Do Changes in Personality Predict Life Outcomes?

Amanda J. Wright and Joshua J. Jackson

Department of Psychological and Brain Sciences, Washington University in St. Louis

The Big Five personality traits predict many important life outcomes. These traits, although relatively stable, are also open to change across time. However, whether these changes likewise predict a wide range of life outcomes has yet to be rigorously tested. This has implications for the types of processes linking trait levels and changes with future outcomes: distal, cumulative processes versus more immediate, proximal processes, respectively. The present study used seven longitudinal data sets \((N = 81,980)\) to comprehensively examine the unique relationship that changes in the Big Five traits have with static levels and changes in numerous outcomes in the domains of health, education, career, finance, relationships, and civic engagement. Meta-analytic estimates were calculated and study-level variables were examined as potential moderators of these pooled effects. Results indicated that changes in personality traits are sometimes prospectively related to static outcomes—such as health status, degree attainment, unemployment, and volunteering—above and beyond associations due to static trait levels. Moreover, changes in personality more frequently predicted changes in these outcomes, with associations for new outcomes emerging as well (e.g., marriage, divorce). Across all meta-analytic models, the magnitude of effects for changes in traits was never larger than that of static levels and there were fewer change associations. Study-level moderators (e.g., average age, number of Big Five waves, internal consistency estimates) were rarely associated with effects. Our study suggests personality change can play a valuable role in one’s development and highlights that both cumulative and proximal processes matter for some trait-outcome associations.

Keywords: Big Five, personality change, personality prediction, life outcomes, meta-analysis

Supplemental materials: https://doi.org/10.1037/pspp0000472.supp

Defined as relatively enduring individual differences in how people feel, think, and behave (Roberts & Yoon, 2022), personality is a ubiquitous feature of human nature that influences countless aspects of life (Beck & Jackson, 2022). The Big Five, the most popular theoretical framework for studying personality, consists of five factors believed to capture the basic dimensions of personality (Goldberg, 1990). These traits—extraversion, agreeableness, conscientiousness, neuroticism, openness—predict an extensive number of life outcomes, with associations emerging across the lifespan (Hill et al., 2011; Turiano et al., 2015), across multiple assessment methods (Jackson et al., 2015; Smith et al., 2008), and decades in advance (Friedman et al., 2010). As such, personality traits constitute one of the best psychological predictors of broad life outcomes (Roberts et al., 2007).

A considerable amount of research has established that personality traits are not immutable and can change across time (e.g., Bleidorn et al., 2022; Roberts & Nickel, 2021; Wright & Jackson, 2023). The discovery of this mutable property, paired with their predictive utility, has led to an interest in purposefully eliciting changes in these traits through interventions (Allemand & Flückiger, 2022; Bleidorn et al., 2020; Stieger et al., 2020). These endeavors are implicitly guided by the belief that if the mean levels of the Big Five traits predict beneficial outcomes, then changes in these traits should lead to changes in behaviors related to life outcomes. As a result, changes in personality traits should also predict life outcomes. However, this assumption has yet to be systematically tested using all Big Five traits with a broad variety of outcomes (cf. Mroczek & Spiro, 2007, for an example with a specific outcome), resulting in a scarcity of tests for the predictive utility of personality trait change. Importantly, if changes in personality are not broadly associated with life outcomes, even though their respective trait levels are, this has both theoretical and practical implications. First, it not only suggests that the distinct processes linking static trait levels and changes in traits to outcomes vary in their influence, such that distal, cumulative effects (trait levels) matter more than more immediate, proximal processes (changes in traits), but also that changes in traits might not be associated with behavioral changes. Second, it further suggests that

Amanda J. Wright
https://orcid.org/0000-0001-8873-9405

The authors have no known conflicts of interest to disclose. All code to reproduce analyses, the data codebook, and supplemental materials are available at https://osf.io/hdms9/. All data are freely available via application and/or data use agreements at the links specified in each data set’s Participants subsection in the methods. Some findings from this article were presented at the 2022 Society for Personality and Social Psychology conference during a poster session.

Amanda J. Wright played a lead role in formal analysis, visualization, and writing—original draft and an equal role in conceptualization and writing—review and editing. Joshua J. Jackson played a lead role in supervision, a supporting role in formal analysis and writing—original draft, and an equal role in conceptualization and writing—review and editing. Correspondence concerning this article should be addressed to Amanda J. Wright, Department of Psychological and Brain Sciences, Washington University in St. Louis, 1 Brookings Drive, St. Louis, MO 63130, United States. Email: ajwright@wusl.edu
intervention efforts may be misguided, as any changes in traits being elicted would likely not lead to long-term changes in behaviors associated with outcomes.

The present study examines whether changes in the Big Five traits are prospectively associated with both static levels and changes in future life outcomes. In doing so, we integrate seven longitudinal data sets to serve as a comprehensive test of these effects with a large sample size ($N = 81,980$). By predicting both fixed levels and changes in distal outcomes across time, we are better able to determine not only if changes in personality are meaningfully related to life outcomes but also investigate the reasons why trait changes are consequential. In the present study, to thoroughly investigate if changes in personality predict life outcomes, special attention is given to model specification choices, such as accounting for static trait levels and minimizing intercept–slope covariance; multiple data sets are used to serve as an indicator of the robustness of effects; initial levels of outcomes are controlled for to better isolate and quantify effects attributable to personality change; and various study-level moderators are included to examine how change associations may vary across different data sets and under which conditions change may matter most.

Why Personality Change Should Be Associated With Life Outcomes

As the most widely used metric of measuring personality traits, there is an abundance of studies investigating what trait levels of the Big Five predict (Beck & Jackson, 2022; Ozer & Benet-Martínez, 2006; Soto, 2021; Wright & Jackson, 2022). For example, higher levels of extraversion predict subjective well-being (DeNeve & Cooper, 1998) and social status (Jensen-Campbell et al., 2002; Penner et al., 2003). Agreeableness predicts volunteering (Penner et al., 1995), religiosity (MacDonald, 2000), disease processes (Miller et al., 1996), and relationship satisfaction (Jensen-Campbell et al., 2002; Robins et al., 2002; Wright & Jackson, 2022). Conscientiousness, one of the traits most strongly related to various outcomes, predicts occupational performance and success (Anderson et al., 2001; Barrick et al., 2001), longevity (Jackson et al., 2015), engagement in health and risk behaviors (Bogg & Roberts, 2004; Friedman et al., 2014; Wright et al., 2022), and relationship quality (Solomon & Jackson, 2014). Neuroticism likewise has many strong associations: higher levels predict lower well-being (Smith & Sprio, 2002), psychopathology (Trull & Sher, 1994; Wright & Jackson, 2022), interpersonal problems (Karney & Bradbury, 1997), and poorer relationship quality (Donnellan et al., 2005). Lastly, openness predicts political views (Saucier, 2000; Van Hiel et al., 2004), occupational interests (Barrick et al., 2003; Larson et al., 2002), and creativity (Peterson & Seligman, 2004). These associations tend to replicate well and are robust against many background factors, and predict at similar levels across gender, socioeconomic status, and race/ethnicity (Beck & Jackson, 2022; Soto, 2021).

As demonstrated above, a lot of research has been dedicated to investigating average levels of the Big Five traits and their predictive utility. However, a considerable amount of research has also established the tendency of mean levels of the Big Five traits to change across time (Bleidorn et al., 2022; Roberts & DelVecchio, 2000; Roberts & Mroczek, 2008). These mean-level shifts often reflect normative maturation processes or the influence of external factors (Lodi-Smith & Roberts, 2007), such as life events (Denissen et al., 2019; Specht et al., 2011) or life experiences (Jayawickreme et al., 2021). Furthermore, there are individual differences in the changes in these mean levels. These individual differences reflect unique changes in a person’s levels of a trait that differ from the normative changes observed at the population level. For example, although most people tend to decline on neuroticism as they age, others increase in neuroticism while some do not change at all (Mroczek & Spiro, 2003).

Given that personality trait levels predict many outcomes and that they change over time, then presumably the mechanisms that link personality to life outcomes (i.e., behavioral/state expressions of personality; Wrzus & Roberts, 2017) will also change. As a result, changes in traits should yield changes in trait-relevant processes. For example, if conscientiousness is positively associated with physical health due to cumulative, long-term effects of having a healthy lifestyle (Bogg & Roberts, 2004), then increases in conscientiousness should correspond with more frequent engagement in health-promoting behaviors (Takahashi et al., 2013). This increased engagement in health behaviors, resulting from the increases in conscientiousness, should then similarly be associated with better health—specifically, health aspects that can be impacted over relatively shorter periods of time compared to decades-long processes. That is, changes in personality traits, and thus changes in the behaviors associated with traits, reflect more proximal processes by virtue that their effect has occurred over a shorter time period. In contrast, trait levels of personality reflect cumulative effects that can take decades to materialize.

However, one alternative hypothesis is that changes in personality traits do not reflect meaningful changes in observable behaviors (Oltmanns et al., 2020). For instance, measured changes in personality traits may instead reflect changes in thoughts or feelings (e.g., internal self-perceptions) that do not lead to nor co-occur with changes in behavior. Additionally, changes in personality may be an artifact of changes in measurement. This can result from changes in the content or meaning of items/indicators used to assess personality, shifting personality structures (e.g., structural changes that occur with age; Beck et al., 2023), or differences in the reference group people use to evaluate their own traits (Crédé et al., 2010; Lenhausen et al., 2022). Unfortunately, there are few tests of whether changes in traits reflect more than just changes in internalized thoughts or self-perceptions, as personality change is usually operationalized through a single method. While there is some evidence that self-reported changes in personality are observable and detected by external sources (Oltmanns et al., 2020; Stieger et al., 2020)—thus validating that changes in trait measures likely reflect changes in behavioral processes—additional tests are needed. If changes in personality traits are associated with life outcomes, particularly changes in life outcomes, it would further provide evidence that trait changes reflect meaningful changes in personality-relevant behavioral processes.

Evidence of Personality Change Prediction

At first blush, a number of studies identify associations between changes in personality traits and some outcomes. Many of these focus on static levels and changes in health-related outcomes. For static outcomes, changes in extraversion, agreeableness, and conscientiousness are associated with self-rated health (Turiano, Pitzer, et al., 2012); changes in conscientiousness and self-control
(related to conscientiousness) are associated with health limitations and knowledge (Richmond-Rakerd et al., 2021; Turiano, Pitzer, et al., 2012); changes in neuroticism, agency (related to extraversion and openness; Entringer et al., 2022), and hostility (related to agreeableness) are associated with markers of physical health (e.g., adiposity, high blood pressure; Human et al., 2013; Siegler et al., 2003); and changes in hostility are associated with substance use, exercise, and dietary habits (Hampson et al., 2010; Siegler et al., 2003). Moreover, changes in neuroticism and conscientiousness predict mortality (Martin et al., 2007; Mrozcek & Spiro, 2007).

For changes in health-related outcomes, changes in all Big Five traits have been found to be associated with changes in some self-rated general or physical health variable, though associations with conscientiousness emerge most frequently (Human et al., 2013; Letzring et al., 2014; Magee, Heaven, & Miller, 2013; Siegler et al., 2003; Takahashi et al., 2013). Similarly, changes in all Big Five traits, particularly neuroticism, extraversion, and conscientiousness, are associated with mental health or well-being-related outcomes (Chow & Roberts, 2014; Hounkpatin et al., 2018; Human et al., 2013; Kandler et al., 2015; Magee et al., 2013; Magee, Millar, & Heaven, 2013; Soto, 2015). Changes in markers of physical health (e.g., body mass index [BMI], disease burden, biomarkers) are associated with changes in self-control, neuroticism (impulsivity facet), extraversion (positive emotional autonomy facet), and openness (Richmond-Rakerd et al., 2021; Sutin et al., 2011, 2013).

Then, changes in preventative health behaviors such as physical activity are associated with changes in all Big Five traits (Jokela et al., 2018; Takahashi et al., 2013). Associations with changes in risk behaviors such as substance use also emerge, particularly for neuroticism, conscientiousness, and related traits (e.g., impulsivity), but they are slightly less consistent across studies (Allen et al., 2015; Jokela et al., 2018; Littlefield et al., 2009, 2012; Turiano, Whiteman, et al., 2012).

A number of studies have examined career-related outcomes as well. For static career outcomes, job satisfaction is associated with changes in extraversion, self-control, and neuroticism (Allemand et al., 2019; Converse et al., 2018; Hoff et al., 2021; Roberts et al., 2003), and sometimes agreeableness (Hoff et al., 2021). Similarly, general career satisfaction is associated with changes in extraversion, conscientiousness, and sometimes neuroticism (Hoff et al., 2021). Work engagement and involvement are associated with changes in self-control and extraversion, respectively (Allemand et al., 2019; Roberts et al., 2003). Then, occupational prestige is associated with changes in agreeableness (Hoff et al., 2021) and occupational attainment with changes in extraversion and neuroticism (Roberts et al., 2003). For changes in career outcomes, changes in job satisfaction are associated with changes in neuroticism and extraversion (Scollon & Diener, 2006), much like static job satisfaction is. Additionally, changes in work-life perception are associated with changes in hostility (Siegler et al., 2003).

Next, for static financial outcomes, changes in neuroticism and, somewhat inconsistently, extraversion, conscientiousness/self-control, and hostility are associated with income (Converse et al., 2018; Hoff et al., 2021; Siegler et al., 2003). Changes in self-control are associated with financial security, credit scores, and financial problems (Richmond-Rakerd et al., 2021; Roberts et al., 2003). Changes in neuroticism are also associated with financial security (Roberts et al., 2003). Fewer studies have examined changes in financial outcomes, but changes in hostility are associated with changes in economic-life perception (Siegler et al., 2003).

For education outcomes, there are fewer examples in general, and changes in these outcomes are often not explicitly analyzed. For static outcomes, though, one study found that educational attainment is associated with changes in self-control (Converse et al., 2018). Although, another study with two samples found that increases in educational attainment, whereas conscientiousness (related to self-control) had no associations (Hoff et al., 2021). Moreover, these findings did not replicate in one sample. For educational achievement (e.g., grade point averages), one study found that this is associated with changes in conscientiousness (Noffle & Robins, 2007).

As for other domains, such as relationship, family, social, or civic engagement, studies have examined both static outcomes and changes in outcomes. For static outcomes, changes in self-control are associated with relationship satisfaction, conflict, and parenting satisfaction (Allemand et al., 2019; Converse et al., 2018). Changes in self-control are also associated with social support and loneliness (Converse et al., 2018; Richmond-Rakerd et al., 2021), and changes in hostility are associated with social isolation (Siegler et al., 2003). Then, for changes in relationship outcomes, changes in neuroticism are associated with changes in relationship satisfaction, closeness, insecurity, and conflict (Lavner et al., 2018; Lehntart & Neyer, 2006; Mund & Neyer, 2014; O’Meara & South, 2019; Scollon & Diener, 2006). Changes in the other Big Five traits have similarly sometimes been associated with changes in relationship satisfaction (Lavner et al., 2018; Lehntart & Neyer, 2006; O’Meara & South, 2019; Scollon & Diener, 2006). Moreover, changes in agreeableness and conscientiousness are associated with changes in engagement with relationships and children, respectively (Lodi-Smith & Roberts, 2012). Then, changes in neuroticism predict future changes in relationship importance and changes in agreeableness predict future changes in relationship importance, insecurity, closeness, contact, and conflict (Mund & Neyer, 2014). Changes in social support are associated with changes in all Big Five traits (Allemand et al., 2015). Last, changes in overall civic engagement are associated with changes in agreeableness and conscientiousness (Lodi-Smith & Roberts, 2012).

Potential Challenges in Personality Change Prediction

In sum, the previous findings lend support for changes in personality traits being associated with numerous life outcomes, similar to static levels of personality. However, despite the number of papers finding that changes in traits are associated with static levels and changes in some outcomes, there are reasons to believe

---

1 For outcomes that are often contingent upon age, cumulative in nature, and/or occur relatively infrequently, it could be argued that changes in these outcomes are inherently being examined if no participants have yet experienced the event(s) at the measurement occasion that is the reference point by which one’s (future) amount of personality change is determined by. This is the case for many studies that first assess personality in childhood or adolescence. Outcomes, such as educational attainment, mortality, employment, marriage, divorce, and so forth, likely fall in this category. If some individuals in a sample have already experienced these events prior to that measurement occasion, though, then the values for these outcomes at that time point must be controlled for in order to accurately conclude that any associations due to changes in traits actually reflect personality processes that are independent of any mechanisms rather attributable to trait levels.
these findings overstate or do not accurately capture the associations between personality change and outcomes. First, many studies find inconsistent results. This includes results that both fail to replicate and those that are contradictory across studies. For example, the changes in traits that are associated with educational attainment (e.g., Converse et al., 2018; Hoff et al., 2021), substance use (e.g., Jokela et al., 2018; Littlefield et al., 2009; Turiano, Whitteman, et al., 2012), BMI (e.g., Jokela et al., 2018; Sutin et al., 2011), and relationship satisfaction (e.g., Lehnart & Neyer, 2006; O’Meara & South, 2019; Scollon & Diener, 2006) vary across studies. Furthermore, some associations do not even replicate within the same study (e.g., Hoff et al., 2021). A single investigation that can estimate an average effect using multiple samples would provide an indicator of which effects are robust.

Second, these inconsistent effects may arise from important study- or data set-level moderators, such as average age, number of personality assessments, years between assessments (thus likely affecting the amount of detectable change), and properties of the trait measures. For example, on average, childhood decreases in agreeableness predict alcohol, cigarette, and marijuana use in adolescence (Hampson et al., 2010), but similar changes in agreeableness in college students fail to predict future smoking or drinking in midlife (Siegl et al., 2003). Examining when change-outcome associations emerge across the lifespan can inform when these changes in personality may matter most or under which conditions they are most detectable. For example, it could be the case that, at some point in the lifespan, the cumulative effects stemming from one’s static levels of personality traits are harder to negate or contribute additional influence to—that is, proximal processes may become less influential. This could potentially happen in middle age when personality is most stable (Bleidorn et al., 2022) and thus fewer changes occur. Levels of personality thus continue exerting their normal influence on outcomes and do so without a great degree of proximal processes occurring. Alternatively, change-outcome associations could emerge less frequently as people age, as the cumulative effects of static levels continue to strengthen and are thus harder to oppose or allow additional influences. Directly testing if the average age of a sample is associated with change associations can perhaps shed light on why inconsistent results have emerged in past studies and theoretically inform when change may matter most.

Third, many of these studies examined personality change across only two waves (Allemand et al., 2015, 2023; Allen et al., 2015; Chow & Roberts, 2014; Human et al., 2013; Letzring et al., 2014; Lodi-Smith & Roberts, 2012; Magee, Heaven, & Miller, 2013; Magee, Miller, & Heaven, 2013; Martin et al., 2007; Noffle & Robins, 2007; Richmond-Rakerd et al., 2021; Roberts et al., 2003; Siegl et al., 2003; Takahashi et al., 2013; Turiano, Pitzer, et al., 2012; Turiano, Whitteman, et al., 2012). As a result, the measurement of change in these studies is likely unreliable or covaries with occasion-specific error. Three or more assessments of personality help distinguish error from true change to provide a more reliable assessment of personality trait change. The number of personality assessments will also affect the degree to which change can be reliably captured and the length of time between assessments will influence the amount of change that can occur (Hopwood et al., 2022). For design characteristics such as these, having too little or too much, can unduly impact the conclusions one draws from analyses (Hopwood et al., 2022). Ideally, one would be able to test if these factors influence associations between changes in traits and outcomes by having data sets that vary in these characteristics.

Fourth, static levels of personality traits are not always controlled for, which can result in associations of outcomes with personality “change” being driven by the covariation between change and mean levels. When static trait levels are controlled for associations due to change often weaken (e.g., Noffle & Robins, 2007) or associations due to level are largely all that emerge (e.g., Converse et al., 2018; Jokela et al., 2018). Relatedly, setting the intercept at the initial time point can result in the estimated associations simply quantifying the initial status covariances between variables as opposed to the desired effects of Big Five trait changes (Klimstra et al., 2013). Thus, this should be avoided when modeling change for the purpose of examining its predictive utility.

Fifth, some studies used some form of a cross-lagged panel model (e.g., Hounkpatin et al., 2018; Kandler et al., 2015; Lehnart & Neyer, 2006; Scollon & Diener, 2006; Soto, 2015). Similarly, many studies used dual or bivariate latent change score models (e.g., Hounkpatin et al., 2018; Littlefield et al., 2012; Lodi-Smith & Roberts, 2012; Magee, Heaven, & Miller, 2013; Magee, Miller, & Heaven, 2013; Mund & Neyer, 2014; Takahashi et al., 2013) and, when two latent change models are combined, these lead to a cross-lagged panel model (Mund & Neyer, 2014). These types of models are problematic for reasons such as a high likelihood of finding spurious associations, modeling two-wave residualized change, and a failure to separate variance at different levels (Lucas, 2023).

Sixth, in all but four of the reviewed studies (Allemand et al., 2023; Martin et al., 2007; Mroczek & Spiro, 2007; Mund & Neyer, 2014), the measurement of the outcome overlaps with the measurement of personality. This has resulted in studies finding contradictory results about the precise direction of effects (e.g., Kandler et al., 2015; Soto, 2015) or conducting further tests that suggest the final trait level measures are what are actually driving the change-outcome associations (e.g., Converse et al., 2018). When predicting an outcome is the goal of a study, models that cannot adequately address the direction of influence (i.e., cross-lagged panel models) should be avoided and a distal outcome be used. For the most robust test of effects, life outcomes should be measured after the assessment of personality change to minimize occasion-specific variance between personality change and the outcome.

Last, one way to further validate that effects between changes in personality and a distal outcome are not instead due to preexisting personality-outcome associations is to control for initial levels of the outcome. This conservative test ensures the association between the personality change and distal outcome quantifies effects due to newer changes in functioning that reflect more proximal personality processes. If this is not done, it is possible that all or some portion of the effect attributed to changes in traits rather reflects long-term, cumulative processes that stem from trait levels. Thus, the isolation of these proximal mechanisms that are indicative of newer changes in functioning, and explicitly testing these associations, provides a rigorous test of the predictive utility of personality change.

The Present Study

Previous studies of personality change do not provide a systematic, rigorous examination of if changes in traits are broadly associated with an array of life outcomes (i.e., health, education, career, finance, relationships, civic engagement) and the nature of
their associations with specific outcomes. The present study uses seven longitudinal data sets to examine the predictive utility of Big Five trait change. To combat potential issues in past research, we (a) incorporate multiple data sets to test how robust effects are; (b) test if associations are moderated by various study-level variables, which can inform under which conditions personality change may matter most and provide insight as to why inconsistent results sometimes emerged in past studies; (c) have personality data spanning up to nine waves so as to reliably assess changes in personality; (d) control for static trait levels when making outcome predictions; (e) use longitudinal growth models that combat the weaknesses of cross-lagged panel models; (f) predict an always-distal outcome relative to one’s final personality trait measures so as to minimize occasion-specific variance; and (g) predict both static levels and changes in outcomes to better elucidate the nature of the processes linking personality with future life outcomes.

Method

Participants

In this article, we use data from \( N = 81,980 \) participants from seven longitudinal panel data sets. To be included in the study, a participant must have had at least two waves of Big Five personality data as well as data for distal outcomes (see Table 1, for demographic and design information per study as well as for all data sets combined). The number of individuals with only two waves was 1,907, three waves was 5,631, four waves was 5,212, five waves was 8,860, six waves was 17,188, seven waves was 10,813, eight waves was 1,919, nine waves was 2,639, 10 waves was 4,438, 11 waves was 5,350, 12 waves was 1,375, 13 waves was 4,183, and 14 waves was 12,465. For the Big Five personality variables, the number of individuals with two waves was 36,529; three waves was 22,839; four waves was 18,306; five waves was 779; six waves was 898; seven waves was 834; eight waves was 1,783; and nine waves was 12. The institutional review board (IRB) at Washington University in St. Louis deemed this project exempt from IRB approval because this project involves accessing publicly available data sets and thus does not meet federal definitions under the jurisdiction of an IRB (IRB ID No.: 202208037).

German Socioeconomic Panel Study

The German Socioeconomic Panel (GSOEP) study (Socio-Economic Panel, 2019) is an ongoing longitudinal study conducted by the German Institute of Economic Research (Deutsches Institut für Wirtschaftsforschung Berlin) collecting data on individuals in more than 11,000 German households. Data are freely available by application at https://www.diw.de/de/soep. Data collection began in 1984 and continues annually, with the latest release in 2021. Data from the years 2005 to 2018 were used in the present study. Through years 2005–2017, the Big Five were assessed every 4 years. Questions regarding each of the outcomes are typically assessed annually. A list of prior publications using the GSOEP data can be found at https://www.diw.de/en/diw_01.c.800183.en/our_soep_publications.html#_801829.

Household Income and Labour Dynamics in Australia Study

The Household Income and Labour Dynamics in Australia (HILDA) study (Watson & Wooden, 2012) is an ongoing longitudinal study collecting data on more than 17,000 individuals in Australian households. Data are freely available by application at https://melbourneinstitute.unimelb.edu.au/hilda/for-data-users. Data collection began in 2001 and has continued annually, with the latest release in 2021. Data from the years 2005 to 2018 were used in the present study. The Big Five are assessed every 4 years, whereas questions regarding the outcomes are typically assessed annually. Prior publications using these data can be found at https://melbourneinstitute.unimelb.edu.au/hilda/publications/journal-articles.

Table 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>GSOEP</th>
<th>HILDA</th>
<th>HRS</th>
<th>LISS</th>
<th>MIDUS</th>
<th>NLSY–CYA</th>
<th>SHP</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>24,472</td>
<td>15,471</td>
<td>14,988</td>
<td>10,236</td>
<td>4,149</td>
<td>7,322</td>
<td>5,342</td>
<td>81,980</td>
</tr>
<tr>
<td>Age (( M ))</td>
<td>51.44</td>
<td>47.48</td>
<td>69.67</td>
<td>50.30</td>
<td>55.11</td>
<td>25.05</td>
<td>51.08</td>
<td>51.22</td>
</tr>
<tr>
<td>Age (( SD ))</td>
<td>16.71</td>
<td>18.05</td>
<td>10.56</td>
<td>17.78</td>
<td>14.00</td>
<td>5.62</td>
<td>17.86</td>
<td>18.63</td>
</tr>
<tr>
<td>% female</td>
<td>54</td>
<td>53</td>
<td>60</td>
<td>55</td>
<td>55</td>
<td>50</td>
<td>56</td>
<td>55</td>
</tr>
<tr>
<td>Number of waves total (( M ))</td>
<td>9.07</td>
<td>10.60</td>
<td>5.69</td>
<td>7.71</td>
<td>2.73</td>
<td>4.89</td>
<td>10.32</td>
<td>7.96</td>
</tr>
<tr>
<td>Number of waves total (( SD ))</td>
<td>3.38</td>
<td>3.56</td>
<td>1.21</td>
<td>3.86</td>
<td>0.44</td>
<td>1.00</td>
<td>1.27</td>
<td>3.66</td>
</tr>
<tr>
<td>Number of waves total (range)</td>
<td>2–14</td>
<td>2–14</td>
<td>2–7</td>
<td>2–13</td>
<td>2–3</td>
<td>2–6</td>
<td>3–12</td>
<td>2–14</td>
</tr>
<tr>
<td>Number of waves Big Five (( M ))</td>
<td>2.73</td>
<td>3.04</td>
<td>2.77</td>
<td>4.50</td>
<td>2.63</td>
<td>2.92</td>
<td>2.00</td>
<td>2.98</td>
</tr>
<tr>
<td>Number of waves Big Five (( SD ))</td>
<td>0.87</td>
<td>0.90</td>
<td>0.74</td>
<td>2.22</td>
<td>0.48</td>
<td>0.58</td>
<td>0.01</td>
<td>1.23</td>
</tr>
<tr>
<td>Number of waves Big Five (range)</td>
<td>2–4</td>
<td>2–4</td>
<td>2–5</td>
<td>2–9</td>
<td>2–3</td>
<td>2–5</td>
<td>2–3</td>
<td>2–9</td>
</tr>
<tr>
<td>Number of years between Big Five waves (( M ))</td>
<td>4.12</td>
<td>4.17</td>
<td>4.13</td>
<td>1.74</td>
<td>9.87</td>
<td>3.77</td>
<td>5.90</td>
<td>3.65</td>
</tr>
<tr>
<td>Number of years between Big Five waves (( SD ))</td>
<td>0.45</td>
<td>0.91</td>
<td>0.74</td>
<td>0.83</td>
<td>1.68</td>
<td>1.26</td>
<td>0.40</td>
<td>2.00</td>
</tr>
<tr>
<td>Number of years between Big Five waves (range)</td>
<td>4–12</td>
<td>4–12</td>
<td>2–12</td>
<td>1–11</td>
<td>9–19</td>
<td>2–10</td>
<td>1–6</td>
<td>0–19</td>
</tr>
</tbody>
</table>

Note. \( N \) = sample size; GSOEP = German Socioeconomic Panel; HILDA = Household Income and Labour Dynamics in Australia; HRS = Health and Retirement Study; LISS = Longitudinal Studies for the Social Sciences; MIDUS = Midlife in the United States; NLSY–CYA = National Longitudinal Survey of Youth 1979–Children and Young Adults; SHP = Swiss Household Panel Study.
Health and Retirement Study

Health and Retirement Study (HRS; Juster & Suzman, 1995) is an ongoing longitudinal study of more than 35,000 individuals from households in the United States. Data are freely available at https://hrs.isr.umich.edu. Data collection began in 1992 and continues biennially, with the latest release in 2020. Data from the years 2006 to 2018 were used in the present study. Generally, the Big Five are assessed every 4 years for an individual, although a small number (35 people) had an assessment gap of only 2 years for one wave. Questions regarding the outcomes are typically assessed every 2 years. Prior publications using this data can be found at https://hrs.isr.umich.edu/publications/bibliography/.

Longitudinal Studies for the Social Sciences

Longitudinal Studies for the Social Sciences (LISS; Scherpenzeel & Das, 2010) is an ongoing longitudinal study of Dutch-speaking individuals from 5,000 households in the Netherlands. Data are freely available through application at https://statements.centerdata.nl/liss-panel-data-statement. Data collection began in 2007 and has continued annually, with the latest release in 2022. Data from the years 2008 to 2020 were used in the present study. The LISS survey includes questions for Big Five traits and outcomes annually. Prior publications can be found at https://www.centerdata.nl/publicaties.

Midlife in the United States Study

The Midlife in the United States (MIDUS) study (Brim et al., 2004; Ryff et al., 2021) is an ongoing longitudinal study of more than 10,000 adults in the United States. Data are freely available at http://www.icpsr.umich.edu. In this article, we used data from MIDUS I, II, and III. Data for MIDUS I were collected in 1995–1996, data for MIDUS II were collected in 2004–2006, and data for MIDUS III in 2013–2014. The variables in our study were assessed at each of the three waves. A list of prior publications can be found at https://midus.wisc.edu/findings/index.php.

National Longitudinal Survey of Youth 1979–Children and Young Adults

The National Longitudinal Survey of Youth 1979–Children and Young Adults (NLSY–CYA; Bureau of Labor Statistics, 2020) is an ongoing longitudinal study of the offspring of individuals in the original NLSY 1979 (NLSY79). Data are freely available at https://www.nlsinfo.org/investigator/pages/login. The NLSY79 consists of data collected on more than 12,500 individuals in the United States since 1979. The NLSY–CYA includes the biological children of the NLSY79 participants and data collection began in 1986 and continues biennially, with the latest release in 2018. Data from the years 2006 to 2016 were used in the present study. Questions for the variables in our study are assessed every 2 years. A list of prior publications can be found at https://nlsinfo.org/bibliography-start.

Swiss Household Panel Study

The Swiss Household Panel (SHP) “Living in Switzerland” (Tillmann et al., 2022) study is an ongoing longitudinal study of more than 10,000 individuals from households in Switzerland. Data are freely available with application at https://forsbase.unil.ch/project/study/public-overview/156320/. Data collection began in 1999 and has continued annually, with the latest release in 2021. Data from the years 2009 to 2019 were used in the present study. The Big Five were assessed annually in the years 2009–2011, with another assessment occurring in 2015. Outcomes are assessed annually. Prior publications can be found at https://forscenter.ch/publications/scientific-publications/.

Measures

Big Five

The predictors in this study are the Big Five personality traits (Goldberg, 1990). Measures and items varied across data sets, but a full list of items and psychometric information per data set can be found in supplemental File S1 and Table S1, respectively. Internal consistency estimates were calculated using the psych package in R (Revelle, 2021). The original measurement scale was used for each Big Five trait per study. For GSOP, HILDA, and NLSY, the items are measured on a 1–7 Likert scale. For HRS and MIDUS, the items are measured on a 1–4 Likert scale. Items in LISS are measured on a 1–5 Likert scale, whereas items in SHP are measured on a 0–10 Likert scale. All traits were scored and composited such that higher values indicate greater levels of the trait. Neuroticism was coded as emotional instability.

Outcomes

For every data set, with the exception of MIDUS,2 each outcome was assessed at least one wave after an individual’s final wave of personality data, so as to minimize occasion-specific variance. Across all outcomes and data sets, the average prediction interval (i.e., length of time between an individual’s final personality assessment and their predicted outcome measure) was 2.35 years and ranged from 0 to 13 years. In all models, each numeric outcome was standardized. Dichotomous outcomes were treated as dummy-coded factors. Unless otherwise noted, every outcome was present in each data set. See supplemental Tables S2 and S3 for concordance of outcomes across data sets, descriptive information for initial and final outcome values in each data set (supplemental Table S2), and descriptive information for the prediction intervals for each outcome per data set (supplemental Table S3).

Health

Outcomes in the health domain were self-reported health status, BMI, number of reported physical health problems, number of reported mental/emotional problems, number of reported health limitations for activities of daily living, and whether the participant reported engaging in any exercise. BMI was either a provided variable in the data sets or was calculated using height and weight variables. Items for individual health problems/limitations were dichotomous (1 indicating having the listed health problem/limitation, 0 indicating not having the listed health problem/limitation). Special attention was paid to try and use similar items across data sets to ensure the resulting variables were comparable.

2 As the MIDUS study only has three waves of data (for all measures), we opted to use all three waves of personality data, when available for a participant, even though this meant it would overlap with the assessment of each outcome. Thus, the final personality measure sometimes co-occurs with the outcome measure in this data set only.
A participant’s total scores for these variables were created by summing their reported health problems/limitations. For reported physical and mental problems, these outcomes were not available in LISS and SHP. For health limitations, this outcome was not available in NLSY. The exercise variable was dichotomous, with 1 indicating the participant exercised and 0 indicating they did not.

**Relationships.** Outcomes in the relationship domain included a participant’s marital status, divorce status, and their reported number of marriages. The variables for marital and divorce status were dichotomous (1 indicated yes, 0 indicated no). For the divorce status variable, the data sets were filtered for participants who, at some point during the study, reported being married at least once. Thus, the association of the traits and divorce status wave was conditional upon a participant reporting being married at least once in the study. For the number of marriages variable, the data sets were filtered for participants who reported being married at least once at their initial wave. Thus, the association of traits and number of marriages at the final wave was conditional upon an individual having already been married at their first measurement occasion. This served to differentiate this variable from the outcome of marital status, which included all participants.

**Education.** The outcome in the education domain was whether the participant had a 4-year college degree (or the country’s equivalent degree level) or higher (i.e., bachelor’s level and higher). The university degree status variable was coded such that 0 indicated no 4-year college degree (or the country’s equivalent degree) and 1 indicated a bachelor’s, master’s, PhD, or a professional degree (e.g., MD, JD; or the country’s equivalent degree[s]).

**Career.** The outcome in the occupational domain was the unemployment status of the participant. The variable was coded such that 1 indicated they were unemployed and 0 indicated they were not. A response of 0 could indicate the participant was employed or retired.

**Income.** The financial outcome was annual salary. Salary was measured in the original currencies for each data set as this variable was standardized for analyses regardless.

**Civic Engagement.** The civic engagement outcome was a person’s volunteer status in the past year. It was coded such that 1 indicated yes and 0 indicated no to reporting volunteering.

**Transparency and Openness**

Within this methods section, we report how we determined our final sample size through inclusion criteria, all measures used along with their psychometric properties, and we follow the APA Style Journal Article Reporting Standards (Kazak, 2018). Data are freely accessible at all links specified in each respective study’s Participants subsection. All raw data were downloaded directly from the data repositories for each respective data set. The codebook and code for cleaning data, composing/constructing variables, and all analyses are available at https://osf.io/hdms9l. Data were analyzed using R, Version 4.2.0 (R Core Team, 2021) and the package brms (Bürkner, 2017). This study’s design and its analyses were not preregistered.

**Analytic Plan**

The central analytic plan consists of a series of Bayesian models, which can be separated into three phases. We describe each phase in detail below.

**Phase 1**

First, we obtained individual trajectories of change for the traits across time using multilevel models. Separate models were fit for each Big Five trait in each study. Additionally, as these models served as the building blocks for Phase 2 models that included outcome data, separate Phase 1 models were fit per outcome such that each model only included trait assessments that occurred prior to an individual’s final outcome measure. The outcomes themselves were not included in Phase 1 models—it is just that the data for each Phase 1 model only included waves that occurred prior to each person’s final outcome assessment. As this can vary by outcome and person, separate models were fit for each trait, outcome, and data set combination. This resulted in 370 models.

Time was measured in years and scaled such that a one-unit change captured the timespan of 10 years per study. In order to have these models serve as a building block for the models including the outcome variables, the middle time point (determined as the nearest whole wave number for a participant’s median wave) was set as the intercept as this decreases the correlation between initial status and slope (Klimstra et al., 2013). The middle wave for each participant had a value of 0 for time and years prior to that point was negative, whereas years afterward were positive. For example, 4 years prior to a participant’s median wave would be coded as −4, the median wave would be coded as 0, and 4 years after the median wave would be coded as 4. An example equation can be demonstrated via the following:

\[
\text{Level 1:} \quad Trait_{ij} = b_{0j} + b_{1j} \text{time}_{ij} + e_{ij}, \tag{1}
\]

\[
\text{Level 2:} \quad b_{0j} = \gamma_{00} + U_{0j}, \quad b_{1j} = \gamma_{10} + U_{1j}, \tag{2}
\]

where subscript \(i\) is for each assessment point, nested within participants, and subscript \(j\) is for each participant. The parameter \(\gamma_{00}\) represents the predicted trait value at the average median wave (i.e., the intercept); the parameter \(\gamma_{10}\) represents the average slope for a one-unit change in the \(\text{time}_{ij}\) variable (quantified as change over a 10-year period); the parameter \(U_{0j}\) represents the person-specific deviation from the average intercept value; and the parameter \(U_{1j}\) represents the person-specific deviation from the average slope value. All models used weakly informative and regularized priors. The prior for the intercept (i.e., \(\gamma_{00}\)) was a normal distribution centered around the nearest whole integer of the average of the Big Five trait in each data set with a standard deviation of 1; the prior for regression coefficient (i.e., \(\gamma_{10}\)) was a normal distribution with a mean of 0 and standard deviation of 1; the prior for the Level 2 standard deviation parameters (i.e., random effects \(U_{0j}\) and \(U_{1j}\)) was a half Cauchy distribution with a location of 0 with a scale of 2; and the prior for the Level 1 residual (i.e., sigma) was an exponential distribution with a parameter value of 1. Maximum a posteriori (MAP) probability estimates were extracted from each model’s posterior distribution along with 95% credible intervals (CIs).

MAP estimates, derived from a Markov chain Monte Carlo estimation process, can be interpreted similarly to traditional, frequentist parameter estimates derived from a maximum likelihood
(ML) estimation process. Likewise, the 95% CIs convey a range of plausible values similar to traditional 95% confidence intervals from frequentist models using ML estimation, but the 95% CIs rather identify the plausible values based on the empirical posterior distribution as opposed to a theoretical sampling distribution (i.e., what traditional confidence intervals from frequentist models do). Moreover, due to the large sample sizes and the use of weakly informative and regularized priors in the present study, results obtained from either Bayesian or frequentist approaches are expected to be very similar.

**Phase 2**

After estimating Phase 1 models, the person-level intercepts and slopes were extracted from each model and integrated with the outcome data for use in a series of multiple regression models. The person-level intercepts and slopes were obtained by adding the fixed effect estimate to each individual’s random effect across all samples in the posterior (all models had 8,000). The fixed effects are the MAP probability estimates from the posterior distribution with the highest probability densities, or the peak (i.e., mode) of the posterior distribution. Thus, across all samples in the posterior distribution, each person’s deviation from this fixed effect was added to it to obtain their own person-specific parameter values. Then, the median value across all samples was calculated to give the resulting person-level parameter that would be used in Phase 2 models. These person-level intercepts and slopes were then standardized. As a reminder, our numeric outcomes were also standardized. Thus, the regression coefficients for individual predictors can be compared amongst each other and further treated as correlation coefficients. A separate model was run for each Big Five trait, outcome, and data set.

When predicting static levels of our distal outcomes, the dependent variable was the final outcome measure for a participant and the independent variables were the person-level intercepts and slopes from each Phase 1 model. This resulted in a total of 430 models. An example equation can be demonstrated via the following:

\[
\text{Outcome}_j = b_0 + b_1 \text{Level}_j + b_2 \text{Change}_j + e_j,
\]

where \(\text{Outcome}_j\) is a participant’s final outcome measure; \(\text{Level}_j\) is a participant’s intercept value from Phase 1 model; and \(\text{Change}_j\) is a participant’s slope value from Phase 1 model. The parameter \(b_0\) is the average outcome value; \(b_1\) quantifies the association between the effects of trait levels on each outcome, in units of 1 SD from the average trait-level value per one-unit change; and \(b_2\) quantifies the association between the effects of changes in traits on each outcome, also in units of 1 SD from the average slope value per one-unit change. These models served as an initial test of if changes in the Big Five traits predicted future outcomes, above and beyond the effects due to static trait levels.

Then, a more conservative test of if changes in traits predict future outcomes is to see if they do so after controlling for initial outcome values—thus effectively predicting changes in an outcome. This allows one to test not only if changes in traits are broadly associated with life outcomes, but further if changes in traits are really capturing those proximal effects associated with newer changes in behavior, and thus likely changes in these outcomes. For numeric outcomes, the initial outcome value was similarly standardized in all models. This resulted in an additional 430 models. An example equation can be demonstrated via the following:

\[
\text{Outcome}_j = b_0 + b_1 \text{Level}_j + b_2 \text{Change}_j + b_3 \text{Outcome}.\text{initial}_j + e_j,
\]

where \(\text{Outcome}.\text{initial}_j\) is a participant’s initial outcome measure and parameter \(b_3\) quantifies the association between the effects of a participant’s initial outcome value predicting their future and final outcome value. For numeric outcomes, this is in units of 1 SD from the average initial outcome value per one-unit change. For dichotomous outcomes and number of marriages, this is in units of a one-unit increase from the minimum outcome value. All models again had weakly informative priors. For all Phase 2 models, the priors for intercepts and regression coefficients were normal distributions centered around 0 with a standard deviation of 1. Numeric outcomes were modeled with a Gaussian distribution, dichotomous outcomes were modeled with a Bernoulli distribution, and the outcome of number of marriages was modeled with a Poisson distribution. MAP probability estimates were extracted from each model along with 95% CIs.

**Phase 3**

Last, we obtained meta-analytic summaries for each of the level and change effects from Phase 2 models. This consisted of extracting the parameters and standard errors from each Phase 2 model in order to obtain a weighted average of the effects—or the “true” effect. Using these values, multilevel models were run, with estimates nested in the data sets. The “true” effect can be represented with the following equation:

\[
\hat{\theta}_k \sim N(\mu, \sigma_k^2 + \tau^2),
\]

where \(\hat{\theta}_k\) is the observed effect size in study \(k\); \(N\) indicates the parameters were sampled from a normal distribution; \(\mu\) represents the weighted, pooled “true” effect size of the \(k\) study-level effect size distributions; \(\sigma_k^2\) is the variance of the effect size distribution for study \(k\); and \(\tau^2\) is the variance of the distribution of “true” effects and quantifies the between-study heterogeneity. For each trait and outcome combination, this “true” effect was estimated via the following equation:

\[
\text{Estimate}_k | SE_{jk} = \theta_{0k},
\]

\[
\theta_{0k} = \mu_{00} + U_{0k},
\]

where \(\text{Estimate}_k\) represents the level or change association parameter \(j\) from study \(k\), weighted by the standard error of parameter \(j\) from study \(k\) (\(SE_{jk}\)); \(\theta_{0k}\) represents the observed effect size from study \(k\), assumed to represent the “true” effect in the study; \(\mu_{00}\) represents the pooled “true” effect size; and \(U_{0k}\) represents the study \(k\)-specific deviation from the “true” pooled effect size.

These models were run for each trait and outcome combination for both level (65) and change (65) effects, resulting in 130 models for the results from Phase 2 models predicting static levels of outcomes (i.e., did not control for the initial outcome) and 130 models for the results from Phase 2 models predicting changes in outcomes (i.e., controlled for the initial outcome).
Then, as a final step, we performed meta-regressions whereby study-level variables including average age, prediction interval for each outcome, years between Big Five waves, number of Big Five waves, and reliability of trait measures (i.e., Cronbach’s α value) were included to view their associations with the pooled average effects. This resulted in an additional 65 models per combination of parameter (i.e., change or level), trait, outcome, and if the initial outcome variable was included. Study-level variables were standardized. Priors for Phase 3 models were normal distributions centered around 0 with a standard deviation of 1 for the intercepts (and regression coefficients when study-level variables were included) and were half Cauchy distributions with a location of 0 and scale of 1 for the study-level random effect.

**Results**

**Individual Differences in Personality Change**

First, we examined the degree to which individual-level random effects were present for the slopes of the Big Five traits in each of our seven data sets. In general, while there were normative patterns of personality development observed in each data set, we additionally found large individual differences in these changes such that different people change at different rates (Table 2). The amount of variability in slopes for the traits not only varied widely across data sets but also across traits within the same data set (see supplemental Tables S4–S10, for all model estimates).

Next, we investigated if changes in the Big Five traits were prospectively associated with static levels and changes in outcomes, independent of any trait-level associations. Generally, changes were associated with numerous outcomes, above and beyond static trait levels. This was especially true after controlling for initial outcome values, thus predicting changes in outcomes. In the Results section, we (a) briefly report the general trends in the effects that emerged for each trait and outcome across individual data sets and (b) present the meta-analytic effects and discuss the meaningful associations. We first do this for the models in which we predicted static outcome variables and then do this for the models in which we predicted changes in outcomes (i.e., controlled for initial outcome values). Finally, we describe the study-level moderators of the meta-analytic effects. Full results from all models with the individual data sets are available in supplemental Tables S11–S23 and Figures S1–S13 for models that predicted static levels of outcomes and supplemental Tables S24–S36 and Figures S14–S26 for models that predicted changes in outcomes.

As a reminder, for all effects, including both static levels and changes in traits, they are interpreted as the effect on the outcome, in standard deviation (SD) units, of having mean levels and changes in traits 1 SD above and beyond the average level and slope, respectively. Numeric outcomes were also standardized. Thus, the magnitude of level and change effects are in correlation units and can be directly compared. The effects for dichotomous outcomes and number of marriages are presented as odds ratios and relative risk ratios, respectively, which can also be directly compared. For numeric outcomes, if the absolute values of the credible intervals of the level and change effects overlap to any degree, they are considered similar in magnitude. For dichotomous outcomes and number of marriages, if the credible intervals of the estimate or its reciprocal overlap to any degree, they are considered similar in magnitude.

### Changes in Personality Predicting Static Life Outcomes

**Individual Data Set Trends**

Overall, out of the possible 430 effects for trait-level associations, 263 emerged (61%). In comparison, out of 430 possible effects for change associations, 147 emerged (34%). In terms of magnitude of the effects, out of 118 paired level and change associations (i.e., same outcome, same trait, same data set), 75 had similar magnitudes (63.6%); 40 level associations were larger than their paired change association (33.9%); and only three change associations were larger than their paired level association (2.5%). Regarding the direction of effects (i.e., positive or negative in direction), paired level and change associations were in the same direction 80% of the time and were in opposite directions 20% of the time. Supplemental Tables S37 and S38 contain further details of this descriptive information for each trait and outcome, separated by health outcomes (supplemental Table S37) and nonhealth outcomes (supplemental Table S38). For traits, the most numerous change effects were found for conscientiousness (47% of possible effects). For outcomes, the most numerous change associations were found for health status (74% of possible effects). Supplemental Tables S39 and S40 contain summary information about the results from all models, organized by data set and trait.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Extraversion</th>
<th>Agreeableness</th>
<th>Conscientiousness</th>
<th>Neuroticism</th>
<th>Openness</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSOEP</td>
<td>0.83 [0.80, 0.85]</td>
<td>0.77 [0.74, 0.79]</td>
<td>0.76 [0.73, 0.78]</td>
<td>0.96 [0.93, 0.99]</td>
<td>0.89 [0.86, 0.91]</td>
</tr>
<tr>
<td>HILDA</td>
<td>0.33 [0.31, 0.35]</td>
<td>0.36 [0.34, 0.38]</td>
<td>0.39 [0.36, 0.41]</td>
<td>0.38 [0.35, 0.41]</td>
<td>0.38 [0.35, 0.41]</td>
</tr>
<tr>
<td>HRS</td>
<td>0.23 [0.20, 0.25]</td>
<td>0.14 [0.11, 0.17]</td>
<td>0.19 [0.17, 0.21]</td>
<td>0.22 [0.20, 0.25]</td>
<td>0.17 [0.14, 0.19]</td>
</tr>
<tr>
<td>LISS</td>
<td>0.32 [0.31, 0.33]</td>
<td>0.26 [0.25, 0.27]</td>
<td>0.29 [0.27, 0.30]</td>
<td>0.39 [0.37, 0.40]</td>
<td>0.22 [0.21, 0.24]</td>
</tr>
<tr>
<td>MIDUS</td>
<td>0.10 [0.08, 0.12]</td>
<td>0.07 [0.04, 0.10]</td>
<td>0.11 [0.09, 0.12]</td>
<td>0.13 [0.11, 0.15]</td>
<td>0.11 [0.09, 0.13]</td>
</tr>
<tr>
<td>NLSY</td>
<td>0.73 [0.62, 0.83]</td>
<td>0.60 [0.47, 0.72]</td>
<td>0.48 [0.31, 0.60]</td>
<td>0.74 [0.62, 0.85]</td>
<td>0.30 [0.03, 0.51]</td>
</tr>
<tr>
<td>SHP</td>
<td>1.55 [1.29, 1.87]</td>
<td>0.57 [0.35, 0.94]</td>
<td>0.67 [0.54, 0.89]</td>
<td>0.35 [0.13, 0.77]</td>
<td>1.04 [0.64, 1.71]</td>
</tr>
</tbody>
</table>

*Note.* Random effects are presented in units of standard deviation (i.e., as opposed to variance). All traits are in their original units. Est. = maximum a posteriori estimate; CI = 95% credible interval; GSOEP = German Socioeconomic Panel; HILDA = Household Income and Labour Dynamics in Australia; HRS = Health and Retirement Study; LISS = Longitudinal Studies for the Social Sciences; MIDUS = Midlife in the United States; NLSY = National Longitudinal Survey of Youth; SHP = Swiss Household Panel Study.
by traits (supplemental Table S39) and outcomes (supplemental Table S40).

**Meta-Analytic Effects**

Next, we conducted a meta-analytic summary of the effects from the individual models to obtain an estimate of the average effect across all data sets for each outcome and trait. Across all outcomes, thus out of 65 possible effects, 24 trait-level associations emerged and only eight associations for changes in traits emerged. The most numerous change associations were found for conscientiousness and zero change effects were found for agreeableness. When both were present, level and change associations for a given trait and outcome were never in opposite directions (i.e., positive vs. negative). See supplemental Table S41 for all summary information per trait and supplemental Table S42 for all summary information organized by outcome.

When viewed in terms of specific outcomes, several change effects emerged (Tables 3 and 4). At the meta-analytic level, there were associations with outcomes in health, education, career, finance, and civic engagement domains. For health status (Table 3; Figure 1), increasing in conscientiousness and openness predicted higher than average ratings. Specifically, increasing in these traits, 1 SD more than the average slope predicted higher levels of health status to the degree of 0.08 SDs for conscientiousness and 0.04 SDs for openness. In comparison, increasing 1 SD more than average in neuroticism was associated with 0.05 SDs lower than average health status. For this outcome, it is notable that changes in conscientiousness predicted health status above and beyond static levels, whereas the reverse could not be said for static levels of this trait. This suggests that any proximal mechanisms were relatively more influential than cumulative effects were, at least for conscientiousness. However, for neuroticism, both level and change effects emerged, and the magnitude of the point estimate for the level effect was 5x that of the change effect. This suggests that cumulative effects for neuroticism might be difficult to overcome—even if you do change in a beneficial direction for neuroticism.

Fewer change associations emerged for other health-relevant outcomes. A notable exception was health limitations (Table 3), where increasing 1 SD more than the average slope in openness predicted fewer than average health limitations, to the degree of 0.04 SDs fewer. Here, we again see a case where the cumulative effects of level were more influential than the proximal effects of change.

For the education and occupation outcomes, changes in conscientiousness and neuroticism were the important predictors. Greater than average changes in conscientiousness positively predicted university degree status, such that individuals who increased 1 SD more than average had 1.10 greater odds of having a university degree (Table 4). Interestingly, the effects for changes in conscientiousness were mostly consistent across all data sets, whereas the effect of static levels was quite heterogeneous (Figure 2). This high degree of variability in level effects seemed to be true for agreeableness and openness as well. For unemployment status, increasing 1 SD more than average in neuroticism predicted a greater likelihood of being unemployed, such that these individuals had 1.06 greater odds of being unemployed (Table 4; Figure 3). Though and similar to health status, the cumulative effects of neuroticism seem to be more consequential than proximal effects do for unemployment, as the level effect was quite larger in magnitude.

The estimates for unemployment status demonstrate how single-study effects are often heterogeneous. For instance, for extraversion and agreeableness, the associations for level and change would sometimes alternate in direction across data sets (Figure 3). Similarly, for salary, greater than average increases in conscientiousness predicted a higher than average salary, to the degree of 0.03 SDs higher (Table 4). In comparison, static levels of conscientiousness had no average pooled effect but did emerge a few times in the individual data sets (e.g., HILDA, LISS, NLSY).

Last, for the civic engagement outcome, increasing 1 SD more than the average slope in extraversion predicted a 1.05 greater odds of reporting volunteering relative to individuals that changed at the average rate in this trait. Again, though, this change effect was less substantial than the effect of having 1 SD higher than average static levels of this trait.

Overall, change effects emerged most frequently for conscientiousness, closely followed by neuroticism and openness, and never emerged for agreeableness. Health status was the outcome most frequently associated with changes in traits and no change effects were found for relationship outcomes. When both trait-level and change associations were present for a given trait and outcome, they were never in opposite directions. Additionally, level effects were usually larger than their respective change association, particularly for neuroticism, but sometimes were similar in magnitude.

**Changes in Personality Predicting Changes in Life Outcomes**

**Individual Data Set Trends**

Across all models, a participant’s initial outcome value predicted their final outcome value, indicating this is the best predictor of someone’s future outcomes. Out of the possible 430 effects for trait-level associations, 203 emerged (47%)—compared to 61% in the previous models. Then, 136 change associations emerged (32%)—compared to 34% in the previous models. In terms of magnitude of the effects, out of 96 paired level and change associations, 73 had similar magnitudes (76%); 18 level associations were larger (19%); and five change associations were larger (5%). Regarding the direction of effects, paired level and change associations were in the same direction 79% of the time and were in opposite directions 21% of the time. Supplemental Tables S43 and S44 contain further details of this descriptive information for each trait and outcome, separated by health outcomes (supplemental Table S43) and nonhealth outcomes (supplemental Table S44). For traits, the most change effects were again found for conscientiousness (49% of possible effects). For outcomes, the most change associations were again found for health status (86% of possible effects). Supplemental Tables S45 and S46 contain summary information about the results from all models, organized by traits (supplemental Table S45) and outcomes (supplemental Table S46).

**Meta-Analytic Effects**

In the meta-analytic models, after controlling for initial outcome values, the number of level associations decreased whereas the number of change associations increased. Furthermore, changes in traits were now meaningfully associated with new outcomes. Out of 65 possible effects, 17 trait-level associations emerged, compared to...
24 in the previous models. In comparison, 15 change associations now emerged, compared to eight in the previous models. The most numerous change associations were again found for conscientiousness and agreeableness again had zero effects. Additionally, when both level and change effects were present for a given trait and outcome, they were now similar in magnitude majority of the time. See supplemental Table S47 for all summary information per trait and supplemental Table S48 for all summary information organized by outcome.

Below, we describe the meaningful effects for the outcomes that were newly predicted by changes in traits. However, in general, all change-outcome associations that were meaningful in the previous models were also present, even after controlling for the initial outcome. This suggests that the prior associations of changes in traits with static outcomes likely did capture effects due to proximal personality processes, as these associations similarly emerged when predicting changes in the outcomes. Furthermore, some of those outcomes had new change associations emerge as well (e.g., health status, unemployment status; Tables 3 and 4).

First, a new change association emerged for number of reported mental problems, such that increasing 1 SD more than average in extraversion predicted a 0.04 SD decrease in mental problems, controlling for the initial number reported (Table 3). Then, increasing 1 SD more than average in conscientiousness predicted an increase in the likelihood of being married, to the degree of 1.11 greater odds, controlling for initial marital status (Table 4; Figure 4). This change effect was equivalent in magnitude to the effect of static levels.

Increasing 1 SD more than average in openness predicted an increase in the likelihood of being divorced, controlling for initial divorce status (Table 4). This effect was to the degree of 1.07 greater odds. Furthermore, although the point estimate for the level effect is larger than the change estimate, their credible intervals overlap, suggesting they are not meaningfully different. Thus, similar to marital status, this is likely a case whereby cumulative and proximal effects may matter to a somewhat similar degree for linking traits with outcomes.

Overall, when predicting changes in outcomes, change effects emerged even more frequently—both for outcomes previously associated with changes as well as for new outcomes. Most change associations were again found for conscientiousness and again never emerged for agreeableness. Changes in health status were the outcome most frequently associated with changes in traits and no robust effects were found for changes in BMI, number of physical health problems, exercise status, or number of marriages. Level and
Moderators of Meta-Analytic Effects

When examining the potential effects of study-level variables on the pooled average effects from the meta-analyses, not many associations emerged. We restricted our discussion of these effects to those with the pooled average effects for changes in traits. But, as a broad overview of the trait-level effects, in the models that did not control for the initial outcome variables, seven effects emerged (2.2% of possible effects). Specifically, four for the number of Big Five waves, two for average age per study, and one for internal consistency of a Big Five trait. Majority of the effects were with the trait openness. See supplemental Tables S49 and S50 for the estimates of each study-level variable with all trait and outcome combinations for health outcomes (supplemental Table S49) and nonhealth outcomes (supplemental Table S50). Then, in the models that did control for the initial outcome, eight trait-level effects emerged (2.5%—four for average age per study, two for internal consistency of a Big Five trait, and two for the number of Big Five waves. See supplemental Tables S51 and S52 for the estimates of these study-level variable with all trait and outcome combinations for health outcomes (supplemental Table S51) and nonhealth outcomes (supplemental Table S52).

For change associations, in the models that did not control for initial outcome variables, three effects emerged (0.9%—all for average age per study (supplemental Tables S53 and S54). More associations emerged when controlling for the initial outcome variable. Specifically, five effects emerged (1.5%—three for average age per study and two for internal consistency of a Big Five trait (supplemental Tables S55 and S56). Finally, an effect of average age per study was found for agreeableness and conscientiousness with health limitations, such that studies with average age per study higher than average had negative associations between conscientiousness with health limitations, such that studies with average age per study (supplemental Table S52).
(b = −0.09, 95% CI [−0.16, −0.02]) and conscientiousness (b = −0.24, 95% CI [−0.45, −0.05]) was associated with decreases in the number of one’s health limitations (supplemental Table S55). Then, another age effect was found for openness and marital status, such that studies with higher than average ages had a larger positive association (OR = 1.10, 95% CI [1.03, 1.18]). This indicates that in these studies, increasing 1 SD more than average was associated with an increase in the likelihood of being married if a participant was not initially married (supplemental Table S56). Last, an effect was found for the reliability of conscientiousness and university degree status, such that studies with higher than average Cronbach’s α values for conscientiousness had an even larger association (OR = 1.10, 95% CI [1.04, 1.17]). This indicates that in these studies, increasing 1 SD

**Figure 1**

*Distributions of Individual Data Set and Meta-Analytic Estimates for Levels and Changes in the Big Five Traits Predicting Static Levels of Health Status*

*Note.* The individual data sets are listed in alphabetical order and the pooled average effects from the meta-analyses are below the individual data set associations for each trait. The effects for changes in traits are dark gray and the effects for static trait levels are light gray. The horizontal lines for each effect delineate the 95% credible interval bounds and the solid dot indicates where the maximum a posteriori estimate is. Estimates are in correlation units. GSOEP = German Socioeconomic Panel; HILDA = Household Income and Labour Dynamics in Australia; HRS = Health and Retirement Study; LISS = Longitudinal Studies for the Social Sciences; MIDUS = Midlife in the United States; NLSY = National Longitudinal Survey of Youth; SHP = Swiss Household Panel Study.
more than average was associated with an increase in the likelihood of obtaining a university degree if a participant did not initially have one (supplemental Table S56).

**General Summary**

Overall, although static trait levels are more frequent predictors of future life outcomes, changes in personality traits sometimes matter as well. This is especially true when controlling for one’s initial outcome value, thus predicting changes in an outcome. Indeed, changes in traits were more robust predictors of outcomes than trait levels were after controlling for initial outcome values, such that the number of change associations increased while the number of level associations decreased. These findings highlight that changes in traits are important for one’s development, especially when predicting newer changes in functioning.

Across all meta-analytic models, change effects emerged most frequently for conscientiousness and never emerged for agreeableness. This suggests that proximal effects for conscientiousness may matter a lot—similar to the cumulative effects due to the static levels of these traits. In comparison, who someone “is” in terms of their typical agreeableness levels is likely more consequential than...
any changes they may exhibit in this trait. Health status was the outcome most frequently associated with changes in traits, whereas no effects were found for BMI, number of physical health problems, exercise status, or number of marriages. When both trait-level and change associations were present for a given trait and outcome in the meta-analytic results, they were never in opposite directions, indicating that the processes linking levels and changes in traits to an outcome are complementary with one another. In terms of the magnitude of effects, change associations were never larger than their respective level association. This suggests that although changes in traits do sometimes matter for some outcomes, who someone “is” is probably more consequential, on average.

These general findings were not always found in individual data sets, though. Heterogeneity at the study level demonstrates the need to integrate a number of large data sets together to make broad claims. However, those claims may not always generalize to future work, depending on the sample and design of the study. For instance, average age per study was the most frequent moderator of the pooled average meta-analytic effects. Though, the two study-
level variables that had moderating effects for the change associations—average age per study and internal consistency of the Big Five traits—rarely emerged, suggesting most change associations in the present study are invariant to these basic sample and design characteristics.

**Discussion**

In this article, we examined whether changes in the Big Five traits prospectively predict a multitude of life outcomes, above and beyond their respective trait levels. Changes in personality were associated with numerous outcomes, with changes in all Big Five traits except agreeableness yielding robust predictive validity. These change effects jointly predicted outcomes along with static levels of personality, thus replicating past work indicating that the levels of the Big Five traits are robust predictors of future life outcomes (Beck & Jackson, 2022; Ozer & Benet-Martínez, 2006; Roberts et al., 2005; Soto, 2021). The change associations were nearly always invariant with respect to study-level variables, except for a small number of effects emerging for average study age and internal consistency estimates of the Big Five.
Despite the appreciable number of personality change effects, the meta-analytic effects for this change were never larger than the standard predictive validity of static personality traits, though sometimes they were similar in magnitude. The larger magnitude of the effects for static personality traits, on average, suggests a greater importance of long-term, cumulative processes linking personality with outcomes. However, the existence of change effects still underscores the utility of short-term, proximal processes. We discuss the implications of our findings below with regard to the effects of changes in traits compared to those of static trait levels, connections with past literature, and potential pathways in which these change-outcome associations may arise.

**Effects of Changes in Personality Traits Compared to Levels**

Across all models, there were a larger number of effects for the static levels of the Big Five traits compared to changes in the traits. For almost every outcome, at least one association for the level of a Big Five trait emerged, replicating past work that has shown trait levels of the Big Five traits are associated with many life outcomes (Beck & Jackson, 2022; Ozer & Benet-Martinez, 2006; Soto, 2019, 2021; Wright & Jackson, 2022). Despite the more numerous associations for trait levels, though, a comparable number of effects for changes in the traits emerged as well—particularly when predicting changes in the outcomes. For our personality level and change effects, there were three broad patterns worth noting.

First, for all meta-analytic effects, when both trait-level and change effects with an outcome were present, they were always in the same direction (i.e., positive or negative). This is expected if the mechanisms relating trait changes with outcomes reflect the same processes relating trait levels with life outcomes. For example, high levels of conscientiousness are associated with health via health behaviors (Hampson et al., 2007; Lodi-Smith et al., 2010; Wright et al., 2022) and physiological mechanisms (O’Súilleabháin et al., 2021; Wright & Jackson, 2022). Increases in conscientiousness should thus lead to changes in these intermediary processes, such as an increase in health behaviors (Takahashi et al., 2013), thereby leading to greater health-promoting practices. The downstream effects of these behavioral changes associated with certain traits could explain why changes in traits similarly predict these outcomes, similar to how changes in conscientiousness predict vital health outcomes such as mortality risk (Martin et al., 2007).

Second, another typical pattern in our meta-analytic results was that, when both effects were present, the magnitude of the change associations was never larger than the magnitude of the trait-level associations. This general finding that static levels of traits more strongly predicted outcomes than changes in traits did highlights differences between more immediate (proximal effects) and more distal (cumulative personality effects). Static levels of personality likely reflect more distal, cumulative processes whereas changes in personality reflect more proximal processes. For example, both levels and changes in neuroticism were negatively predictive of future health status. The detrimental effect that neuroticism can have on one’s health can stem from many sources (Friedman, 2019), including greater likelihood to engage in negative health behaviors (Wright et al., 2022), such as smoking and drinking (Turiano, Whiteman, et al., 2012), cognitive decline (Terracciano et al., 2014, 2017), higher levels of inflammation (Graham et al., 2018; Sutin et al., 2010; Wright et al., 2022), and greater comorbidity of physical health problems with psychopathology (Kotov et al., 2010). The negative consequences of some of these risk factors take time to emerge, such that, for example, smoking a single cigarette will not immediately lead to cancer. Instead, neuroticism’s effect is due to a person continually having these levels of neuroticism across many years, resulting in them cumulatively performing negative health behaviors that ultimately lead to poorer health. The cumulative effects of certain factors associated with static levels of personality traits thus take time to emerge—processes that have likely been in place for decades.

In contrast, some health behaviors can be detrimental in a more proximate time frame, such that their effects do not take decades to materialize. For instance, risky driving, binge drinking, and many other negative health behaviors could also lead to poorer health outcomes, even if not performed over a long time frame. Changes in personality, in comparison to static levels of personality, reflect, by definition, newer changes in functioning. If these changes are associated with outcomes, then they reflect processes that are closer in time compared to more distal pathways. Given that most outcomes are long-term processes in and of themselves (e.g., marriage, education, and salary are not something that one can easily change day to day), it is thus not surprising that static levels of personality often out-predicted changes in traits.

As for the third and final pattern, when changes in the Big Five traits were associated with outcomes, there was generally also an effect of level. This suggests there are unique mechanisms linking levels and changes of traits with certain outcomes—namely, cumulative and proximal processes—and that both mechanisms matter. To illustrate how both mechanisms can be important, we will use the (negative) associations of levels and changes in conscientiousness with predicting unemployment status. The cumulative effects of conscientiousness on job attainment can start from a very young age, as the benefits reaped from this trait cannot only produce tangible and relatively immediate outcomes (e.g., higher grades, better job performance in entry-level jobs; Bakker et al., 2012; Dudley et al., 2006; Noffle & Robins, 2007; O’Connor & Paunonen, 2007) but also make it easier to have access to other paths and opportunities (Hill et al., 2019) that allow for continued success such as higher class ranks, more opportunities for internships to gain relevant work experience, a more competitive résumé, and being able to perform the necessary behaviors to obtain and keep a job (Bakker et al., 2012; Brown & Hirschi, 2013; Dudley et al., 2006; Roberts et al., 2003). This type of developmental branching highlights that taking certain paths at an early point in life can make it easier or limits one’s ability to take other paths in life (Strøbye, 1997), truly emphasizing the impact of these long-term, cumulative effects (Hill & Jackson, 2016). However, these are not the only effects that matter.

For instance, while someone who worked hard, received high grades and test scores throughout primary and secondary school, and was able to obtain a competitive résumé are in a better position to continue this success by being a desirable job candidate and being more likely to keep a job (Brown & Hirschi, 2013; Roberts et al., 2003), this does not mean that someone in a different position cannot achieve the same end outcome. That is, even changes enacted over
the short-term can be associated with these same beneficial outcomes. Someone who is stably low or average in conscientiousness is less likely to have the academic/work record or regular behaviors associated with performing a job well (e.g., responsibility, organization skills; Brown & Hirschi, 2013) than someone high on this trait is (O’Connor & Paunonen, 2007; Roberts et al., 2003). However, even changes in the final year or two of secondary school in which one decides to focus on their schoolwork, study hard, keep up with deadlines, earn higher grades, and/or seek out relevant work-related opportunities to their career of interest can have a substantial impact on their future chances of getting a job. Similarly, changes one enact to be more organized, reliably show up to work and complete their necessary duties, and finish tasks in a careful, thorough manner can make them more likely to keep their job and remain employed. Indeed, past work has found that job attainment and changes in conscientiousness are positively associated (Roberts et al., 2003)—highlighting that changes in traits are associated with changes in important outcomes, even though levels of the traits are as well.

The importance of both cumulative and proximal processes points to a few interesting implications. First, it somewhat opposes theories of personality that heavily emphasize situations or context as being the sole determinants of the consequences of one’s personality (Mischel & Shoda, 1998). Although much can be gained by considering the dynamic nature of personality and how it interacts with one’s environment, it seems to be a fruitless effort to continue doing so while ignoring the undeniable impact that who one “is” on average has on their future outcomes. Second, the malleability of personality and its tendency to change over time is indisputable (Bledorn et al., 2022; Wright & Jackson, 2023), but it appears that the degree to which one changes does not, on average, overpower the impact that their previous level of personality has on their outcomes. While someone can change in their personality, and this change can sometimes be meaningfully associated with outcomes, it will likely never fully negate the effects associated with their previous static levels. Third, to maximize predictive validity when using personality, a multitrait assessment, lifespan approach is needed. According to the differential pathways hypothesis, the pathways that explain why personality traits impact future outcomes may differ at various points throughout the lifespan (Hill et al., 2019). Indeed, past work taking this lifespan approach has found differential predictive validity when using childhood versus adult-based personality (Wright & Jackson, 2022). The present study shows that this matters at different points in adulthood as well for some traits and outcomes.

Outcome-Specific Pathways

In our study, meta-analytic results for health outcomes showed fewer change associations than past work examining health markers (e.g., Magee, Heaven, & Miller, 2013; Takahashi et al., 2013; Turiano, Pitzer, et al., 2012). Changes in personality were only associated with self-reported health status, sometimes health limitations, and rarely number of reported mental problems. Most effects interestingly emerged for changes in openness, followed by neuroticism, and then conscientiousness and extraversion. When present, though, effects for conscientiousness were largest in magnitude. The directions of effects were in line with what past work has found for trait changes (Chow & Roberts, 2014; Human et al., 2013; Jokela et al., 2018; Letzring et al., 2014; Magee, Heaven, & Miller, 2013; Mroczek & Spiro, 2007; Siegler et al., 2003; Sutin et al., 2013; Takahashi et al., 2013; Turiano, Pitzer, et al., 2012) as well as for the respective trait levels. The lack of finding associations for changes in the Big Five traits with BMI is also consistent with past work focusing on broad Big Five domains (Sutin et al., 2011).

While there were no conflicting directions of associations with past work on health outcomes, there were sometimes discrepancies in the presence or absence of effects. For instance, changes in openness are typically not associated with self-reported health status, whereas changes in the other Big Five traits are frequently associated with this outcome. Static levels of openness are generally associated with better health and lower mortality risk (Iwasa et al., 2008; Jackson et al., 2015; Taylor et al., 2009)—likely due to factors such as intelligence and academic attainment (Deary et al., 2008; Gottfredson & Deary, 2004)—but the ability of changes in openness to predict higher self-reported health status is less likely to occur through this pathway. This is due to the relatively stable nature of cognitive ability (i.e., less likely to suddenly change unless there is some external or biological factor mediating change; Rönnlund et al., 2005; Schalke et al., 2013) and how its effects would take place over years of more stable environments, higher socioeconomic conditions, and benefits derived from a good education and career associated with cognitive ability (Gottfredson & Deary, 2004; Hart et al., 2007; Petrill et al., 2004). In contrast, changes in openness could indicate that people are experiencing sudden health changes, which are then reflected by lower ratings of self-reported health status. Additionally, individuals with higher cognitive ability tend to report more active lifestyles and a higher frequency of engagement in activities (Hultsch et al., 1999), and diversity in activity engagement has been shown to reduce the risk of cognitive impairment (Carlson, 2011). Thus, given the association between openness and cognitive ability (DeYoung et al., 2005; Parisi et al., 2009), and past work showing that diversity in activity engagement helps link these two constructs (Jackson et al., 2020), changes in openness could reflect a newer, more active and engaged lifestyle and improve health in that manner.

For our education outcome, only effects for changes in conscientiousness were associated with degree attainment while past work has found it with changes in neuroticism and extraversion, but not conscientiousness (Hoff et al., 2021). It is likely that the experience of being in school forces one to develop habits that are typical of the conscientious person, such as being responsible, keeping up with deadlines, staying organized, etc. (Brandt et al., 2019). However, not every person can successfully do this. That is, individuals who fail to display these necessary behaviors, which hinders them from successfully completing school, do not go on to receive their degree. It has been theorized that the many repeated instances of these new behaviors and incorporation of them into one’s daily lifestyle underlay the observed trait development, as these behavioral manifestations are important to the trait (Wrzus & Roberts, 2017). Thus, individuals who regularly engage in these behaviors increase in conscientiousness and are also more likely to obtain their degree. Indeed, past work has found that changes in conscientiousness predict educational achievement (Nofle & Robins, 2007), suggesting that individuals who do manage to increase in this trait do so via changing behaviors that promote educational success.
For the financial outcome, only effects for changes in conscientiousness were associated with salary, which is somewhat in line with past research (Converse et al., 2018; Hoff et al., 2021). However, unlike Hoff et al. (2021), we did not find that changes in extraversion or neuroticism were related to income—although, this effect only replicated in one sample in their study. Notably, though, there were effects of static levels for neuroticism in our study—the largest across all traits—both when controlling for initial salary and when not controlling for it. The lack of association for changes in neuroticism could arise for a couple reasons. Mainly, it could be that static levels just matter more. That is, cumulative processes overpower any effects rather due to proximal mechanisms. Considering the way in which one obtains a high salary, this makes sense. The process of attending school, getting good grades, obtaining a degree, starting a job, and receiving promotions and/or raises to have a high salary takes many years. As for the lack of extraversion effects, in our study, neither levels nor changes in this trait were associated with salary. However, it is worth noting that these associations did emerge when considering the individual data sets. As the exact content of each extraversion scale was not identical across studies, it could be that multiple aspects of extraversion were assessed, not all of which are related to salary, and past research has found this with association when using an extraversion scale capturing more social vitality facets (Hoff et al., 2021).

For relationship outcomes, results for changes in traits were relatively nonexistent. No results were found when simply predicting static levels of a distal outcome—but two associations emerged when predicting changes in marital and divorce status. Prior to marriage, it is likely one is in a (relatively) long-term relationship in which regular commitment and responsibility are expected, and the behaviors enacted in a marriage that are believed to lead to changes in maturity-related traits are present as well in these serious intimate relationships (Neyer & Asendorpf, 2001; Robins et al., 2002). If one is successful at these behaviors that make a relationship successful, then it is understandable that those who adopt these behaviors into a regular routine experience increases in conscientiousness and have a relationship that proceeds to marriage. Similarly, if someone in a marriage begins to explore other interests, perhaps through seeking new hobbies, diversifying their friend group, pursuing a new career, and so forth, then increases in openness might be expected to follow (Jackson et al., 2020). If these changes are associated with an individual realizing their life is not aligned with the new life they wish to pursue, or their spouse does not like the newly enacted changes to their personality or lifestyle, especially as most spouses have similar scores on openness (McCrae, 1996) and this is associated with marital satisfaction (O’Rourke et al., 2011), then divorce might be likely.

Last, no changes in any of the “mature” traits were associated with volunteer work, and static levels of these traits mostly were not either (Lodi-Smith & Roberts, 2012). However, levels of neuroticism were robustly associated with a lower likelihood of volunteering as well as decreases in one’s volunteering activity, whereas levels of agreeableness predicted a greater likelihood of volunteering. It appears that, on average, static levels of the maturity-related traits (i.e., agreeableness, neuroticism) are more important for predicting if one volunteers. It may be the case that having certain levels of these traits is important for leading one to volunteer early on, and it is this early engagement in volunteering that then predicts subsequent volunteering (Marta et al., 2014; cf. Mike et al., 2014).

Moderators of Meta-Analytic Effects

Past work has discussed how various design characteristics have implications for inferences drawn from analyses and their theoretical contributions (Hopwood et al., 2022); thus, we conducted an empirical examination of these factors in our study. Generally, our effects were invariant with respect to these variables, possibly due to the somewhat similar sample and design profiles across the data sets, but there were three study-level moderators sometimes associated with meaningful deviations from the average pooled effects.

For change associations, data sets with older participants generally had associations that were larger in magnitude when a study age moderation effect was present. As the only data set in our study that deviated from the average age by having older than average participants, this indicates that HRS generally had stronger effects than the average pooled effects for these associations. Accordingly, this also indicates that NLSY had weaker associations, or sometimes even associations in the opposite direction of the pooled average effect, as they had an average age that was younger than the overall average age per study.

Somewhat surprisingly, the number of waves in a study was not associated with change effects, but rather was only associated with static trait-level associations. When present, these effects generally indicated that level associations were larger in magnitude than the pooled effects for studies with a greater than average number of Big Five waves. These findings suggest that the reliability of change assessments may be similar across different numbers of waves of assessment—or at least comparable enough to where its predictive utility is not affected.

Last, internal consistency estimates of the Big Five were moderators of both level and change associations. For university degree status, associations for openness and conscientiousness had the effect of further magnifying the pooled average effect in studies with measures that had higher internal consistency values, suggesting that the “true” effect might be larger when measured with better assessments that more accurately capture the trait. For marital status and salary, it was the case that having higher than average internal consistency estimates for openness and extraversion, respectively, exacerbated the magnitude of the pooled effects. This suggests that the data sets with lower than average internal consistency estimates (NLSY and SHP) were perhaps underestimating or misrepresenting the true association due to poorer measurement of the traits. Though, and similar to number of Big Five waves, change-outcome associations appear to be mostly robust against using measures with poorer properties.

Limitations and Future Directions

First, while a strength of our study was our use of multiple longitudinal data sets, it is advantageous to have as many data sets as possible when examining pooled average effects. Although most outcomes were present in all seven data sets, three outcomes had pooled effects estimated from only five or six data sets. Furthermore, when examining moderators of these pooled average effects, five to seven data points are likewise not the ideal amount of information. The pooled average effects highly resembled the
individual associations present in the two largest data sets—and although greater weight should be attributed to these larger samples—a larger number of data sets to average across would likely be ideal to obtain more holistic estimates.

Second, the countries included in our study all fit the traditional “WEIRD” description. Thus, generalizations to samples from countries that do not share these key similarities might not be warranted and future research should incorporate more diverse samples into this type of work.

Third, given that we restricted our prediction to focusing on prospective outcomes, we did not look at potential bidirectional associations between the outcomes and traits. Given the stress placed on the importance of environmental factors and life experiences for personality development theory, it is likely that there are reciprocal associations occurring between traits and outcomes that serve to influence one another across time. For some of our outcomes, this is less likely to occur (e.g., university degree) compared to others (e.g., self-rated health). However, for those outcomes that do change more, it could be the case that it is not changes in the traits that are causally driving these associations, but rather changes in the outcome that elicit changes in the traits. When controlling for initial outcomes, we indeed did find associations between changes in traits and changes in outcomes. However, we are still not able to precisely disentangle the directions of associations. The source of these associations is valuable to know both theoretically and practically (e.g., in the case of interventions), but the present study only focuses on descriptive associations rather than causal pathways.

Fourth, our study emphasized personality changes that occur over longer periods of time, but changes that occur at shorter time scales could be just as important to consider as it is theorized these are what lead to long-term personality change (Wrzus & Roberts, 2017). A future direction is to examine if changes at these shorter time scales are also related to outcomes, above and beyond trait levels. However, it would be important to tease apart if these are true changes and not simply variability or fluctuations from one’s daily or weekly average levels, though.

Last, the ability of personality traits to predict changes in outcomes was likely somewhat hampered by the tendency for some outcomes to not change much, on average, from the initial to final time points. This typically affected some outcomes more than others (e.g., number of marriages) and affected some outcomes in some data sets more than others (e.g., the older HRS sample relative to the youngest NLSY sample for university degree attainment). Thus, these results should perhaps not be interpreted as a firm conclusion that changes in traits do not matter for these outcomes.

Conclusion

This study showed that changes in most of the Big Five personality traits are prospectively related to numerous outcomes, above and beyond associations due to static levels of traits. This is especially true when predicting newer changes in functioning, highlighting the role of proximal personality processes. These results indicate that personality trait change does sometimes matter—at least in the long-term—as meaningful associations were found over a minimum of a decade of time. Trait levels do appear to have more of an impact on outcomes, though, in terms of more numerous and stronger associations. Overall, our findings suggest that personality trait change has a valuable role in one’s personality development and environmental interactions, with the processes relating this change to future outcomes emerging independently of those connecting static levels to future outcomes.

References


Brandt, N. D., Mike, A., & Jackson, J. J. (2019). Do school-related experiences impact personality? Selection and socialization effects of


high-risk status. Psychosomatic Medicine, 65(5), 738–745. https://doi.org/10.1097/01.PSY.00008853.25140.9C


Accepted May 15, 2023