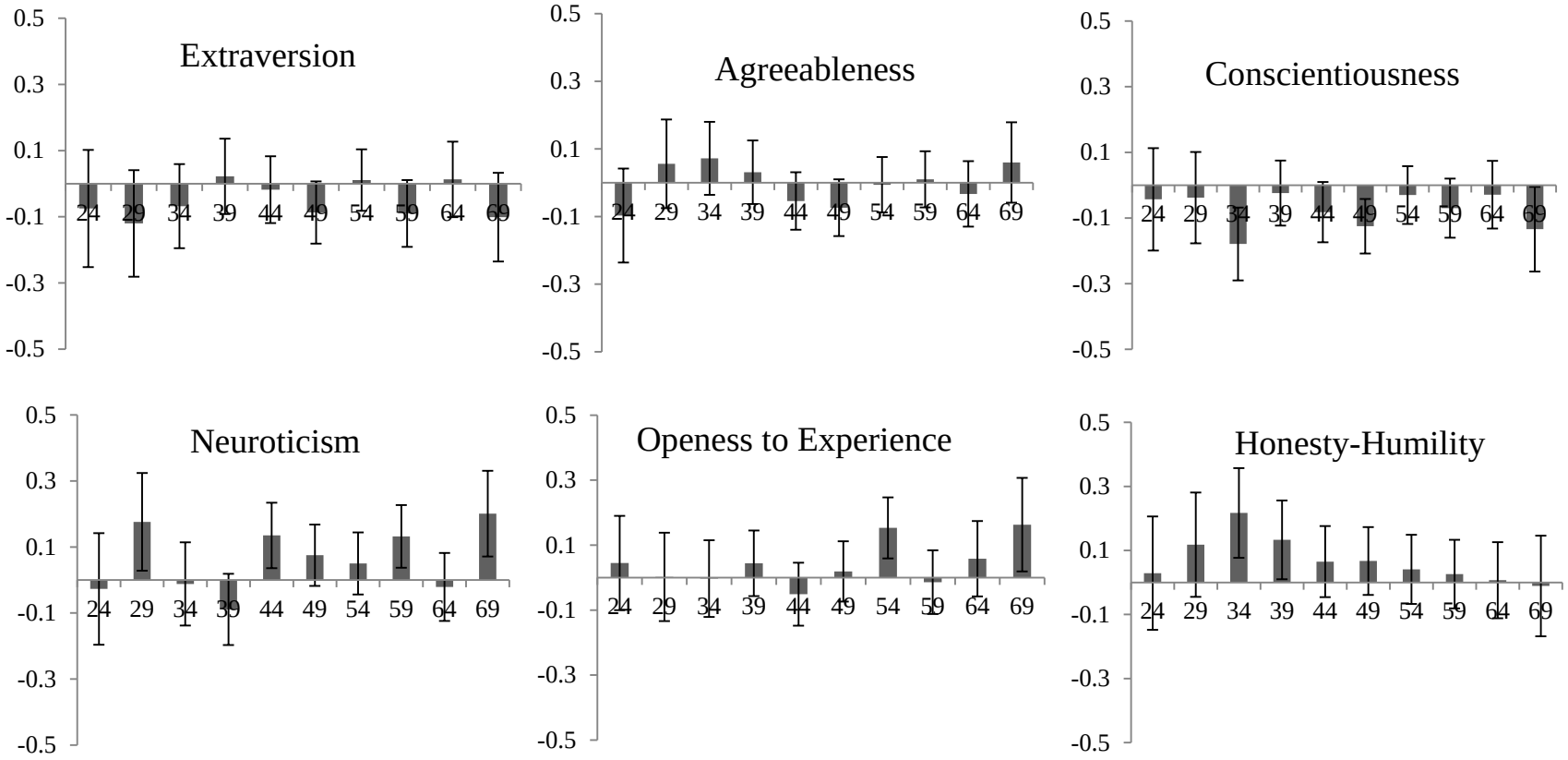


Figure 5.3. Estimated cohort effects between each of the cohorts as seen in Figure 1. Notes: Within each panel the bars represent the difference between the mean-level of the given trait at the participants' age as estimated by the latent growth trajectory from the preceding cohort and the intercept of the adjacent cohort. For example, the first bar in the panel showing cohort effects for Extraversion indicates the difference between the level of Extraversion at age 24 as estimated by the latent growth model from the youngest cohort - representing ages 19 through to 24 - and the level of Extraversion at age 24 based on the intercept of the subsequent cohort – representing ages 24 through to 29. Error bars represent 95% confidence intervals.

Discussion

The present study investigated patterns of normative change in markers of Big-Six personality traits across the adult life span. We did so by employing a series of Cohort-Sequential Latent Growth Models assessing patterns of change in personality traits in a large, heterogeneous longitudinal sample of adult New Zealanders. Overall, our analyses identify theoretically meaningful patterns of both consistency and change in personality across the adult life span. Our findings thus provide support for predictions from both the Neo Socioanalytic perspective (Roberts et al., 2006) and the Five Factor theory (McCrae & Costa, 1999) of personality development.

To sum up the key trends, Extraversion, Neuroticism, and Openness to Experience tended to decrease as people aged, and all followed cubic trajectories with different rates of decrease during different periods of adulthood. The developmental trajectory of Agreeableness was characterised by an initial decrease up to about 30 years of age, followed by relative stability across the rest of the life span. Conscientiousness increased at a decelerating rate, between 19 years of age and about 45 years of age, before settling at stable levels for the remainder of the life span. Finally, Honesty-Humility increased in a consistent linear fashion across the adult life span.

Normative Change in Personality Traits

Consistent with previous longitudinal findings (e.g., Lucas & Donnellan, 2011, Roberts et al., 2006), our models indicated that levels of Extraversion tend to decrease as people age. This decrease followed a cubic trajectory with a more extreme decline occurring up to about 30 years of age, followed by further slight decline during middle age. In older adulthood, particularly in early 60’s and 70’s the decrease in Extraversion accelerated again. This pattern suggests that Extraversion changes at different rates during different periods of

the life span. On the one hand, the decrease in Extraversion observed in younger cohorts – up to about 30 years of age – highlights the importance of early adulthood in personality development, as suggested by the Five Factor theory (Costa & McCrae, 2006). On the other hand, the continuous decrease across adulthood, accelerating in later years, supports the notion of continuous change (Roberts et al., 2006).

Moreover, the cubic pattern of development in Extraversion is broadly consistent with the meta-analytic findings documented by Roberts and colleagues (2006). Their meta-analysis suggested that Extraversion, specifically the Social Vitality facet, generally decreased across the life span with a sharp decrease in the 20's, followed by a period of stability through to the late 50's, followed by another decrease through to the late 60's. The short-form scale used to assess personality in the present study (the Mini-IPIP6; Sibley et al., 2011) is likely to be more closely related to the Sociability facet analysed by Roberts et al (2006; rather than the assertiveness, or social dominance facets). Similarly, in more recent analyses with a more restricted age range, Soto and colleagues (2011) suggested that the Activity facet of Extraversion decreases in adolescence and early adulthood. Likewise, Soto and John (2012) have suggested that Gregariousness (and not Assertiveness) decreases across age. Our findings, therefore, provide further evidence replicating and extending the finding that people tend to become less sociable across adulthood, and that this decrease is most pronounced between the ages of 19 and 30, and then again in later adulthood, after the age of 60.

Agreeableness, on the other hand, showed an unexpected developmental pattern. Our findings did not support the expected increase in Agreeableness with increasing age (Lucas & Donnellan, 2009; Lucas & Donnellan, 2011; Terraciano et al., 2005). Instead, Agreeableness showed an initial decrease between the ages of 19 and 29, followed by relative stability over the rest of the adult life span. This finding is particularly interesting in light of the *maturity*

principle (Caspi et al., 2005). One of the key premises of the maturity principle is that Agreeableness tends to show a consistent increase with increasing age. That is, as people age they adopt social roles and invest in social institutions associated with maturity. The success in these social roles (e.g., school, employment, family) places increasing demands on people to behave in ways associated with Agreeableness (e.g., interpersonal cooperation). Thus, as people mature, particularly during young adulthood and towards middle age, they should experience normative increases in levels of Agreeableness. Our findings, however, suggest almost the opposite. That is, the young adults in our sample—those who aged from 19 to 24 and from 24 to 29 years of age during our study—showed a systematic decrease in Agreeableness. In interpreting this trend it is important to keep in mind that Agreeableness as assessed by the Mini-IPIP6 in this study is more closely related to emphatic concern for others, rather than the cooperation, politeness, or the compliance facets of the trait (e.g. DeYoung et al., 2007). This suggests that it may be facets of Agreeableness related to emphatic concern that decrease in younger adulthood, up to about the age of 30, while other facets related to cooperation, such as politeness, may show the increasing pattern predicted by the *maturity principle*. Even so, the present finding is still surprising in light of the literature on personality development in adulthood (e.g., Soto & John, 2012).

The recent findings of Van der Akker and colleagues (2014) may also provide insight into the finding that Agreeableness (at least as related to empathy) tends to decrease over the lifespan. Van der Akker et al. observed a temporary counter of the expected maturation effect whereby there was an initial decrease in the expected trait (i.e. benevolence), rather than a consistent linear increase. However, the pattern indicated by our models is characterised by a decrease in young adulthood followed by relative stability in Agreeableness (i.e. empathic concern), rather than a subsequent increase. It may be the case that the observed decrease in empathic concern in young adulthood indicates a substantive effect. Indeed, the demands

associated with the social roles invested in during that period (pre-30 years old), such as career competition, as well as economic pressure in expectations of starting a family or buying a first home, may lead to decreases in empathic concern over that period. We would suggest that these may be similar pressures that would lead to a decrease in sociability, but not necessarily to decreases in facets of Extraversion related to social dominance or assertiveness. Indeed, the opposite would be expected, whereby these competition and maturity pressures would lead to increases in social dominance or assertiveness, as observed in previous literature (e.g., Roberts et al., 2006, Soto & John, 2012). These predictions, along with the prediction of an increase in Conscientiousness due to the same types of pressures, are the tenets of the *maturity principle* (Caspi et al., 2005; Roberts et al., 2006). The decrease in empathic concern and sociability, on the other hand, may be a reflection of the typical family or career demands whereby people become more self-focused and less concerned with broad social circles.

The *maturity principle* is a formalisation of the robust findings suggesting that Agreeableness, Conscientiousness, and Emotional Stability increase with increasing age (Caspi et al., 2005; Roberts et al., 2006). Although the decrease in Agreeableness was unexpected in light of these predictions, we did observe a systematic increase in Conscientiousness across adulthood. Conscientiousness increased at a decelerating rate from the age of 19 through to about 45 years old. Our findings thus suggest that as people get older they become more conscientious, particularly in young adulthood - between 19 and 29 years of age. This may be due to the demands and expectations of the social roles invested in and associated life events experienced, such as entering the job market, or the responsibility associated with starting a family, during this part of the life span.

We did not find evidence for a curvilinear pattern of normative change in Conscientiousness, as there was no specific or more dramatic decrease in older age relative to

other parts of the lifespan. Thus, our models do not support the decrease in Conscientiousness during older age observed in previous research (Lucas & Donnellan, 2011; Specht et al., 2011; Terraciano et al., 2005). Moreover, the *la dolce vita effect,* a model of trait changes in old age (Marsh et al., 2012), states that Conscientiousness should decrease in old age, as people become more relaxed and satisfied with their achievements, reducing their needs for industriousness or preoccupation with productivity. Our findings do not support these predictions. However, it is worth noting at this point that given the results of the power simulations reported in the supplementary documents, these 'null' findings should be interpreted tentatively. Further, some of the cohort-specific—or within cohort-latent slopes did identify a decreasing trend in older age—such as between 59 and 74. These cohort-specific growth factors did not reach significance, however. Thus, we are cautious about making firm assertions about a lack of decrease in Conscientiousness in older age. In addition, Conscientiousness as assessed by the Mini-IPIP (Donnellan et al., 2006) and the Mini-IPIP6 (Sibley et al., 2011) may be more closely related to orderliness than to industriousness or productivity. It is possible that as people progress towards old age they become less preoccupied with industriousness and productivity (Marsh et al., 2012), leading to decreases in this facet of Conscientiousness, while their need for orderliness may remain unchanged. This issue, however, is a matter of open debate that needs to be addressed through further facet-level research, and is beyond the scope of the present study.

We observed the expected decrease in Neuroticism over the adult life span. Specifically, this pattern of normative change followed a cubic function, with the most pronounced decease in Neuroticism occurring during the mid-to-late 30's and through to middle age. This finding supports the suggestion that as people mature and engage in socially prescribed roles associated with normative maturation, their levels of Neuroticism decrease in response to the demands of these roles (Caspi et al., 2005; Roberts et al., 2006). For people in

their late 30's, 40's and 50's this might mean that the relevant social context, such as that of the career or family demands, allows for greater levels of emotional stability. For example, through this period of the adult life span people may enjoy more stable career contexts than are found in younger adulthood. Similarly, people in this age-range tend to be establishing families and households. This may contribute to a decline in Neuroticism to the extent that establishing such roles and structures promotes stability (e.g. parenting and family). That said, the comparison of our two modelling approaches, and the tentative assessment of the cohort effects suggest that this pattern may also be at least in part due to cohort differences across these age ranges, rather than a developmental effect.

Keeping in mind the power restrictions mentioned above, our models seem to suggest an unexpected trend of increasing Neuroticism for those who aged between 69 and 74 years (i.e., our oldest cohort). While this may be simply noise in the data, the possibility that we are detecting a reliable trend is worth mentioning. Older age is typically a period of considerable physiological and cognitive change, usually associated with a decline in capability and health. Similarly, this is a period of life where roles and social networks tend to change often due to sickness and death. The combination of these factors may, in turn, be associated with domain-specific change in personality, such as the increase in Neuroticism, which has been shown to increase in relation to negative life events more generally (e.g., Löckenhoff et al., 2009; Milojev, Osborne, & Sibley, 2014). Interestingly, the increase in Neuroticism in older age observed in this study mirrors the recent findings on the developmental trajectory of self-esteem (Orth, Trzesnewski, & Robins, 2010). These researchers found a quadratic developmental pattern of self-esteem across adulthood, characterised by increases through young and middle age, followed by a distinct decrease after about the age of 60 years old. To the extent that self-esteem is related to personality traits such as Neuroticism (or Emotional Stability) it may be the case that the same processes argued to lead to the increases in

Neuroticism may lead to decreases in self-esteem. The question of the potential causal relationship between these two variables, however, remains open.

In terms of Openness to Experience, our models suggest a decreasing trend across the adult life span, with a slight acceleration in this decrease occurring in young adulthood – 19 through to about mid 30's—and older adulthood—early 60's onwards. However, as was the case with Neuroticism, the comparison of models and the tests of cohort effects suggest that this pattern may be in part due to cohort differences, with the older cohorts showing lower levels of Openness to Experience. We did detect some significant within-cohort change, with Openness tending to decrease as people aged from 19 to 24. Keeping in mind the aforementioned power restrictions, the non-significant decreasing patterns observed between 29 and 44, and between 39 and 44, do seem consistent with a systematic decrease during this period.

Interestingly, we did not observe strong evidence for a decrease in Openness to Experience among the elderly. While the overall pattern of change does suggest an accelerating decrease from the early-60's onwards, the cohort-specific models do not support this suggestion, with no significant change in older cohorts. In this regard, we failed to find direct support for the predictions associated with the *la dolce vita effect* model which would predict a decrease in Openness to Experience in old age (Marsh et al., 2012).

The most striking pattern of change across the adult life span was that of Honesty-Humility. As hypothesized, we observed a consistent linear increase in Honesty-Humility, indicating a systematic effect of aging across the adult life span. Indeed, all of the within-cohort latent slopes between the ages of 19 and 64 were significant. The consistent developmental change in Honesty-Humility provides an interesting comparison to the more complex developmental patterns observed for other personality traits. As discussed above,

our findings of the normative change in Agreeableness seemingly bring into question some of the predictions of the *maturity principle*. However, we argue that in the context of the six HEXACO personality traits it is Honesty-Humility which may represent those qualities that most facilitate maturation-related role success—fairness (honesty) and a lack of self-entitlement (humility). Therefore, according to the Neo-Socioanalytic perspective, and the *maturity principle* (Caspi et al., 2005; Roberts et al., 2006), levels of Honesty-Humility should increase with increasing age. The inclusion of this trait in the model informs the understanding of personality development and the dynamics of patterns and principles that are well established in the literature.

An interesting finding across most of our models, with the exception of Neuroticism, Openness to Experience, and Honesty-Humility is the general lack of cohort effects. However, some differences in the patterns of normative change were observed across the different periods of the life span. For example, in relation to Extraversion, Agreeableness, Conscientiousness, and even Openness to Experience, our findings highlight the importance of early adulthood in personality development (Soto, *in press*; Soto et al., 2011). In this way, unlike some previous research (e.g., Lucas & Donnellan, 2011), we *did* find evidence that may be tentatively interpreted as suggesting a critical period for personality development. This critical period would seem to occur in a window around about age 30. This is important in light of the suggestion that most of personality development occurs before the age of 30 (McCrae & Costa, 1997). We are not suggesting this as definitive proof that the arguments forwarding intrinsic maturation as processes driving personality development are correct. Indeed, another important aspect of our models is that all of the observed patterns of normative change, with the possible exception of Agreeableness, indicate a developmental process that continues throughout adulthood. For example, the cubic pattern of change in Extraversion identifies the importance of both early adulthood and old age in personality

development. Our findings thus also provide strong support for the propositions of the Neo-Socioanalytic perspective (Roberts *et al*., 2006). How can this be reconciled? In our view, the current findings highlight the importance of an integrated approach to understanding personality development across the life span—an approach that incorporates both consistency and systematic change in personality traits.

Limitations and Future Research

The present investigation has a number of strengths. The longitudinal nature of our sample and heterogeneity in terms of participants' age allowed for analyses of change across the adult life span (see also Specht et al., 2011). Furthermore, the size and the focus on providing a representative sample of adult New Zealanders facilitates investigations of normative changes, conceptualised as patterns of change in personality that generalize broadly. Moreover, our use of Cohort-Sequential Latent Growth Models (Preacher, Wichman, MacCallum, & Briggs, 2008; Prinzie & Onghena, 2005) provides a valuable way of estimating the developmental patterns of mean-level change across the adult life span. Indeed, this strategy has allowed us to detect novel findings regarding predicted patterns of change, and inform the debate about the nature of personality development.

With these strengths in mind, there are also a number of limitations inherent in the present investigation. While the modelling approach allows us to investigate both change over time and change due to cohort differences, the issue of formally disentangling cohort effects remains. Our estimation strategy provides a possible avenue that may help to quantify the extent to which cohort effects contaminate estimates of mean-level change; however, our approach provides only an approximation.[23] Furthermore, the models employed here are

[23] We hope that future researchers will help to address this limitation and aid in developing a formal equation or proof allowing the estimation of the proportion of variance in the change trajectories that is due to cohort differences versus developmental change or ageing. Intuitively, it would seem to us that this might be an equation that would take the general form of the sum of the ratios of the predicted point for the growth

necessarily simplistic in some regards, notably because they estimate latent change trajectories across cohorts using means of the personality scales at each time point. That is, rather than estimating latent factors for each personality trait at each time point, we used the means of the respective scales in our estimations. This may present issues in terms of measurement error. We opted for this approach due to the issue of statistical power. Finally, it is important to note the necessary limitation in the demographic scope of the present investigation. That is, our analyses estimated sample level change trajectories in the Big Six personality traits across the adult life span, and did not investigate more fine-grained individual differences in the developmental trajectories, such as gender differences. Investigations of such possible demographic differences in the life span trajectories of trait development are an important goal for future research in this field.

Another limitation that needs to be mentioned, and one that we could go as far as to characterize as the bane of longitudinal studies, is that of systematic attrition. Previous research with the NZAVS (Satherley et al., 2015) indicates that Conscientiousness and Honesty-Humility were associated with sample retention. That is, those people who are higher on Conscientiousness and those higher on Honesty-Humility are more likely to remain in the panel study over time. This presents a concern about the representativeness of effects observed for these two traits to the broader population. Such systematic attrition is an inevitable component of representative panel designs and a problem which must be acknowledged through all longitudinal personality research.

Finally, as mentioned in previous sections, it is important to emphasize that we measured Big-Six personality traits using short-form marker items. The use of longer and more comprehensive measures was simply not possible due to space constraints in the

---

parameter at age *x* divided by the predicted point at that same age based solely on the cross-sectional cohort data, for all values of age. However, it is beyond the scope of the current paper to formally propose and evaluate a general equation of this nature.

questionnaire, funding constraints, and the sheer number of diverse constructs (and researchers) competing for space in the NZAVS each year. We used the Mini-IPIP6 (Sibley et al., 2011). This is a short-form measure extending the original Mini-IPIP (Donnellan et al., 2006) to include marker items for the sixth dimension of Honesty-Humility. Firstly, there is the obvious problem associated with the use of short-form scales in terms of the potentially increased Type I and Type II errors (Credé et al., 2012; Kruyen, Emons, & Sijtsma, 2013). Although this caveat should be kept in mind, we also emphasize that both the Mini-IPIP and the Mini-IPIP6 have been extensively validated (Donnellan et al., 2006; Milojev & Sibley, 2014; Milojev et al, 2013; Sibley et al, 2011; Sibley, 2012; Sibley & Pirie, 2013). As evidenced below, however, the use of short-form scales carries with it inevitable costs, and as large scale longitudinal research in the field of personality and beyond gains traction, these costs will and do face the entire field.

A more focused caveat that should be kept in mind relates to the breadth of the constructs covered. For example, as assessed by the Mini-IPIP6, Extraversion is more closely related to the sociability aspect than the social dominance or assertiveness facet of the trait. Given the suggestions of differential developmental patterns of mean-level change in social vitality which tends to decrease across the life span, and social dominance which tends to increase (Roberts et al., 2006), it may be argued that our findings only tell half of the story of the development of Extraversion. Further, our measure of Conscientiousness largely measures orderliness rather than industriousness or productiveness. This distinction may help to explain why we failed to detect the expected decrease in Conscientiousness in old age (Lucas & Donnellan, 2011; Marsh, et al., 2012; Terraciano et al., 2005). Indeed, it may be the case that differential developmental patterns may be expected for these two aspects of Conscientiousness—with orderliness increasing in adulthood, and productiveness decreasing in old age (as posited by the *la dolce vita* effect; Marsh, et al., 2012). Similarly, our measure

of Agreeableness largely assesses empathic concern, rather than respectfulness or politeness, while our measure of Honesty-Humility largely represents humility or low self-entitlement. The present findings thus suggest that humility consistently increases across adulthood. Empathic concern, on the other hand, seems to show a decrease in younger cohorts (between the ages of 19 and 30).

The present study has limitations with respect to the representation of personality traits that are inherent in the use of short form scales. While these trade-offs are often necessary in large-scale longitudinal research, future research should seek to address the possibility of different developmental patterns associated with different facets of personality traits—and can hopefully build on the priors provided by our models and findings. Indeed, previous research has already identified distinct developmental patterns for different personality trait facets. This highlights the importance of facet-level research more generally (Soto et al., 2011; Soto & John, 2012). Keeping this point in mind, our findings nevertheless contribute to the research corpus and help to improve our estimation of the development of personality as a field.

Concluding Comments

To conclude, we presented a series of Cohort-Sequential Latent Growth Models assessing developmental patterns of normative change in Big-Six personality traits. Our analyses provide novel evidence contributing to the growing literature mapping the developmental patterns of stability and change in personality across adulthood. Our findings directly test competing predictions derived from the Neo-Socioanalytic theory (Roberts et al., 2006), and the intrinsic maturation perspective of the Five Factor theory (Costa & McCrae, 2006). We find support for key aspects of both theories; some aspects of personality change are most pronounced among people in young adulthood and during a window around the age

of 30; other aspects of personality show continual development and change across the adult life span.

To sum up, Extraversion decreased as people aged, with the most pronounced declines occurring in young adulthood and then again in old age. Agreeableness, as indexed using our measure relating to empathy, decreased in young adulthood. Conscientiousness increased among young adults then levelled off and remained fairly consistent for the rest of the adult life span. Neuroticism decreased as people aged. Openness to Experience also decreased in a relatively linear fashion as people aged. Honesty-Humility showed a pronounced and consistent increase across the adult life span. This consistent increase in Honesty-Humility is likely to be consistent with predictions regarding Agreeableness. Our findings highlight the importance of considering recently developed personality models beyond the Big Five and also in both younger and older adulthood. Our findings also emphasize the need for an integrated perspective to explain personality development across the adult life span—one incorporating genetic influences, early development, and transactional processes. We hope that our analyses will encourage future research assessing personality development across longer time-frames, and also further research refining the use of Cohort-Sequential Latent Growth Modelling for developmental research more generally.

Authors' Note. This research was supported by a Templeton World Charity Foundation Grant (ID: 0077). This research forms part of the Petar Milojev's PhD thesis, supervised by Chris Sibley. Mplus syntax for the models reported here are posted on the NZAVS website. Data are also available upon request from the corresponding author, or any member of the NZAVS advisory board. www.psych.auckland.ac.nz/uoa/NZAVS

Chapter VI

General Discussions

Over the years, the central question in developmentally oriented personality research has changed from “whether” to “how” personality changes across the adult life span (Roberts et al., 2006). This research interest in the developmental patterns of change reflects the competing views on personality traits and trait development – the trait perspectives emphasising consistency (e.g., the Five Factor theory; McCrae & Costa, 1999; McCrae et al., 1999, 2000), and the transactional life span perspectives emphasising consistency but also meaningful change across the life span (e.g., Neo-Socioanalytic perspective; Roberts & Caspi, 2003; Roberts et al., 2008). Thorough investigations of the patterns of stability and change in personality traits, however, are methodologically demanding, requiring large longitudinal samples with adequate heterogeneity, and application of appropriate modelling strategies (Specht et al., 2011). Largely due to such demands, but also driven by the dominant position emphasising trait consistency (McCrae et al., 1999, 2000, Roberts et al., 2006) such investigations have been, in relative terms, lacking. Recent years, however, have seen a re-uptake of interest and methodological rigour in longitudinal investigations of change and changeability in personality traits (Lucas & Donnellan, 2011; Specht et al., 2011; Wortman et al., 2012) contributing important information to the debate about personality development.

The present thesis sought to further contribute to this growing literature and aid in establishing a consensus on the normative patterns of personality development in adulthood. Through four systematic studies, this thesis investigated the stability (Milojev et al., 2013; Milojev & Sibley, 2014) and change (Milojev, Osborne, & Sibley, 2014; Milojev & Sibley, 2016) in personality traits across the adult life span. These investigations were based on the longitudinal nationally representative sample of adult New Zealanders, from the New

Zealand Attitudes and Values Study (NZAVS). The four studies thus comprise a thorough longitudinal investigation of change and stability in a large, heterogeneous, longitudinal sample, adding potentially generalizable findings to the relatively few extant studies employing samples with similar characteristics (Lucas & Donnellan, 2011; Specht et al., 2011; Wortman et al., 2012). In addition, the studies within this thesis are consistent in terms of the model of personality investigated and the method of assessment thereof. Specifically, the present thesis investigated patterns of stability and change in the Big Six personality traits – Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility. By keeping abreast with the recent developments in the literature concerned with personality structure – namely the six factor HEXACO model (Ashton & Lee, 2007) – the work here presented provides further contribution to and extension of extant literature on personality development. Indeed, as seen in previous chapters, important insights may be gained by considering the sixth factor of Honesty-Humility.

Major Findings

The studies within this thesis investigated both the rank-order stability and the mean-level change in personality traits. As discussed previously, rank-order stability, or differential stability, refers to the longitudinal stability in individual differences in personality traits, or the rank-ordering of a given group of people. Mean-level change, on the other hand, refers to changes in the absolute levels of a given personality trait over time. At sample-level mean-level change is assumed to represent patterns of change that generalise to most people most of the time. This aspect of personality stability is often referred to as normative change (Caspi et al., 2005; Roberts et al., 2008). Within this thesis, study 1 (Milojev et al., 2013) and study 2 (Milojev & Sibley, 2014) were concerned with rank-order stability of the Big Six, while study 3 (Milojev et al, 2014) and study 4 (Milojev & Sibley, 2016) were concerned with mean-level change.

Study 1 was the foundational investigation, allowing for the three studies to follow. Specifically, a Bayesian Structural Equation Modelling approach was employed investigating the rank-order stability of the Big Six personality traits as assessed by the Mini-IPIP6 (Sibley et al., 2011). The importance of this study is twofold. Firstly, given the conceptualisation of personality as "*relatively enduring styles of thinking, feeling, and acting"* (McCrae & Costa, 1997, p.509) a high level of rank-order stability must be expected over moderate time periods – such as a 1-year test-retest period used in study 1. Indeed, all of the proposed working definitions of personality, including that proposed by Mischel (1968), centre on a very high level of stability. As such, a test of rank-order stability of the Big Six over a one year test-re-test period in a nationally-representative sample of adults provides and important test of the construct. Secondly, by testing the rank-order stability of the Big Six personality traits as assessed by the Mini-IPIP6, the study provided a further test of the psychometric properties of the scale. Establishing the validity of the construct and the test-retest reliability of the scales was crucial to the subsequent longitudinal investigations. The finding of very high stability of the Big Six (all stability estimates were $\beta > .90$) thus provides support for the use of the Mini-IPIP6 in longitudinal investigations and supports the necessary expectation of high temporal consistency, or rank-order stability, of personality traits.

While the message of study 1 is that personality is highly stable over a 1-year period, the message of study 2 was that the stability of personality traits varies across the adult life span. Study 2 employed an SEM approach estimating latent interactions to investigate the patterns of differential stability in the Big Six personality traits across the cohorts of the adult life span from 20 to 80 years old. This investigation is particularly important in light of the recent evidence of systematic variation in personality stability across the life span (Lucas & Donnellan, 2011; Specht et al., 2011; Wortman et al., 2012) and is pertinent to the discussion regarding the causes of developmental change in personality traits. Specifically, while the

arguments from the Five Factor Theory suggest that personality should stabilise through younger adulthood, the rank-order stability should remain largely unchanged over the rest of the life span (McCrae & Costa, 1999). On the other hand, the Neo Socioanalytic perspective (Roberts et al., 2008) would argue that rank-order stability should vary across the life span in a way that reflects relative stability or change in people's lives in terms of social roles and institutions engaged in. Supporting this perspective, recent longitudinal investigations have shown that rank-order stability varies across the life span in a specific quadratic – inverted U – pattern (Lucas & Donnellan, 2011; Specht et al., 2011; Wortman et al., 2012). Study 2 thus provides a much needed replication of these recent findings, in a nationally representative sample of New Zealanders. The findings indicate that 2-year rank-order stability of the Big Six personality traits, with the exception of Agreeableness, follows a quadratic distribution across the cohorts of the adult life span. That is, personality tends to stabilise through younger adulthood reaching peak stability in middle age, around the age of 50. Beyond that, however, the rank-order stability systematically decreases towards older age. These patterns thus seem to represent a 'typical' pattern of stability of personality traits across the adult life span that is consistent with the Neo-Socioanalytic life span perspective on personality development.

Therefore, studies 1 and 2 systematically investigated differential stability of the Big Six personality traits. In general, as expected, personality is highly stable across a one and two-year test-re-test period. However, rank-order stability of personality is not uniform across the life span. In fact, personality tends to be less stable in younger adulthood, increasing in stability towards middle age, and destabilising thereafter towards older age. With this potential for changeability in mind, studies 3 and 4 focused on investigating the mean-level change in the personality traits.

The overwhelming message of study 3 is the consistency of personality in the face of a large-scale natural disaster – the 2010/2011 earthquakes in Canterbury, New Zealand. Study 3 investigated mean-level changes in the Big Six personality traits before and after the Christchurch earthquakes, among those who were affected and those who were not affected by the devastating event. Systematic change in personality in relation to various life-events is crucial for the arguments of the Neo-Socioanalytic perspective (Roberts et al., 2008). The transactional processes proposed would predict that personality changes in meaningful ways in the face of life-events that signify a change in social roles pertinent to identity formation. Indeed, recent work by Specht and colleagues (2011) has shown some evidence for such changes. However, changes in personality associated with major, potentially traumatic non-normative life-events have been scarcely studied for obvious methodological reasons. Given the magnitude of such events a considerable amount of change in the lives of the people affected can be safely assumed. Having said that, the Neo-Socioanalytic perspective suggests that a natural disaster would result in substantive changes in personality, only if such an event indicated substantive changes in the social roles people identify with. In the normative population, however, this may not be the case. Indeed the findings of study 3 indicate a remarkable level of consistency in personality among those affected by the Christchurch earthquakes. That is, no significant changes in personality were observed pre and post the earthquakes, with the exception of Emotional Stability (the reverse of Neuroticism). Specifically, those affected by the earthquakes showed a small decrease in Emotional Stability pre and post the event, relative to those not affected by the earthquakes. This finding is consistent with similar domain specific changes seen in extant literature (e.g., Löckenhoff et al., 2008; Specht et al., 2011). Thus, study 3 indicates that personality is highly consistent in the context of a large-scale natural disaster, with only small domain-specific changes in Emotional Stability. The findings of the study serve a further purpose of informing the

potential changes in wellbeing of Cantabrians affected by the earthquake, and indeed people affected by other similar events. Specifically, the decrease in Emotional Stability may indicate an increased risk of personality-based vulnerability to depression (Zuroff, 1994). Clearly, future follow-ups on these changes will be of great importance.

The final study of the thesis, study 4, investigated the patterns of normative change in the Big Six personality traits across the adult life span. This study employed a 6-year Cohort-Sequential Latent Growth Model as a framework for estimating developmental patterns of change across the adult life span. Investigating such patterns is of particular interest in developmental literature (Roberts et al., 2006). However, comprehensive longitudinal investigations covering a broad range of the adult life span in samples representative of the general population are scarce (Lucas & Donnellan, 2011; Specht et al., 2011; Wortman et al., 2012). The major problem with extant longitudinal investigations of change is the entanglement of the effects of ageing – or the effects of time – and cohort differences. Study 4 sought to provide a methodological approach that may aid in the efforts of addressing this issue. This study thus produced an interesting and novel pattern of findings.

Consistent with extant literature, Extraversion decreased with increasing age. Specifically, Extraversion deceased in younger adulthood, stabilised through middle age, and decreased again in older adulthood. Agreeableness showed an interesting pattern, decreasing in younger adulthood, up to about the age of 30, followed by relative stability across the rest of the life span. Conscientiousness showed the expected increase in younger adulthood through to middle age, remaining stable thereafter. Neuroticism and Openness to Experience decreased with increasing age. However, the models suggested that this effect may be partially due to cohort differences rather than the effects of ageing. Finally, Honesty-Humility showed a consistent linear increase across the adult life span. Indeed, these findings provide evidence of both consistency and systematic change in personality traits. In addition, these

findings highlight the importance of early adulthood as a period of development, lending some support to the suggestions that most, if not all developmental changes occur prior to around the age of 30 (Costa & McCrae, 1997). However, study 4 also indicates that personality change is an ongoing process throughout adulthood. Indeed, rather than providing concrete evidence for either the trait perspective of the Five Factor theory (McCrae & Costa, 1999) or the Neo-Socioanalytic perspective (Roberts et al., 2008), study 4 suggests that an integrated perspective is necessary to account for the developmental patterns of personality change in adulthood.

Normative personality development in adulthood

So, are we any closer to reaching a general consensus on personality development in adulthood? The simple answer is "Yes". However, the current state of evidence on patterns of normative development of personality is clearly more nuanced and in places inconsistent.

In terms of developmental patterns of differential stability across the adult life span, the finding of the quadratic - inverted U - pattern is largely consistent across the recent studies - including study 2 of this thesis (Lucas & Donnellan, 2011; Specht et al., 2011; Wortman et al., 2012). Importantly, unlike most of the earlier literature concerned with patterns of stability across the life span these studies were based on large longitudinal samples with considerable heterogeneity in terms of the range of the adult life span represented. Thus, a relative consensus can be assumed in terms of the 'typical' developmental patterns of differential stability across the adult life span. That is, personality traits tend to be relatively unstable, or changeable, in younger adulthood, systematically stabilising towards middle age. In general, personality traits are at their most stable in middle age, particularly when people are in their late 40s and early 50s. Following this period of stabilisation, the differential stability of personality traits tends to decrease towards older age.

These patterns are broadly consistent with the Neo-Socioanalytic perspective on personality development (Roberts et al., 2008). That is, personality tends to be more stable in periods of the life-span where people's lives are more stable in terms of the consistency of social roles invested in – such as family and professional engagements. In a normative population, middle age is the period of life that is associated with most stability. Conversely, periods of life that are associated with greater normative changes are those periods where people's personality is more unstable. Younger adulthood is a period of considerable change in terms of identity formation through the social roles and investments - such as finishing school and formal education and entering the professional environment, or starting one's own family and having children. Older age is also a period of normative changes in terms of changing social institutions and roles invested in – people's children may leave home to start families of their own; older age is normatively associated with retirement, often bereavement, and changing social networks, as well as biological and cognitive changes.

However, while there is general consensus across the recent studies on rank-order stability, there are still some inconsistencies. For example, study 2 of the present thesis indicates that the stability of Agreeableness is relatively uniform across the adult life span, showing a slight linear decrease rather than a curvilinear pattern observed in extant literature (Lucas & Donnellan, 2011; Specht et al., 2011; Wortman et al., 2012). These inconsistencies highlight the importance of future follow-ups as well as the importance of cross-study consistency in the way personality is assessed.

The state of evidence regarding mean-level developmental changes in personality traits in adulthood is more complex. One seemingly robust finding that found some support in the present research is that Extraversion tends to decrease with age (Lucas & Donnellan, 2011). However, literature suggests that the developmental pattern of change in Extraversion may be more nuanced in terms of the different facets of the trait (Roberts et al., 2008). This is

corroborated by the findings of study 4. That is, mean-levels of the sociability facet of Extraversion seem to decrease across the adulthood following the aforementioned cubic trajectory. On the other hand, as suggested by the findings of Roberts and colleagues (2006) the social dominance and assertiveness facets of the trait may increase across the adult life span. Similarly, extant literature (Lucas & Donnellan, 2009; Roberts et al., 2006, Specht et al., 2011), supported by our findings, suggest that mean-levels of Neuroticism decrease with increasing age. However, the findings of study 4 indicate that this pattern may, at least in part, be due to cohort differences – with older cohorts reporting lower levels of the trait – rather than a developmental effect of ageing.

The findings in this thesis –specifically the findings of study 4 – contribute to the literature suggesting that Conscientiousness tends to increase with age (Lucas & Donnellan, 2009; Lucas & Donnellan, 2011; Roberts et al., 2008). In particular, the increase in Conscientiousness tends to be most pronounced in young adulthood, gradually decelerating towards middle age, and levelling off at about the age of 45. The increase in levels of Conscientiousness with increasing age supports the suggestions that those traits that facilitate performance of social roles generally associated with maturation should increase with increasing age. The findings here reported, tentatively suggesting that most age-related change in Conscientiousness occurs in younger adulthood, support such processes.

While older age tends to be associated with lower levels of Openness to Experience (Lucas & Donnellan, 2011; Roberts et al., 2006) the findings of reported in study 4 suggest that this decrease may be due to cohort differences. Indeed, while some evidence was found for a systematic decrease in Openness to Experience in younger adulthood, the models suggest that the decrease in this trait across the rest of adulthood is at least in part due to older cohorts reporting lower levels Openness, rather than a developmental effect of ageing.

Given the findings presented in this thesis, further inconsistencies are evident in relation to normative changes in Agreeableness. While extant literature suggests that levels of Agreeableness should increase with increasing age (Lucas & Donnellan, 2009; Lucas & Donnellan, 2011; Terraciano et al., 2005) the present findings indicate that Agreeableness tends to decrease through younger adulthood, remaining relatively stable thereafter. However, as discussed above, the specific interpretation of this finding suggests that it is the facet of Agreeableness relating to empathic concern that decreases in younger adulthood. In light of this evidence, the *maturity principle* (Caspi et al., 2005) may suggest that the other facets of the Agreeableness trait – those related to politeness and cooperation – increase with maturation. Taken together with the findings that the sociability facet of Extraversion also tends to decrease in early adulthood, these findings seem to indicate a substantive effect whereby people tend to become more self-focused and less concerned with broad social networks during this period. Interestingly, it is Honesty-Humility that shows a robust developmental pattern whereby increasing age is systematically associated with increasing levels of the trait. The two facets of Honesty-Humility - fairness (honesty) and a lack of self-entitlement (humility) – arguably represent those traits that would facilitate success in maturity-related social roles. Thus, while the empathic concern facet of agreeableness decreases across younger adulthood – a period of normative pressure in terms of starting careers or families – Honesty-Humility shows a reliable increase that would be predicted by the *maturity principle*. These novel findings highlight the importance of developments in personality structure – such as the HEAXCO model (Ashton & Lee, 2007, 2009) – as well as the need for future facet level research.

The inconsistencies between the work in this thesis and extant literature are evidence that a consensus on normative patterns of mean-level change in personality traits is still far from achieved. Indeed, the findings of studies 3 and 4 suggest a complex picture of both

consistency and change, and highlight the need for an integrated perspective on developmental change in personality in adulthood. This highlights the importance of continuing research on personality change and development, in adequate samples and over longer time-frames, as well as a consideration of the specific facets of personality traits.

Methodological notes

As discussed in previous chapters, developmentally oriented research in personality is methodologically challenging, both in principle and in practice. Most notably, investigating the patterns of rank-order stability and mean-level change across the life span necessitates longitudinal research (e.g., Specht et al., 2011). While cross-sectional studies can be informative in inferring mean-level differences in personality across the life span (e.g., Anusic, Lucas, & Donnellan, 2012; Donnellan & Lucas, 2008; Lucas & Donnellan, 2009; McCrae et al., 2000; Sristave et al., 2003; Terraciano et al., 2005), longitudinal approaches are necessary for direct investigations of stability and change. To that effect, this thesis presents a series of four longitudinal studies investigating stability and change in personality over a variety of time-frames. Longitudinal studies in research on personality development by no means represent a new direction (see e.g., Roberts et al., 2006). However, given the necessity of longitudinal approaches in the study of stability and change, they have been relatively underdeveloped until recently (e.g., Lucas & Donnellan, 2011; Specht et al., 2011). Using the NZAVS panel sample, the work comprising the present thesis represents an important contribution to the longitudinal work on the developmental patterns of change and changeability in personality.

As discussed previously, the NZAVS sample is one of the few nationally representative longitudinal panel samples used in personality research. This sample presents heterogeneity of participant characteristics that allows for investigations of stability and

change across the broad range of the adult life span. With notable exceptions in recent research (Lucas & Donnellan, 2011; Specht et al., 2011; Wortman et al., 2012), most extant longitudinal research tends to be limited in the range and periods of life span represented (see Roberts et al, 2008). With the focus on life span development of personality, this thesis, and in particular studies 2 and 4, provide a comprehensive investigation of stability and change across a broad range of the adult life span. The importance of this is most notably highlighted in terms of the developmental patterns of rank-order stability, allowing for the crucial replication of the recently emerging developmental trends – the inverted U distribution (Lucas & Donnellan, 2011; Specht et al., 2011; Wortman et al., 2012). Furthermore, as the NZAVS sample is a nationally representative sample of adult New Zealanders, this provides a further advantage to the present work. Specifically, as discussed in previous chapters, the focus of the present work is on identifying normative patterns of adult personality development. That is, patterns of stability and change across the life span that can be generalised, or that apply to most people in a normative population. Given this interest in identifying generalizable patterns of development, the use of a nationally representative sample - such as the NZAVS sample - is of central importance.

Another methodological note regarding sample characteristics concerns sample size. Specifically, a large sample size is necessary to allow sufficient statistical power to investigate subtle effects in terms of change in personality. While the NZAVS panel sample is not as large as some recently reported (e.g., n = 20000 in Specht et al., 2011), it is indeed a large sample affording considerable power in longitudinal analyses as discussed in the earlier chapters. Moreover, relating specifically to study 4, yearly assessments in the NZAVS allow the identification of and the power for sophisticated growth models. On a related note, comprehensive investigations of stability and change in personality require sophisticated modelling approaches. For instance, as previously discussed, estimates of personality stability

tend to be notably attenuated due to measurement error (Ferguson, 2010). Thus, studies 1 and 2 employed SEM frameworks in investigating rank-order stability of the Big Six personality traits. Furthermore, study 2 employed an SEM framework estimating latent interactions to assess the variability of the stability estimates across age. Indeed, the work in this thesis has strived throughout to employ and in places develop the most appropriate and sophisticated methods that allow the most accurate investigation and estimation of the effects of interest. While simplicity of analysis was sometimes the optimal approach, such as the repeated-measures ANOVA in study 3, study 4 put forward a 6-year Cohort-Sequential Latent Growth Modelling framework (Preacher, Wichman, MacCallum, & Briggs, 2008; Prinzie & Onghena, 2005) as an analytic method for investigating patterns of mean-level change across the life span. These methodological characteristics contribute to the utility of this thesis, and it is also a hope that these approaches will prove useful in encouraging future developments in developmentally oriented longitudinal research.

Finally, the method of personality assessment throughout the studies in this thesis is worth mentioning. Through all four studies the Big Six personality traits were assessed using the Mini-IPIP6 (Sibley et al., 2011). As discussed in detail in previous chapters, the Mini-IPIP6 is a short-form measure based on the original Mini-IPIP assessing the Big Five (Donnellan et al., 2006) and including additional items assessing the sixth factor of Honesty-Humility (Ashton & Lee, 2009). The Mini-IPIP6 has been extensively validated (Sibley, 2012; Sibley et al., 2011; Sibley & Pirie, 2013), including the establishment of high test-retest reliability in study 1. While the breadth of representation of the traits in this measure is by necessity limited, as discussed in detail in previous chapters, two points about the Mini-IPIP6 are particularly notable within the present discussion. Firstly, it is a short-form scale, assessing the Big Six personality traits by 4 items each. Secondly, it is based on the International Personality Item Pool (Goldberg, 1999) and as such the scales are in the public

domain and accessible to a broad body of researchers and interested parties. It is indeed worthy of note that the extensive work on personality development in this thesis was entirely based on publically available short form measures of personality. This note will hopefully highlight the accessibility of such research and encourage future endeavours.

A Note on Power

The specific analytic strategies employed in the four studies comprising this thesis were presented in detail in the corresponding chapters (chapters 2 through 5). This section presents the post-hoc power analyses for studies 1 to 3 in one place. Note that the power analyses for growth models in study 4 are presented in Appendix A, within the published supplementary materials. Monte Carlo simulations were used to estimate the power as pertaining to the models in study 1 and 2. Post Hoc power analysis for study 3 was only conducted for the specific effect associated with change in Emotional Stability. The same procedure as used in study 4 was used in the Monte Carlo simulations presented here. Specifically, the adapted procedure outlined by Muthén and Muthén (2002) and used by Sibley & Milojev (2014) was used to estimate obtained power to detect the focal effect given the known sample size.

Study 1 used a SEM framework to estimate stability of the six personality variables. The parameter of interest in this case is the stability estimate – that is, the regression coefficient whereby the latent personality trait at Time 1 predicts the latent personality trait at Time 2. Monte Carlo simulations were thus conducted to estimate power to detect the regression coefficient between two latent variables. Specifically, a data structure was simulated with two latent variables, estimated by four items each. The factor loadings were constrained to equality at .80 and the variances of the two factors were constrained to 1.0. The residual variances of the manifest items were correlated across the measurement points.

These correlations were constrained to equality at .10. The focal parameter was systematically varied between .05 and .95. The number of observations was fixed to the actual sample size in the analysis – n = 4422. The simulations were run with 10000 replications. The results of these simulations are presented in Figure 6.1. As can be seen in the figure, given the sample size, the power to detect even very small magnitudes of the focal coefficient quickly reaches 1.0 (or 100%).

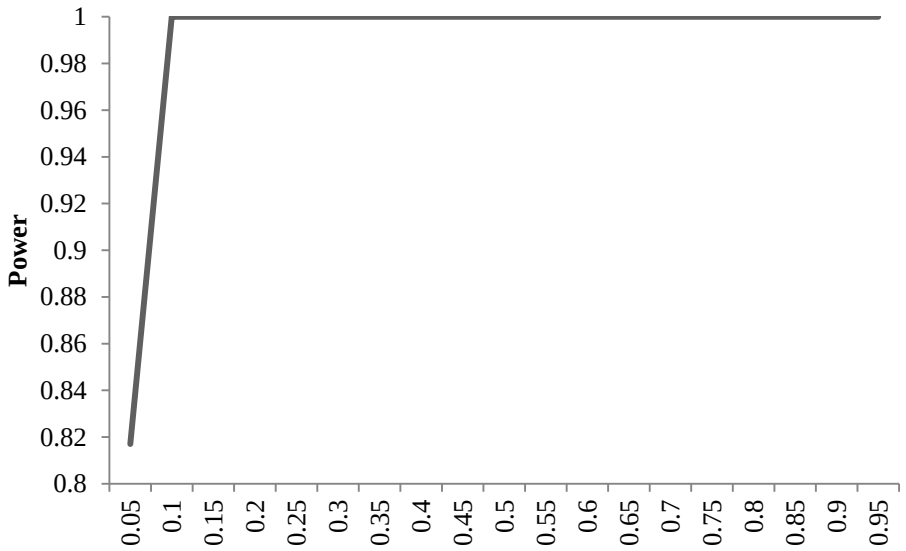


Figure 6.1. Power function for the focal parameter in study 1 – the regression coefficient between two latent factors estimated by four manifest items each (i.e. personality trait at time 1 and at time 2).

Study 2 employed a SEM framework estimating latent personality variables at Time 1 and at Time 3 (2011). A moderated regression structure was specified whereby the latent personality variable at Time 3 was predicted by the corresponding latent personality trait at Time 1, age, the quadratic effect of age, and gender. More importantly, to test whether the stability estimate varies across age – that is, if age moderates the effect of Time 1 latent personality trait on its corresponding factor at Time 3 – latent interactions were estimated between personality at Time 1 and gender, age, and age squared. The coefficient of interest in relation to the post-hoc power estimates in this case is the effect of the latent interaction with age. Monte Carlo simulations were conducted to estimate the power to detect a latent

interaction with effects of varying magnitudes, given the actual sample size from study 2. A data structure was simulated with two latent variables estimated by four manifest items each. The factor loadings were all constrained to .80 and the variances of the two factors were fixed at 1.0. The residual variances of the manifest items were correlated across the measurement points. These correlations were constrained to equality at .10. The parameter representing the stability estimate was fixed to .70 based on the estimates observed in study 2, indicating high stability. A continuous covariate was also estimated with a mean of 0 and a variance of 1.0. The effect of the covariate on the outcome factor was fixed to a relatively small value of .05 as suggested by the models in study 2. The latent interaction between the predictor factor and the continuous covariate was simulated. The effect of the interaction was the focal parameter in these simulations and was varied from .001 to .050. These small values were decided on based on the observed parameter estimates in study 2. The number of observations was fixed to the sample size used in study 2 – n = 3902. Due to the computational demands the simulations were run with 1000 replications. The results of these simulations are presented in Figure 6.2. The power for the latent interaction shows a cubic function. Most importantly, the latent interactions as here specified seem to be particularly difficult to detect. The identification of these types of effects is evidently demanding; however, the models in study 2 detected significant effects to show predicted patterns.

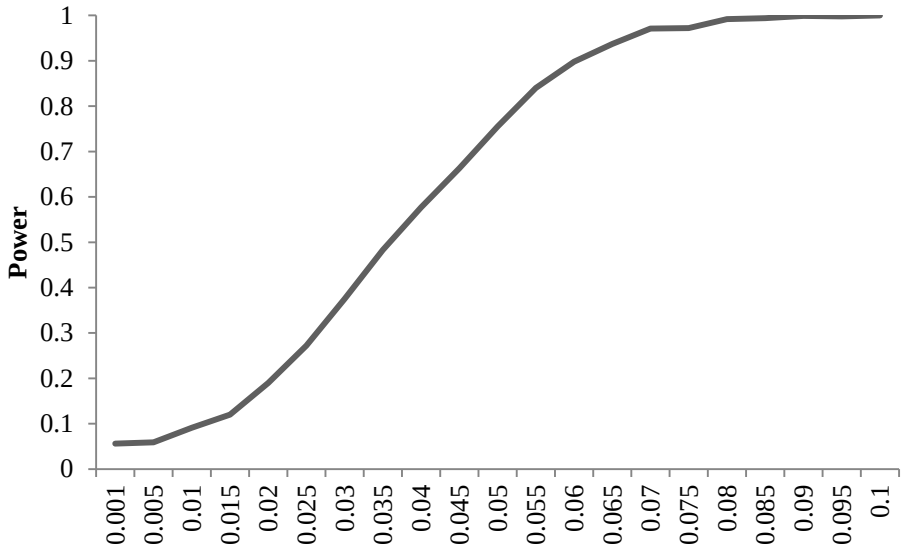


Figure 6.2. Power function for the focal parameter in study 1 – the latent interaction between the latent personality trait and continuous covariate in a model predicting a latent personality trait at a later point in time.

To assess personality changes associated with the Christchurch earthquakes, study 3 involved multiple mixed-models repeated-measures ANCOVA's. The only significant effect of interest was the interaction between time – before to after the earthquakes – and whether or not one was affected by the earthquake – using both the subjective self-report measure and the objective measure based on place of residence. Post-hoc power analyses were conducted for this specific effect of interest. The power was estimated for the effect based on the subjective assessment of being affected by the earthquake given the observed effect size (F (1, 3552) = 9.422, p = .002, η^2_p = .003), the known sample size (n = 3914) and the 9 covariates. The analysis indicated power at .89 (89%) indicating sufficient power to detect this interaction. Post-Hoc power analyses for the effect based on the objective assessment of being affected by the earthquakes (F(1, 3587) = 4.465, p = .035, η^2_p = .001) revealed power at .60 (60%), considerably lower than for the subjective option.

Future questions

The present thesis put forward a comprehensive investigation of personality development in adulthood. The body of work comprising this thesis suggests that personality has very high levels of temporal consistency over moderate time frames. However, this stability is not uniform across all periods of the adult life span. Specifically, personality traits tend to be less stable in younger adulthood, stabilising towards middle age, and systematically reducing in rank-order stability towards older age. Further, systematic mean-level changes in personality do occur in adulthood. Indeed, evidence suggests that both younger and older adulthood are important periods in personality development. In relation to life-experiences, the mean-level change in personality is likely to be domain-specific and is more likely to be related to normative life-events associated with maturation across the life span (Roberts et al., 2008) than non-normative events such as large scale natural disasters. The findings of the studies here presented provide considerable insights, informing the debate about the causes of developmental changes and stability in personality. That is, the findings largely support the arguments of the Neo-Socioanalytic life-span perspective (Roberts et al., 2008); however, the patterns of mean-level changes across the life-span also support for the Five Factor trait perspective on personality development (McCrae et al., 1999, 2000).

With this in mind, this thesis leaves unanswered questions that should inspire and be addressed in future research. Most obviously, there is the question of the length of the time periods over which personality change and stability are analysed. The longest time-frame in the present work was the assessment of change over 5 years – or 6 yearly assessments – in study 4. While 5 years is a notable amount of time that may see a number of normative developmental changes in people's lives, it may not be long enough for a stable construct such as personality traits to show notable change. As such, future research must look to build on extant longitudinal literature (e.g., Lucas & Donnellan, 2011; Specht et al., 2011,

Wortmatn et al., 2012) and the work with studies such as the NZAVS, to investigate personality change and stability over longer time-frames. Indeed, the present work highlights the utility and importance of large, representative panel studies in developmentally oriented personality literature, and beyond.

The debate regarding the perspectives on personality traits and personality development is still far from settled. Reaching a consensus on the developmental patterns of personality change and changeability across the life span is indeed crucial for theoretical and practical purposes. However, it is unclear how much further investigations of broad, normative developmental patterns can inform this debate. The answers are more likely to be found in more nuanced approaches. Investigations of the life-events and experiences that are associated with personality change, such as in study 3 and in the recent work of Specht and colleagues (2011) may present a fruitful avenue. One important development in such research will have to be the extension of longitudinal follow-ups in order to ascertain the true longitudinal effects of life-experiences. For example, a longitudinal follow-up of study 3 is anticipated in order to investigate whether the initial decrease in Emotional Stability observed among those affected by the Christchurch earthquakes indicates a longitudinal change, or if the levels may return to the pre-earthquake levels after a certain period. Furthermore, research efforts should be directed at identifying the life-experiences associated with domain-specific personality change. Presumably these life-experiences will be those that signify notable changes in social roles and social institutions invested in (Roberts et al., 2008). However, it may be the case that some normative events may be more strongly related to one personality trait than others. For example, Specht and colleagues (2011) observed that commencing a professional career had a domain-specific effect in increasing the levels of Conscientiousness. This is indeed an important avenue of future research.

An interesting caveat, not discussed in previous chapters, relates to the limitations inherent in the use of self-report measures – such as the Mini-IPIP6 – in most of personality development research, including the work presented here. In the context of large-scale longitudinal research the use of such scales is indeed necessary. However, several issues pertinent to the conceptualisation of personality and personality change remain. Indeed social desirability biases (Paulhus, 1991) or methodological issues such as interpretations and understanding of instructions and protocols may compromise the validity of such personality scales (Piedmont, McCrae, Riemann, & Anglietner, 2000).

Beyond this, however, is the question of interpreting change on self-report personality scales as change in dispositional personality traits. It may conceivably be the case that the change in personality traits scores observed is a function of the participants' belief of how they have changed in light of life events and changes in social roles that have occurred (Robins et al. 2001). For instance, a young adult entering the job market with a new job and starting a career may, at the time of response, evaluate themselves as more conscientious – that is more task oriented, more ordered. However, this reported change in these self-reports may be likened to changes in *characteristic adaptations* – goals and strategies relevant to the individual's current context (DeYoung, 2015) – rather than changes in the underlying broad bandwidth personality trait. This problem of measurement, particularly pertinent to investigations of personality trait change and development, is widespread in personality psychology and studies relying on questionnaire based designs. Indeed, reliance on self-report questionnaire measures may be, to an extent compromising the validity of our measures and of our findings. However, one consistent feature observed in longitudinal studies employing self-report questionnaire based personality measures, including the work presented in this thesis, offers some reason for confidence. Through all of the four studies presented in the preceding chapters, while evidence of systematic and substantive change and

instability was observed across the adult life span, the most consistent finding across all six personality traits is the temporal consistency or rank-order stability across different periods of time. This finding provides some evidence that we are indeed measuring a stable construct that is a personality trait. Despite this, future research must invest effort into developing design frameworks that will allow replication of the findings here observed using personality trait measures that do not rely on self-report responses – such as peer ratings. Unfortunately such methods are hardly compatible with the need for large scale longitudinal panel samples such as that utilised here. Thus, the two approaches must complement one another.

While personality traits are the focus of the developmental research presented here, there are other factors of individual differences that warrant attention. Individual differences in terms of life goals (Roberts, O'Donnell, & Robins, 2004), and, more broadly, values (e.g., Schwartz, 1992) may show similar developmental patterns. Moreover, to the extent that values and goals represent more proximal characteristic adaptations (DeYoung, 2015) this line of research would open the possibility of addressing the extent to which these factors may drive personality change or vice versa. In general terms, the developmental research such as that presented here should look to extend its focus to a broader conceptualisation of the individual in the bid to allow comprehensive understanding of the "whole person" (McAdams & Pals, 2006, p. 204).

Closing Words

General consensus on the development of personality in adulthood remains elusive. The ongoing debate as to the causes and the predicted patterns of consistency and change in personality reflects the debate about the very nature of personality traits – are they a matter of temperament and genetics (McCrae & Costa, 1999) or are they a function of transactional processes between a person's dispositions and the their environment (Roberts et al., 2009)?

One question pertinent to this debate is "Why should we care?" Simply put, personality traits are important. Not only are personality traits predictive of important life outcomes such as marital satisfaction, occupational achievement, and even mortality (e.g., Caspi et al, 2005; Roberts et al., 2007, Ozer & Benet-Martinez, 2006), but they are also "the explicit focus of socialisation efforts" in terms of providing children with personality traits and structures (Roberts, 2009). Moreover, personality traits are arguably the central component of who people are and what they are like. As such, the patterns of consistency and change across adulthood relate directly to the question of if and how people change.

Perhaps the most salient message of this thesis is how far the personality research still has to go in terms of comprehensively understanding personality traits and their development. The work here presented goes a long way in both aiding the development of a consensus on patterns of personality development in adulthood, but also raising further questions in relation to proposed processes. On the one hand, there is now a reasonable level of consensus as to the consistency of personality and the systematic variation in the stability of personality traits across the adult life span. On the other hand, the normative patterns of mean-level change in personality traits remain a contested issue with inconsistencies across studies. Indeed, it is hard to say whether developmental patterns of change are more aligned with the trait-perspectives (McCrae & Costa, 1999) or the transactional life-span perspective on personality (Roberts et al., 2006). One thing is certain – personality is consistent, changeable, and changing. Furthermore, as a field, we have a lot of work to do in identifying the exact patterns and dynamics behind the patterns of change and changeability. It is hoped that the approaches employed and developed through this thesis will encourage further methodological refinement and facilitate further research on this important topic.

13 References

Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt, Rinehart and Winston.

Allport, G. W. (1955). *Becoming; basic considerations for a psychology of personality* (Vol. 20). Yale University Press.

Anusic, I., Lucas, R.E., & Donnellan, M.B. (2012). Cross-sectional age differences in personality: evidence from nationally representative samples from Switzerland and the United States. *Journal of research in Personality, 46*, 116-120.

Ardelt, M. (2000). Still stable after all these years? Personality stability theory revisited. *Social Psychology Quarterly, 63(4)*, 392-405.

Asendorpf, J.B. (1992). Beyond stability: Predicting inter-individual differences in intra-individual change. *European Journal of Personality, 6*, 103-117.

Ashton, M.C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review, 11*, 150-166.

Ashton, M.C., & Lee, K. (2009). The HEXACO-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment, 91*, 340-345.

Ashton, M. C., Lee, K., & Goldberg, L. R. (2004). A hierarchical analysis of 1,710 English personality-descriptive adjectives. *Journal of Personality and Social Psychology, 87*, 707-721.

Ashton, M. C., Lee, K., Perugini, M., Szarota, P., de Vries, R. E., Di Blas, L., et al. (2004). A six-factor structure of personality-descriptive adjectives: Solutions from

psycholexical studies in seven languages. *Journal of Personality and Social Psychology*, *86*, 356-366.

Asparouhov, T., & Muthén, B. (2010). *Bayesian analysis of latent variable models using Mplus.* Mplus working paper.

Baltes, P. B. (1997). On the incomplete architecture of human ontogeny. *American Psychologist, 52,* 366–380.

Baltes, P. B., Lindenberger, U., & Staudinger, U. M. (1998). Life-span theory in developmental psychology. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology* (5th ed., Vol. 1, pp. 1029–1143). New York: Wiley.

Bandura, A. (1999). Social cognitive theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 154–196). New York: Guilford Press.

Baumgarten, F. (1933). *Die Charaktereigenshaften. Beitrage zur Charakter- und Personlichkeitsforschung: Monogr. 1* . Bern: A. Francke

Beck, A.T. (1983). Cognitive therapy of depression: New perspectives. In P.J. Clayton, & J.E. Barrett (Eds.), *Treatment of depression: Old controversies and new approaches* (pp. 265-290). New York: Raven Press.

Blatt, S.J., & Zuroff, D.C. (1992). Interpersonal relatedness and self-definition: Two prototypes for depression. *Clinical Psychology Review, 12(5),* 527-562.

Borgatta, E. F. (1964). The structure of personality characteristics. *Behavioural Science, 12,* 8-17

Brand, C. R. (1984). Personality dimensions: an overview of modern trait psychology. In *Psychology Survey (*5), J. Nicholson, H. Beloff. Leicester (Eds.), UK: Br. Psychol. Soc.

Campbell, W.K., Bonacci, J.K., Shelton, J., Exline, J.J., & Bushman, B.J. (2004). Psychological entitlement: interpersonal consequences and validation of a self-report measure. *Journal of Personality Assessment, 83,* 29-45.

Cattell, R. B. (1947). Confirmation and clarification of primary personality factors. *Psychometrika, 12,* 197-220

Cattell, R. B. (1948). The primary personality factors in women compared with those in men. *Br. J. Psychol. 1,* 1 14-30

Cattell, R.B. (1957). *Personality and motivation structure and measurement.* New York: Harcourt, Brace and World.

Cattell, R. B., Eber, H. W., & Tatsuoka, M . M. (1970). *Handbook for the sixteen personality factor questionnaire.* Champaign, IL: Inst. Pers. Ability Testing

Caspi, A., & Roberts, B.W. (1999). Personality development across the life course: The argument for change and continuity. *Psychological Inquiry: An International Journal for the Advancement of Psychological Theory.m, 12(2),* 49-66. DOI: 10.1207/S15327965PLI1202_01

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology, 56,* 453–484.

Cheung, G.W., & Rensvold, R.B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modelling: A multidisciplinary Journal, 9 (2),* 233-255.

Chmielewski, M, & Watson, D. (2009). What is being assessed and why it matters: the impact of transient error in trait research. *Journal of Personality and Social Psychology, 97,* 186-202.

Costa, P. T., & McCrae, R. R. (1988a). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology,* 54**,** 853–863.

Costa, P. T., & McCrae, R. R. (1988b). From catalogue to classification: Murray's needs and the five-factor model. *Journal of Personality and Social Psychology,* 55**,** 258–265.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory and NEO Five-Factor Inventory professional manual.* Odessa, FL: Psychological Assessment Resources.

Costa, P. T., & McCrae, R. R. (1994). *Set like plaster? Evidence for the stability of adult personality.* In T. F. Heatherton & J. L. Weinberger, (Eds.), *Can personality change? (pp. 21–40).* Washington, DC: American Psychological Association Books.

Costa Jr, P. T., & McCrae, R. R. (1997). Stability and change in personality assessment: the revised NEO Personality Inventory in the year 2000. *Journal of personality assessment, 68*(1), 86-94.

Costa, P. T., Jr., & McCrae, R. R. (2006). Age changes in personality and their origins: Comment on Roberts, Walton, and Viechtbauer (2006). *Psychological Bulletin, 132,* 26–28.

Costa, P. T., McCrae, R. R., & Dye, D. A. (1991). Facet scales for agreeableness and conscientiousness: a revision of the NEO personality inventory. *Personality and Individual Differences, 12*(9), 887-898.

Credé, M., Harms, P.D., Niehorster, S., & Gaye-Valentine, A. (2012). An Evaluation of the consequences of using short measures of the Big Five personality traits. *Journal of Personality and Social Psychology, 102(4),* 874-888.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York, NY: Holt, Rinehart and Winston.

DeYoung, C.G. (2010). Toward a theory of the Big Five. *Psychological Inquiry, 21(1),* 26-33.

DeYoung, C. G. (2015). Cybernetic big five theory. *Journal of Research in Personality, 56,* 33-58.

DeYoung, C.G., Quilty, L.C., & Peterson, J.B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology, 93(5),* 880-896.

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,* 558–566.

Donnellan, M. B., & Robins, R. W. (2009). The development of personality across the life span. In P. J. Corr & G. Matthews (Eds.), *Cambridge handbook of personality* (pp. 191–204). New York, NY: Cambridge University Press.

Donnellan, M.B., Oswald F.L., Baird B.M, & Lucas, R.E. (2006). The Mini-IPIP scales: Tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment, 18 (2),* 192-203.

Duckitt, J., & Sibley, C.G. (2010). Personality, ideology, prejudice, and politics: A dual-process motivational model. *Journal of Personality, 78(6),* 1861-1894.

Emmonds, R.A. (1995). Levels and domains of personality: An introduction. *Journal of Personality, 63 (3),* 341-364.

Erikson, E. H. (1950). *Childhood and society*. New York: Norton.

Eysenck, H. J. (1970). *The structure of human personality*. London: Methuen. 3rd. ed.

Ferguson, C.J. (2010). A meta-analysis of normal and disordered personality across the life dpan. *Journal of Personality and Social Psychology, 98(4),* 659-667.

Fiske, D. W. (1949). Consistency of the factorial structures of personality ratings from different sources. 1. *Abnormal Social Psychology. 44*, 329-44

Galton, F. (1884). Measurement of character. *Fortnightly Review, 36*, 179-185.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2003). *Bayesian data analysis* (2nd ed.). New York, NY: Chapman & Hall/CRC.

Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.), *Review of personality and social psychology* (Vol. 2, pp. 141- 165). Beverly Hills, CA: Sage.

Goldberg, L. R. (1990). An alternative "Description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology,59,* 1216-1229.

Goldberg, L. R. (1992). The development of markers of the Big-Five factor structure. *Psychological Assessment, 4,* 26-42.

Goldberg, L.R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe, Vol. 7* (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. Jr., (2003). A very brief measure of the Big-Five personality domains. Journal of Research in Personality, 37, 504–528.

Guilford, J. P. (1975). Factors and factors of personality. *Psychol. Bull. 82*, 802- 1 4

Hogan, R. (1983). Socioanalytic theory of personality. In *1982 Nebraska Symposium on Motivation: Personality-Current Theory and Research*, ed. M . M. Page, pp. 55-89. Lincoln, NE: Univ. Nebraska Press

Jang, K.L., Livesley, W.J., & Vemon, P.A. (1996). Heritability of the Big Five personality dimensions and trait facets: A twin study. *Journal of Personality, 64(3),* 577-591.

John, O.P., & Srivastava, S. (1999). The Big-Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102-138). New York: Guilford Press.

Klages, L. (1926). *The science of character*. (Trans!. 1932). London: Allen & Unwin

Kogan, N. (1990). Personality and aging. In J. E. Birren & S. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 330–346). San Diego: Academic Press.

Kruyen, P.M., Emons, W.H.M., Sijtsma, K. (2013). On the shortcomings of shortened tests: A literature review. *International Journal of Testing, 13,* 223-2484.

Kruschke, J. K., Aguinis, H., & Joo, H. (2012). The time has come: Bayesian methods for data analysis in the organizational sciences. *Organizational Research Methods, 15,* 722-752.

Lee, K., & Ashton, M.C. (2007). Factor analysis in personality research. In R.W. Robins, R.C. Fraley, & R.F. Krueger (Eds.) *Handbook pf Research Methods in Personality Psychology* . New York: Guilford Pres.

Lee, K., Ashton, M. C., & de Vries, R. E. (2005). Predicting workplace delinquency and integrity with the HEXACO and five-factor models of personality structure. *Human Performance*, *18*, 179-197.

Lee, K., Gizzarone, M., & Ashton, M. C. (2003). Personality and the likelihood to sexually harass. *Sex Roles*, *49,* 59-69.

Levinson, D. J. (1978). *The seasons of a man's life*. New York: Ballantine Books.

Lewis, M. (1999). On the development of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality theory and research* (pp. 327–346). New York: Guilford Press.

Löckenhoff, C.R., Terraciano, A., Bienvenu, O., Patriciu, N.S., Nestadt, G., McCrae, R., Eaton, W.W., & Costa Jr., P.T. (2009). Self-reported extremely adverse life events

and longitudinal changes in Five-Factor Model personality traits in an urban sample. *Journal of Traumatic Stress, 22(1),* 53-59.

Loehlin, J., McCrae, R.R., Costa Jr., P.T., & John, O.P. (1998). Heritabilities of common and measure-specific components of the Big Five personality factors. *Journal of Research in Personality, 32,* 431-453.

Lucas, R.E. (2007). Adaptation and the set-point model of subjective well-being: Does happiness change after major life events? *Current Directions in Psychological Science, 16(2),* 75-79.

Lucas, R.E., & Donnellan, M.B. (2011). Personality development across the life span: Longitudinal analyses with a national sample from Germany. *Journal of Personality and Social Psychology, 101(4),* 847-861.

Magnus, K., Diener, E., Fujita, F., & Pavot, W. (1993). Extraversion and neuroticism as predictors of objective life events: A longitudinal analysis. *Journal of Personality and Social Psychology, 35(5),* 1046-1053.

Marsh, H. W., Nagengast, B., & Morin, A. J. S. (2013). Measurement invariance of Big-Five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and La Dolce Vita effects. *Developmental Psychology, 49,* 1194-1218

McAdams, D.P. (1995). What do we know when we know a person? *Journal of Personality, 63 (3),* 366-396.

McCartney, K., Harris, M.J., & Bernieri, F. (1990). Growing up and growing apart: A developmental meta-analysis of twin studies. *Psychological Bulletin, 107(2),* 226-237.

McCrae, R. R., & Costa, E T., Jr. (1985). Comparison of EPI and psychoticism scales with measures of the five-factor model of personality. *Personality and Individual Differences, 6,* 587-597.

McCrae, R. R., & Costa, P. X, Jr. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52,* 81-90.

McCrae, R.R., & Costa, P.T. (1997). Personality trait structure as a human universal. *American Psychologist, 52,* 509-516.

McCrae, R. R., & Costa, P. T., Jr. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality theory and research* (Vol. 2, pp. 139–153). New York: Guilford Press.

McCrae, R. R., & Costa, P. T., Jr. (2008). The Five-Factor Theory of personality. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 159–181). New York, NY: Guilford Press.

McCrae, R. R., Costa, P. T., Jr., Ostendorf, F., Angleitner, A., Høebı´cˇkova´, M., Avia, M. D.…Smith, P. B. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology, 78,* 173–186.

McCrae, R. R., Costa, P. T., Jr., Pedroso de Lima, M., Simoes, A., Ostendorf, F., Angleitner, A., et al. (1999). Age differences in personality across the adult life span: Parallels in five cultures. *Developmental Psychology, 35,* 466–477.

McGue, M., Bacon, S., & Lykken, D.T. (1993). Personality stability and change in early adulthood: A behavioural genetic analysis. *Developmental Psychology, 29(1),* 96-109.

Muthén, B. (2010). Bayesian analysis in Mplus: a brief introduction. *Mplus working paper*.

Muthén, L. K., & Muthén, B. O. (2002). How to Use a Monte Carlo Study to Decide on Sample Size and Determine Power. *Structural Equation Modelling: A Multidisciplinary Journal, 9,* 599-620.

Milojev, P., & Sibley, C.G. (2014). The stability of adult personality varies across age: Evidence from a two-year longitudinal sample of adult New Zealanders. *Journal of Research in Personality, 51,* 29-37.

Milojev, P., Osborne, D., & Sibley, C. G. (2014). Personality resilience following a natural disaster. *Social Psychology and Personality Science,* 5(7), 760-768.

Milojev, P. Osborne, D., Greaves, L.M., Barlow, F.K., & Sibley, C.G. (2013). The Mini-IPIP6: Tiny yet highly stable markers of Big Six personality. *Journal of Research in Personality, 47,* 936-944.

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

Mischel, W., & Shoda, Y. (1995). A cognitive-Affective System Theory of personality: Reconceptualising situations, disposition, dynamics, and invariance in personality structure. *Psychological review, 102(2),* 246-268.

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

Neyer, F. J., & Asendorpf, J. B. (2001). Personality–relationship transaction in young adulthood. *Journal of Personality and Social Psychology, 81,* 1190–1204.

Nolen-Hoeksema, S., & Morrow, J. (1991). A prospective study of depression and posttraumatic stress symptoms after a natural disaster: The 1989 Loma Prieta earthquake. *Journal of Personality and Social Psychology, 61(1),* 115-121.

Norman , W. T. (1963). Toward an adequate taxonomy of personality attributes: replicated factor structure in peer nomination personality ratings. *Journal of Abnormal Social Psychology 66*,574-83

Ogle, M., C., Rubin, D.C., & Siegler, I.C. (2013). Changes in neuroticism following trauma exposure. *Journal of Personality.*

Orth, U, Trzesniewski, K.H., & Robins, R.W. (2010). Self-esteem development from young adulthood to old age: A cohort-sequential longitudinal study. *Journal of Personality and Social Psychology, 98 (4),* 645-658.

Osborne, D., & Sibley, C.G. (2013). After the disaster: Using the Big Five to predict changes in mental health among survivors of the 2011 Christchurch earthquake. *Disaster Prevention and Management, 22(5),* 456-466.

Paulhus, D.L. (1991). Measurement and control of response bias. In. J.P. Robertson, P.R. Shaver, & L.S. Wrightsman (Eds.), *Measurement of personality and social psychological attitudes* (pp 17-59). San Diego: Academic Press.

Phifer, J.F. (1990). Psychological distress and somatic symptoms after natural disaster: Differential vulnerability among older adults. *Psychology and Aging, 5(3),* 412-420.

Piedmont, R.L. McCrae, R.R., Riemann, R. & Angleitner, A. (2000). On the invalidity of validity scales: Evidence from self-reports and observed ratings in volunteer samples. *Journal of Personality and Social Psychology, 78,* 582-593.

Prinzie, P., & Onghena, P. (2005). Cohort sequential design. In B. Everitt & D. Howell (Eds.), *Encyclopaedia of statistics in behavioral science. Vol. 1* (pp. 319-322). New York: Wiley.

Rammstedt, B., & John, O.P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality, 41,* 203-212.

Rantanen, J.m Metsäpelto, R.-L., Pulkkinen, L., & Kokko, K. (2007), Long-term stability of the Big Five personality traits in adulthood. *Scandinavian Journal of Psychology, 48,* 511-518.

Roberts, B.W. (2009). Back to the future: *Personality and Assessment* and personality development. *Journal of Research in Personality, 43,* 137-145.

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.

Roberts, B. W., & Caspi, A. (2003). The cumulative continuity model of personality development: Striking a balance between continuity and change in personality traits across the life course. In U. Staudinger & U. Lindenberger (Eds.), *Understanding human development: Lifespan psychology in exchange with other disciplines* (pp. 183–214). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Roberts, B. W., Caspi, A., & Moffitt, T. (2001). The kids are alright: Growth and stability in personality development from adolescence to adulthood. *Journal of Personality and Social Psychology, 81,* 670–683.

Roberts, B. W., Wood, D., & Caspi, A. (2008). The development of personality traits in adulthood. *Handbook of personality: Theory and research, 3*, 375-398.

Roberts, B.W., O'Donnell, M., & Robins, R.W. (2004). Goal and personality development. *Journal of Personality and Social Psychology, 8,* 402-416.

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

Robert, B.W., & Wood, D. (2006). Personality development in the context of the Neo-Socioanalytic model of personality. In D.K. Mroczek & T.D. Little (Eds.), *Handbook of personality development* (pp. 11-39). Mahwah, NJ, US; Lawrence Erlbaum Associates Publishers.

Roberts, B.W., Wood, D., & Caspi, A. (2008). The development of personality traits in adulthood. In O.P. John, R.W., Robins, & L.A. Pervin (eds.), *Handbook of Personality: Theory and research* (3rd ed., pp. 375-398). New York, NY: 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,* 166-184.

Robins, R.W., Fraley, R.C., Roberts, R.W., & Trzesniewski, K.H. (2001). A longitudinal study of personality change in young adulthood. *Journal of Personality, 69 (4),* 617 – 640.

Salmond, C., & Crampton, P, (2012). Development of New Zealand's deprivation index (NZDep) and its uptake as a national policy tool. *Canadian Journal of Public Health, 103,* 7-11.

Salmond, C, P Crampton and J Atkinson (2007), *NZDep2006 index of deprivation*. Wellington: University of Otago (Department of Public Health). Retrieved from http://www.wnmeds.ac.nz/academic/dph/research/socialindicators.html.

Satherley, N., Milojev, P., Greaves, L.M., Huang, Y., Osborne, D., Bulbulia, J., & Sibley, C.G. (2015). Demographic and psychological predictors of panel attrition: Evidence from the New Zealand Attitudes and Values Study. *PloS one*, *10(3),* e0121950.

Schmitt, D.P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology, 94(1),* 168-182.

Scollon, C. N., & Diener, E. (2006). Love, work, and changes in extraversion and neuroticism over time. *Journal of Personality and Social Psychology, 91,* 1152–1165. doi:10.1037/0022-3514.91.6.1152

Sibley, C. G. (2012). The Mini-IPIP6: Item Response Theory analysis of a short measure of the big-six factors of personality in New Zealand. *New Zealand Journal of Psychology*, *41,* 21-31.

Sibley, C. G. (2014) Sampling procedure and sample details for the New Zealand Attitudes and Values Study. *NZAVS Technical Documents, e01.* Available at: http://www.psych.auckland.ac.nz/uoa/NZAVS

Sibley, C. G., Luyten, N., Purnomo, M., Moberly, A., Wootton, L. W., Hammond, M. D., Sengupta, N., Perry, R., West-Newman, T., Wilson, M. S., McLellan, L., Hoverd, W. J., & Robertson, A. (2011). The Mini-IPIP6: Validation and extension of a short measure of the Big-Six factors of personality in New Zealand. *New Zealand Journal of Psychology, 40,* 142-159.

Sibley, C. G., & Milojev, P. (2014). Power estimation of slope growth factors in the NZAVS using Monte Carlo simulation. *NZAVS Technical Documents, e19.* Available at: http://www.psych.auckland.ac.nz/uoa/NZAVS

Sibley, C. G., & Pirie, D. J. (2013). Personality in New Zealand: Scale norms and demographic differences in the Mini-IPIP6. *New Zealand Journal of Psychology, 42,* 13-30.

Smith, G. M. (1967). Usefulness of peer ratings of personality in educational research. *Educational and Psychological Measurement, 27,* 967-984.

Soldz, S, & Valliant, G.E. (1999). The Big Five personality traits and the life course: A 45-year longitudinal study. *Journal of Research in Personality, 33,* 208-232.

Soto, C. J. (in press). The Little Six personality dimensions from early childhood to early adulthood: Mean-level age and gender differences in parents' reports. *Journal of Personality.*

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. *Journal of Personality, 80, 881-914.*

Specht, J., Egloff, B., & Scmukle, S.C. (2011). Stability and change of personality across the life course: The impact of age and major life events in mean-level and rank-order

stability of the Big Five. *Journal of Personality and Social Psychology, 101(4),* 862-882.

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. *European Journal of Personality, 28,* 216-230.

Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of Personality and Social Psychology, 84,* 1041–1053.

Terracciano, A., McCrae, R. R., Brant, L. J., & Costa, P. T., Jr. (2005). Hierarchical linear modelling analyses of the NEO-PI-R scales in the Baltimore Longitudinal Study of Aging. *Psychology and Aging, 20,* 493–506.

Tupes, E. C., & Christal, R. E. (1961). Recurrent personality factors based on trait ratings. *USAF ASD Tech . Rep*, 61-97

Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety with an emphasis on self-report. In *Anxiety and the Anxiety Disorders* (pp. 68 1-706.), A. Tuma, J. Maser (Eds.), Hillsdale, NJ: Erlbaum

Terracciano, A., Costa, P. T., Jr., & McCrae, R. R. (2006). Personality plasticity after age 30. *Personality and Social Psychology Bulletin, 32,* 999–1009.

Thurnstone, L.L. (1934). The vectors of mind. *Psychological Review, 41,* 1-32.

Trzesniewski, K.H., Donnellan, M.B., & Robins, R.W. (2003). Stability of self-esteem across the Life Span. *Journal of Personality and Social Psychology, 24(1),* 205-220.

Twenge, J. M., Konrath, S., Foster, J. D., Campbell, W. K., & Bushman, B. J. (2008). Egos inflating over time: A cross-temporal meta-analysis of the Narcissistic Personality Inventory. Journal of Personality, 76, 875–901.

Van den Akker, A. L., Deković, M., Asscher, J. 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, 736-750.*

White, P., Gunston, J., Salmond, C., Atkinson, J., & Crampton, P. (2008). Atlas of socioeconomic deprivation in New Zealand: NZDep2006. Wellington Ministry of Health.

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, 27(4),* 867-874.

Yuan, Y., & MacKinnon, D. P. (2009). Bayesian mediation analysis. *Psychological Methods, 14,* 301-322.

Zelli, A., & Dodge, K. A. (1999). Personality development from the bottom up. In D. Cervone & Y. Shoda (Eds.), *The coherence of personality: Social– cognitive bases of consistency, variability, and organization* (pp. 94–126). New York: Guilford Press.

Zuroff, D.C. (1994). Depressive personality styles and the Five Factor Model of personality. *Journal of Personality Assessment, 63(3),* 253-265.

Zuroff, D.C., Mongrian, M., & Santor, D.A. (2004). Conceptualizing and measuring personality vulnerability to depression: Comment on Coyne and Whiffen (1995). *Psychological Bulletin, 130(3),* 489-511.

Appendices

Appendix A

This Appendix presents the integral supplementary file accompanying the paper presented in Chapter 5 – study 4. As referred to in the sad chapter the following supplementary file s present the detailed NZAVS sampling procedure (Supplementary file 1), Monte Carlo Simulations for power analyses (Supplementary file 2), sample *MPlus* syntax for estimation of the Cohort-Sequential Latent Growth Models – both the single-group and the multi-group estimation procedure (Supplementary file 3), tests of measurement invariance of the Mini-IPIP6 (Supplementary file 4), and the bivariate correlations and descriptive statistics for the variables used in the analyses (Supplementary file 5).

Supplementary file 1 – NZAVS participants and sampling procedure

The New Zealand Attitudes and Values Study (NZAVS) is an ongoing effort that has been conducting an annual longitudinal panel survey of adult New Zealanders since 2009 (Time 1) through 2010 (Time 2), 2011 (Time 3), 2012 (Time 4) ,2013 (Time 5), and 2014 (Time 6). The Time 1 (2009) NZAVS contained responses from 6518 participants sampled from the 2009 New Zealand electoral roll. The electoral roll is publicly available for scientific research and in 2009 contained 2,986,546 registered voters. This represented all citizens over 18 years of age who were eligible to vote regardless of whether they chose to vote, barring people who had their contact details removed due to specific case-by-case concerns about privacy. That year postal questionnaires were sent to 40,500 registered voters or roughly 1.36% of all registered voters in New Zealand. The overall response rate (adjusting for the address accuracy of the electoral roll and including anonymous responses) was 16.6%.

The Time 2 (2010) NZAVS contained responses from 4442 participants - 4423 from the initial Time 1 sample and additional 19 respondents who could not be matched to the Time 1 participant database (a retention rate of 67.9% over one year). The Time 3 (2011) NZAVS contained responses from 6884 participants - 3918 retained from the initial Time 1 sample (a 60.1% retention rate over two years), and 2966 new participants. To boost sample size at Time 3 and compensate for sample attrition, a booster sample was recruited through an unrelated survey posted on the website of a major New Zealand newspaper in 2011. A total of 3208 participants registered an initial expression of interest in being contacted to participate in the NZAVS via this survey. A total of 2961 participants completed the questionnaire when subsequently contacted (92.4%). This yielded a total sample size for the Time 3 (2012) NZAVS of 6884 (3918 retained from Time 1, 3 additions retained from Time 2, and 2961 recruited from the newspaper website at Time 3).

The Time 4 (2012) NZAVS contained responses from 12,182 participants - 6805 retained from one or more previous wave, 5377 new additions from booster sampling. The sample retained 4051 participants from the initial Time 1 sample (a retention rate of 62.2% over three years); the sample retained 5762 participants from the full Time 3 sample (a retention rate of 83.7% from the previous year). To boost sample size at Time 4 and increase sample diversity for subsequent waves, five independent booster samples using different sample frames were also conducted. The first sample frame consisted of a randomly selected sample of 20,000 people from the 2012 New Zealand Electoral Roll. A total of 2431 participants responded to this booster sample (response rate = 12.34% when adjusting for the 98.5% accuracy of the 2012 electoral roll). The second sample frame consisted of a regional booster of 10,000 people randomly selected from people listed in the 2012 Electoral Roll who

lived in the Auckland region. A total of 890 participants responded to this booster sample (adjusted response rate = 9.04%).

The third sample frame consisted of 3,000 people randomly selected from the 2012 Electoral Roll who lived in the Christchurch region. A total of 333 participants responded to this booster sample (adjusted response rate = 13.52%). The fourth sample frame consisted of 9000 respondents selected from meshblock area units across the country that were moderate-to-high in deprivation according to the index developed by Salmond, Crampton and Atkinson (2007). Regions with levels of deprivation were selected using scores on the decile-ranked NZ Deprivation index from 6-10 (with 10 being the most deprived). This sample frame used scaled weighting so that people in increasingly deprived regions were increasingly more likely to be selected (with random sampling of people within regions that had a given level of deprivation). A total of 767 participants responded to this booster sample (adjusted response rate = 9.73%). The fifth sample frame consisted of 9,000 people randomly selected from those who indicated on the 2012 Electoral Roll that they were of Maori ethnicity (ethnic affiliation as Maori is listed on the role, but other ethnic affiliations are not). A total of 690 participants responded to this booster sample (adjusted response rate = 7.78%).

The Time 5 (2013) NZAVS contained responses from 18,264 participants (10,502 retained from one or more previous wave, 7,581 new additions from a booster sample, and 181 unmatched participants or unsolicited opt-ins). The sample retained 3,934 participants from the initial Time 1 (2009) NZAVS of 6518 participants (a retention rate of 60.4% over four years). The sample retained 9,844 participants from the Time 4 (2012) sample (a retention rate of 80.8% from the previous year). Participants were posted a copy of the questionnaire with a second postal follow-up two months later. Participants who provided an email address were also simultaneously emailed and invited to complete an online version if they preferred. As described in the Time 4 procedure, we offered a prize draw for participation, non-respondents were emailed and phoned multiple times, and all participants were posted a Season's Greetings card from the NZAVS research team and informed that they had been automatically entered into a bonus seasonal grocery voucher prize draw. We also posted our yearly pamphlet summarizing key research findings published during the current wave of the study. To boost sample size and increase sample diversity for subsequent waves, two booster samples were also conducted. As with previous booster samples, sampling was conducted without replacement (i.e. all people included in previous sample frames were identified and removed from the 2014 roll). The first sample frame consisted of 70,000 people aged from 18-60 randomly selected from the 2014 New Zealand Electoral Roll. The New Zealand electoral roll contains participants' date of birth (within a one-year window), and we limited our frame to people who were 60 years or younger due to our aim of retaining participants for the following 15 years. A total of 7,489 participants responded to this booster sample (response rate = 10.9% when adjusting for the 98.6% accuracy of the 2014 electoral roll). The second sample frame consisted of 1,500 people who were listed as having Maori ancestry aged from 18-60 randomly selected from the 2014 New Zealand Electoral Roll. A total of 92 participants responded to this booster sample (response rate = 6.2% when adjusting for electoral roll accuracy).

The Time 6 (2014) NZAVS contained responses from 15,822 participants (15,740 retained from one or more previous wave, and 82 unmatched participants or unsolicited opt-ins). The sample retained 3,727 participants from the initial Time 1 (2009) NZAVS of 6,518 participants (a retention rate of 57.2% over five years). The sample retained 14,875 participants from the full Time 5 (2012) sample (a retention rate of 81.5% from the previous year). Participants were posted a copy of the questionnaire, with a second postal follow-up

two months later. Participants who provided an email address were also emailed and invited to complete an online version if they preferred. As described in the Time 5 procedure, we offered a prize draw for participation, non-respondents were emailed and phoned multiple times, and all participants were posted a Season's Greetings card from the NZAVS research team and informed that they had been automatically entered into a bonus seasonal grocery voucher prize draw. We also emailed participants an online pamphlet containing a series of video interviews with the researchers summarizing different research findings.

For illustrative purposes the sample details and retention rates are presented in Table A.1.1 below.

Table A.1.1. Summary sample details for the first five yearly waves of the NZAVS (2009-2013).

	Time 1	Time 2	Time 3	Time 4	Time 5	Time 6
Sample size during wave	6518	4442	6884	12182	18264	15822
Additions during wave	-----	20	2965	5377	7699	82
Retained from wave 1 (Time 1, %)	-----	67.9	60.1	62.2	60.4	57.2
Retained from previous wave (%)	-----	67.9	79.5	83.7	80.8	81.5

Note: Additions to a particular wave include booster sampling, opt-ins, and un-matched participants.

Supplementary File 2 - A Note on Statistical Power

The modelling approach here proposed – specifically the multi-group cohort-sequential model - is highly demanding. Moreover, given the relative novelty of work in this area, researchers may well be wondering about the statistical power of a Cohort-Sequential Growth Modelling approach. Specifically, the power to detect the focal parameter of interest - the mean of the latent slope in each of the estimated Latent Growth Models – given the relatively small cohort size. Our analysis of power is somewhat different to what most readers may be familiar with, as our sample size is know well in advance (for waves of data already collected, and for that matter, can be accurately projected into the future given knowledge about rates of sample attrition).

We conducted Monte Carlo simulations to assess our power to detect latent slopes of various sizes, given the known sample sizes across out age-cohorts. We used the procedure promoted by Sibley and Milojev (2014) for use with the NZAVS, as adapted from Muthén and Muthén (2002). Based on the work of Sibley & Milojev (2014) we simulated a data structure specifying a LGM with 5 equidistant time points. As such, the Monte Carlo simulation uses a somewhat simpler approach given that the Cohort-Sequential Growth Models here presented are estimated based on time-varying indicators. The means of the manifest, continuous variables at each time point were fixed at 0 and their variances fixed at 1. The mean of the latent intercept was fixed to 4, and the variance of the latent intercept was fixed to 1. The variance of the latent slope was fixed to .10. The correlation between the intercept and the slope was fixed to .10. The parameter of interest was the mean of the latent slope. This parameter was systematically varied between .005 and .10. As such, we were able to obtain a function of the obtained power to detect latent rates of change of different magnitudes given the known sample size. Four versions of these simulations were conducted at different specified sample sizes. Two of the sample sizes were based on the actual cohort-specific sample sizes and selected to represent the small (n = 300) and the large (n = 500) cohort size. The other two sample sizes were selected to represent a large sample size in light of our cohorts (n = 700), and a very large sample size (n = 5000) for explorative purposes. As such, the simulations were conducted for sample sizes at n = 300, n = 500, n = 700, and n = 5000 observations.

The results of these power simulations are presented in Figure A.2.1. The figure presents functions for the estimated power (*y-axis*) across the range of effect sizes (*x-axis*), at the four different sample sizes. As can be seen in the figure, the specified LGMs have relatively little power to detect small effect sizes of the latent slope across the two cohort sizes. The function showing lowest power across the effect sizes, unsurprisingly, was the function for the smallest sample size (n = 300). The obtained power functions suggest that the interpretations of the Cohort-Sequential Growth Models as presented below must be made cautiously, particularly with respect to the small values of the slope in the smaller cohorts. Furthermore, these functions, coupled with the Monte Carlo simulations based on the NZAVS sample (Sibley & Milojev, 2014) highlight the importance of the relative simplicity of the modelling approach as described above.[24] Quite simply, we have adequate power, but even with our longitudinal sample size (n = 4511) care is needed not to step beyond the bounds of a reasonable modelling approach given we were working with only five annual waves of data.

[24] For Monte Carlo simulations of more advanced Latent Growth Model within a simulated NZAVS data structure see Sibley & Milojev (2014), available at: http://www.psych.auckland.ac.nz/en/about/our-research/research-groups/new-zealand-attitudes-and-values-study/nzavs-tech-docs.html.

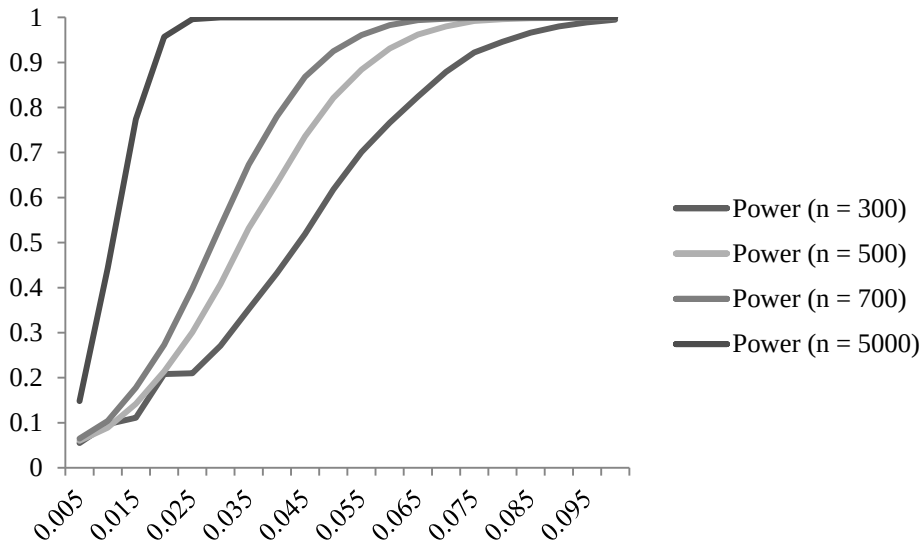


Figure A.2.1. Estimated power for detecting the men of the latent slope across a range of sample sizes. The estimated power (*y axis*) refers to the proportion of Monte Carlo replications for which the null hypothesis that a parameter is equal to zero is rejected at the .05 level (Muthén & Muthén, 2002). The effect size for the focal parameter – the mean of the slope – was systematically varied between .005 and .10 (*x-axis*). The simulations were conducted at two different sample sizes based on the sizes of the birth cohorts presented in Table 1: 1) small cohort size (n = 300), 2) large cohort size (n = 500); as well as 3) a large sample size (n = 700), and 4) a very large sample size (n = 5000). These are represented as separate functions in the figure.

Supplementary File 3 – Sample *MPlus* syntax

This supplementary document presents the selected *MPlus* syntax used to implement the models presented in this paper. The syntax is annotated for the readers' reference (*note:* annotations are preceded with "!" as per *MPlus* language specifications). Please note that the full syntax for the models presented here, as well as worked examples of these models with the associated data will be presented online – on the NZAVS web page and the first author's domain – upon the acceptance of the paper.

Single-Group Cohort-Sequential Latent Growth Model

```
DATA:

FILE IS S CSLGM DATA.txt;

VARIABLE:

IDVARIABLE IS subnum;
MISSING ARE ALL (9999);


NAMES ARE

Y_T1
Y_T2
Y_T3
Y_T4
Y_T5
Y_T6
TS1
TS2
TS3
TS4
TS5
TS6
SUBNUM;

USEVARIABLE ARE

y_T1 y_T2 y_T3 y_T4 y_T5 y_T6
TS1 TS2 TS3 TS4 TS5 TS6;

TSCORES = TS1 TS2 TS3 TS4 TS5 TS6;
!Defines TSCORES for time-varying growth models
!in this case: participants' age

ANALYSIS:

PROCESSORS = 4; TYPE = RANDOM; ESTIMATOR = MLR;

MODEL:
!LGM specified based on the TSCORES (age)
!estimating linear, quadratic and cubic growth curves
```

```
i s q c| y_T1-y_T6 AT TS1-TS6;

y_T1 (a);
y_T2 (a);
y_T3 (a);
y_T4 (a);
y_T5 (a);
y_T6 (a);

[i] (ix);
[s] (sx);
[q] (qx);
[c] (cx);

!Estimating conditional mean-level estimates for the given trait
MODEL CONSTRAINT:

new(Age18 Age19
   Age20 Age21 Age22 Age23 Age24 Age25 Age26 Age27 Age28 Age29
   Age30 Age31 Age32 Age33 Age34 Age35 Age36 Age37 Age38 Age39
   Age40 Age41 Age42 Age43 Age44 Age45 Age46 Age47 Age48 Age49
   Age50 Age51 Age52 Age53 Age54 Age55 Age56 Age57 Age58 Age59
   Age60 Age61 Age62 Age63 Age64 Age65 Age66 Age67 Age68 Age69
   Age70 Age71 Age72 Age73 Age74);

Age18 = ix + (sx * ((18-45)/10)) + (qx * ((18-45)/10)^2) + (cx * ((18-45)/10)^3);
Age19 = ix + (sx * ((19-45)/10)) + (qx * ((19-45)/10)^2) + (cx * ((19-45)/10)^3);
Age20 = ix + (sx * ((20-45)/10)) + (qx * ((20-45)/10)^2) + (cx * ((20-45)/10)^3);
Age21 = ix + (sx * ((21-45)/10)) + (qx * ((21-45)/10)^2) + (cx * ((21-45)/10)^3);
Age22 = ix + (sx * ((22-45)/10)) + (qx * ((22-45)/10)^2) + (cx * ((22-45)/10)^3);
Age23 = ix + (sx * ((23-45)/10)) + (qx * ((23-45)/10)^2) + (cx * ((23-45)/10)^3);
Age24 = ix + (sx * ((24-45)/10)) + (qx * ((24-45)/10)^2) + (cx * ((24-45)/10)^3);
Age25 = ix + (sx * ((25-45)/10)) + (qx * ((25-45)/10)^2) + (cx * ((25-45)/10)^3);
Age26 = ix + (sx * ((26-45)/10)) + (qx * ((26-45)/10)^2) + (cx * ((26-45)/10)^3);
Age27 = ix + (sx * ((27-45)/10)) + (qx * ((27-45)/10)^2) + (cx * ((27-45)/10)^3);
Age28 = ix + (sx * ((28-45)/10)) + (qx * ((28-45)/10)^2) + (cx * ((28-45)/10)^3);
Age29 = ix + (sx * ((29-45)/10)) + (qx * ((29-45)/10)^2) + (cx * ((29-45)/10)^3);
Age30 = ix + (sx * ((30-45)/10)) + (qx * ((30-45)/10)^2) + (cx * ((30-45)/10)^3);
Age31 = ix + (sx * ((31-45)/10)) + (qx * ((31-45)/10)^2) + (cx * ((31-45)/10)^3);
Age32 = ix + (sx * ((32-45)/10)) + (qx * ((32-45)/10)^2) + (cx * ((32-45)/10)^3);
Age33 = ix + (sx * ((33-45)/10)) + (qx * ((33-45)/10)^2) + (cx * ((33-45)/10)^3);
Age34 = ix + (sx * ((34-45)/10)) + (qx * ((34-45)/10)^2) + (cx * ((34-45)/10)^3);
Age35 = ix + (sx * ((35-45)/10)) + (qx * ((35-45)/10)^2) + (cx * ((35-45)/10)^3);
Age36 = ix + (sx * ((36-45)/10)) + (qx * ((36-45)/10)^2) + (cx * ((36-45)/10)^3);
Age37 = ix + (sx * ((37-45)/10)) + (qx * ((37-45)/10)^2) + (cx * ((37-45)/10)^3);
Age38 = ix + (sx * ((38-45)/10)) + (qx * ((38-45)/10)^2) + (cx * ((38-45)/10)^3);
Age39 = ix + (sx * ((39-45)/10)) + (qx * ((39-45)/10)^2) + (cx * ((39-45)/10)^3);
Age40 = ix + (sx * ((40-45)/10)) + (qx * ((40-45)/10)^2) + (cx * ((40-45)/10)^3);
Age41 = ix + (sx * ((41-45)/10)) + (qx * ((41-45)/10)^2) + (cx * ((41-45)/10)^3);
Age42 = ix + (sx * ((42-45)/10)) + (qx * ((42-45)/10)^2) + (cx * ((42-45)/10)^3);
```

```
Age43 = ix + (sx * ((43-45)/10)) + (qx * ((43-45)/10)^2) + (cx * ((43-45)/10)^3);
Age44 = ix + (sx * ((44-45)/10)) + (qx * ((44-45)/10)^2) + (cx * ((44-45)/10)^3);
Age45 = ix + (sx * ((45-45)/10)) + (qx * ((45-45)/10)^2) + (cx * ((45-45)/10)^3);
Age46 = ix + (sx * ((46-45)/10)) + (qx * ((46-45)/10)^2) + (cx * ((46-45)/10)^3);
Age47 = ix + (sx * ((47-45)/10)) + (qx * ((47-45)/10)^2) + (cx * ((47-45)/10)^3);
Age48 = ix + (sx * ((48-45)/10)) + (qx * ((48-45)/10)^2) + (cx * ((48-45)/10)^3);
Age49 = ix + (sx * ((49-45)/10)) + (qx * ((49-45)/10)^2) + (cx * ((49-45)/10)^3);
Age50 = ix + (sx * ((50-45)/10)) + (qx * ((50-45)/10)^2) + (cx * ((50-45)/10)^3);
Age51 = ix + (sx * ((51-45)/10)) + (qx * ((51-45)/10)^2) + (cx * ((51-45)/10)^3);
Age52 = ix + (sx * ((52-45)/10)) + (qx * ((52-45)/10)^2) + (cx * ((52-45)/10)^3);
Age53 = ix + (sx * ((53-45)/10)) + (qx * ((53-45)/10)^2) + (cx * ((53-45)/10)^3);
Age54 = ix + (sx * ((54-45)/10)) + (qx * ((54-45)/10)^2) + (cx * ((54-45)/10)^3);
Age55 = ix + (sx * ((55-45)/10)) + (qx * ((55-45)/10)^2) + (cx * ((55-45)/10)^3);
Age56 = ix + (sx * ((56-45)/10)) + (qx * ((56-45)/10)^2) + (cx * ((56-45)/10)^3);
Age57 = ix + (sx * ((57-45)/10)) + (qx * ((57-45)/10)^2) + (cx * ((57-45)/10)^3);
Age58 = ix + (sx * ((58-45)/10)) + (qx * ((58-45)/10)^2) + (cx * ((58-45)/10)^3);
Age59 = ix + (sx * ((59-45)/10)) + (qx * ((59-45)/10)^2) + (cx * ((59-45)/10)^3);
Age60 = ix + (sx * ((60-45)/10)) + (qx * ((60-45)/10)^2) + (cx * ((60-45)/10)^3);
Age61 = ix + (sx * ((61-45)/10)) + (qx * ((61-45)/10)^2) + (cx * ((61-45)/10)^3);
Age62 = ix + (sx * ((62-45)/10)) + (qx * ((62-45)/10)^2) + (cx * ((62-45)/10)^3);
Age63 = ix + (sx * ((63-45)/10)) + (qx * ((63-45)/10)^2) + (cx * ((63-45)/10)^3);
Age64 = ix + (sx * ((64-45)/10)) + (qx * ((64-45)/10)^2) + (cx * ((64-45)/10)^3);
Age65 = ix + (sx * ((65-45)/10)) + (qx * ((65-45)/10)^2) + (cx * ((65-45)/10)^3);
Age66 = ix + (sx * ((66-45)/10)) + (qx * ((66-45)/10)^2) + (cx * ((66-45)/10)^3);
Age67 = ix + (sx * ((67-45)/10)) + (qx * ((67-45)/10)^2) + (cx * ((67-45)/10)^3);
Age68 = ix + (sx * ((68-45)/10)) + (qx * ((68-45)/10)^2) + (cx * ((68-45)/10)^3);
Age69 = ix + (sx * ((69-45)/10)) + (qx * ((69-45)/10)^2) + (cx * ((69-45)/10)^3);
Age70 = ix + (sx * ((70-45)/10)) + (qx * ((70-45)/10)^2) + (cx * ((70-45)/10)^3);
Age71 = ix + (sx * ((71-45)/10)) + (qx * ((71-45)/10)^2) + (cx * ((71-45)/10)^3);
Age72 = ix + (sx * ((72-45)/10)) + (qx * ((72-45)/10)^2) + (cx * ((72-45)/10)^3);
Age73 = ix + (sx * ((73-45)/10)) + (qx * ((73-45)/10)^2) + (cx * ((73-45)/10)^3);
Age74 = ix + (sx * ((74-45)/10)) + (qx * ((74-45)/10)^2) + (cx * ((74-45)/10)^3);
```

OUTPUT:

SAMPSTAT
CINTERVAL

Multi-Group Cohort Sequential Growth Model

```
DATA:

FILE IS M CSLGM DATA.txt;

VARIABLE:

IDVARIABLE IS subnum;
MISSING ARE ALL (9999);

NAMES ARE

Y_T1
Y_T2
Y_T3
Y_T4
Y_T5
TS1
TS2
TS3
TS4
TS5
SUBNUM
COHORTNB

USEVARIABLE ARE

y_T1 y_T2 y_T3 y_T4 y_T5
TS1 TS2 TS3 TS4 TS5;

USEOBSERVATIONS ARE (TSUM GE 3);

TSCORES = TS1 TS2 TS3 TS4 TS5;
!Defines TSCORES for time-varying growth models
!in this case: date/time of response

GROUPING IS CohortNB (8=1936-1940,
                      9=1941-1945,
                      10=1946-1950,
                      11=1951-1955,
                      12=1956-1960,
                      13=1961-1965,
                      14=1966-1970,
                      15=1971-1975,
                      16=1976-1980,
                      17=1981-1985,
                      18=1986-1990,
                      19=1991-1995);
!Specifying 5-year birth cohorts
```

```
ANALYSIS:

PROCESSORS = 12; TYPE = RANDOM; ESTIMATOR = MLR;

MODEL:
!Estimating a linear LGM based on the TSCORES
i s | y_T1-y_T5 AT TS1-TS5;


[i];
[s];


MODEL 1936-1940:

[i] (ia);
[s] (sa);

i (1);
s (2);
i WITH s (3);

MODEL 1941-1945:

[i] (ib);
[s] (sb);

i (1);
s (2);
i WITH s (3);

MODEL 1946-1950:

[i] (ic);
[s] (sc);

i (1);
s (2);
i WITH s (3);

MODEL 1951-1955:

[i] (id);
[s] (sd);

i (1);
s (2);
i WITH s (3);

MODEL 1956-1960:
```

[i] (ie);
[s] (se);

i (1);
s (2);
i WITH s (3);

MODEL 1961-1965:

[i] (if);
[s] (sf);

i (1);
s (2);
i WITH s (3);

MODEL 1966-1970:

[i] (ig);
[s] (sg);

i (1);
s (2);
i WITH s (3);

MODEL 1971-1975:

[i] (ih);
[s] (sh);

i (1);
s (2);
i WITH s (3);

MODEL 1976-1980:

[i] (ii);
[s] (si);

i (1);
s (2);
i WITH s (3);

MODEL 1981-1985:

[i] (ij);
[s] (sj);

i (1);

```
s (2);
i WITH s (3);

MODEL 1986-1990:

[i] (ik);
[s] (sk);

i (1);
s (2);
i WITH s (3);

MODEL 1991-1995:

[i] (il);
[s] (sl);

i (1);
s (2);
i WITH s (3);

!Specifying conditional mean-level estimates for the given trait
MODEL CONSTRAINT:

NEW (a09_Oct a10_Oct a11_Oct
a12_Oct a13_Oct a14_Oct);

NEW (b09_Oct b10_Oct b11_Oct
b12_Oct b13_Oct b14_Oct);

NEW (c09_Oct c10_Oct c11_Oct
c12_Oct c13_Oct c14_Oct);

NEW (d09_Oct d10_Oct d11_Oct
d12_Oct d13_Oct d14_Oct);

NEW (e09_Oct e10_Oct e11_Oct
e12_Oct e13_Oct e14_Oct);

NEW (f09_Oct f10_Oct f11_Oct
f12_Oct f13_Oct f14_Oct);

NEW (g09_Oct g10_Oct g11_Oct
g12_Oct g13_Oct g14_Oct);

NEW (h09_Oct h10_Oct h11_Oct
h12_Oct h13_Oct h14_Oct);

NEW (i09_Oct i10_Oct i11_Oct
i12_Oct i13_Oct i14_Oct);
```

```
NEW (j09_Oct j10_Oct j11_Oct
j12_Oct j13_Oct j14_Oct);

NEW (k09_Oct k10_Oct k11_Oct
k12_Oct k13_Oct k14_Oct);

NEW (kjDIFF jiDiff ihDIFF hgDIFF
gfDIFF feDIFF edDIFF dcDIFF
cbDIFF baDIFF);

!MODEL 1986-1990
k09_Oct = ik + (sk * 0.25);
k10_Oct = ik + (sk * 1.25);
k11_Oct = ik + (sk * 2.25);
k12_Oct = ik + (sk * 3.25);
k13_Oct = ik + (sk * 4.25);
k14_Oct = ik + (sk * 5.25);

!MODEL 1981-1985:
j09_Oct = ij + (sj * 0.25);
j10_Oct = ij + (sj * 1.25);
j11_Oct = ij + (sj * 2.25);
j12_Oct = ij + (sj * 3.25);
j13_Oct = ij + (sj * 4.25);
j14_Oct = ij + (sj * 5.25);

!MODEL 1976-1980
i09_Oct = ii + (si * 0.25);
i10_Oct = ii + (si * 1.25);
i11_Oct = ii + (si * 2.25);
i12_Oct = ii + (si * 3.25);
i13_Oct = ii + (si * 4.25);
i14_Oct = ii + (si * 5.25);

!MODEL 1971-1975
h09_Oct = ih + (sh * 0.25);
h10_Oct = ih + (sh * 1.25);
h11_Oct = ih + (sh * 2.25);
h12_Oct = ih + (sh * 3.25);
h13_Oct = ih + (sh * 4.25);
h14_Oct = ih + (sh * 5.25);

!MODEL 1966-1970
g09_Oct = ig + (sg * 0.25);
g10_Oct = ig + (sg * 1.25);
g11_Oct = ig + (sg * 2.25);
g12_Oct = ig + (sg * 3.25);
g13_Oct = ig + (sg * 4.25);
g14_Oct = ig + (sg * 5.25);
```

```
!MODEL 1961-1965:
f09_Oct = if + (sf * 0.25);
f10_Oct = if + (sf * 1.25);
f11_Oct = if + (sf * 2.25);
f12_Oct = if + (sf * 3.25);
f13_Oct = if + (sf * 4.25);
f14_Oct = if + (sf * 5.25);

!MODEL 1956-1960
e09_Oct = ie + (se * 0.25);
e10_Oct = ie + (se * 1.25);
e11_Oct = ie + (se * 2.25);
e12_Oct = ie + (se * 3.25);
e13_Oct = ie + (se * 4.25);
e14_Oct = ie + (se * 5.25);

!MODEL 1951-1955:
d09_Oct = id + (sd * 0.25);
d10_Oct = id + (sd * 1.25);
d11_Oct = id + (sd * 2.25);
d12_Oct = id + (sd * 3.25);
d13_Oct = id + (sd * 4.25);
d14_Oct = id + (sd * 5.25);

!MODEL 1946-1950
c09_Oct = ic + (sc * 0.25);
c10_Oct = ic + (sc * 1.25);
c11_Oct = ic + (sc * 2.25);
c12_Oct = ic + (sc * 3.25);
c13_Oct = ic + (sc * 4.25);
c14_Oct = ic + (sc * 5.25);

!MODEL 1941-1945
b09_Oct = ib + (sb * 0.25);
b10_Oct = ib + (sb * 1.25);
b11_Oct = ib + (sb * 2.25);
b12_Oct = ib + (sb * 3.25);
b13_Oct = ib + (sb * 4.25);
b14_Oct = ib + (sb * 5.25);

!MODEL 1936-1940
a09_Oct = ia + (sa * 0.25);
a10_Oct = ia + (sa * 1.25);
a11_Oct = ia + (sa * 2.25);
a12_Oct = ia + (sa * 3.25);
a13_Oct = ia + (sa * 4.25);
a14_Oct = ia + (sa * 5.25);

!Difference tests for end and start of sequential cohort points
```

```
kjDIFF = k14_Oct - j09_Oct;
jiDiff = j14_Oct - i09_Oct;
ihDIFF = i14_Oct - h09_Oct;
hgDIFF = h14_Oct - g09_Oct;
gfDIFF = g14_Oct - f09_Oct;
feDIFF = f14_Oct - e09_Oct;
edDIFF = e14_Oct - d09_Oct;
dcDIFF = d14_Oct - c09_Oct;
cbDIFF = c14_Oct - b09_Oct;
baDIFF = b14_Oct - a09_Oct;

OUTPUT:

SAMPSTAT;
CINTERVAL;
```

Supplementary File 4 – Tests of Measurement Invariance across Age

This supplementary material presents tests of measurement invariance for the six personality dimensions – Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness to Experience, and Honesty-Humility – each estimated by their corresponding four items as described in the main text. Please note these tests are a direct extension of the tests of measurement invariance presented in Milojev and Sibley (2014). Each of the six dimensions was tested separately. For each dimension we tested temporal measurement invariance – imposing configural, metric, and finally scalar invariance for the given trait at time 1 and at time 6. Subsequently, we tested measurement models specifying strict scalar temporal invariance across age – that is, across the 5-year age-cohorts as described in the main text. The respective model fit indices are presented in Table A.4.1 below. Comparative tests including the Chi Square difference tests and the ΔCFA (see Cheung & Rensvold, 2002) are presented in Tables A.4.2 and A.4.3.As can be seen in Table A.4.1, the measurement specifications show reasonable model fit for strict scalar invariance over time and across age-groups for Extraversion, Conscientiousness, and Neuroticism. However, fit indices for Agreeableness, Openness to Experience and Honesty-Humility indicate that the measurement of these dimensions of personality may differ somewhat for different age cohorts, particularly with regard to the test of scalar invariance. This concern is more clearly seen in Tables A.4.2 and A.4.3, with significant differences across all comparisons. Indeed, as can be seen in Table A.4.3, tests of scalar invariance across age groups indicate violations of invariance with ΔCFA > .01 for all six traits. These tests highlight the high demands such measurement specifications place on the data. More importantly, they indicate the need for caution in the interpretations of the models presented in this study.

Table A.4.1. Fit indices for tests of Measurement Invariance.

Model	χ^2	*df*	RMSEA	CFI	TLI	SRMR
Extraversion						
1. Configural over time	559.568	*15*	.060	.953	.912	.035
2. Metric over time	583.970	*19*	.054	.951	.928	.037
3. Scalar over time	638.083	*22*	.052	.946	.932	.037
4. Between age-groups configural	734.506	*176*	.060	.950	.913	.044
5. Between age-groups metric	793.082	*229*	.053	.950	.932	.051
6. Between age-groups scalar	1063.602	*303*	.054	.932	.931	.062
Agreeableness						
1. Configural over time	1346.613	*15*	.093	.842	.706	.060
2. Metric over time	1410.013	*19*	.085	.835	.757	.100
3. Scalar over time	1479.213	*22*	.080	.828	.780	.100
4. Between age-groups configural	1605.388	*176*	.096	.836	.714	.085
5. Between age-groups metric	1606.957	*229*	.083	.842	.788	.082
6. Between age-groups scalar	1811.640	*303*	.075	.827	.824	.083
Conscientiousness						
1. Configural over time	143.996	*15*	.029	.981	.965	.021
2. Metric over time	159.127	*19*	.027	.980	.970	.026
3. Scalar over time	195.789	*22*	.028	.975	.968	.029
4. Between age-groups configural	309.775	*176*	.029	.981	.967	.041
5. Between age-groups metric	395.999	*229*	.029	.977	.969	.049
6. Between age-groups scalar	641.987	*303*	.036	.953	.952	.061
Neuroticism						
1. Configural over time	239.589	*15*	.038	.975	.953	.025
2. Metric over time	266.217	*19*	.036	.972	.959	.031
3. Scalar over time	305.134	*22*	.035	.968	.959	.032
4. Between age-groups configural	419.613	*176*	.040	.971	.950	.040
5. Between age-groups metric	471.344	*229*	.035	.971	.961	.044
6. Between age-groups scalar	845.832	*303*	.045	.936	.935	.066
Openness to Experience						
1. Configural over time	1644.972	*15*	.103	.828	.680	.081
2. Metric over time	2088.742	*19*	.103	.782	.679	.120
3. Scalar over time	2166.091	*22*	.098	.774	.713	.121
4. Between age-groups configural	2030.872	*176*	.110	.806	.660	.102
5. Between age-groups metric	2072.223	*229*	.096	.807	.740	.102
6. Between age-groups scalar	2355.417	*303*	.088	.785	.782	.109
Honesty-Humility						
1. Configural over time	1702.329	*15*	.105	.874	.766	.062
2. Metric over time	1908.953	*19*	.099	.859	.793	.101
3. Scalar over time	2247.925	*22*	.099	.834	.789	.117
4. Between age-groups configural	2086.047	*176*	.111	.852	.740	.102
5. Between age-groups metric	1968.334	*229*	.093	.865	.818	.081
6. Between age-groups scalar	2681.826	*303*	.095	.815	.812	.120

Notes. Strict Scalar Factorial invariance over time tested between age groups.

Table A.4.2. Chi Square difference tests (with the MLR corrections applied) for the test of measurement invariance across time.

	$\chi^2_{(df)}$	*p*	*ΔCFI*
Extraversion			
Metric-Configural	12.59 (4)	.013	-.002
Scalar-Metric	46.30 (3)	<.001	-.005
Agreeableness			
Metric-Configural	97.82 (4)	<.001	-.007
Scalar-Metric	27.69 (3)	<.001	-.007
Conscientiousness			
Metric-Configural	11.5 (4)	.021	-.001
Scalar-Metric	90.34 (3)	<.001	-.005
Neuroticism			
Metric-Configural	24.18 (4)	<.001	-.003
Scalar-Metric	38.39 (3)	<.001	-.004
Openness to Experience			
Metric-Configural	482.39 (4)	<.001	-.046
Scalar-Metric	29.74 (3)	<.001	-.008
Honesty-Humility			
Metric-Configural	231.70 (4)	<.001	-.015
Scalar-Metric	356.49 (3)	<.001	-.025

Table A.4.3. Chi Square difference tests (with the MLR corrections applied) for the test of measurement invariance across age groups.

	$\chi^2_{(df)}$	*p*	*ΔCFI*
Extraversion			
Metric-Configural	59.449 (53)	.252	.000
Scalar-Metric	272.550 (74)	<.001	-.018
Agreeableness			
Metric-Configural	82.94 (53)	.005	.006
Scalar-Metric	153.18 (74)	<.001	-.015
Conscientiousness			
Metric-Configural	85.91 (53)	.003	-.004
Scalar-Metric	266.25 (74)	<.001	-.024
Neuroticism			
Metric-Configural	54.45 (53)	.419	.000
Scalar-Metric	398.693 (74)	<.001	-.035
Openness to Experience			
Metric-Configural	108.360 (53)	<.001	.001
Scalar-Metric	216.805 (74)	<.001	-.022
Honesty-Humility			
Metric-Configural	698.01 (53)	<.001	.013
Scalar-Metric	724.39 (74)	<.001	-.050

Notes. Strict Scalar Factorial invariance over time tested between age groups.

Supplementary File 5 – Bivariate correlations and descriptive statistics

Table A.5.1. Descriptive statistics and bivariate correlations for the Big Six personality (Extraversion, Agreeableness, Conscientiousness, Openness to Experience, Neuroticism, and Honesty-Humility) traits across the six repeated assessments.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
1 T1.EXTRAVERSION																																				
2 T2.EXTRAVERSION	**.75***																																			
3 T3.EXTRAVERSION	**.74***	**.77***																																		
4 T4.EXTRAVERSION	**.73***	**.75***	**.78***																																	
5 T5.EXTRAVERSION	**.71***	**.75***	**.76***	**.79***																																
6 T6.EXTRAVERSION	**.71***	**.74***	**.76***	**.77***	**.80***																															
7 T1.AGREEABLENESS	.21*	.17*	.17*	.16*	.16*	.15*																														
8 T2.AGREEABLENESS	.17*	.19*	.15*	.15*	.14*	.13*	**.61***																													
9 T3.AGREEABLENESS	.16*	.17*	.22*	.19*	.18*	.18*	**.59***	**.63***																												
10 T4.AGREEABLENESS	.16*	.17*	.18*	.20*	.18*	.18*	**.58***	**.62***	**.63***																											
11 T5.AGREEABLENESS	.16*	.16*	.16*	.17*	.18*	.17*	**.59***	**.63***	**.63***	**.65***																										
12 T6.AGREEABLENESS	.16*	.16*	.16*	.17*	.17*	.19*	**.59***	**.63***	**.62***	**.64***	**.68***																									
13 T1.CONSCIENTIOUSNESS	.01	.02	.02	.34*	.02	.03	.15*	.11*	.12*	.11*	.12*	.12*																								
14 T2.CONSCIENTIOUSNESS	.03*	.05*	.04*	.04*	.04*	.05*	.11*	.17*	.13*	.12*	.12*	.13*	**.69***																							
15 T3.CONSCIENTIOUSNESS	.03	.05*	.08*	.05*	.04*	.06*	.11*	.12*	.15*	.10*	.10*	.09*	**.68***	**.69***																						
16 T4.CONSCIENTIOUSNESS	.03*	.04*	.05*	.05*	.06*	.06*	.10*	.09*	.10*	.15*	.10*	.08*	**.67***	**.70***	**.70***																					
17 T5.CONSCIENTIOUSNESS	.03	.04*	.03*	.05*	.06*	.06*	.10*	.11*	.10*	.09*	.14*	.09*	**.65***	**.69***	**.70***	**.72***																				
18 T6.CONSCIENTIOUSNESS	.04*	.05*	.03*	.05*	.04*	.06*	.09*	.11*	.08*	.10*	.10*	.14*	**.64***	**.67***	**.68***	**.71***	**.74***																			
19 T1.NEUROTICISM	-.09*	-.07*	-.07*	-.07*	-.07*	-.06*	-.03	-.03	-.03	-.02*	-.01	.00	-.11*	-.11*	-.11*	-.11*	-.13*	-.13*																		
20 T2.NEUROTICISM	.08*	-.10*	-.09*	-.06*	-.09*	-.07*	-.03	-.05*	-.03*	-.02	-.02	-.01	-.11*	-.16*	-.14*	-.15*	-.17*	-.17*	**.89***																	
21 T3.NEUROTICISM	-.10*	-.11	-.02*	-.14*	-.15*	-.16*	-.02	-.04*	-.08*	-.02	-.04*	-.02	-.13*	-.15*	-.20*	-.16*	-.16*	-.17*	**.69***	**.69***																
22 T4.NEUROTICISM	-.09*	-.09*	-.14*	-.13*	-.13*	-.12*	-.02	-.04*	-.02	-.04*	-.04*	-.02	-.13*	-.16*	-.16*	-.18*	-.16*	-.16*	**.63***	**.67***	**.70***															
23 T5.NEUROTICISM	-.09*	-.09*	-.14*	-.12*	-.16*	-.14*	-.03	-.04*	-.02	-.02	-.06*	-.02	-.12*	-.14*	-.16*	-.15*	-.19*	-.16*	**.64***	**.66***	**.79***	**.72***														
24 T6.NEUROTICISM	-.09*	-.09*	-.13*	.11*	-.13*	.16*	-.02	-.04*	-.02	-.02	-.04*	-.05*	-.13*	-.15*	-.15*	-.15*	-.17*	.20*	**.62***	**.65***	**.67***	**.69***	**.72***													
25 T1.OPENNESS	.25*	.21*	-.19*	.21*	.19*	.16*	.23*	.19*	.16*	.17*	.19*	.18*	.01	.00	-.01	.00	.00	-.01	-.01	-.01	-.03*	-.03*	-.02	-.02												
26 T2.OPENNESS	.21*	.23*	.21*	.20*	.19*	.17*	.19*	.25*	.20*	.18*	.22*	.20*	-.02	.04*	.00	.01	.01	-.01	-.01	-.02	-.04*	.00	-.02	-.01	**.66***											
27 T3.OPENNESS	.21*	.21*	.21*	.19*	.19*	.18*	.16*	.20*	.26*	.18*	.17*	.18*	-.02	.03	-.02	-.04*	-.16*	-.06*	-.02	-.03	-.07*	-.03*	-.02	-.02	**.65***	**.68***										
28 T4.OPENNESS	.20*	.20*	.19*	.22*	.19*	.18*	.16*	.17*	.21*	.24*	.19*	.18*	-.01	.02	-.05*	.00	-.14*	-.04*	-.01	-.01	-.04*	-.04*	-.03*	-.02	**.65***	**.66***	**.70***									
29 T5.OPENNESS	.21*	.20*	.19*	.20*	.22*	.19*	.18*	.19*	.19*	.19*	.25*	.19*	-.03	.00	-.05*	-.03*	-.01	-.03*	-.01	.00	-.03*	-.04*	-.06*	-.04*	**.64***	**.66***	**.70***	**.71***								
30 T6.OPENNESS	.19*	.18*	.17*	.18*	.18*	.19*	.15*	.15*	.18*	.18*	.18*	.23*	-.02	.01	-.06*	-.04*	-.04*	-.01	-.01	-.02	-.04*	-.04*	-.05*	-.05*	**.64***	**.67***	**.71***	**.71***	**.74***							
31 T1.HONESTY_HUMILITY	-.09*	-.09*	-.08*	-.10*	-.06*	-.10*	.16*	.15*	.15*	.15*	.17*	.17*	.08*	-.07*	.08*	.07*	.07*	.07*	-.19*	-.18*	-.17*	-.16*	-.20*	-.16*	.02	.01	.00*	.01	.01	.02					
32 T2.HONESTY_HUMILITY	-.12*	-.14*	-.13*	-.13*	-.10*	-.13*	.15*	.19*	.15*	.15*	.16*	.17*	.06*	.10*	.08*	.07*	.06*	.08*	-.15*	-.18*	-.16*	-.15*	-.16*	-.15*	.01	.02	.01	.01	.02	.03	**.72***				
33 T3.HONESTY_HUMILITY	-.10	-.11*	-.10*	-.10*	-.07*	-.07*	.16*	.17*	.20*	.17*	.18*	.18*	.07*	.07*	.09*	.07*	.08*	.07*	-.14*	-.14*	-.18*	-.15*	-.16*	-.15*	.02	.01	.05*	.03*	.02	.04*	**.71***	**.75***			
34 T4.HONESTY_HUMILITY	-.07*	-.09*	-.08*	-.10*	-.08*	-.08*	.16*	.18*	.18*	.20*	.18*	.19*	.05*	.06*	.07*	.09*	.07*	.07*	-.12*	-.14*	-.14*	-.17*	-.16*	-.15*	.02	.00	.05	.06*	.05*	.05*	**.68***	**.71***	**.74***		
35 T5.HONESTY_HUMILITY	-.08*	.09*	-.07*	-.07*	-.07*	-.06*	.18*	.18*	.17*	.17*	.21*	.20*	.06*	.05*	.06*	.04*	.08*	.06*	-.13*	-.15*	-.15*	-.15*	-.18*	-.16*	.03	.02	.04*	.06*	.08*	.07*	**.66***	**.69***	**.71***	**.74***	
36 T6.HONESTY_HUMILITY	-.06*	-.09*	-.06*	-.08*	-.07*	-.07*	.18*	.18*	.18*	.18*	.20*	.23*	.06*	.05*	.07*	.07*	.07*	.09*	-.10*	-.14*	-.14*	-.15*	-.16*	-.17*	.03	-.01	.03*	.06*	.06*	.07*	**.68***	**.70***	**.71***	**.73***	**.76***
M	4.02	3.96	3.93	3.93	3.93	3.92	5.29	5.27	5.37	5.33	5.33	5.36	5.15	5.13	5.03	5.07	5.06	5.06	3.38	3.45	3.36	3.38	3.39	3.36	4.77	4.73	4.91	4.89	4.88	4.88	5.16	5.15	5.15	5.21	5.27	5.33
SD	1.15	1.12	1.16	1.15	1.14	1.15	0.97	0.96	0.95	0.94	0.94	0.94	1.06	1.01	1.02	1.01	1.01	1.01	1.1	1.09	1.12	1.1	1.1	1.1	1.11	1.09	1.11	1.09	1.09	1.1	1.3	1.28	1.23	1.23	1.22	1.2

* p < .05; Test-retest correlations presented in bold font.

Appendix B

This appendix presents the MPlus syntax developed for conducting the analyses presented in Study 1 (Milojev *et al*., 2013) and Study 2 (Milojev & Sibley, 2014)). The syntax presented here are adjusted for the focus on only the variables used in the analyses. The full syntax with the full list of variable names are available from the NZAVS website (http://www.psych.auckland.ac.nz/en/about/our-research/research-groups/new-zealand-attitudes-and-values-study/nzavs-information-for-researchers.html). Furthermore, the worked examples of the following syntax and the associated data files are available from the author's webpage (http://www.psych.auckland.ac.nz/en/about/our-research/research-groups/new-zealand-attitudes-and-values-study/Petar-Materials.html), or on contacting the author.

Please note that the sample syntax developed for the Single-Group and Multi-Group Cohort-Sequential Latent Growth Models used in Study 4 (Milojev & Sibley, 2016) are presented in the Appendix A.

Study 1
TITLE:

The New Zealand Attitudes and Values Study
Supporting Information for Research Collaboration
NZAVS Longitudinal Data

This is the base syntax file for importing and analysing anonymized longitudinal data from the New Zealand Attitudes and Values Study in Mplus. This is designed to complement the SPSS Base Dataset containing the same variables. This dataset contains the complete list of all variables in the cross-sectional datasets for each timepoint. For a full list of the variables and item labels contained in the study please visit the NZAVS website or contact me directly.

Updated February 2013

http://www.psych.auckland.ac.nz/uoa/NZAVS
Petar Milojev; p.milojev@auckland.ac.nz
Dr. Chris Sibley, c.sibley@auckland.ac.nz

==========

This is the base script for stability analyses of the Mini-IPIP6 reported in:

Milojev, P., Osborne, D., Greaves, L. M., Barlow, F. K., & Sibley, C. G. (2013). The Mini-IPIP6: Tiny yet highly stable markers of Big Six personality. Journal of Research in Personality.

Model syntax and annotation prepared by Petar Milojev.

==========

DATA:

FILE IS Milojev et al. DATA.txt;

VARIABLE:

IDVARIABLE IS subnum;
MISSING ARE ALL (-9999);

NAMES ARE
!Mini-IPIP6 Marker Items at Time 1 for
!Extraversion, Agreeableness, Conscientiousness,
!Openness to Experiance, and Honesty-Humility
PE1T1
PE2RT1
PE3RT1
PE4T1

PA1T1
PA2RT1
PA3T1
PA4RT1
PC1T1
PC2T1
PC3RT1
PC4RT1
PNU1T1
PNU2RT1
PNU3T1
PNU4RT1
PO1T1
PO2RT1
PO3RT1
PO4RT1
PNA1RT1
PNA2RT1
PH3RT1
PH4RT1
!Mini-IPIP6 Marker Items at Time 2 for
!Extraversion, Agreeableness, Conscientiousness,
!Openness to Experiance, and Honesty-Humility
PE1T2
PE2RT2
PE3RT2
PE4T2
PA1T2
PA2RT2
PA3T2
PA4RT2
PC1T2
PC2T2
PC3RT2
PC4RT2
PNU1T2
PNU2RT2
PNU3T2
PNU4RT2
PO1T2
PO2RT2
PO3RT2
PO4RT2
PNA1RT2
PNA2RT2
PH3RT2
PH4RT2
SUBNUM; !questionnaire number;

```
USEVARIABLE ARE

PE1T1 PE2rT1 PE3rT1 PE4T1
PA1T1 PA2rT1 PA3T1 PA4rT1
PC1T1 PC2T1 PC3rT1 PC4rT1
PNu1T1 PNu2rT1 PNu3T1 PNu4rT1
PO1T1 PO2rT1 PO3rT1 PO4rT1
PNa1rT1 PNa2rT1 PH3rT1 PH4rT1

PE1T2 PE2rT2 PE3rT2 PE4T2
PA1T2 PA2rT2 PA3T2 PA4rT2
PC1T2 PC2T2 PC3rT2 PC4rT2
PNu1T2 PNu2rT2 PNu3T2 PNu4rT2
PO1T2 PO2rT2 PO3rT2 PO4rT2
PNa1rT2 PNa2rT2 PH3rT2 PH4rT2;

!USEOBSERVATIONS ARE (PE1T1 > 0 AND PE1T2 > 0);

DEFINE:


ANALYSIS:
!Specifying Bayesian estimation methods
COVERAGE = .15; PROCESSORS = 4;
ESTIMATOR = BAYES; CHAINS IS 6;
FBITERATIONS IS 1000; THIN = 10;
ALGORITHM = GIBBS; STVALUES = UNPERTURBED;

MODEL:
!Measurement component estimating the six latent personality factors
!at Time 1
E1 BY PE4T1 PE1T1 PE2rT1 PE3rT1;
A1 BY PA1T1 PA2rT1 PA3T1 PA4rT1;
C1 BY PC1T1 PC2T1 PC3rT1 PC4rT1;
N1 BY PNu1T1 PNu2rT1 PNu3T1 PNu4rT1;
O1 BY PO4rT1 PO1T1 PO2rT1 PO3rT1;
H1 BY PH3rT1 PNa1rT1 PNa2rT1 PH4rT1;
!Measurement component estimating the six latent personality factors
!at Time 1
E2 BY PE4T2 PE1T2 PE2rT2 PE3rT2;
A2 BY PA1T2 PA2rT2 PA3T2 PA4rT2;
C2 BY PC1T2 PC2T2 PC3rT2 PC4rT2;
N2 BY PNu1T2 PNu2rT2 PNu3T2 PNu4rT2;
O2 BY PO4rT2 PO1T2 PO2rT2 PO3rT2;
H2 BY PH3rT2 PNa1rT2 PNa2rT2 PH4rT2;

!Structural component specifying the estimation
!of the stability coefficients
E2 ON E1 (e);
A2 ON A1 (a);
```

```
C2 ON C1 (c);
N2 ON N1 (n);
O2 ON O1 (o);
H2 ON H1 (h);

MODEL CONSTRAINT:

!Specifications of additional model constraints
!modelling difference calcualtions between stability coefficients
NEW (e_a e_c e_n e_o e_h
    a_c a_n a_o a_h
    c_n c_o c_h
    n_o n_h
    o_h);

e_a = e - a;
e_c = e - c;
e_n = e - n;
e_o = e - o;
e_h = e - h;
a_c = a - c;
a_n = a - n;
a_o = a - o;
a_h = a - h;
c_n = c - n;
c_o = c - o;
c_h = c - h;
n_o = n - o;
n_h = n - h;
o_h = o - h;
```

Study 2
TITLE:

The New Zealand Attitudes and Values Study
Supporting Information for Research Collaboration
NZAVS Full Longitudinal Data

This is the base syntax file for importing and analysing anonymized longitudinal data from the New Zealand Attitudes and Values Study in Mplus. This is designed to complement the SPSS Base Dataset containing the same variables. This dataset contains the complete list of all measures collected during each wave of the NZAVS. For a full list of the variables and item labels contained in the study please visit the NZAVS website or contact me directly.

Updated 08-March-2014

http://www.psych.auckland.ac.nz/uoa/NZAVS
Petar Milojev: p.milojev@auckland.ac.nz
Dr. Chris Sibley, c.sibley@auckland.ac.nz

==========

This script outlines the model developed by Milojev and Sibley (2014) testing cohort differences in the stability of latent Big-Six personality markers over a two-year period. This version of the syntax tests the model for extraversion. (The model can be adapted to test for cohort differences in the stability of different personality dimensions simply by interchanging the manifest indicators for extraversion with those of a different personality subscale. Note also that estimation settings are turned up well beyond what is actually needed, in line with our standard recommendations for analysis of NZAVS data.

Milojev, P., & Sibley, C. G. (in press). The stability of adult personality varies across age: evidence from a two-year longitudinal sample of adult New Zealanders. Journal of Research in Personality.

Syntax prepared by Petar Milojev

==========

DATA:

FILE IS Extraversion Trimmed.txt;
!LISTWISE IS ON; !Include this statement to apply listwise deletion.
!NGROUPS = XXX; !Number of groups in multi-group analysis

VARIABLE:

!IDVARIABLE IS subnum;
MISSING ARE ALL (-9999);
!TSCORES = TSCR01 TSCR02, etc !Defines TSCORES for time-varying growth models
!AUXILIARY = (m)xxx !Specifies missing data covariate in LGM
!AUXILIARY = GenT35 (DCAT)(DCON); !specifies auxiliary variables in a mixture model

```
!WEIGHT IS STATE_VAR;            !Sample weight correction factor (integer weight)
!CLUSTER = AU06T01;              !Specifies Census Area Units as a Level 2 unit in MRCM
!BETWEEN = STATE_VAR;            !Specifies BETWEEN (Level 2) variables in MRCM
!WITHIN = STATE_VAR;             !Specifies WITHIN (Level 1) variables in MRCM
!CENSORED ARE STATE_VAR(bi);     !Used to test a Censor-Inflated Regression Model
(bi)
!GROUPING IS STATE_VAR;          !(1 = Group1 2 = Group 2, etc) used in multi-group

NAMES ARE   !Item content and database keys for the variable list
        !are available on the NZAVS technical details webpage

PE1T1
PE2RT1
PE3RT1
PE4T1
PE1T3
PE2RT3
PE3RT3
PE4T3
GENT1
AGE
AGESQ;

!=========================================================

USEVARIABLE ARE
PE1T1 PE2rT1 PE3rT1 PE4T1
PE1T3 PE2rT3 PE3rT3 PE4T3
GenT1
Age AgeSq;! AgeCB;

ANALYSIS:
COVERAGE = .33;
ESTIMATOR = ML;
TYPE = RANDOM;
PROCESSORS IS 24; STARTS=24; STITERATIONS = 100;
ALGORITHM = INTEGRATION;
MODEL:
!Specifying the measurement model
!item loading constrained to equality across time
E1 BY PE1T1(1)
    PE2rT1(2)
    PE3rT1(3)
    PE4T1(4);
E2 BY PE1T3(1)
    PE2rT3(2)
    PE3rT3(3)
    PE4T3(4);
!Item varaincs constrained to equality across time
PE1T1(5);
```

```
PE2rT1(6);
PE3rT1(7);
PE4T1(8);
PE1T3(5);
PE2rT3(6);
PE3rT3(7);
PE4T3(8);
!Item intercepts constrained to equality across time
[PE1T1](9);
[PE2rT1](10);
[PE3rT1](11);
[PE4T1](12);
[PE1T3](9);
[PE2rT3](10);
[PE3rT3](11);
[PE4T3](12);
!Item residuals allowed to correlate
PE1T1 WITH PE1T3;
PE2rT1 WITH PE2rT3;
PE3rT1 WITH PE3rT3;
PE4T1 WITH PE4T3;
!Specifying random interaction effects
E1Gen | E1 XWITH GenT1;
E1Age | E1 XWITH Age;
E1Agesq | E1 XWITH AgeSq;
!E1AgeCB | E1 XWITH AgeCB;
!Structural model
E2 ON E1(a)
     GenT1(c)
     E1Gen(ac)
     Age(b)
     AgeSq(bb)
     !AgeCB(bbb)
     E1Age(ab)
     E1Agesq(abb);
     !E1AgeCB(abbb);

MODEL CONSTRAINT:

New (za zb zab
    zbb zabb
    !zbbb zabbb
    zac
    m
    A20 A21 A22 A23 A24 A25 A26 A27 A28 A29
    A30 A31 A32 A33 A34 A35 A36 A37 A38 A39
    A40 A41 A42 A43 A44 A45 A46 A47 A48 A49
    A50 A51 A52 A53 A54 A55 A56 A57 A58 A59
    A60 A61 A62 A63 A64 A65 A66 A67 A68 A69
    A70 A71 A72 A73 A74 A75 A76 A77 A78 A79
```

```
    A80

    s20 s21 s22 s23 s24 s25 s26 s27 s28 s29
    s30 s31 s32 s33 s34 s35 s36 s37 s38 s39
    s40 s41 s42 s43 s44 s45 s46 s47 s48 s49
    s50 s51 s52 s53 s54 s55 s56 s57 s58 s59
    s60 s61 s62 s63 s64 s65 s66 s67 s68 s69
    s70 s71 s72 s73 s74 s75 s76 s77 s78 s79
    s80 2080diff);
    !men women);

!Specifies standardisation of parameters:
za = a * (SQRT(.953)/SQRT(.868));
zb = b * (SQRT(226.280)/SQRT(.868));
zbb = bb * (SQRT(76179.251)/SQRT(.868));
!zbbb = bbb*
zab = ab * (SQRT(.953)*SQRT(226.280)/SQRT(.868));
zabb = abb * (SQRT(.953)*SQRT(76179.251)/SQRT(.868));
!zabbb = abbb *
zac = ac * (SQRT(.953)*SQRT(.236)/SQRT(.868));

m = 50.278;        !specifies mean used for centering age

!specifies condiitonal values of age
A20 = 0 - (m - 20)/SQRT(226.280);
A21 = 0 - (m - 21)/SQRT(226.280);
A22 = 0 - (m - 22)/SQRT(226.280);
A23 = 0 - (m - 23)/SQRT(226.280);
A24 = 0 - (m - 24)/SQRT(226.280);
A25 = 0 - (m - 25)/SQRT(226.280);
A26 = 0 - (m - 26)/SQRT(225.665);
A27 = 0 - (m - 27)/SQRT(225.665);
A28 = 0 - (m - 28)/SQRT(225.665);
A29 = 0 - (m - 29)/SQRT(225.665);
A30 = 0 - (m - 30)/SQRT(225.665);
A31 = 0 - (m - 31)/SQRT(225.665);
A32 = 0 - (m - 32)/SQRT(225.665);
A33 = 0 - (m - 33)/SQRT(225.665);
A34 = 0 - (m - 34)/SQRT(225.665);
A35 = 0 - (m - 35)/SQRT(225.665);
A36 = 0 - (m - 36)/SQRT(225.665);
A37 = 0 - (m - 37)/SQRT(225.665);
A38 = 0 - (m - 38)/SQRT(225.665);
A39 = 0 - (m - 39)/SQRT(225.665);
A40 = 0 - (m - 40)/SQRT(225.665);
A41 = 0 - (m - 41)/SQRT(225.665);
A42 = 0 - (m - 42)/SQRT(225.665);
A43 = 0 - (m - 43)/SQRT(225.665);
A44 = 0 - (m - 44)/SQRT(225.665);
A45 = 0 - (m - 45)/SQRT(225.665);
```

```
A46 = 0 - (m - 46)/SQRT(225.665);
A47 = 0 - (m - 47)/SQRT(225.665);
A48 = 0 - (m - 48)/SQRT(225.665);
A49 = 0 - (m - 49)/SQRT(225.665);
A50 = 0 - (m - 50)/SQRT(225.665);
A51 = 0 - (m - 51)/SQRT(225.665);
A52 = 0 - (m - 52)/SQRT(225.665);
A53 = 0 - (m - 53)/SQRT(225.665);
A54 = 0 - (m - 54)/SQRT(225.665);
A55 = 0 - (m - 55)/SQRT(225.665);
A56 = 0 - (m - 56)/SQRT(225.665);
A57 = 0 - (m - 57)/SQRT(225.665);
A58 = 0 - (m - 58)/SQRT(225.665);
A59 = 0 - (m - 59)/SQRT(225.665);
A60 = 0 - (m - 60)/SQRT(225.665);
A61 = 0 - (m - 61)/SQRT(225.665);
A62 = 0 - (m - 62)/SQRT(225.665);
A63 = 0 - (m - 63)/SQRT(225.665);
A64 = 0 - (m - 64)/SQRT(225.665);
A65 = 0 - (m - 65)/SQRT(225.665);
A66 = 0 - (m - 66)/SQRT(225.665);
A67 = 0 - (m - 67)/SQRT(225.665);
A68 = 0 - (m - 68)/SQRT(225.665);
A69 = 0 - (m - 69)/SQRT(225.665);
A70 = 0 - (m - 70)/SQRT(225.665);
A71 = 0 - (m - 71)/SQRT(225.665);
A72 = 0 - (m - 72)/SQRT(225.665);
A73 = 0 - (m - 73)/SQRT(225.665);
A74 = 0 - (m - 74)/SQRT(225.665);
A75 = 0 - (m - 75)/SQRT(225.665);
A76 = 0 - (m - 76)/SQRT(225.665);
A77 = 0 - (m - 77)/SQRT(225.665);
A78 = 0 - (m - 78)/SQRT(225.665);
A79 = 0 - (m - 79)/SQRT(225.665);
A80 = 0 - (m - 80)/SQRT(225.665);

!specifies conditinal stability estimates
s20 = za + (zab * A20) + (zabb * A20^2);
s21 = za + (zab * A21) + (zabb * A21^2);
s22 = za + (zab * A22) + (zabb * A22^2);
s23 = za + (zab * A23) + (zabb * A23^2);
s24 = za + (zab * A24) + (zabb * A24^2);
s25 = za + (zab * A25) + (zabb * A25^2);
s26 = za + (zab * A26) + (zabb * A26^2);
s27 = za + (zab * A27) + (zabb * A27^2);
s28 = za + (zab * A28) + (zabb * A28^2);
s29 = za + (zab * A29) + (zabb * A29^2);
s30 = za + (zab * A30) + (zabb * A30^2);
s31 = za + (zab * A31) + (zabb * A31^2);
s32 = za + (zab * A32) + (zabb * A32^2);
```

```
s33 = za + (zab * A33) + (zabb * A33^2);
s34 = za + (zab * A34) + (zabb * A34^2);
s35 = za + (zab * A35) + (zabb * A35^2);
s36 = za + (zab * A36) + (zabb * A36^2);
s37 = za + (zab * A37) + (zabb * A37^2);
s38 = za + (zab * A38) + (zabb * A38^2);
s39 = za + (zab * A39) + (zabb * A39^2);
s40 = za + (zab * A40) + (zabb * A40^2);
s41 = za + (zab * A41) + (zabb * A41^2);
s42 = za + (zab * A42) + (zabb * A42^2);
s43 = za + (zab * A43) + (zabb * A43^2);
s44 = za + (zab * A44) + (zabb * A44^2);
s45 = za + (zab * A45) + (zabb * A45^2);
s46 = za + (zab * A46) + (zabb * A46^2);
s47 = za + (zab * A47) + (zabb * A47^2);
s48 = za + (zab * A48) + (zabb * A48^2);
s49 = za + (zab * A49) + (zabb * A49^2);
s50 = za + (zab * A50) + (zabb * A50^2);
s51 = za + (zab * A51) + (zabb * A51^2);
s52 = za + (zab * A52) + (zabb * A52^2);
s53 = za + (zab * A53) + (zabb * A53^2);
s54 = za + (zab * A54) + (zabb * A54^2);
s55 = za + (zab * A55) + (zabb * A55^2);
s56 = za + (zab * A56) + (zabb * A56^2);
s57 = za + (zab * A57) + (zabb * A57^2);
s58 = za + (zab * A58) + (zabb * A58^2);
s59 = za + (zab * A59) + (zabb * A59^2);
s60 = za + (zab * A60) + (zabb * A60^2);
s61 = za + (zab * A61) + (zabb * A61^2);
s62 = za + (zab * A62) + (zabb * A62^2);
s63 = za + (zab * A63) + (zabb * A63^2);
s64 = za + (zab * A64) + (zabb * A64^2);
s65 = za + (zab * A65) + (zabb * A65^2);
s66 = za + (zab * A66) + (zabb * A66^2);
s67 = za + (zab * A67) + (zabb * A67^2);
s68 = za + (zab * A68) + (zabb * A68^2);
s69 = za + (zab * A69) + (zabb * A69^2);
s70 = za + (zab * A70) + (zabb * A70^2);
s71 = za + (zab * A71) + (zabb * A71^2);
s72 = za + (zab * A72) + (zabb * A72^2);
s73 = za + (zab * A73) + (zabb * A73^2);
s74 = za + (zab * A74) + (zabb * A74^2);
s75 = za + (zab * A75) + (zabb * A75^2);
s76 = za + (zab * A76) + (zabb * A76^2);
s77 = za + (zab * A77) + (zabb * A77^2);
s78 = za + (zab * A78) + (zabb * A78^2);
s79 = za + (zab * A79) + (zabb * A79^2);
s80 = za + (zab * A80) + (zabb * A80^2);

2080diff = s20 - s80;
```

```
!men = za * (zac * 1);
!women = za * (zac * 0);

OUTPUT:
CINTERVAL SAMPSTAT
```

Appendix C

This appendix presents the MPlus syntax implementing the Monte Carlo simulations used to estimated statistical power pertaining to the models presented in Study 1 (Milojev *et al*., 2013), Study 2 (Milojev & Sibley, 2014), and Study 4 (Milojev & Sibley, 2015). This syntax is also available from the author's webpage (http://www.psych.auckland.ac.nz/en/about/our-research/research-groups/new-zealand-attitudes-and-values-study/Petar-Materials.html), or on contacting the author.

Study 1
TITLE:

Monte Carlo Simulations used to estimate the statistical power for the SEM stability model of the Mini-IPIP6 supplementing the analysis in:
Milojev, P., Osborne, D., Greaves, L. M., Barlow, F. K., & Sibley, C. G. (2013). The Mini-IPIP6: Tiny yet highly stable markers of Big Six personality. Journal of Research in Personality.

```
MONTECARLO:

NAMES ARE
x1-x4 y1-y4;

NOBSERVATIONS = 4422; !Known Sample size
NREPS = 10000; !number of replications in simulation
SEED = 648135; !Random seed
CLASSES = C(1);
GENCLASSES = C(1);

ANALYSIS:

TYPE = MIXTURE;
ESTIMATOR = ML;

!Simulating data structure
MODEL MONTECARLO:

%OVERALl%

!Factor loadings
f1 BY x1-x4*.80;
f2 BY y1-y4*.80;

!Factor and item variances
f1@1; f2@1;
x1-x4*.4;
y1-y4*.4;

!Residual correlations across time
x1 WITH y1 *.10;
x2 WITH y2 *.10;
x3 WITH y3 *.10;
x4 WITH y4 *.10;

!The focal parameter
!varied in simulations between .05 to .95
f2 ON f1*.05;

!Specifying model
```

```
MODEL:

%OVERALl%

f1 BY x1-x4*.80;
f2 BY y1-y4*.80;

f1@1; f2@1;
x1-x4*.4;
y1-y4*.4;

x1 WITH y1 *.10;
x2 WITH y2 *.10;
x3 WITH y3 *.10;
x4 WITH y4 *.10;

f2 ON f1*.05;
```

Study 2
TITLE:

Monte Carlo Simulations used to estimate the statistical power for the SEM moderated stability model of the Mini-IPIP6 supplementing the analysis in:
Milojev, P., & Sibley, C. G. (2014). The stability of adult personality varies across age: evidence from a two-year longitudinal sample of adult New Zealanders. Journal of Research in Personality.
MONTECARLO:

```
NAMES ARE
x1-x4 y1-y4 z1;

NOBSERVATIONS = 3902; !Known sample size
NREPS = 1000; !Number of replications in simulation
SEED = 648135;

ANALYSIS:

TYPE = RANDOM; ESTIMATOR = ML;
ALGORITHM = INTEGRATION;
PROCESSORS = 4; STARTS = 2;

!Simulating data structure
MODEL MONTECARLO:

!Factor loadings
f1 BY x1-x4*.80;
f2 BY y1-y4*.80;

!Factor and item variances
f1@1; f2@1;
x1-x4*.4;
y1-y4*.4;

!Residual item correlations across time
x1 WITH y1 *.10;
x2 WITH y2 *.10;
x3 WITH y3 *.10;
x4 WITH y4 *.10;

!mean and variance of continuous covariate
[z1*0]; z1*1;

!simulating latent interaction
intxn | f1 XWITH z1;

!specifying regression structure
f2 ON f1*.70;
f2 ON z1*.05;
```

```
!Focal parameter varied between .001 to .10
f2 ON intxn*.001;

!Specifying model
MODEL:

f1 BY x1-x4*.80;
f2 BY y1-y4*.80;

f1@1; f2@1;
x1-x4*.4;
y1-y4*.4;

x1 WITH y1 *.10;
x2 WITH y2 *.10;
x3 WITH y3 *.10;
x4 WITH y4 *.10;

[z1*0]; z1*1;

intxn | f1 XWITH z1;

f2 ON f1*.70;
f2 ON z1*.05;
f2 ON intxn*.001;
```

Study 4
TITLE:

Monte Carlo simulations used to estimate power to detect a significant latent slope given the known sample sizes across the 5-year birth cohorts.
These simulations supplement the models presented in:

Milojev, P., & Sibley, C.G. (2016). Normative personality trait development in adulthood: A six-year cohort-sequential growth model. *Journal of Personality and Social Psychology, manuscript in press.*

MONTECARLO:

```
NAMES ARE y1-y5;
NOBSERVATIONS = 300;!Known sample size - varied between 30, 500, and 700
NREPS = 10000; !Number of replications in simulation
SEED = 754473;
CLASSES = C(1);
GENCLASSES = C(1);
```

ANALYSIS:

```
TYPE = MIXTURE;
ESTIMATOR = ML;

!simulating data structure
MODEL MONTECARLO:

%OVERALL%

!Linear latent growth model with 5 time points
i s | y1@0 y2@1 y3@2 y4@3 y5@4;

!manifest item means and variances
[y1-y5@0]; y1-y5*.5;

!mean and variance of the latent intercept
[i*4]; i*1.0;
!Focal parameter - mean of slope,
!varied between .005 to .10
[s*.005];
!variance of the latent slope
s*.10;

!covariance between latent slope and latent intercept
i WITH s*.10;


!specifying model
```

```
MODEL:

%OVERALL%

i s | y1@0 y2@1 y3@2 y4@3 y5@4;

[y1-y5@0]; y1-y5*.5;

[i*4]; i*1.0;
[s*.005];
s*.10;


i WITH s*.10;
```