Title

Integrating Prospective Longitudinal Data: Modeling Personality and Health in the Terman Life Cycle and Hawaii Longitudinal Studies

Journal

Developmental Psychology, 50(5)

ISSN

0012-1649

Authors

Kern, Margaret L
Hampson, Sarah E
Goldberg, Lewis R
et al.

Publication Date

2014-05-01

DOI

10.1037/a0030874

Peer reviewed

Developmental Psychology
2014, Vol. 50, No. 5, 1390–1406

0012-1649/14/$12.00 DOI: 10.1037/a0030874

Integrating Prospective Longitudinal Data: Modeling Personality and Health in the Terman Life Cycle and Hawaii Longitudinal Studies

Margaret L. Kern
University of Pennsylvania

Sarah E. Hampson and Lewis R. Goldberg
Oregon Research Institute, Eugene, Oregon

Howard S. Friedman
University of California, Riverside

The present study used a collaborative framework to integrate 2 long-term prospective studies: the Terman Life Cycle Study and the Hawaii Personality and Health Longitudinal Study. Within a 5-factor personality-trait framework, teacher assessments of child personality were rationally and empirically aligned to establish similar factor structures across samples. Comparable items related to adult self-rated health, education, and alcohol use were harmonized, and data were pooled on harmonized items. A structural model was estimated as a multigroup analysis. Harmonized child personality factors were then used to examine markers of physiological dysfunction in the Hawaii sample and mortality risk in the Terman sample. Harmonized conscientiousness predicted less physiological dysfunction in the Hawaii sample and lower mortality risk in the Terman sample. These results illustrate how collaborative, integrative work with multiple samples offers the exciting possibility that samples from different cohorts and ages can be linked together to directly test life span theories of personality and health.

Keywords: integrative data analysis, collaborative studies, life span perspective, child personality, five-factor model

Supplemental materials: http://dx.doi.org/10.1037/a0030874.supp

A single research study necessarily has limitations, including its design, the sample characteristics, the measures used, the length of follow-up (if any), and the historical period in which it takes place. However, when findings replicate across studies, we gain confidence in those findings, in the theories on which they are based, and indeed in psychological science as a whole. Although possible causal pathways often emerge from cross-sectional and short-term studies, longitudinal studies are especially valuable for understanding developmental processes, and considerable time, effort, and research funds have been spent during the past century developing longitudinal datasets that include information about individual differences in personality and later health outcomes (Greenhoot & Dowsett, 2012). Friedman, Kern, Hampson, and Duckworth (2014) suggest that existing studies can be integrated to test theories of personality, health, and development across the life span. An important question is the extent to which studies can indeed be combined. In this article, we give an overview of integrative methods for longitudinal studies and then examine the potential for—and challenges to—integrating and comparing findings from two life-span studies: the Terman Life Cycle Study and the Hawaii Personality and Health Longitudinal Study.

Integrating Findings Across Studies

There are a number of strategies for combining findings across studies, the most common being sequential independent replication, in which one study finds a relation between two constructs, and then other independent studies are performed (either by the same researchers or others) to replicate and extend those findings. For example, Friedman and colleagues (1993) found that children rated high on conscientiousness were at a lower risk of dying at any given age than children rated low on conscientiousness. This finding was then replicated in other studies with quite diverse participants (e.g., Christensen et al., 2002; Weiss & Costa, 2005; Wilson, Mendes de Leon, Bienias, Evans, & Bennett, 2004).

A second strategy involves meta-analytically combining standardized effect sizes from multiple studies to find an overall average effect and moderators of this effect (i.e, aggregate data meta-analysis, or AD). For example, by combining the findings from 20 studies that had examined conscientiousness and mortality, an overall significant protective effect of high conscientiousness was determined (Kern & Friedman, 2008). Conclusions from meta-analytic work are more robust than conclusions from any single study, with greater precision and increased statistical power, although meta-analyses remain limited by the characteristics of the studies

This article was published Online First December 10, 2012.

Margaret L. Kern, Department of Psychology, University of Pennsylvania; Sarah E. Hampson and Lewis R. Goldberg, Oregon Research Institute, Eugene, Oregon; Howard S. Friedman, Department of Psychology, University of California, Riverside.

This research was supported by Grant AG020048 from the National Institute on Aging.

Correspondence concerning this article should be addressed to Margaret L. Kern, Department of Psychology, University of Pennsylvania, 3701 Market Street, Philadelphia, PA 19104. E-mail: mkern@sas.upenn.edu

that are included (Cohn & Becker, 2003; Rosenthal & DiMatteo, 2001; Tak, Meijer, Manoharan, de Jonge, & Rosmalen, 2010).

A newly emerging third strategy involves integrative techniques, in which data are directly combined. Methods include (a) pooled data analysis, in which data from two or more independent samples are combined as a single dataset and analyzed as a single study (Curran & Hussong, 2009); (b) longitudinal item-response modeling, in which responses are aligned to an assumed underlying scale using item response theory (IRT) methods, and then change over age is examined (McArdle, Grimm, Hamagami, Bowles, & Meredith, 2009); and (c) meta-analysis of individual participant data (IDA; or mega-analysis), in which, like in aggregate level (AD) meta-analysis, a comprehensive review of existing studies is performed, but the raw participant data from each study, rather than standardized effect sizes, are aggregated (Cooper & Patall, 2009; Stewart & Clarke, 1995). Such techniques are challenging and often not feasible, but when possible, these integrative methods offer multiple advantages including replication across samples, increased statistical power, broader psychometric assessments of constructs, more extended periods of the life span, and greater emphasis on data sharing and collaboration (Hofer & Piccinin, 2010).

Finding Measure Similarity

Without some degree of measurement similarity, data cannot be integrated (Bauer & Hussong, 2009). Ideally, constructs are measured with the same measures and consistent procedures across multiple samples. Such an approach is increasingly used in biogenetic, economic, and medical research (e.g., Eldevik et al., 2010; Hallahan et al., 2011; Manichaikul et al., 2012; Olgiati et al., 2012; Salimi-Khorshidi, Smith, Keltner, Wager, & Nichols, 2009; Serretti, Cusin, Rausch, Bondy, & Smeraldi, 2006). Similarly, the pharmaceutical community has established internationally accepted standards and procedures for harmonizing drug-related research, reducing duplication while promoting public health worldwide (International Conference on Harmonisation, 2010). However, in archived social-behavioral studies, the same measures more often than not are unavailable. Longitudinal studies might have a handful of items that tap any particular construct. Even within a single study, measure equivalence can be a problem, as the same constructs may be measured differently across assessments, as newer measures are validated; as the underlying theories, the main goals of the study, and the investigators change over time; and as the cohort reaches different developmental milestones.

New methods have been developed to address intra- and inter-study measurement variance. Perhaps the most straightforward (though often not easy) method of creating similarity involves harmonization, in which variables are aligned or recoded to match across studies. For example, Cooper et al. (2011) harmonized physical capability measures across eight United Kingdom cohort studies and tested age and gender interactions. Schaap et al. (2011) harmonized demographics, socioeconomic status (SES), chronic disease, anthropometry, physical performance, grip strength, pain, self-perceived health, and hospitalization items across five European cohort studies, with high-quality overlap. Other variables were of questionable quality or could not be harmonized. Bath, Deeg, and Poppelaars (2010) harmonized data from two longitudinal cohort studies in the Netherlands and the United Kingdom. To align items, variables were transformed and recoded, mostly as dichotomous variables. Twenty-six variables could be harmonized. The authors noted numerous issues that necessarily must be considered in harmonization, including sample differences, the time interval from baseline to follow-up, selective and nonselective attrition, and phrasing of questions and response categories.

Alternatively, statistical methods, such as latent factor analysis, IRT, and nonlinear factor analysis, now make it possible to establish equivalence and align items through latent approaches (Bauer & Hussong, 2009; Curran et al. 2008). For example, in latent factor analysis, potentially relevant items with broad distributional variance are chosen to represent this latent construct in each sample. In a multigroup analysis, factor loadings, mean values, and unique variances for the common items are then equated across the two samples to test for weak, strong, or strict invariance, respectively. In many cases, if factor loadings can be matched for at least half of the variables, partial invariance can be established (Reise, Widaman, & Pugh, 1993). As the number of invariant variables increases, the more confident one can be that the scales are measuring the same construct in the two samples. Stemming from testing in education, IRT was originally used with dichotomous items to establish similarities based on the difficulty and discriminating power of each item. More recent IRT formulations have incorporated generalized linear mixed effects modeling, allowing any outcome distribution within the exponential family for parametric models, as well as numerous nonparametric model variants (e.g., Rijmen, Tuerlinckx, De Boeck, & Kuppens, 2003). Nonlinear factor analysis allows both categorical and continuous variables to be included and compared (Bauer & Hussong, 2009).

These techniques assume that a continuous construct exists that can be measured along a latent scale. A measure or item assessed in a particular sample is assumed to tap part of this underlying construct, and the goal is to empirically place the items along this ruler, providing a basis for comparing the samples directly. This concept of an underlying latent construct has been used extensively in educational testing services, such as the Graduate Record Exam (GRE). Computer adaptive testing (CAT) techniques create a large bank of items and align people across an underlying dimension, and then the exam exploits this supposed latent distribution so that different items across different samples are comparable.

A third sample can be used to bridge samples (Martin & Friedman, 2000; McArdle et al., 2009), essentially defining this underlying ruler. A new group of participants, matched as closely as possible on age, gender, and other key characteristics, completes the measures from each archival sample. Item invariance is established between each sample and this third sample. The third sample thus acts as a structural bridge between the two main samples. For example, Martin and Friedman (2000) had participants complete the Revised NEO Personality Inventory (NEO-PI-R) and personality questions that were included in the original archival Terman study. Invariance across the archival and contemporary samples was established by confirming that items loaded on the same factors for both samples. In the contemporary sample, scales based on the archival items were highly correlated with four of the five NEO personality traits, thus aligning the archival items along the NEO factors and creating modern interpretable scales that could be derived from these 50- to 70-year-old archival data. Such bridging of studies may be particularly important for aligning

studies across the life span, using overlapping items and age or time periods as structural bridges.

To date, most studies in developmental psychology that consider links between personality and health have been conducted in independent samples. We suggest that more rapid progress can be made by creatively exploiting existing datasets, and such collaborative integration is needed to move the field from simple second-generation personality–health models to complex third-generation models (Friedman et al., 2014). Dynamic processes can be tested, providing a stronger theoretical foundation for effective interventions. Drawing on missing data designs, multilevel modeling, and Bayesian-based techniques, within-person trajectories and between-group variables can be combined, increasing power and precision (Riley, Simmonds, & Look, 2007; Sutton, Kendrick, & Coupland, 2008). That is, by linking studies from different cohorts and developmental ages, using areas of overlap as structural bridges between studies, we can potentially piece together life-span processes. Because we cannot randomly assign persons to personality traits or health conditions, some of the most informative research designs will involve longitudinal studies of widely varying samples.

In the current study, we applied this collaborative framework to two relatively large prospective studies, the Terman Life Cycle Study and the Hawaii Personality and Health Longitudinal Study, to examine the extent to which these samples could be directly integrated. We highlight both possibilities and potential pitfalls and then suggest directions for future collaborative efforts.

The Current Study

One of the greatest benefits of attempting to integrate data across studies is the explicit testing of measure comparability. Whereas meta-analyses rely on standardized effects that are based on each investigator's conceptualization of the constructs, integrative analyses require returning to the items themselves in order to theoretically and empirically establish equivalence. In the current study, we examined a model in which childhood personality traits predict midlife health (see Figure 1). The model we tested here was purposely simple and not meant to be a full model of personality and health. Our goal was to demonstrate the potential for and challenges to linking studies; by starting with a simple framework, we can then build more complex models, linking pieces of the personality–health puzzle together across the entire life span.

In health psychology, recent studies have increasingly used a five-factor model of personality traits as a framework for structuring and understanding personality–health relations (Smith & Williams, 1992). There are essentially two five-factor models, one stemming from early lexical studies (e.g., Goldberg, 1990, 1993) and the other stemming from later work by Costa and McCrae and captured by the NEO-PI-R personality questionnaire (Costa & McCrae, 1992). In the present study, we followed the lexical tradition, in which the five factors are typically labeled as (I) Extraversion, (II) Agreeableness, (III) Conscientiousness, (IV) Emotional Stability, and (V) Intellect/Imagination.

For the purpose of this study, we defined health as a multidimensional construct that includes both physical and mental components. Multiple pathways link personality and health, including health behaviors, social relationships, situational selection, underlying genetic or biological differences, and physiological changes (Kern & Friedman, 2010). Characteristics of the sample may moderate personality, behavioral, and health relations. In this study, we examined health behaviors and educational attainment as mediators linking child personality and midlife health using two longitudinal samples.

The Terman Life Cycle Study was begun in 1921–1922 by Lewis M. Terman. Participants were followed prospectively throughout their lives until death and were assessed on thousands of psychosocial variables, including child and adult personality traits, social relationships, health-related behaviors, and lifelong health status. The Hawaii Personality and Health Longitudinal Study was begun in 1959 by John M. Digman. Attempts to recontact participants began in 1998, and follow-up efforts continue today. Paralleling the Terman archive, there are data on child and adult personality traits, social relationships, health-related behaviors, and adult health status. Importantly, in both samples teachers assessed the child participants on personality traits, and 40 years later participants completed various health and behavioral measures. The two studies have followed thousands of participants across many decades, have similar measures and structures relevant to personality and health theory, and together offer an example of the possibilities for integrating longitudinal studies. The Terman sample has complete life-span data (childhood through mortality), whereas the Hawaii sample has more health-related measures, including medical/physiological measures at midlife.

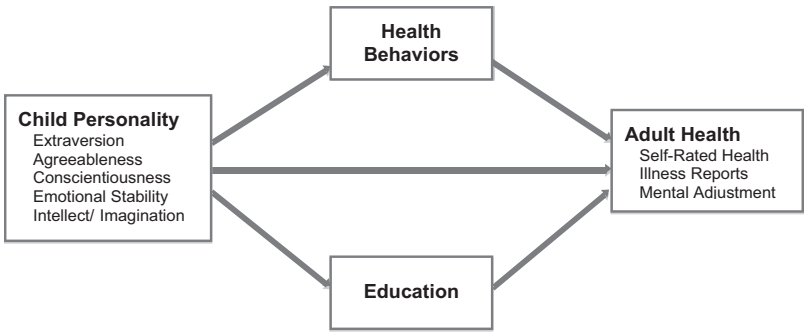

Figure 1. Conceptual model. In this model, child personality traits predict adult health, partially mediated through educational attainment and health behaviors (for simplicity, we focus in this article on alcohol abuse, but other behaviors could be included).

Working within a single dataset involves a large commitment in terms of gathering the data and understanding the intricacies of the sample, which only becomes more complicated when more studies are added. Each dataset is unique and complicated, and we drew on each investigator's unique expertise with the datasets to inform the models and analyses. Rather than a complete harmonization of variables assessed in the two samples, we aligned a minimal number of variables to specifically examine a theoretical model of personality and health across the two samples. We detail our harmonization process, testing the extent to which we can align the studies and then use unique aspects of each study (Hofer & Piccinin, 2010). We had three goals: (a) to determine whether equivalent measures of child personality traits could be established in the Terman and Hawaii samples; (b) to test a theoretical model of childhood personality, education, behavior, and adult health, using parallel analyses (in each sample individually) and integrative methods (directly pooling the data); and (c) to use similar items as a bridge and then extend analyses to unique aspects of each sample, linking child personality to mortality in the Terman sample and to physiological health measures in the Hawaii sample.

Method

Participants

Sample 1: The Terman Life Cycle Study. In the Terman sample, teachers across California were asked to identify the youngest and most intelligent students in their classes; those nominated were tested using the Stanford-Binet Intelligence Scale and were included in the study if their IQ was 135 or greater. Other children were added through 1928, yielding a total sample of 1,528 participants (856 males [M] and 672 females [F]). On average, participants were born in 1910 ($SD = 3.7$ years) and were 11 years old at the first assessment. Most came from a middle- to upper-middle-class background. Almost all were Caucasian; about 51% were English, Scottish, or Irish; 15.7% were German; 10.5% were Jewish (mostly Eastern European); and only a few Oriental, Mexican, African American, or other ancestries were represented (Terman et al., 1925). Follow-up efforts began in 1936, and by 1940 approximately 98% of the sample had been successfully contacted (Terman & Oden, 1947). Participants were followed prospectively throughout their lives, completing written assessments every 5–10 years, with the last formal assessment in 1999. In addition, our research team has supplemented the extensive archival information by collecting death certificates, and we have refined and independently validated various psychosocial measures, including measures of child personality traits, social relationships, health-related behaviors, and health status (Friedman, Kern, & Reynolds, 2010). The Terman study is the longest longitudinal study with multiple repeated assessments that has ever been conducted.

The current study was limited to participants with data on child personality traits and midlife health information. Of the total sample, 252 participants (141 M, 111 F) were excluded because they were missing child personality ratings, and 191 participants (101 M, 90 F) were excluded because they were missing midlife health information, leaving a final sample of 1,085 individuals (614 M, 471 F). Those missing personality data were born later, $t(1,526) = 10.23$, $p < .0001$; were more likely to report illnesses at midlife, $t(1,307) = 3.21$, $p = .001$; and came from a slightly higher socioeconomic background, $t(1,205) = 3.78$, $p = .0002$. Participants missing health data completed fewer years of education than those with health data, $t(1,499) = -9.88$, $p < .0001$; had a slightly lower childhood IQ, $t(1,526) = -2.33$, $p = .02$; and were more likely to die or be lost to follow-up at a younger age, $t(1,526) = -23.87$, $p < .0001$. These differences may narrow the relevant ranges somewhat and thus attenuate relations but are unlikely to introduce bias.

Sample 2: The Hawaii Personality and Health Longitudinal Study. In the Hawaii sample, elementary school teachers in Grades 1, 2, 5, and 6 were recruited by John M. Digman and asked to describe the children in their classes on personality traits, which were later matched to the five-factor structure (Goldberg, 2001). Starting in 1998 (age 41–50), the now middle-aged members (age 50–60) were being located and followed up (Hampson et al., 2001). Participants complete assessments periodically and are invited to attend a half-day session at a medical clinic, where they complete a battery of physical, medical (including blood tests), personality, social, cognitive, and behavioral measures. Of the original Hawaii cohort, 2,320 were potentially available for recruitment, and 1,928 have been located (83%). Of those, only 45 refused continued participation. The sample includes a mix of socioeconomic backgrounds. The adult sample is composed of Japanese Americans (37%), Native Hawaiian or part-Hawaiian Americans (21%), those of European ancestry (18%), and those of Filipino, Chinese, Okinawan, Korean, or other ancestry (24%).

The initial cohort included six subsamples based on school grade and location in Hawaii (see Goldberg, 2001, for details); for the present study, the subsamples were combined into a single sample. The current study was limited to participants with child personality and midlife health data. Of the initial cohort, 2,221 participants (1,176 M, 1,045 F) had childhood personality data. Of those, 1,051 participants (559 M, 492 F) were excluded because they were missing midlife health data, leaving a final sample of 1,170 participants (617 M, 553 F). Those excluded at midlife were described by their teachers as less conscientious, $t(2,219) = -4.21$, $p < .0001$, and less intellectual/imaginative, $t(2,219) = -2.95$, $p = .003$, as children.

Measures

We selected items that fit within the model noted in Figure 1, with child personality traits as the main predictors, educational attainment and alcohol abuse as potential mediators, and indicators of adult health status, illness, and mental adjustment as the main outcomes. In addition, we included physiological measures (in the Hawaii sample) and mortality information (in the Terman sample) as outcomes unique to each sample.

Child personality. In the Terman sample, as part of the initial 1922 assessment (average age = 11 years) parents and teachers were asked to rate their child, compared to average children of the same age, on 25 traits. Traits were assessed on a 13-point scale ranging from very low levels of the trait to very high levels. In our original study (Friedman et al., 1993), six personality factors were created using the average of parent and teacher assessments: cheerfulness, conscientiousness, energy, motivation/self esteem, sociability, and permanency of moods. In the Hawaii sample, teachers (Grades 1, 2, 5, or 6) were presented with a set of 39 to 63 traits (depending on the subsample), and they were given the names of all children in their class on pieces of cardboard. Similar

to a Q-sort, teachers sorted the children into a nine-step quasi-normal distribution for each trait. A five-factor structure was recovered, creating measures of extraversion, agreeableness, conscientiousness, emotional stability, and intellect/imagination (Goldberg, 2001). In addition, 11 middle-level clusters of traits, such as activity level, sociability, and perseverance, were developed. To establish comparable measures across the two samples, we returned to the original 25 trait ratings by teachers in the Terman sample and to the 39 traits common across the subsamples in the Hawaii sample (see the Appendix for traits and descriptions, and the Results section for the harmonization procedure).

Educational attainment. In the Terman sample, at each assessment participants indicated the highest level of education achieved and any additional schooling completed during each interlude. In our prior studies, a continuous total educational attainment score was constructed in terms of years of schooling. In the Hawaii sample, at the first adult assessment (1999) participants indicated the highest grade or year of school completed on a 9-point scale. To harmonize an index of educational level, the more detailed information in the Terman sample was reduced to the Hawaii scale.[1] Thus, in both samples education was coded as 1 = *eighth grade or less,* 2 = *completed junior high school,* 3 = *some high school,* 4 = *high school graduate/GED,* 5 = *some technical school,* 6 = *technical/nursing school graduate,* 7 = *some college or community college,*8 = *college graduate,* and 9 = *postgraduate or professional degree.* Educational information was available for all 1,085 participants in the Terman sample and for 1,136 participants (601 M, 535 F) in the Hawaii sample.

Alcohol abuse. As a marker of unhealthy behavior, we selected alcohol abuse.[2] In the Terman sample, the 1950 assessment (average age = 40 years) asked participants to indicate their typical alcohol use on a four-point scale (1 = *none or rare,* 2 = *moderate,* 3 = *fairly heavy,* 4 = *problem drinking*). In the Hawaii sample, considerably more information was available. To harmonize an index of alcohol abuse across the two samples, we reduced the more detailed information in the Hawaii sample to the Terman scale. At the first adult assessment (average age = 45 years), participants indicated how often in the past month they drank any alcoholic beverages and how many drinks they had when they drank. On the basis of this information, participants were coded as 1 (*none or very rarely consumed alcohol in the past month*), 2 (*moderate; some drinking in the past month, but no binge drinking, defined as more than two drinks at a time*), or 3 (*fairly heavy; periods of drinking three or more drinks at a time*). In addition, several questions in the first and third assessments asked about problem drinking, with questions about whether the participants felt any concern about the amount they drank or had experienced problems stemming from alcohol use. If participants were rated as fairly heavy drinkers and reported problems or concerns, then they were classified as 4 (*problem drinking*).[3] Alcohol data were available for all 1,085 participants in the Terman sample and for 1,014 participants (540 M, 474 F) in the Hawaii sample.

Adult health. Several measures were selected from each sample to assess physical and mental health when the participants were about 50 years old.

Self-rated health. As is typically reported in most studies of health, both samples indicated on a 5-point scale their general health in comparison with others of the same age and gender (1 = *poor,* 2 = *fair,* 3 = *good,* 4 = *very good,* 5 = *excellent*). Self-rated health data were available for all participants in both samples.

Illness. In the Terman sample, each assessment asked participants to report any major illnesses experienced since the last assessment. Reports of heart disease, stroke, cancer, breathing problems, major infections, injury, and other major illnesses through the 1960 assessment were summed (0 = *none,* 4 = *four or more reported*). In the Hawaii sample, the first adult assessment asked participants to report whether they had ever been treated for heart attacks, heart disease, stroke, cancer, diabetes, thyroid disease, high blood pressure, high cholesterol, migraines, chronic fatigue syndrome, or any other major illnesses. Conditions were summed (0 = *none,* 4 = *four or more reported;* the variable was reversed to create the composite health factor). Illness information was available for all 1,085 participants in the Terman sample and for 1,166 participants (614 M, 552 F) in the Hawaii sample.

Mental adjustment. In the Terman sample, the adult questionnaires asked the Terman participants whether they had experienced any tendency toward nervousness, worry, special anxieties, or nervous breakdown in recent years, and if so, the "nature of such difficulties," how the difficulties had been handled, and the status of their present condition. On the basis of this information and considered in the light of the total case history (i.e., all information available including personal correspondence with participants and their families), Terman and his associates classified each participant into one of three categories: (a) serious maladjustment, (b) some maladjustment, or (c) satisfactory adjustment (Terman & Oden, 1959). Individuals in the first category were those who showed marked signs of anxiety, depression, personality maladjustment, psychopathic personality problems, or suffered a nervous breakdown. The second category contained individuals who experienced feelings of inferiority or inadequacy, anxiety, or emotional conflict but were still able to function. The third category contained those who were able to cope normally with everyday problems and who were judged to be essentially typical in terms of their emotional make-up.

To approximate a similar measure in the Hawaii sample, we compiled variables that related to mental health from the five adult questionnaires. Questions concerned emotional problems over the past 4 weeks that limited functioning, depression (based on the Center for Epidemiological Studies Depression Scale), ratings on emotional health as compared to other individuals, reports of psychosomatic problems, and signs of posttraumatic stress disorder. Five trained research assistants rated each participant as *seriously maladjusted* (1), *some maladjustment* (2), or *satisfactory*

[1] Alternatively, the coarser Hawaii data could be mapped to the more nuanced Terman variable, as there is essentially a "years of education" metric in educational attainment. That is, *some high school* could be coded 10, *high school graduate,* 12, and *college graduate,* 16. The coarser scaling used here most likely attenuates the education relations and correlations with intellect.

[2] Based on the data available in each dataset, alcohol use information was the most straightforward health behavior to harmonize, as similar questions were available at the same age in both samples. Despite some evidence for modest cardiovascular protective effects of limited alcohol consumption, high alcohol use and abuse has repeatedly been documented as a significant health threat in Western societies.

[3] Some questions indicated an allergic reaction to alcohol. As many of the participants were of Asian descent, such reactions were not coded as problem drinking.

Table 1
Final Child Personality Factors, With the Traits That Were Included in Each Sample

Factor	Terman traits	α	Hawaii traits	α
Extraversion	Cheerfulness	.78	Assertive	.84
	Fondness for large groups		Energetic	
	Leadership		Gregarious	
	Physical energy		Socially confident	
	Popularity with other children		Lethargic (r)	
			Seclusive (r)	
			Submissive (r)	
Agreeableness	Generous	.74	Considerate	.83
	Sympathetic		Self-minimizing	
			Rude (r)	
			Spiteful (r)	
Conscientiousness	Conscientiousness	.83	Careful of personal belongings	.90
	Desire to excel		Careless of others' property	
	Prudence/forethought		Conscientious	
	Truthfulness		Persevering	
	Will power and perseverance		Planful	
			Fickle (r)	
			Irresponsible (r)	
Emotional Stability	Permanency of moods	Single item	Concerned about acceptance (r)	.72
			Jealous (r)	
			Touchy (r)	
Intellect	Common sense	.78	Curious	.81
	Intellect		Imaginative	
	Knowledgeable		Original	
	Originality		Verbal	

Note. See Appendix for trait descriptions. (r) indicates reverse-scored trait. α = coefficient alpha. Child personality was measured in 1922 (average age = 11 years) in the Terman sample and between 1959 and 1967 (Grades 1, 2, 5, or 6) in the Hawaii sample.

adjustment (3). Interrater reliability (*r*, complete agreement across all five raters) was .96, and discrepant cases were determined through discussion. Mental adjustment data were available for 1,084 participants (613 M, 471 F) in the Terman sample and 1,169 participants (617 M, 552 F) in the Hawaii sample.

Measures unique to each sample. The Terman sample is unique in that we have now collected life-span mortality data (death certificates) for 90% of the participants, but detailed physiological information is not available. The Hawaii sample does not include mortality information, but physiological data have been collected on a growing subsample. If the two studies can be linked, we can potentially extend analyses to unique aspects of the samples, essentially allowing one study to fill in missing pieces from the other study. We examined this possibility by including health measures unique to each sample.

In our prior studies (Friedman et al., 1993, 1995), higher levels of childhood conscientiousness were related to a lower risk of dying at any given age. The conscientiousness–mortality relation was partially mediated by unhealthy behaviors. In the present study, age of death information was available for 984 participants (791 M, 572 F). The remaining 101 participants were censored at the last age that we heard from them. Some are known to still be alive.

In the Hawaii study, global health status at midlife (age 50) was evaluated by a composite variable of physiological dysregulation comprising multiple widely used biomarkers of the cardiovascular and metabolic systems indicative of allostatic load (Seeman, McEwen, Rowe, & Singer, 2001). Previously, Hampson, Goldberg, Vogt, Hillier, and Dubanoski (2009) demonstrated that this measure is related to lower self-rated health, depressive symptoms, and several health behaviors. Physiological dysregulation scores were derived from 10 clinically assessed biomarkers (systolic and diastolic blood pressure, total cholesterol, triglycerides, total-to-HDL cholesterol, urinary protein, fasting blood glucose, waist-to-hip ratio, body mass index, and use of blood pressure or cholesterol lowering medications). Using this information, we standardized within gender and summed the markers to create dysregulation scores. Higher scores indicate more physiological dysregulation. Physiological data were available for 619 participants (305 M, 314 F).

Data Analyses

The primary challenge was to harmonize measures across the two studies. For child personality, we conceptually aligned items in each sample to the five-factor model, iterating between theoretical and empirical consideration. Traits were then standardized and averaged to create composite personality factors. Health, education, and alcohol variables were recoded to align response categories across the two samples, as noted under the Measures section. A composite health variable was created by standardizing and summing self-rated health, mental adjustment, and illness reports (reverse scored).

Next, we conducted parallel analyses in the two samples. Using hierarchical linear regression in SAS (Version 9.2) software, the composite personality factors were regressed on the health composite variable separately in each sample, controlling for age and sex. Education and alcohol use were then added to the models as mediators. To allow for missing values on some items and to more appropriately address mediation (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002), we estimated the full model as a structural

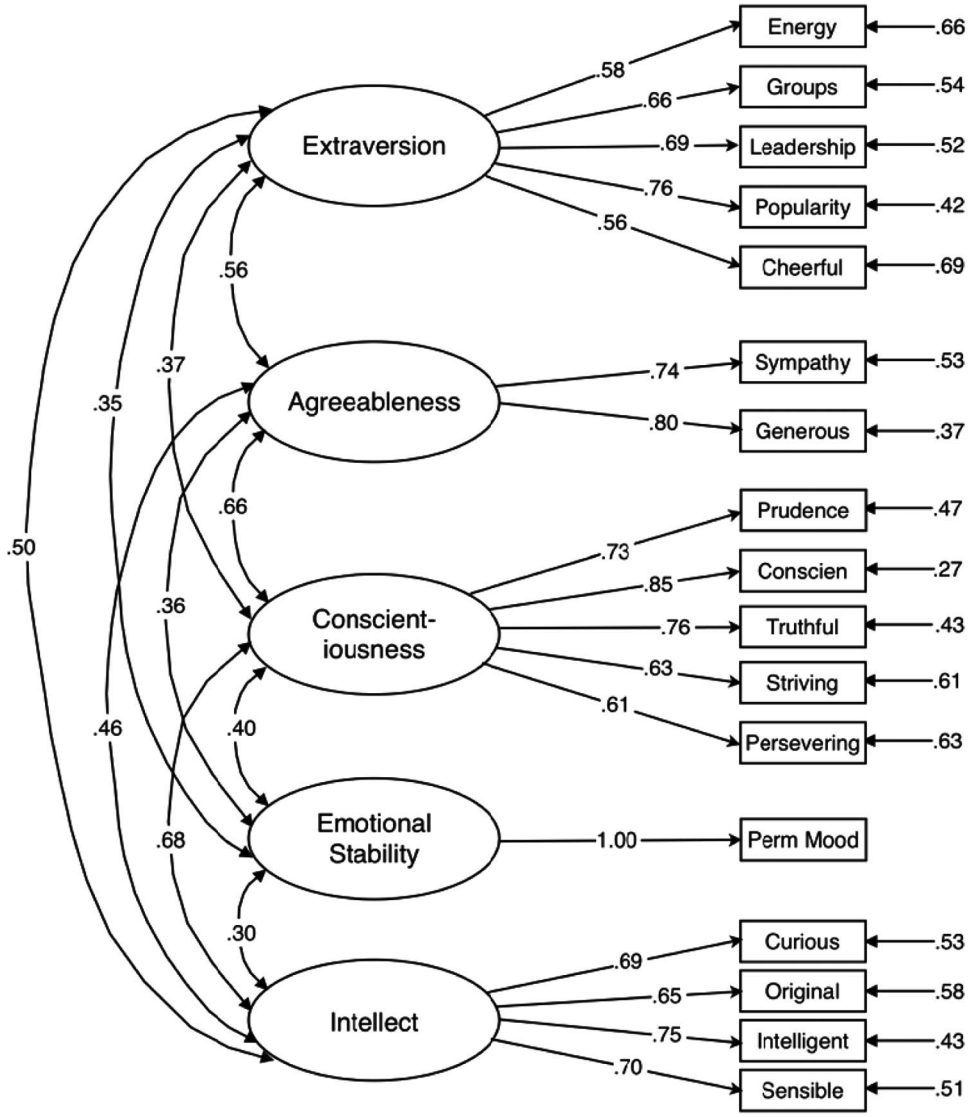

Figure 2. Final child personality (average age = 10 years) measurement model in the Terman sample. See Appendix for traits and descriptions. Standardized path estimates are presented. Emotional stability was evaluated as a single observed variable. Model was estimated in R (Version 2.15.0, package lavaan). Conscien = conscientiousness; Perm mood = permanency of moods.

equation model (SEM) in R (Version 2.15.0), using the lavaan package (Rosseel, 2012), with latent variables used for the personality and health variables. Using the composite harmonized variables, we then directly pooled the samples, creating a single larger dataset, and tested the SEM as a multigroup analysis, with factor loadings and regression pathways constrained to equality, but allowing intercepts and means to vary by group.[4]

Finally, using the aligned personality, education, and alcohol use variables, we extended analyses to the unique aspects of each sample, with personality predicting physiological dysregulation in the Hawaii sample using linear regression, and personality predicting mortality risk in the Terman sample using Cox regression (survival analysis). Analyses were conducted in SAS, using the composite personality and health variables.

All analyses controlled for age and gender. To assess model fit, we relied primarily upon the RMSEA (root mean square error of approximation), a population-based index that is not affected by sample size when the sample size is at least 200 (Curran, Bollen, Chen, Paxton, & Kirby, 2003) and which includes confidence intervals (95% CI). An RMSEA of less than .05 is considered a close fit, one of .05 to .08 is considered acceptable, and one of .08 to .10 is considered mediocre (Loehlin, 2004). We also examined the Tucker

[4] Although the personality factors were aligned to the five factors in each sample, the factors comprised varying numbers of items, so factor loadings cannot be constrained across groups. For simplicity, we used the aligned composite factors, rather than more sophisticated multigroup comparisons with differing individual items.

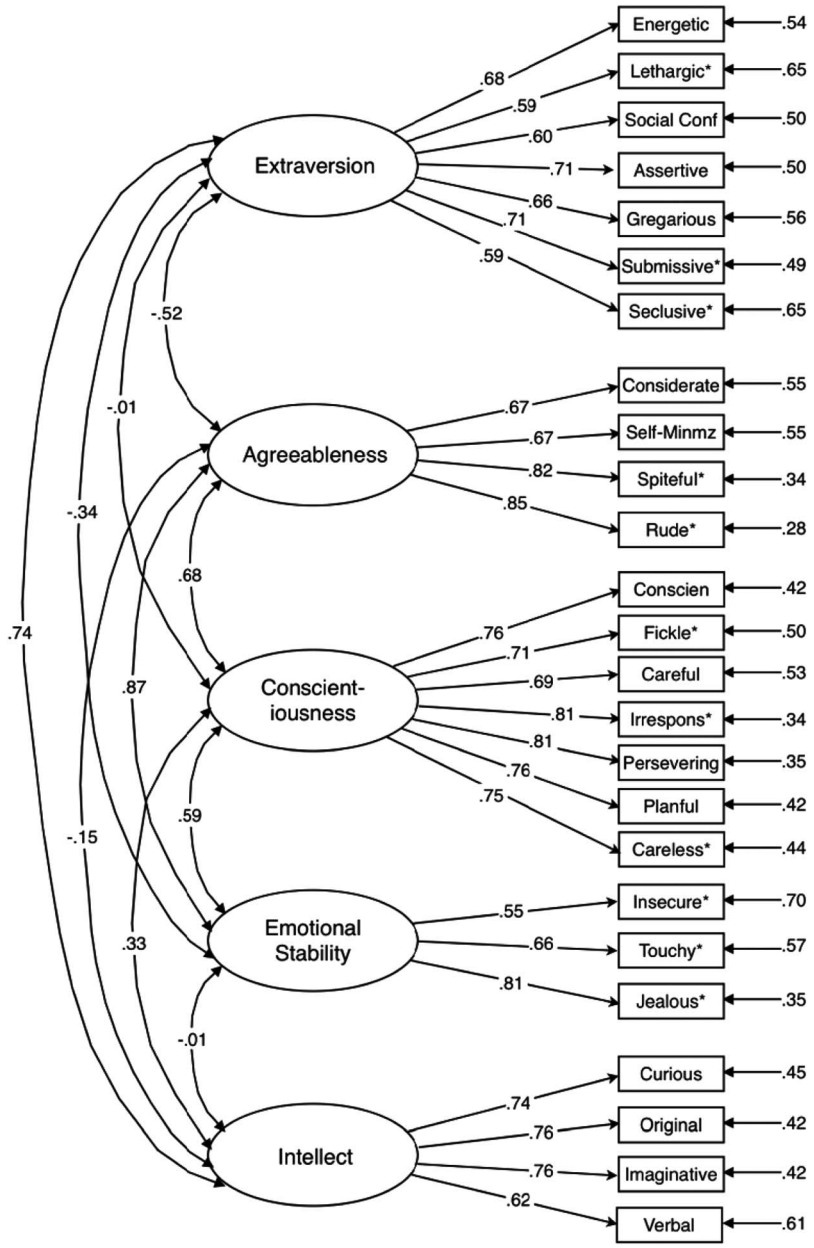

Figure 3. Final child personality (Grades 1, 2, 5, or 6) measurement model in the Hawaii sample. See Appendix for traits and descriptions. * indicates reversed-scored items. Standardized path estimates are presented. Model was estimated in R (Version 2.15.0, package lavaan). Social conf = socially confident; Self-Minimz = self-minimizing; Conscien = conscientious; Irrespons = irresponsible.

Lewis index (TLI), a population-based index in which the best-fitting model is closest to 1.00. For model estimation, the lavaan package allows missing values based on a missing at random assumption using the missing = "ml" option in the fit statement.[5]

Results

Establishing Comparable Child Personality Factors

The first part of the analyses aimed to establish comparable measures of child personality traits across the two cohorts. We considered the teacher assessments on 25 traits in the Terman sample and the teacher assessments on 39 traits consistent across all subsamples in the Hawaii sample. Using the Big Five personality model as a theoretical base, we identified traits relevant to each of the five constructs and then used factor analysis and intertrait correlations to create more pure measures of the factors. Then, in a final step, we tested the measurement model in each

[5] SAS and R scripts are available in the Appendix in the online supplemental material.

Table 2
Correlations Between the New Big Five Personality Factors and Prior Measures of Personality Traits in Each Sample

Personality trait	E	A	C	S	I
			Terman sample		
Sociability	**.69**	.26	.17	.14	.18
Cheerfulness	**.53**	.31	.18	.27	.36
Conscientiousness	.16	.47	**.74**	.33	.39
Permanency of mood	.26	.25	.29	**.76**	.23
High motivation	.32	.23	**.55**	.19	**.62**
Energetic	.42	.09	.06	.11	.10
			Hawaii sample		
Extraversion	**.80**	−.28	−.10	−.09	.33
Agreeableness	−.32	**.74**	.21	**.53**	−.11
Conscientiousness	.04	.43	**.83**	.16	.07
Emotional stability	.02	.26	.22	**.70**	.04
Intellect	.42	−.04	.24	.03	**.86**
Activity level	**.81**	−.24	.10	−.06	**.52**
Sociability	**.81**	−.28	−.06	−.11	.42
Self-assertion	**.81**	**−.71**	−.24	−.42	**.54**
Resilience	**.54**	.21	.49	.39	**.61**
Antagonistic	.39	**−.94**	**−.58**	**−.65**	.13
Impulsivity	.37	**−.73**	**−.70**	**−.59**	.10
Mannerliness	−.03	**.76**	**.78**	**.53**	.20
Carefulness	−.04	**.52**	**.83**	.35	.13
Perseverance	.07	**.53**	**.96**	.44	.31
Insecurity	.23	**−.68**	**−.50**	**−.97**	.03
Imagination	**.58**	−.10	.29	.04	**.97**

Note. See Goldberg (2001) and Friedman et al. (1993) for details on prior measures. Correlations > .50 are bolded. E = extraversion; A = agreeableness; C = conscientiousness; S = emotional stability; I = intellect. Child personality was measured in 1922 (average age = 11 years) in the Terman sample and between 1959 and 1967 (Grades 1, 2, 5, or 6) in the Hawaii sample.

sample to recover the five-factor structure and to confirm similar relations across the two samples.

First, the trait names and descriptions for each sample were matched across samples. Five trained raters indicated whether each of the 39 Hawaii traits were completely comparable, mostly comparable, somewhat comparable, or not at all comparable to each of the Terman traits, based on the descriptions provided (see the Appendix for traits and descriptions). For example, *energetic* in the Hawaii sample was rated as completely comparable to *amount of physical energy* in the Terman sample. Twenty-one traits from the Terman sample and 26 traits from the Hawaii sample were identified as being at least somewhat comparable.

Next, working with one sample at a time, we factor analyzed the 21 and 26 traits (for the Terman and Hawaii samples, respectively) using oblique rotations. Because the traits were all measured by teacher assessments, we expected a general evaluative "halo" effect such that factors were expected to be interrelated. A five-factor solution was recovered in each sample, although in the Terman sample, only one item tapped emotional stability (*permanency of moods*).

Iterating between theoretical consideration of the traits and empirical item interrelations (intertrait correlations and coefficient alpha reliability estimates), we examined each empirically derived factor, first in terms of the initial placements by the trained raters and later in terms of the scientific literature on the Big Five factor representation (e.g., Goldberg, 1990, 2001). A few traits identified by the raters were not considered pure indicators of the factors. Specifically, in the Hawaii sample, *restless* and *fidgets* were identified by the initial raters as part of Extraversion, but both personality theory and the correlations between these traits and the main factors suggest that these traits are a blend of low conscientiousness, high extraversion, and high neuroticism, rather than pure measures of extraversion. In the Terman sample, *appreciation of beauty*, *sensitivity to approval*, and *freedom from vanity* were identified as part of Agreeableness, but because they do not fit within our usual conceptions of the factor, these traits were not included in the final factors. In addition, a few traits that past studies have suggested should be included in the factors were not initially identified by the raters. We examined the extent to which the initially excluded traits correlated with each factor. In the Hawaii sample, *fearful* (reverse scored) was added to Emotional Stability, and *seclusive* (reverse scored) was added to Extraversion.[6] The final traits that were included in each factor and their reliabilities are summarized in Table 1.

We then used confirmatory factor analysis in each sample to evaluate the final personality measurement models. A five-factor model fit best in each sample[7] (Terman sample: TLI = .843, RMSEA = .086, 95% CI [.081, .091]; Hawaii sample: TLI = .813, RMSEA = .092, 95% CI [.089, .095]). Figures 2 and 3 present the final measurement models with standardized factor loadings for the Terman and Hawaii samples, respectively.

Finally, we computed correlations between our new factors and the personality variables used in our prior studies (see Table 2). In our prior work with each sample, the personality measures have been used elsewhere to predict numerous adult outcomes, supporting their validity. The measures used here were slightly altered: In the Terman sample, the original measures included parent reports, and here we only included teacher reports; in the Hawaii sample, slightly different traits for each factor were included to best align with the Terman items. Thus, the factors used here are similar to brief scales, and it is helpful to confirm that they correlate with our other measures to support their validity. In the Terman sample, conscientiousness and permanency of moods correlated .74 and .76 with the new Conscientiousness and Emotional Stability factors.[8] Sociability and motivation related, respectively, to the new Extraversion and Intellect factors. In the Hawaii sample, correlations with the prior five factors ranged from .70 for eEotional Stability to .86 for Intellect.

[6] One could question why items were identified by the raters and later rejected. The raters were psychology students who were trained to consider overlap in definitions and common word use, not placement within the five-factor model. They were not personality experts. Our final theoretical considerations involved decisions by an expert in the five-factor approach to personality (L. R. Goldberg).

[7] For comparison, we also examined a two-factor model, with the Conscientiousness, Agreeableness, and Emotional Stability traits on Factor 1 and the Extraversion and Intellect traits as Factor 2, and a four-factor model, in which Emotional Stability was excluded, as only a single item was available for the Terman sample. For both samples, the five-factor model was superior.

[8] In the original study (Friedman et al., 1993), although permanency of mood was reported as a single item, the study used the average of parent and teacher ratings, whereas the value reported here is based only on teacher assessments (for comparability with the Hawaii study).

Table 3
Descriptive Statistics and Intervariable Correlations in the Terman Sample, the Hawaii Sample, and the Combined Terman–Hawaii Sample for the Full Sample and Separately by Sex

Variable	Mean (*SD*)	E	A	C	S	I
			Terman sample (N = 1,085)			
Sex[a]	614 M, 471 F	.16***	.09**	.15***	−.01	−.04
Year of birth	1910 (3.10)	.01	−.07*	−.12***	−.02	−.02
Males	1910 (3.16)	−.08*	−.12***	−.14***	−.03	−.05
Females	1910 (2.98)	.09*	−.02	−.13***	.000	.03
Health composite[b]	9.20 (2.09)	.02	.02	.02	.06*	−.02
Males	9.42 (1.98)	.08*	.03	.04	.10*	−.02
Females	8.92 (2.19)	−.01	.03	.03	.009	−.03
Self-rated health	4.22 (0.72)	.03	−.004	.03	.05	.001
Males	4.29 (0.70)	.08*	.01	.05	.10*	.004
Females	4.13 (0.73)	.01	−.004	.03	−.02	−.01
Illnesses	0.89 (0.88)	.03	.02	.06	.02	.02
Males	0.80 (0.83)	−.02	.03	.04	.001	.01
Females	1.00 (0.94)	.05	−.02	.03	.05	.05
Mental adjustment	2.57 (0.64)	.04	.06	.07*	.10**	−.02
Males	2.59 (0.64)	.06	.07	.06	.11**	−.04
Females	2.55 (0.64)	.02	.05	.09*	.09	.000
Education	7.69 (1.44)	.01	−.003	.11***	.06	.07*
Males	7.80 (1.47)	.05	.05	.12**	.05	.06
Females	7.55 (1.40)	−.02	−.05	.12**	.07	.08
Alcohol abuse	1.84 (0.61)	.08**	−.05	−.11***	−.02	−.04
Males	1.95 (0.62)	.12**	.003	−.07	−.003	−.03
Females	1.71 (0.58)	.10*	−.07	−.11*	−.05	−.06
			Hawaii sample (N = 1,170)			
Sex[a]	617 M, 553 F	−.05	.21***	.24***	.01	−.07*
Year of birth	1955 (1.61)	.01	−.01	−.02	.02	−.001
Males	1955 (1.60)	.01	−.001	−.04	.04	.02
Females	1955 (1.62)	.02	−.04	−.01	−.001	−.02
Health composite[b]	9.66 (2.16)	.04	.05	.10***	.04	.05
Males	9.72 (2.02)	.02	−.003	.07	−.01	.02
Females	9.59 (2.30)	.05	.12**	.16***	.07	.09*
Self-rated health	3.38 (0.97)	.07*	.04	.10***	.03	.08**
Males	3.34 (0.95)	.02	.001	.06	.01	.01
Females	3.43 (0.99)	.12**	.07	.14**	.05	.17***
Illnesses	0.83 (1.00)	−.02	−.03	−.07*	−.03	−.01
Males	0.82 (0.98)	−.02	.002	−.04	.02	−.003
Females	0.84 (1.02)	−.02	−.07	−.12**	−.07	−.02
Mental adjustment	2.73 (0.57)	−.002	.03	.05	.02	.03
Males	2.78 (0.51)	.01	−.01	.05	.001	.02
Females	2.67 (0.63)	−.02	.11*	.10*	.05	.02
Education	6.56 (1.89)	.04	.13***	.23***	.10***	.18***
Males	6.43 (1.93)	.05	.13**	.23***	.10*	.23***
Females	6.71 (1.83)	.05	.10*	.20***	.10*	.13**
Alcohol abuse	2.18 (1.12)	.06*	−.06	−.09**	.01	.01
Males	2.44 (1.17)	.03	.02	.01	.01	−.03
Females	1.89 (0.98)	.08	−.04	−.10*	.04	.03
			Combined sample (N = 2,255)			
Sex[a]	1,231 M, 1,024 F	.05*	.15***	.20***	.003	−.06**
Year of birth	1934 (22.81)	.001	−.01	−.01	−.000	−.002
Males	1933 (22.97)	.09**	−.07*	−.06	−.01	.01
Females	1935 (22.58)	.11***	.06	.03	.01	−.02
Health composite[b]	9.44 (2.13)	.03	.03	.06**	.05*	.02
Males	9.57 (2.00)	.06	.01	.05	.05	.000
Females	9.28 (2.27)	.004	.08**	.11***	.05	.03
Self-rated health	3.78 (0.95)	.05*	.02	.06**	.03	.04*
Males	3.81 (0.96)	−.01	.04	.07*	.04	−.001
Females	3.75 (0.94)	.11***	.02	.07*	.02	.10**
Illnesses	0.86 (0.95)	.004	−.01	−.01	−.01	.005
Males	0.81 (0.91)	−.02	.02	−.002	.01	.003
Females	0.91 (0.99)	.02	−.05	−.05	−.02	.01

(*table continues*)

Table 3 (*continued*)

Variable	Mean (*SD*)	E	A	C	S	I
Mental adjustment	2.65 (0.61)	.02	.04*	.06**	.06**	.003
Males	2.69 (0.59)	.05	.02	.05	.06*	−.01
Females	2.61 (0.63)	−.02	.08**	.10**	.07*	.01
Education	7.12 (1.78)	.03	.07**	.17***	.08***	.13***
Males	7.12 (1.85)	.01	.11***	.18***	.07*	.14***
Females	7.11 (1.69)	.05	.02	.15***	.08*	.11***
Alcohol abuse	2.01 (0.91)	.06**	−.05*	−.09***	.001	−.01
Males	2.18 (0.95)	.08**	−.01	−.03	−.001	−.02
Females	1.80 (0.81)	.07*	−.04	−.09**	.01	−.01

Note. Child personality was measured in 1922 (average age = 11 years) in the Terman sample and between 1959 and 1967 (Grades 1, 2, 5, or 6) in the Hawaii sample. Adult variables were measured in the 1950 and 1960 assessments: ages 40 to 50 years in the Terman sample and ages 41 to 50 years (average = 45 years) in the Hawaii sample. Pearson correlations are presented. E = extraversion; A = agreeableness; C = conscientiousness; S = emotional stability; I = intellect.

[a] *N*s given for males (M) and females (F): 0 = male, 1 = female. [b] The composite health variable is the sum of self-rated health, illness (reverse scored), and mental adjustment.

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

In sum, using the original traits assessed by teachers in each sample, we rationally harmonized the traits and then empirically established similar factor structures. The pattern of correlations suggests that the new variables are tapping similar constructs to our prior studies, with the additional benefit of capturing a similar five-factor structure in the two samples. Through this intensive process, we can be more confident that we are comparing similar constructs across the two studies.

Integrating the Samples

The second part of the analyses aimed to (a) test the personality–health model (see Figure 1) in each sample individually and (b) test the personality–health model using the combined sample, pooled on the harmonized composite variables.

First, we tested the personality–health model individually in each sample, initially excluding and then including the education and alcohol abuse variables, controlling for age and sex. In the Terman sample, higher conscientiousness and lower agreeableness predicted higher educational attainment (conscientiousness: $\beta = .22$, $p = .003$; agreeableness: $\beta = -.16$, $p = .02$), and higher levels of extraversion predicted increased alcohol use ($\beta = .22$, $p = .001$), but personality, education, and alcohol use did not significantly predict midlife health.[9] In the Hawaii sample, higher conscientiousness, lower extraversion, and high intellect predicted greater educational attainment (conscientiousness: $\beta = .16$, $p = .04$; extraversion: $\beta = -.19$, $p = .04$; intellect: $\beta = .28$, $p < .001$), higher extraversion and lower intellect predicted increased alcohol use (extraversion: $\beta = .25$, $p = .009$; intellect: $\beta = -.15$, $p = .04$), and education predicted better health ($\beta = .28$, $p < .001$).

To more directly compare model differences across samples, we directly pooled the data from the two samples on the harmonized composite items, creating a single larger dataset ($N = 2{,}255$; 1,231 M, 1,024 F). Descriptive statistics and interitem correlations for the Terman sample, the Hawaii sample, and the combined Terman–Hawaii sample are summarized in Table 3. Figure 4 presents the final SEM with the significant standardized pathways provided. The full model that included education and alcohol abuse was superior to the personality-only model, and provided acceptable fit to the data (RMSEA = .055, 95% CI [.049, .061]). In the personality and health model, childhood conscientiousness was the only significant predictor, with higher levels of conscientiousness relating to better adult health ($\beta = .12$, $p = .001$). When alcohol abuse and education were added to the model, the direct path between conscientiousness and health was reduced, but remained significant ($\beta = .09$, $p = .02$). Interestingly, extraversion was related to better health, but only when alcohol and education were included in the model, suggesting a suppressive effect.

Extending the Model to Unique Elements

A promising possibility for integrative methods is that by carefully linking multiple datasets through common variables, we can use these variables to bridge the studies and then extend the analyses to unique aspects of each sample, effectively piecing together life-span pathways. We used the now-parallel child personality factors to examine mortality risk in the Terman sample and markers of physiological dysfunction in the Hawaii sample. These analyses extended our prior studies with each independent sample (Friedman et al., 1993; Hampson et al., 2009) by using the newly linked personality factors.

The individual traits were standardized and summed to create composite personality factors. In the Terman sample, the five factors were used to predict mortality risk, using Cox proportional hazards regression analysis (survival analysis). Personality variables were included, and then education and alcohol abuse were added to the models. All analyses controlled for age and sex. Replicating our prior studies, the Conscientiousness factor predicted lower mortality risk (hazard ratio [HR] = .90, 95% CI [.85, .95], $p < .0001$). When education and alcohol abuse were added to the model, the effects of conscientiousness were slightly decreased but still significant (HR = .91, 95% CI [.87, .96]), education predicted lower risk (HR = .90, 95% CI [.86, .94], $p < .0001$), and alcohol abuse predicted increased risk (HR = 1.32, 95% CI [1.18, 1.49], $p < .0001$).

[9] This null relation remained when self-rated health, illness reports, and mental adjustment were considered separately as single outcomes.

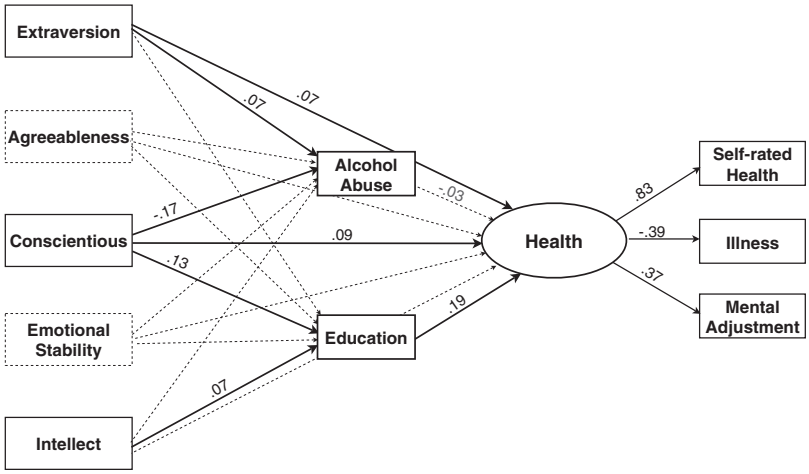

Figure 4. Final estimated structural equation model for the combined sample (N = 2,255, Terman n = 1,085, Hawaii n = 1,170). Standardized estimates are given for the significant pathways, shown as solid lines (dotted lines were nonsignificant). The model was estimated as a multigroup model in R (Version 2.15.0, package lavaan), with factor loadings and regression paths constrained, but means and intercepts allowed to vary by group (RMSEA = .055, 95% CI [.049, .061]). Child (ages 10–11) personality composite factors variables predict midlife (~age 50) health, with partial mediation by education. Harmonized childhood conscientiousness thus helps us understand the prediction of mortality in the Terman data through examination of midlife health in the Hawaii data.

In the Hawaii sample, the five factors were used to predict physiological dysregulation, using standard multiple regression analyses. Personality variables were included, and then education and alcohol abuse were added to the models. The Conscientiousness factor was related to less dysregulation, $\beta = -.12$, $t(617) = -2.03$, $p = .04$. When education and alcohol were added to the models, more education was related to less dysregulation, $\beta = -.22$, $t(543) = -4.99$, $p < .001$; alcohol use was related to less dysregulation, $\beta = -.13$, $t(543) = -2.87$, $p = .004$; and the Conscientiousness factor was no longer significantly related to dysregulation.

Discussion

In this study, we took advantage of integrative data-analytic techniques to combine the data on personality and health from two longitudinal samples. The Terman Life Cycle Study is the longest study that has repeatedly followed individuals throughout their lives from childhood through death. The Hawaii Personality and Health Longitudinal Study includes an ethnically diverse cohort whose participants are being followed from childhood into late adulthood. Previously, individual studies with these samples found that child personality traits, especially conscientiousness, predicted measures of health later in life (self-rated health in the Hawaii sample: Hampson, Goldberg, Vogt, & Dubanoski, 2006, 2007; and longevity in the Terman sample: Friedman et al., 1993, 1995). In this new synthesis of these two studies, child conscientiousness predicted broadly construed adult health, and this effect was partially mediated by education.

The greatest difference between studies was that the education pathway was particularly salient for the Hawaii sample. This is just as it should be. The Terman study purposely included only the brightest children in California, so the sample is restricted in both intellectual level and concomitantly in educational attainment. In contrast, the Hawaii sample included every child in each classroom studied, and therefore it includes the full range of scholastic talent and concomitantly of educational attainment. Not surprisingly, intellect predicted educational attainment in the Hawaii sample but was not related to attainment in the Terman sample. Education is an enormously important factor in health given its relation to socioeconomic status and all the health benefits that flow from one's relative position in the social hierarchy and its associated influences. Intellectually oriented individuals are more likely to pursue educational opportunities, but education itself matters, and these points jump out from an integrative, cross-study perspective.

This study suggests that combining data is both possible and will be fruitful in extending theories of development, personality, and health across the life span. But it also highlights some of the limitations and issues that must be considered, involving characteristics of the sample, measures, and analyses (Bauer & Hussong, 2009; Curran & Hussong, 2009; Hofer & Piccinin, 2009). Sample characteristics include (a) the representativeness of the samples; (b) birth cohort, major historical events, and the socioeconomic context of the time; (c) socioeconomic status and racial, ethnic, and educational differences; and (d) attrition, mortality, and other possible selection effects. Measurement characteristics include (a) the constructs and measures used; (b) changes in constructs or measures at different periods in life; (c) how often the participants are assessed and the interval between assessments; (d) retest effects; and (e) how time is handled in the analyses. Perhaps the greatest benefit of attempting to pool data is that one must intentionally examine measure invariance. Only by establishing conceptual and metric equivalency can one really determine what differences may be due to actual processes as compared to characteristics of the sample or measures. Such differences are often glossed over and

ignored, potentially leading to a science built on invalid assumptions and conclusions. By treating heterogeneity as interesting limits on generalizability rather than as problematic noise, both similarities and differences across samples become important.

Both of the samples are limited by the measures initially included. When the teachers assessed the personality traits of the school children, the present conception of the Big Five personality factor model did not exist. Health status, mental adjustment, and health-related behaviors were assessed with different questions and at different time points. The present study starts to address generalizability across the two samples, but the extent to which our findings generalize to other samples is unknown, especially when other social, cultural, and historical variables are relevant. The two samples are to some extent selective. Still, our process suggests that harmonization is indeed possible in longitudinal studies, and important conclusions emerged across the cohorts and the decades concerning conscientiousness and health.

By finding regions of overlap at both conceptual and measurement levels, relations that are unique to each sample can then be used to fill in missing pieces in life-span models of personality and health. The Hawaii sample does not yet have much data concerning mortality and cause of death, but it has much more midlife physiological health information (including blood tests), whereas the Terman sample lacks detailed physiological measures, but we have gathered extensive information about length of life and cause of death. The two studies, taken alone and then together, confirm the importance of conscientiousness to later health and show the relevance of education and alcohol abuse. Thus, for the first time, we have evidence from a full life-span analysis for the importance of child conscientiousness to adult health and longevity. Creatively using the resources that are available, we can address life-span questions that are impossible in shorter term studies, offering the means necessary to accumulate sound scientific knowledge about the interplay of personality and health.

References

Bath, P. A., Deeg, D., & Poppelaars, J. (2010). The harmonisation of longitudinal data: A case study using data from cohort studies in The Netherlands and the United Kingdom. *Ageing & Society, 30,* 1419–1437. doi:10.1017/S0144686X1000070X

Bauer, D. J., & Hussong, A. M. (2009). Psychometric approaches for developing commensurate measures across independent studies: Traditional and new models. *Psychological Methods, 14,* 101–125. doi:10.1037/a0015583

Christensen, A. J., Ehlers, S. L., Wiebe, J. S., Moran, P. J., Raichle, K., Ferneyhough, K., & Lawton, W. J. (2002). Patient personality and mortality: A 4-year prospective examination of chronic renal insufficiency. *Health Psychology, 21,* 315–320. doi:10.1037/0278-6133.21.4.315

Cohn, L. D., & Becker, B. J. (2003). How meta-analysis increases statistical power. *Psychological Methods, 8,* 243–253. doi:10.1037/1082-989X.8.3.243

Cooper, H., & Patall, E. A. (2009). The relative benefits of meta-analysis conducted with individual participant data versus aggregated data. *Psychological Methods, 14,* 165–176. doi:10.1037/a0015565

Cooper, R., Hardy, R., Sayer, A. A., Ben-Shlomo, Y., Birnie, K., Cooper, C., . . . Kuh, D. (2011). Age and gender differences in physical capability levels from mid-life onwards: The harmonisation and meta-analysis of data from eight UK cohort studies. *PLoS ONE, 6*(11): e27899. doi:10.1371/journal.pone.0027899

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

Curran, P. J., Bollen, K. A., Chen, F., Paxton, P., & Kirby, J. (2003). The finite sampling properties of the RMSEA: Point estimates and confidence intervals. *Sociological Methods & Research, 32,* 208–252. doi:10.1177/0049124103256130

Curran, P. J., & Hussong, A. M. (2009). Integrative data analysis: The simultaneous analysis of multiple data sets. *Psychological Methods, 14,* 81–100. doi:10.1037/a0015914

Curran, P. J., Hussong, A. M., Cai, L., Huang, W., Chassin, L., Sher, K. J., & Zucker, R. A. (2008). Pooling data from multiple longitudinal studies: The role of item response theory in integrative data analysis. *Developmental Psychology, 44,* 365–380. doi:10.1037/0012-1649.44.2.365

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2010). Using participant data to extend the evidence base for intensive behavioral intervention for children with autism. *AJIDD: American Journal on Intellectual and Developmental Disabilities, 115,* 381–405. doi:10.1352/1944-7558-115.5.381

Friedman, H. S., Kern, M. L., Hampson, S. E., & Duckworth, A. L. (2014). A new life-span approach to conscientiousness and health: Combining the pieces of the causal puzzle. *Developmental Psychology, 50,* 1377–1389. doi:10.1037/a0030373

Friedman, H. S., Kern, M. L., & Reynolds, C. A. (2010). Personality and health, subjective well-being, and longevity. *Journal of Personality, 78,* 179–216. doi:10.1111/j.1467-6494.2009.00613.x

Friedman, H. S., Tucker, J. S., Schwartz, J. E., Martin, L. R., Tomlinson-Keasey, C., Wingard, D., & Criqui, M. (1995). Childhood conscientiousness and longevity: Health behaviors and cause of death. *Journal of Personality and Social Psychology, 68,* 696–703. doi:10.1037/0022-3514.68.4.696

Friedman, H. S., Tucker, J., Tomlinson-Keasey, C., Schwartz, J., Wingard, D., & Criqui, M. H. (1993). Does childhood personality predict longevity? *Journal of Personality and Social Psychology, 65,* 176–185. doi:10.1037/0022-3514.65.1.176

Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology, 59,* 1216–1229. doi:10.1037/0022-3514.59.6.1216

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48,* 26–34. doi:10.1037/0003-066X.48.1.26

Goldberg, L. R. (2001). Analyses of Digman's child-personality data: Derivation of Big-Five factor scores from each of six samples. *Journal of Personality, 69,* 709–744. doi:10.1111/1467-6494.695161

Greenhoot, A. F., & Dowsett, C. J. (2012). Secondary data analysis: An important tool for addressing developmental questions. *Journal of Cognition and Development, 13,* 2–18. doi:10.1080/15248372.2012.646613

Hallahan, B., Newell, J., Soares, J. C., Brambilla, P., Strakowski, S. M., Fleck, D. E., . . . McDonald, C. (2011). Structural magnetic resonance imaging in bipolar disorder: An international collaborative mega-analysis of individual adult patient data. *Biological Psychiatry, 69,* 326–335. doi:10.1016/j.biopsych.2010.08.029

Hampson, S. E., Dubanoski, J. P., Hamada, W., Marsella, A. J., Matsukawa, J., Suarez, E., & Goldberg, L. R. (2001). Where are they now? Locating former elementary-school students after nearly 40 years for a longitudinal study of personality and health. *Journal of Research in Personality, 35,* 375–387. doi:10.1006/jrpe.2001.2317

Hampson, S. E., Goldberg, L. R., Vogt, T. M., & Dubanoski, J. P. (2006). Forty years on: Teachers' assessments of children's personality traits predict self-reported health behaviors and outcomes at midlife. *Health Psychology, 25,* 57–64. doi:10.1037/0278-6133.25.1.57

Hampson, S. E., Goldberg, L. R., Vogt, T. M., & Dubanoski, J. P. (2007). Mechanisms by which childhood personality traits influence adult health

status: Educational attainment and healthy behaviors. *Health Psychology, 26,* 121–125. doi:10.1037/0278-6133.26.1.121

Hampson, S. E., Goldberg, L. R., Vogt, T. M., Hillier, T. A., & Dubanoski, J. P. (2009). Using physiological dysregulation to assess global health status: Associations with self-rated health and health behaviors. *Journal of Health Psychology, 14,* 232–241. doi:10.1177/1359105308100207

Hofer, S. M., & Piccinin, A. M. (2009). Integrative data analysis through coordination of measurement and analysis protocol across independent longitudinal studies. *Psychological Methods, 14,* 150–164. doi:10.1037/a0015566

Hofer, S. M., & Piccinin, A. M. (2010). Toward an integrative science of life-span development and aging. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 65*B, 269–278. doi:10.1093/geronb/gbq017

International Conference on Harmonisation. (2010). ICH for better health: Vision. Retrieved from http://www.ich.org/about/vision.html

Kern, M. L., & Friedman, H. S. (2008). Do conscientious individuals live longer? A quantitative review. *Health Psychology, 27,* 505–512. doi:10.1037/0278-6133.27.5.505

Kern, M. L., & Friedman, H. S. (2010). Why do some people thrive while others succumb to disease and stagnation? Personality, social relations, and resilience. In P. S. Fry & C. L. M. Keyes (Eds.), *Frontiers of resilient aging: Life-strengths and well-being in late life* (pp. 162–184). Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511763151.008

Loehlin, J. C. (2004). *Latent variable models: An introduction to factor, path, and structural equation analysis* (4th ed.). Mahwah, NJ: Erlbaum.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7,* 83–104. doi:10.1037/1082-989X.7.1.83

Manichaikul, A., Chen, W. M., Williams, K., Wong, Q. N., Sale, M. M., Pankow, J. S., . . . Tsai Mychaleckyj, J. C. (2012). Analysis of family- and population-based samples in cohort genome-wide association studies. *Human Genetics, 131,* 275–287. doi:10.1007/s00439-011-1071-0

Martin, L. R., & Friedman, H. S. (2000). Comparing personality scales across time: An illustrative study of validity and consistency in life-span archival data. *Journal of Personality, 68,* 85–110. doi:10.1111/1467-6494.00092

McArdle, J. J., Grimm, K. J., Hamagami, F., Bowles, R. P., & Meredith, W. (2009). Modeling life-span growth curves of cognition using longitudinal data with multiple samples and changing scales of measurement. *Psychological Methods, 14,* 126–149. doi:10.1037/a0015857

Olgiati, P., Politis, A., Albani, D., Rodilossi, S., Polito, L., Ateri, E., . . . Serretti, A. (2012). Association of SORL1 alleles with late-onset Alzheimer's disease. Findings from the GIGAS_LOAD study and mega-analysis. *Current Alzheimer Research, 9,* 491–499. doi:10.2174/156720512800492431

Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin, 114,* 552–566. doi:10.1037/0033-2909.114.3.552

Rijmen, F., Tuerlinckx, F., De Boeck, P., & Kuppens, P. (2003). A nonlinear mixed model framework for item response theory. *Psychological Methods, 8,* 185–205. doi:10.1037/1082-989X.8.2.185

Riley, R. D., Simmonds, M. C., & Look, M. P. (2007). Evidence synthesis combining individual patient data and aggregate data: A systematic review identified current practice and possible methods. *Journal of Clinical Epidemiology, 60,* 431–439. doi:10.1016/j.jclinepi.2006.09.009

Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology, 52,* 59–82. doi:10.1146/annurev.psych.52.1.59

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

Salimi-Khorshidi, G., Smith, S. M., Keltner, J. R., Wager, T. D., & Nichols, T. E. (2009). Meta-analysis of neuroimaging data: A comparison of image-based and coordinate-based pooling of studies. *NeuroImage, 45,* 810–823. doi:10.1016/j.neuroimage.2008.12.039

Schaap, L. A., Peeters, G. M. E. E., Dennison, E. M., Zambon, S., Nikolaus, T., Sachez-Martinez, M., . . . Deeg, D. J. H. (2011). European Project on Osteoarthritis (EPOSA): Methodological challenges in harmonization of existing data from five European population-based cohorts on aging. *BMC Musculoskeletal Disorders, 12,* 272–284. doi:10.1186/1471-2474-12-272

Seeman, T. E., McEwen, B. S., Rowe, J. W., & Singer, B. H. (2001). Allostatic load as a marker of cumulative biological risk: MacArthur studies of successful aging. *Proceedings of the National Academy of Sciences of the United States of America, 98,* 4770–4775. doi:10.1073/pnas.081072698

Serretti, A., Cusin, C., Rausch, J. L., Bondy, B., & Smeraldi, E. (2006). Pooling pharmacogenetic studies on the serotonin transporter: A mega-analysis. *Psychiatry Research, 145,* 61–65. doi:10.1016/j.psychres.2005.09.020

Smith, T. W., & Williams, P. G. (1992). Personality and health: Advantages and limitations of the five-factor model. *Journal of Personality, 60,* 395–425. doi:10.1111/j.1467-6494.1992.tb00978.x

Stewart, L. A., & Clarke, M. J. (1995). Practical methodology of meta-analyses (overviews) using updated individual patient data. *Statistics in Medicine, 14,* 2057–2079. doi:10.1002/sim.4780141902

Sutton, A. J., Kendrick, D., & Coupland, C. A. C. (2008). Meta-analysis of individual- and aggregate-level data. *Statistics in Medicine, 27,* 651–669. doi:10.1002/sim.2916

Tak, L. M., Meijer, A., Manoharan, A., de Jonge, P., & Rosmalen, J. G. M. (2010). More than the sum of its parts: Meta-analysis and its potential to discover sources of heterogeneity in psychosomatic medicine. *Psychosomatic Medicine, 72,* 253–265. doi:10.1097/PSY.0b013e3181d714e1

Terman, L. M., Baldwin, B. T., Bronson, E., DeVoss, J. C., Fuller, F., Goodenough, F. L., . . . Yates, D. H. (1925). *Genetic studies of genius: Vol. 1. Mental and physical traits of a thousand gifted children.* Stanford, CA: Stanford University Press.

Terman, L. M., & Oden, M. H. (1947). *Genetic studies of genius: Vol. IV. The gifted child grows up.* Stanford, CA: Stanford University Press.

Terman, L. M., & Oden, M. H. (1959). *Genetic studies of genius: Vol. V. The gifted group at mid-life*. Stanford, CA: Stanford University Press.

Weiss, A., & Costa, P. T. (2005). Domain and facet personality predictors of all-cause mortality among Medicare patients aged 65 to 100. *Psychosomatic Medicine, 67,* 724–733. doi:10.1097/01.psy.0000181272.58103.18

Wilson, R. S., Mendes de Leon, C. F., Bienias, J. L., Evans, D. A., & Bennett, D. A. (2004). Personality and mortality in old age. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 59B,* 110–116. doi:10.1093/geronb/59.3.P110

(*Appendix follows*)

Appendix

Trait Labels and Descriptions

Trait	Description
	Terman sample
Amount of physical energy	High = extraordinary amount of energy, pep, and animation; dynamic and tireless. Low = extreme physical inertia and lack of pep; sluggish and easily fatigued.
Appreciation of beauty	High = extraordinary appreciation of beautiful colors, landscapes, forms, sunsets, flowers, etc.; has natural "taste." Low = practically no appreciation of beauty in things seen; no "taste."
Cheerfulness and optimism	High = extraordinarily cheerful and optimistic; never sees dark side; never worries. Low = usually extremely depressed and pessimistic; looks on dark side of everything; worries constantly.
Common sense	High = possesses common sense and judgment to an extraordinary degree; his advice always highly valuable. Low = extreme lack of common sense and judgment; opinions not taken seriously by anyone.
Conscientiousness	High = extraordinarily conscientious; keen sense of duty; does right for right's sake; always dependable. Low = extreme lack of conscientiousness; no sense of duty; does wrong for any advantage; not dependable.
Desire to excel	High = extraordinary pride in accomplishment and desire to excel; does his upmost to stand first. Low = no pride in accomplishment; no ambition to excel; almost never does his best.
Desire to know	High = extraordinarily strong intellectual curiosity and broad interests; insistent on knowing. Low = extreme lack of intellectual curiosity; mentally inert; few interests; rarely asks questions.
Fondness for large groups	High = extraordinary fondness for large groups; unhappy alone; devoted to parties, picnics, etc. Low = invariably avoids groups; always prefers to be either alone or with one or two close chums.
Freedom from vanity and egotism	High = extraordinarily free from egotism or vanity; shrinks from praise and admiration. Low = extremely egotistical and vain; "fishes" for praise; always showing off.
General intelligence	High = extraordinary all-around intelligence. Low = general intelligence extremely inferior; almost feeble-minded.
Generosity and unselfishness	High = extraordinarily generous, unselfish, and fair-minded. Low = extremely selfish; cares only for own pleasures; takes unfair advantage.
Health	High = extraordinary good health, almost never sick, vigorous. Low = extremely weakly and sickly; extreme lack of vigor.
Leadership	High = extraordinary qualities of leadership; gets others to do his will; not easily influenced. Low = always a follower; never takes initiative; suggestible and easily influenced.
Mechanical ingenuity	High = extraordinary mechanical ingenuity; likes and understands machinery, apparatus, etc.; clever at "fixing" things. Low = extreme lack of mechanical ingenuity; cares nothing for machinery; a blunderer with tools.
Musical appreciation	High = extraordinary musical appreciation. Low = no musical appreciation whatever.
Originality	High = extraordinary ability to think things through for self; original, resourceful, and inventive; excels in reasoning. Low = extreme lack of originality and resourcefulness; always depends on teacher or book.
Permanency of moods	High = moods extraordinarily permanent; almost never goes quickly from joy to sadness or sadness to joy. Low = moods extremely changeable; always alternating between extreme joy and extreme sadness.
Popularity with other children	High = extraordinarily popular; universal favorite; is sought after and has many friends. Low = extremely unpopular; disliked and shunned; a social outcast.
Prudence and forethought	High = extraordinarily prudent; always looks ahead; never sacrifices future goods for present pleasure. Low = extreme lack of prudence; never looks ahead; lives wholly in the present.
Self-confidence	High = extreme self-confidence and self-reliance; always relies on own judgments; courts responsibility. Low = extreme lack of self-confidence; distrusts own judgment; afraid of responsibility.
Sense of humor	High = extraordinarily keen sense of humor; witty; appreciates jokes; sees the funny side in everything. Low = extremely lacking in sense of humor; serious and prosy; never sees the funny side.
Sensitiveness to approval or disapproval	High = extraordinary sensitiveness to approval or disapproval of other children; can't endure to be disliked. Low = utterly indifferent to opinion of other children; does not care in the least to be liked.
Sympathy and tenderness	High = extraordinarily tender and sympathetic; kind on principle; abhors cruelty. Low = extreme lack of tenderness or sympathy; rarely does a kind act; tendency to cruelty.
Truthfulness	High = extraordinarily truthful, honest, and frank; never misleads or misrepresents, however great the temptation. Low = extreme tendency to lying, deceitfulness, and evasiveness; lies for the slightest advantage.
Will power and perseverance	High = extraordinary will power; persistent in overcoming difficulties; extremely steadfast; never gives up. Low = extreme lack of will power; easily discouraged and gives up at slightest difficulty.

(*Appendix continues*)

Appendix (*continued*)

Trait	Description
	Hawaii sample
Adaptable	Copes easily and successfully with new and strange situations; bravely faces up to uncertainty.
Assertive	Bossy; usually attempts to direct the actions of others; is convinced his (her) way is the best way of doing things; shows others "how things should be done."
Careful of personal belongings	Takes good care of things which belong to him (her); becomes concerned when possessions are missing and searches for them; keeps his (her) own things neat, clean, and in order.
Careless of others' property	Seldom or never concerned about the possibility of damage to others' belongings; borrowed things are often returned broken or dirty; loses or forgets to return borrowed things.
Complains about others	Frequently complains about what others are saying about or doing to him (her); asks adult to intervene and change the behavior of others.
Concerned about acceptance	Expresses concern about real or imagined rebuffs and slights from others; unable to take his (her) relationship with others for granted; worries that he (she) may lose friends or that others will not like him (her).
Conscientious	Honest; knows what is right and generally does it even if no one is watching; tells the truth even when this is difficult; does not attempt to deceive others.
Considerate	Thoughtful of others; sensitive to others' feelings; cannot do things which hurt others' feelings; sympathetic when others are in trouble and tries to help.
Curious	Given to wondering and musing about the why and wherefore of things; examines things closely; asks questions.
Eccentric	Has interests and views which are different from those of other children; acts differently from others; not interested in wearing the same clothes or doing the same things as others are wearing and doing; goes his (her) own, rather off-beat, way.
Energetic	Active; full of pep; vigorous; movements are quick, darting.
Esthetically sensitive	Notices and responds with pleasure to beauty in his (her) surroundings; enjoys art and/or music.
Fearful	Has many fears and worries, some of which are unreasonable; easily becomes alarmed or frightened.
Fickle	Changes frequently in interests, opinions, and pursuits; "flighty"; starts one thing and shifts to another.
Fidgets	Finds it difficult to sit still, frequently changing position; often "doodles" or manipulates objects for no reason other than to be doing something.
Gregarious	Likes to be with others and seeks their company; spends as much time with others as possible; dislikes being alone.
Happy	Joyful; has a sunny disposition; enjoys life; gives impression of contentment with the way things are going for him (her).
Imaginative	Has an active, vivid imagination; very fanciful; sees possibilities overlooked by others.
Impulsive	Behavior always seems very close "to the surface"; often acts before the appropriate moment; finds it difficult to hold back; often acts or speaks without thinking of possible consequences.
Irresponsible	Does not take his (her) assigned duties seriously; cannot be depended upon to carry out assigned tasks.
Jealous	Envies and begrudges the accomplishments of others; is disturbed when others are shown special attention or given special favors; shows disappointment or annoyance when others are praised.
Lethargic	Slow moving; seldom or never runs or hurries; unresponsive or slow to react; works slowly.
Mannerly	Has good manners; knows what to say and do when introduced to others; has a sense of "good form" and behaves accordingly; uses "please" and "thank you" properly.
Neat in appearance	Careful about clothes and appearance; dislikes being dirty or disheveled; is usually well groomed.
Nervous habits	Has a great variety of nervous habits (e.g., nail-biting, grimacing, tics, hair twisting, pencil chewing, etc.).
Original	Has remarkably novel and different ideas and/or solutions to problems; thinking and behavior are characterized by unusual approaches.
Outspoken	Speaks his (her) mind without reservation or hesitation; seldom or never hesitates to express views and opinions on any subject.
Persevering	Keeps at his (her) work until it is completed; sees a job through despite difficulties; painstaking and thorough.
Planful	Behavior, including play, is purposeful; attacks work in systematic fashion.
Restless	Constantly or frequently moves about the room; unable to settle down after activity period or recess.
Rigid	Has difficulty adapting to change to new situations; prefers to keep old ways and routines, even where these are obviously inappropriate.
Rude	Insolent and sassy to others; often gives impression he (she) goes out of his (her) way to be discourteous to others.
Seclusive	Dislikes group activities and games; prefers to be by self or in company of one or two others; dislikes being in a crowd.

(*Appendix continues*)

Appendix (*continued*)

Trait	Description
Self-minimizing	Tends to minimize own importance; humble; never brags or shows off; seeks out or is content with less important tasks or positions.
Socially confident	Approaches others without hesitation; shows poise when performing before others.
Spiteful	Deliberately does or says things which annoy or hurt others; says hateful things about others; belittles others.
Submissive	Usually easily led or persuaded by others; seldom or never sticks up for own rights; gives in easily in arguments.
Touchy	Very sensitive to criticism; cries, pouts, or sulks when criticized; does not take well to jokes or pranks on him (her).
Verbally fluent	Speech seems to "pour out," often in a torrent of words, sometimes making it difficult to understand him (her).

Received February 9, 2012
Revision received August 23, 2012
Accepted September 18, 2012 ■