Replicability and 40-Year Predictive Power of Childhood ARC Types

Benjamin P. Chapman and
University of Rochester School of Medicine and Dentistry Department of Psychiatry Laboratory of Personality and Development

Lewis R. Goldberg
Oregon Research Institute

Abstract

We examined three questions surrounding the Undercontrolled, Overcontrolled, and Resilient--or Asendorpf-Robins-Caspi (ARC)--personality types originally identified by Block (1971). In analyses of the teacher personality assessments of over 2,000 children in 1st through 6th grade in 1959-1967, and follow-up data on general and cardiovascular health outcomes in over 1,100 adults recontacted 40 years later, we found: (1) Bootstrapped internal replication clustering suggested that Big Five scores were best characterized by a tripartite cluster structure corresponding to the ARC types; (2) this cluster structure was fuzzy, rather than discrete, indicating that ARC constructs are best represented as gradients of similarity to three prototype Big Five profiles; and (3) ARC types and degrees of ARC prototypicality showed associations with multiple health outcomes 40 years later. ARC constructs were more parsimonious, but neither better nor more consistent predictors than the dimensional Big Five traits. Forty-year incident cases of heart disease could be correctly identified with 68% accuracy by personality information alone, a figure approaching the 12-year accuracy of a leading medical cardiovascular risk model. Findings support the theoretical validity of ARC constructs, their treatment as continua of prototypicality rather than discrete categories, and the need for further understanding the robust predictive power of childhood personality traits for mid-life health.

Keywords

The phrase "personality structure" is typically used to refer to the dimensions reflecting the covariation of personality traits in the population (e.g., Goldberg, 1993; John & Srivastava, 1999; McCrae & Costa, 1997). Personality structure has another meaning, however. It can also refer to the configuration of attributes within an individual (e.g., Allport, 1937; Asendorpf, 2006a,b; Funder, 1991). This so-called "person-centered" approach to

Address correspondence to: Ben Chapman, PhD, MPH Laboratory of Personality and Development Rochester Center for Mind-Body Research. Department of Psychiatry University of Rochester School of Medicine and Dentistry 300 Crittenden Rochester, NY 14608 Phone: 585-275-9431 Fax: 585-273-1082 Ben_Chapman@URMC.Rochester.edu.

[1]The kordax was a lewd Greek dance.

personality dates back at least to the German differential psychologist Stern's (1911) "psychography"—the description of an individual by his or her pattern of traits. Later, Allport (1937) championed a person-centered, or idiographic, view in American psychology. Block (1971; Block & Block, 1980) formalized a distinction between "person-centered" and "variable-centered" research strategies. Person centered approaches focus on the configuration of personality traits within an individual, while variable-centered approaches focus on the covariation of traits in groups of people. Magnusson (cf. 1999) carried out further work on personality development using this person-centered strategy, examining trait configurations in people longitudinally.

Person-centered approaches focus on personality prototypes, which have seen renewed interest in the past 15 years (e.g., Asendorpf, 2003; 2006a, b; Asendorpf & Van Aken, 1999; Asendorpf, Borkenau, Ostendorf, & Van Aken, 2001; Caspi & Silva, 1995; Hart, Burock, London, Atkins, & Bonilla-Santiago, 2005; McCrae, Terracciano, Costa, & Ozer, 2006; Robins, John, Caspi, Moffitt, & Stouthamer-Loeber, 1996; see also the 2002 Special Issue of the European Journal of Personality). Below, we briefly review this aspect of personality psychology, focusing on the so-called Asendorpf-Robins-Caspi (ARC) types. We then address three central questions about ARC types concerning their replicability, categorical versus dimensional nature, and predictive power, in analyses of a large, richly characterized cohort of children followed up over 40 years later at midlife.

The Asendorpf-Robins-Caspi Types

Personality prototypes are generally defined as particular configurations of traits that occur with some regularity in the population (Asendorpf & Van Aken, 1999). Over the years, a number of different personality prototypes have been proposed, as others have noted (e.g., Robins et al., 1996). The persistence of the concept of personality types is underscored by the fact that efforts to understand them date back to ancient Greece. The most wide-ranging and richly descriptive typology of persons may be that of Theophrastus's "Characters" (Rusten & Cunningham, 2002). The characters were 30 different types of cads and misanthropes inhabiting Athens, such as the "Shameless Man," who is described as a boorish exhibitionist, callously negligent of ethics and gleefully contemptuous of the social order (p. 69-70).

Although a number of contemporary typologies exist,[2] we focus on the one that has received the most attention in scientific personality psychology: the Asendorpf-Robins-Caspi (ARC) types. This taxonomy combines a strong theoretical foundation with the empirical rigor of the Big Five. The ARC acronym was coined by Costa and his colleagues (2002), and reflects the lead investigators of three articles (Asendorpf et al., 2001; Caspi & Silva, 1995; Robins et al., 1996) on three of Block's (1971) personality prototypes referred to as Overcontrolled, Undercontrolled, and Resilient. Block first discussed the types (Block, 1971; Block & Block, 1980) in the context of his theory of ego resiliency and ego control. Ego resiliency is

---

[2]The briefest overview of typologies would include the 16 Myers-Briggs Types (Myers & McCaulley, 1985), which have been interpreted in Five Factor Model (FFM) terms by McCrae and Costa (1989); the Diagnostic and Statistical Manual Fourth Edition, Text Revision Axis II classification system (American Psychiatric Association, 2000), the pathological types of which have been translated into extreme standing on FFM dimensions (Costa & Widiger, 2001); Eysenck's (cf., 1991a; b), Grossarth-Maticek's (Grossarth-Maticek, Eysenck, & Vetter, 1988) four health-related (e.g., cancer prone, coronary prone) types, which Eysenck distinguished from classic "Type A" (coronary prone) and "Type B" ("relaxed") types (Friedman & Booth-Kewley, 1987); Temoshok's (1987) similar articulation of a "Type C" or cancer-prone type; and the "Type D" or the Distressed Type (Denollet & Van Heck, 2001) related to coronary events, which has also been profiled from the FFM perspective (Chapman, Duberstein, & Lyness, 2007; DeFruyt & Denollet, 2002). Also, the notion of personality "styles," or bivariate combinations of traits, explored by Costa and Piedmont (2003), is similar to but not exactly the same as prototypes. Hofstee, Barelds, & Ten Berge (2006) report an interesting approach to prototypes. Roth & von Collani (2007) also report evidence for five, rather than three types in Big Five data. See also Herzberg's work (Herzberg & Roth, 2006; Herzberg & Hoyer, 2009) for evidence of further types beyond the ARC clusters, and Pulkinnen (1996) for a hierarchical typology. See also Hofstee (2002) on methodological issues.

the capacity to respond flexibly and adaptively to environmental challenges, while ego control refers to the capacity to restrain emotion and impulse. High standing on ego resiliency and moderate standing on ego-control characterize the Resilient type. The Overcontrolled type is low on ego resiliency and high on ego control, whereas the Undercontrolled type is low on ego resiliency and low on ego control. Block's original work (1980) was concerned primarily with finding some way to usefully condense a rich battery of personality data from Q-sort assessments. He accomplished this through inverse factor analysis, which transposes variables and persons and "factors persons" to identify similar groups of persons.

Interest in these types was initially reinvigorated by Robins and colleagues (1996), who believed that they would generalize to other samples. Hart and colleagues (1997) then replicated the types, with Asendorpf and Van Aken publishing another replication in 1999. Thus, the ARC types might also be called ARCH types, the amended H for Hart, or even BARCH types, with an initial B also added for Block.

In terms of Big Five profiles, Resilients are characterized by low Neuroticism combined with above average standing on the other four factors; high Neuroticism and low Extraversion characterize Overcontrollers; and Undercontrollers evidence high Neuroticism and low Conscientiousness, with some studies also suggesting low Agreeableness. Longitudinal findings suggest that Resilient children and adolescents tend to be well-adjusted by later adolescence or early adulthood. Undercontrollers may experience externalizing problems and Overcontrollers may experience internalizing problem (cf. Caspi & Silva, 1995; Asendorpf & Denissen, 2006).

ARC Research over the Last Decade

The ARC types featured prominently in an international workshop on lifespan development orchestrated by Paul Costa in 2000, and a special edition of the European Journal of Personality (EJP) was devoted to them in 2002. In this issue, which was edited by Jens Asendorpf (2002) and co-edited by Avshalom Caspi and Willem Hofstee, a set of investigators attempted to replicate the ARC types using the same instrument (the NEO-PI-R or NEO-FFI) and the same type of clustering procedures (Asendorpf et al., 2001). Some investigators reported full replication (Schnabel, Asendorpf, & Ostendorf, 2002, with German students), while others reported partial replications (Barbanelli, 2002, with Italian adults; Boehm, Asendorpf, & Avia, 2002; DeFruyt, Mervielde, & Van Leeuwen, 2002, with Flemish children) and others no evidence of the types (Costa, Herbst, McCrae, Samuels, & Ozer, 2002 within U.S. adults).

Prototype research subsequent to the EJP special issue has continued to reflect mixed findings. This may reflect differing interpretations of what constitutes evidence for the ARC constructs, as well as inherent differences in samples and methods (i.e., cluster analysis vs. inverse factoring). Rammstedt and her colleagues (2004) found evidence of ARC types in self-reports but not observer ratings of the Big Five. Methodological debates have occurred over type derivation from inverse factoring of Q-sort data (Asendorpf, 2006a; McCrae et al., 2006) versus clustering, as well as the relative predictive power of types vs. traits (Asendorpf, 2003; Asendorpf & Denissen, 2006; Bergman & Trost, 2006; Costa et al., 2002; Laursen & Hoff, 2006, von Eye & Bogat, 2006; cf. also Hart et al., 2003; 2008). Other longitudinal work has supported Block's original typology (1971), which actually included five types (Morizot & Le Blanc, 2003; 2005). All in all, work in recent years highlights the need to better understand the replicability of ARC types, their status as more categorical or more dimensional entities, and their predictive value.

Current Issues in Type Research

Question 1: Do ARC Types Represent the Optimal Cluster Structure in Big Five Data?

A variety of different samples have yielded evidence supporting the tripartite ARC taxonomy (Asendorpf, 2002). Nevertheless, agreement that such a cluster structure characterizes individuals in Big Five multivariate space is far from universal (Cost et al., 2002; McCrae et al., 2006). Continued replication **is** essential in order to determine whether the ARC cluster structure appears as robustly in persons as the Big Five factor structure does in variables. Thus, our first objective was to identify the empirically optimal cluster structure in what, to our knowledge is the largest sample examined to date: over 2,000 Hawaii school children. Additionally, we implement an elaboration of traditional ARC clustering methods very sensitive to cluster structure over a range of possible cluster solutions.

Question 2: Are ARC Types Discrete or Fuzzy? Is There a Spectrum of Prototypicality?

The usefulness of categories versus dimensions is an important issue in psychology, and this controversy has engendered negative reactions to ARC types. Indeed, this issue has received great consideration in the context of discussion about the Diagnostic and Statistical Manual for Mental Disorders (DSM) – V Axis II personality disorders. Whether maladaptive personality "types" are better represented as dimensions has stimulated a good deal of recent Axis II research, and is summarized in Widiger and Trull (2007). To some extent, the categorical or dimensional representation of the ARC types is a more general extension of the basic issue underlying Axis II categories: Are people classified into one of the putative types sufficiently homogenous, and is there sufficient distinction between the types to warrant categorization?

Our second objective was therefore to examine the utility of a dimensional approach to types, in the spirit of Asendorpf (2006b). This approach assesses the degree to which an individual's Big Five profile is similar to the prototype Big Five profile of each ARC type. The result is a spectrum ranging from high dissimilarity to high similarity to a given prototype. We refer to this dimensional measure as a prototypicality score, and examine evidence for such a continuum using methods of profile comparison (Cronbach & Gleser, 1953).

Question 3: Do ARC Types and/or Prototypicality Scores Display Long-Term Predictive Validity for Health Outcomes? Are They Better Predictors Than Traits?

Various arguments have been marshaled for and against the predictive power of ARC types relative to Big Five traits (see Asendorpf, 2003; Costa et al. 2002; von Eye & Bogat, 2006). A common admonition has been that categories discard potentially relevant information about the differences between members within the same class. As a result, continuous measures typically afford more predictive power. Counter-arguments emphasize (a) the practical meaning of a classification over scores on scales whose units have no inherent meaning; (b) the possibility that ARC types may capture higher order trait interactions and other non-linearities missed in analyses of independent linear main effects of traits; and finally (c) the assumption that traits and types may each be optimal predictors of different outcomes (Asendorpf, 2003; von Eye & Bogat, 2006). We therefore compared the predictive validity of ARC types and prototypicality scores to the dimensions provided by the Big Five factor structure.

Our outcomes were general health and cardiovascular health 40 years after personality measurement. We focused on health because it has been repeatedly shown to be associated with personality traits (Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007), including childhood traits (Friedman, 2000). Health is clearly an important life outcome, yet only one

other sample in addition to the present has been able to capture the longitudinal associations between childhood personality traits and midlife health outcomes (Friedman, Tucker, Tomlinson-Keasey, Schwartz, Wingard, & Criqui, 1993). Moreover, nonlinear functional forms and higher-order interactions involved in personality associations with health may exist but are rarely considered (Smith & Spiro, 2002). Therefore any advantage types have over traits in summarizing such complexities should be evident.

We focus on cardiovascular health in particular, for several reasons. First, cardiovascular diseases, including stroke and diabetes, represent four of the top six leading causes of death (the other two being cancer and accidents; Jemal, Ward, Hao, & Thun, 2005). Hypertension is directly implicated in the pathogenesis of such diseases (Isselbacher et al., 1994). Second, a large literature has shown links between personality and cardiovascular health outcomes (Friedman, 2000; Smith and Spiro, 2002). Third, there are well-developed biopsychosocial models of the pathways linking personality and cardiovascular health (Smith & Ruiz, 2002).

Specifically, we hypothesized that childhood types characterized by low Agreeableness will be at higher risk for incident hypertension and myocardial infarction, as suggested by the literature on hostility and health (Smith & Ruiz, 2002). Second, we hypothesized that childhood types characterized by emotional instability and lower Extraversion will also be at greater risk for myocardial infarction, hypertension, and coronary heart failure. This hypothesis arises from the literature showing Type D personality (combining negative affect and social inhibition) is linked to these cardiovascular problems (Denollet, Pedersen, Vrints, & Conraads, 2006; Denollet, Schiffer, Kwaijtaal, Hooijkaas, Hendriks, Widdershoven, et al., 2009). Third, we expected the incidence of stroke, hypertension, and diabetes to be lower among childhood types displaying higher Conscientiousness, based on previous observations in a national sample (Goodwin & Friedman, 2006).

Method

Participants and Procedures

The participants were 2,403 children in the 1st through 6th grade in the Hawaiian school system in 1959 to 1967; this sample and the original procedures have been described in detail in Digman (1989; 1994), Goldberg (2001), and Hampson & Goldberg (2006). The present analyses included data from 2,215 children in the Oahu and Kauai subsamples for whom a common set of original personality variables were available (Goldberg 2001). In 1999, an intensive 40-year longitudinal follow-up of these individuals began (see Hampson, Dubanoski, Hamada, Marsella, Matsukawa, Suarez, & Goldberg, 2001). The total initial follow-up sample had reached 1,321 persons at the time of data preparation for the present analyses, with continued recruitment and subsequent follow-ups still ongoing. Predictive validity analyses (our third objective) employed data on 1,177 to 1,189 Oahu and Kauai participants with both usable childhood personality and midlife health outcome data. The childhood Oahu/Kauai sample was 47.4% female (n = 1,050) and 52.6% male (n = 1,165).

Measures

Childhood personality traits—Teachers were provided a set of trait descriptors (each with its associated behavioral examples) and asked to rank their students on each trait in turn. The ranking was achieved by sorting the names of students in the class into nine bins for each trait, alternating between children displaying the highest level, then the lowest levels of each trait, then the second highest and second lowest, and so forth. This procedure forced a quasi-normal distribution for each trait, with comparable means and standard deviations across classrooms. It was analogous to a Q-sort, but rather than sorting attributes teachers rank-ordered their students on each different trait.

Midlife general and cardiovascular health—A series of 40-year follow-up questionnaires included items used in epidemiologic studies to assess general and cardiovascular health. The latter consisted of a set of cardiovascular conditions after the stem "has a doctor ever diagnosed you with..." The cardiovascular conditions included hypertension (high blood pressure), heart attack, other heart disease, stroke, hypercholesterolemia (high cholesterol), and diabetes. Answer choices were no; yes, but never treated; yes, treated in the past but not at present; and yes, being treated now. To examine 40-year incidence (that is, any occurrence over the follow-up period; Rothman, Greenland, & Lasch, 2008), we coded each discrete event as having happened (1) or not (0).

General health was assessed by the SF-1, or General Health Item (Benyamini & Idler, 1999; Idler & Benyamini, 1997): "In general, my health is..." Answer choices are the ordinal options poor, fair, good, very good, or excellent. The item anchors the Medical Outcomes Study (MOS) Short Form-36 general health scale (Ware & Sherbourne, 1992), a mainstay of epidemiologic and medical cohort research. Although one-item measures are uncommon in psychology, the SF-1 is regularly used in public health research on the strength of two different meta-analyses showing that it consistently predicts all-cause mortality (Benyamini & Idler, 1999; Idler & Benyamini, 1997). From a classical test theory standpoint, the reliability index of a measure (i.e., the square root of its internal consistency) is a correlation between the measure and true scores on a construct. If one grants that latent risk for death is the "true score" for general health, the item might be said to be strongly reliable based on these meta-analyses. In the sense that one views time to actual death as an important criterion measure, the SF-1 likewise might be said to have excellent criterion validity.

Analytic Procedures

Question 1: Do ARC Types Represent the Optimal Cluster Structure for Big Five Data?—Prior to our main analyses, we derived Big Five scores from the trait variables by conducting a principal components analysis, extracting 5 components, and rotating them obliquely (by promax). To remain consistent with previous work in this sample (Goldberg, 2001), components were used but principal axis factors correlated .98 to .99 with the components. We determined the optimal cluster structure using an elaboration of the usual ARC procedure (based on Breckenridge, 1989; 2000). This involves randomly splitting the sample in half, then separately within each subsample conducting Ward's hierarchical clustering (using Euclidean distance) to derive initial cluster centers. Then K-means clustering is conducted to "clean up" the cluster solution. The K-means algorithm reassigns individuals as necessary to achieve clusters with the "tightest" within-cluster similarity and the greatest between-cluster variability, producing a clearer solution--much as rotating a factor solution optimizes its clarity.

Next, cases from the first subsample were assigned to clusters derived in the second subsample (on the basis of Euclidean distance to those cluster centers). The same cross-classification was then conducted with cases in the second subsample matched to clusters derived in the first. Finally, Cohen's Kappa was used to determine cross-classification agreement. This provides a measure of cluster structure replicability.

We bootstrapped this basic technique, an idea proposed by Barbaranelli (2002). For a given level of cluster solution (e.g., 4 clusters), we drew a bootstrap sample (that is, a random sample with replacement) from the data, performed the above procedure, and recorded the Kappa measurement of classification accuracy. We then repeated this iteratively, resulting in a distribution of Kappas for solutions with that number of clusters. The central tendency of the distribution describes the average replicability of the solution from repeated random split-sample cross-validations, and the standard deviation of the Kappa distribution describes the variability of this average. Distributions with high means and low variances suggest

highly replicable solutions because not only is the Kappa usually high but it varies little about its central tendency. The skewness and kurtosis of the distribution are also informative. A cluster solution producing a Kappa distribution with a high mean, low variance, long left tail (negative skew), and "spiking" around a high mean (leptikurtotic) is superior to a solution producing a distribution with the opposite characteristics. For solutions including two through nine clusters, we conducted 500 bootstrapped split-sample replications, creating empirical distributions of 1,000 Kappas for each number of clusters. We then compared the Kappa distributions with the Komolgorov test of equality of distribution functions.

Question 2: Are ARC Types Discrete or Fuzzy? Is There a Spectrum of Prototypicality?—Each observation was assigned a raw prototypicality score based on Cronbach and Gleser's (1953) D (Euclidean Distance) measure of profile dissimilarity between the observation's Big Five profile, and the Big Five profile of each ARC type. This resulted in three different sample distributions of prototypicality, one for each of the ARC prototypes. If these distributions are relatively normal, they suggest a continuum of variation in the extent to which persons resemble a particular prototype. If they are bimodal, they indicate that persons fall into groups of either high similarity or high dissimilarity to that prototype. In turn, that would support a discrete categorical approach to the types. We also compared the degree of overlap between distributions of prototypicality scores for different prototypes. Large regions of overlap would suggest large proportions of the sample with nearly equal similarity to two or more prototype profiles. Such a taxonomy would be less useful, leading to ambiguous classification of many people.

We examined cluster "fuzziness" by graphically decomposing Big Five space into a series of planes that displayed scatter plots of each Big Five dimension against one of the others. Because obvious separability of the points according to cluster often is not apparent in actual data (Everitt et al., 2001), members of each cluster were depicted in a particular color. To the extent that types are similar to one another, they will be located proximally in multivariate space. To the extent that they are "fuzzy" or lack discrete boundaries, certain regions of one or more planes may be inhabited by members of two or more clusters. Finally, the shape and location of the clusters across planes provides a useful understanding of within- versus between- cluster variability of individuals. To display fuzziness in a discernable way, we used only a random 20 percent of the observations; otherwise, the density of points would prevent easy visualization. To provide a meaningful metric for the gradient of prototypicality, we transformed D scores to percentiles so that higher percentiles reflected greater overall similarity between an individual Big Five profile and a prototype Big Five profile.

Question 3: Do ARC Types and/or Prototypicality Scores Display Long-Term Predictive Validity for Health Outcomes? Are They Better Predictors Than Traits?—Before modeling the associations between personality in childhood and midlife health, we examined adult sample bias by investigating the associations between childhood personality traits and the availability of adult follow-up data. We then predicted midlife health outcomes using probit models for 40-year incidence of each cardiovascular condition, and ordinal probits for the self-assessment of general health. Probit models treat binary data as reflecting an underlying continuous distribution, consistent with the notion that a diagnosed condition reflects underlying pathophyisology exceeding some threshold on a continuum (Isselbacher, Braunwald, Wilson, Martin, Fauci, & Kasper, 1994). This is also consistent with the emphasis in psychology on latent continua. The coefficient of the probit model reflects the shift along the latent outcome continuum in standard deviation (Z-score) units associated with a 1-unit change in each predictor.

We fit these models for (a) categorical ARC types (dummy variables for Overcontrolled and Undercontrolled, against a reference category of Resilient); and (b) ARC prototypicality scores (scaled so that 1 unit reflects a 25% increase in similarity to the prototype; the rectangular distribution of percentile scores is not a problem for probit models). All models also controlled for gender. We found that probit and logit models yielded virtually identical results. The purpose of these models was to examine whether any association existed at all between childhood ARC constructs and health outcomes 40 years later.

The next series of models examined which kind of childhood personality structure—ARC types, ARC prototypicality scores, or Big Five traits—yielded the best overall prediction of midlife health outcomes, again based on probit models. We used two measures of fit for nonlinear models: First, we used the Bayesian Information Criteria (BIC), which is a function of the log likelihood of the model, with an increasing penalty for additional predictors. BIC differences of 10 or greater are considered "very strong" evidence favoring one model over the other (Kass & Rafferty, 1995). Second, for the cardiovascular outcomes, we examined the Area Under the Receiving Operator Curve (AUC) resulting from each model. The AUC is a measure of the percentage of correct positive predictions (i.e., incident cases) across all possible cut-points of the predicted probabilities generated by the model. Randomly guessing about whether children in the Hawaii cohort would be diagnosed with a particular cardiovascular condition sometime over the next 40 years would produce an AUC of .5, whereas perfect prediction of incident cases would yield an AUC of 1. We also computed the 95% confidence intervals (95% CIs) around the AUC for each model. One model has a significantly higher AUC than another if the two 95% CIs do not overlap. If a 95% CI encompasses .5, the predictor set yields no better classification than randomly guessing.

Results

Preliminary Factoring

The principal component solution for teacher-rated trait variables was characterized by five components with rotated variance percentages of 9.8, 8.7, 6.1, 5.0, and 4.4, collectively capturing 87% of the variance in the personality traits. Parallel analysis confirmed that only the first five Eigenvalues differed significantly from those produced by random data. The pattern matrix, depicted in on-line supplement Table S1, revealed loadings in concert with the traditional Big Five factor structure. The first two components corresponded to low Conscientiousness and low Agreeableness, respectively, so scores were reversed for subsequent analyses. Component intercorrelations ranged from -.16 (Extraversion and Neuroticism) to .48 (Conscientiousness and Agreeableness).

Question 1: Do ARC Types Represent the Optimal Cluster Structure for Big Five Data?—Table 1 shows the characteristics of the empirical Kappa distributions for the bootstrap-aggregated cross-validation clustering. Distribution graphs can be found in online Figure S1. The three-cluster solution was clearly superior to its closest competitors by the Komolgorov equality of distribution test, showing a spike around an average Kappa of .83. The Big Five Profiles of the three clusters, listed in Table 2 and depicted in online Figure S2, corresponded to those reported previous for the three ARC types and were so named. Additional information about the means of specific traits by type is available in on-line Table S2.

Question 2: Are ARC Types Discrete or Fuzzy? Is There a Spectrum of Prototypicality?—Figure 1 shows the distributions of Cronbach and Gleser's (1953) D measure of dissimilarity to each ARC prototype profile, for children classified as

Undercontrolled, Overcontrolled, or Resilient. Figure 1 reveals that observations classified into a type did indeed show the greatest average similarity to that type's Big Five profile. However, the degree of overlap in D distributions within each type suggests that non-trivial numbers of individuals were roughly similar to two or even three ARC prototypes. Finally, the distributions in Figure 1 are unimodal and roughly normal, suggesting a continuity of type similarity.

Figure 2 depicts the graphical decomposition of multivariate Big Five space, showing scatter-plots of individuals in each of the bifactor planes. The ARC types appear to have overlapping borders within virtually all planes. This "cluster fuzziness" also suggests a dimensional nature to prototypicality.[3]

Question 3: Do ARC Types and/or Prototypicality Scores Display Long-Term Predictive Validity for Health Outcomes? Are They Better Predictors Than Traits?—Preliminary analyses showed that childhood ARC type was not significantly related to follow-up status, although there was a trend for Undercontrolled children to be less likely to have been followed-up as of the time of this writing (p = .07).

The top portion of Table 3 shows the association between categorical ARC types and midlife health outcomes. Relative to Resilients, the Over- and Under-controlled types reported roughly -.18 SD worse general health scores on the SF-1. Relative to resilients, Undercontrollers were at greater risk for incident hypertension over the 40 year followup (i.e., .25 SD increase on the latent continuum of blood-pressure) and Undercontrollers were at greater risk for stroke (.47 SD increase on the latent continuum of cerebrovascular deterioration). Dimensional prototypicality scores showed a similar pattern, with 25% increases in resilient prototypicality associated with .09 SD increases and .09 decreases on the latent general health and hypertensive continua, respectively. Increased Resilient prototypicality was also associated with decreased cardiovascular disease risk, an association not apparent with the categorical types. Online Figure S3 graphs the predicted probabilities for protypicality scores vs. discrete types.

Table 4 shows for each outcome the aggregate predictive validity of the categorical ARC types, the dimensional prototypicality scores, the Big Five factor scores, and all combinations of these. BICs for each model indicated that the categorical ARC types and the ARC prototypicality scores yielded similar prediction for all outcomes except hypercholesterolemia and cardiovascular disease. For these outcomes, categorical types showed strong evidence of better model fit (i.e., BICs 10 or more less) than traits. Combinations of predictor sets resulted in increasingly higher BICs, consistent with decreasing parsimony of the predictor set.[4]

---

[3]In addition to these, we also examined the spectrum of prototypicality using a multivariate graphic technique known as Chernoff faces (Chernoff, 1973; Raciborski, in press; Tufte, 2001). This technique capitalizes on people's innate ability to discern fine gradients in facial expressions. Individual observations are portrayed as cartoon-like faces, where the geometric characteristics of different facial features are linked to numeric values of different variables. In other words, the angle of smile inflection of one observation's Chernoff face is dictated by its values on variable 1, while the dilation of its pupils is linked to its values on variable 2, and so forth. We constructed Chernoff faces for each prototype profile, linking values on the Big 5 to logically corresponding facial features (i.e., higher Neuroticism = greater frown curvature). We then randomly selected observations 1%, 20%, 40%, 60%, and 80% similar to the prototype and arrayed their Chernoff faces from left to right. The results showed an increasing gradient of subjective similarity to the prototype face, with observations only beginning to resemble prototypes at 60% similarity and only clearly like their prototype at 80% similarity. Further technical information and the graphical array are available in the online supplement (Figure S4).

[4]Exploratory analyses also examined ARC type-Big 5 interactions for health outcomes. Only one appeared robust (p = .016): A 1 SD increase in Extraversion was associated with a trend toward reduced risk for midlife hypertension in childhood Resilients (B (SE) = -.11 (.06), z = -1.77, p = .077), but was associated with significantly *increased* hypertension risk in Undercontrollers (B (SE) = .27 (.11), z = 2.37, p = .018).

AUCs indicated that personality data in any form enhanced predictive accuracy over guessing whether an individual would experience one of the cardiovascular outcomes over the 40-year period. For hypertension and stroke, no one form of personality data (i.e., types, prototypicality scores, traits) was superior to another. However, neither ARC types nor prototypicality scores predicted high cholesterol, diabetes, or cardiac arrest at greater than chance accuracy, whereas Big Five trait data did so. Prototypicality scores appeared superior to types for cardiovascular disease, but was not a better predictor set than the Big Five. Combining multiple predictor sets did not significantly increase predictive power. Childhood personality information appeared to show the greatest predictive accuracy for 40-year stroke incidence, with the combination of traits and types yielding AUCs indicating 72% accuracy.

Discussion

Question 1: Do ARC Types Represent the Optimal Cluster Structure for Big Five Data?

Our analysis implemented a bootstrap cross-validation version of the clustering procedures traditionally used in studies of personality types. The results provided very clear evidence in favor of a tripartite cluster solution, when compared to other structures ranging from two to nine clusters. The three clusters corresponded closely to the theoretical ARC types at both the level of the Big Five factors and that of specific trait descriptors.

An intriguing question concerns previous failures to replicate the ARC typology. There are at least two possible explanations, which are not necessarily mutually exclusive. First, the ARC clusters may not be empirically robust entities, in the sense that they only appear in some samples under some methodological procedures. For instance, samples of prisoners or thrill seekers might show an overrepresentation of undercontrollers and very few overcontrollers, leading to a different cluster structure. Samples of high achievers and stable, successful persons would might an overrepresentation of resilients, possibly yielding a different overall clustering. These possibilities suggest that unlike the Big Five, which are postulated to be a relatively population-invariant variable structure, the ARC person-type structure may depend more on the particular samples. Second, a wide variety of methods exist to examine cluster structure. Split-sample replication techniques relying on only a single or few replications may fail if the random splits yield, by chance, subsamples with different cluster structures. Aggregating over hundreds or thousands of replications takes care of this problem. This suggests that regardless of the specific clustering method used (partitioning, hierarchical, model-based/latent profile techniques, etc.), resampling can be fruitfully used to examine robustness. And it may be the case that oblique, rather than orthogonal, Big Five scores are most likely to yield the ARC structure.

Question 2: Are ARC Types Discrete or Fuzzy? Is There a Spectrum of Prototypicality?

Our analysis yielded some potentially valuable findings about this issue. First, distributions of prototypicality scores were roughly normal and overlapped even within subgroups classified based on categorical ARC type. This means that non-trivial heterogeneity exists within types derived from the best cluster structure. Second, graphical decompositions of Big 5 space revealed that although the types did generally occupy different sectors of each plane, but were characterized by diffuse or fuzzy, rather than discrete, boundaries. In other words, an individual located at the boundary of all three clusters consistently over multiple planes of Big Five space might show little resemblance to any prototype, whereas others might show modest similarity to multiple ARC types. Such findings are similar to those of many investigations of Axis II construct dimensionality (reviewed in Widiger & Trull, 2007). Thus, general recognition of the fuzzy structure of ARC types may not be particularly surprising to some.

All of these considerations suggest that while a tripartite cluster structure may be the most replicable, it is characterized by substantial fuzziness. This conclusion represents a return to the original spirit of ARC types. In Block's (1971) initial work, the types were derived through inverse factoring of Q-sort data, and the loadings of individuals on person factors represented the degree of resemblance that individual bore to the prototype represented by the person-factor. Our findings ultimately reinforce the usefulness of such a "typeness" conception of ARC constructs. Asdendorpf (2006b) also highlighted this idea. Other work on dimensional profile similarity could also be applied (Barret, 2005; McCrae, 1993).

Question 3: Do ARC Types and/or Prototypicality Scores Display Long-Term Predictive Validity for Health Outcomes? Are They Better Predictors Than Traits?

The issue of the predictive validity of ARC types versus traits has been a topic of debate (Asendorpf, 2003; Costa et al., 2002; von Eye & Bogat, 2006). Our first finding was that both categorical types and dimensional prototypicality scores show prospective associations with general health at midlife, as well as with the 40-year incidence of hypercholesterolemia and stroke. Resilient children enjoyed better general health and lower risk of these outcomes relative to Overcontrolled (for hypertension) and Undercontrolled (for stroke) children. Overcontrolled findings are consistent with the Type D personality literature, a profile similar to Overcontrollers, and associated with cardiovascular ill-health (Denollet, Pedersen, et al., 2006; Denollet, Schiffer, et al., 2009). While these observations were partially consistent with hypotheses, some specific disease associations were not observed.

One possible reason is that specific vascular diseases often arise over time out of hypertension and hypercholesterolemia (Isselbacher et al., 2004). For instance, a robust association was observed between the Overcontrolled type and hypertension, but not between this type and other diseases that emerge from hypertension. With age, however, such associations may become apparent. By contrast, dimensional resiliency was associated with both a major risk factor (hypercholesterolemia) and disease outcome (heart disease). The specific disease association in this case may be due to a somewhat earlier lifecourse emergence of heart disease, given high cholesterol (12.3% prevalence ages 45-64 vs. 4.6% ages 18 to 44; CDC, 2009). The emergence of strokes, by contrast, tends to be somewhat later in the lifespan (CDC 2009). Absence of personality associations with diabetes, a common condition by midlife, may simply reflect a failure of personality types to capture salient etiologic factors like high sugar consumption. Finally, stroke risk was noted for Undercontrollers in the presence of non-significant links between Undercontroller and hypertension. This type may therefore be susceptible to other stroke-specific etiologic factors, such as genetic risk for weakening of cerebral blood vessels (Isselbacher et al., 1994).

Effect sizes for categorical outcomes ranged from roughly .20 to .50 SD on the latent continua underlying these outcomes; for 25 percent increases on the Resiliency spectrum, they were roughly half this size. The pattern of findings also indicated that dimensionalizing prototypes did not necessarily afford increased ability to detect associations with specific outcomes, although it did for cardiovascular disease. One reason may be that dimensional variables are only better predictors if they are linearly related (via a probit link, in this case) to the outcome. We did not test non-linear transformations of the dimensional scores, such as squares or squares roots. In contrast to linear relations, outcomes may also show more discontinuous differences or threshold effects corresponding to classes defined by groupings of predictors.

With respect to overall predictive power and incremental validity, we examined two indices: the BIC, a measure of overall model fit strongly rewarding parsimony in the predictor set; and the AUC, or percentage of correct positive predictions generated by the model across all

possible probability cut-points. The BIC indicated no clear difference between categorical ARC types and ARC prototypicality scores, although both evidenced considerably better BICs (i.e., >10 lower) than the traits. This is commensurate with the fact that categorical ARC models involve estimating only two parameters (two dummy variables for three categories), and prototypicality scores condense personality information into three dimensions, compared to five for the trait factors. Combining predictor sets yielded worse BICs, consistent with the increasing number of parameters and apparent redundancy of information.

AUCs revealed that childhood categorical ARC types and ARC prototypicality scores predicted adult cases of hypertension, stroke, and cardiovascular disease better than randomly guessing whether a child would develop one of these conditions. However, only continuous trait data predicted positive occurrences of all cardiovascular outcomes at greater than chance levels. AUC differences across types, prototypicality scores, and traits were not significantly different. As well, no combination of categorical ARC types, prototypicality scores, and traits yielded significant predictive improvements over any one form of personality data alone. The lack of evidence for incremental validity suggests that adding traits and types provides redundant predictive information for these outcomes. This point has been made in earlier work (Costa et al., 2002 with respect to depression. In such instances, one faces the question of whether one wishes to package the criterion-relevant personality information in trait vs. type form. Interpretation or theory-building may be facilitated better by one or the other. However, in other studies, dimensional type status has been uniquely associated with health and mental health factors even after controlling traits (Chapman et al., 2007). Thus, incremental validity may depend on the outcome, and/or nature of the subject sample.

In many respects, it is remarkable to observe any associations between childhood personality and health outcomes over four decades later. The length of this period represents a substantial predictive challenge even for biological markers of disease. For instance, the popular Framingham risk model (Anderson, Wilson, Odell, & Kannel, 1991; Wilson, D'Agostino, Levy, Belanger, Sibershatz, & Kannel, 1998) focuses on predicting cardiovascular events over 12 years. The Framingham model is also based on iterative procedures selecting maximally predictive variables from a large set of candidate factors, including biomarkers and physiological measurements. Our models, by contrast, were simple a priori specifications containing only childhood personality data. The upper AUC obtained from our models for any outcome--.72, for 40-year stroke incidence, based on childhood types and traits--rivaled the Framingham 12-year AUCs across men and women for medically documented coronary heart disease (CHD; AUC's =.69-.77). This level of predictive accuracy is obtained using lipid panels, blood-pressure, smoking, diabetes, and age (Wilson et al., 1998). The maximal personality AUC in our analysis for a roughly equivalent outcome to CHD, self-report of 40-year incident heart disease, was .67 (based on all forms of personality information).

The reasons that childhood personality traits appear to be related to later health have been discussed at length elsewhere (Friedman, 2000; Hampson et al., 2006), but the general consensus implicates life-course pathways leading to variations in health behaviors such as activity, diet, and smoking (Friedman, 2000; Hampson et al., 2006). Nevertheless, personality appears to work through psychophysiological pathways as well (Smith & Spiro, 2002). In particular, ARC types in children are reportedly associated with physiological parameters relevant to cardiovascular health (Hart, Burock, London, Atkins, & Bonilla-Santiago, 2005). An ambiguous association also exists with respect to early personality traits and ultimate attainment of socioeconomic status (SES), which is strongly related to health. There is some evidence that socioeconomic circumstances themselves shape ARC constructs

during adolescence (Hart, Atkins, & Matsuba, 2008), and the interplay of socioeconomic factors, personality traits, and health outcomes over the life-course remains poorly understood (Chapman, Fiscella, Duberstein, Coetta, & Kawachi, 2009; Chapman, Fiscella, Kawachi, & Duberstein, in press).

Summary and Recommendations

Our findings suggest that ARC types represent the optimal cluster structure in teacher ratings of children's personality. The types arise from a fuzzy, rather than discrete, cluster structure, suggesting that individuals fall on a spectrum from lesser to greater resemblance to each type. As a result, ARC constructs appear most accurately represented by a set of spectra, each ranging from great dissimilarity to high similarity to a particular ARC prototypical Big Five profile. Finally, ARC types and prototypicality scores in childhood appear to have long-term relevance for general health, as well as select cardiovascular outcomes such as stroke and heart disease. Big Five factor scores in childhood are associated with a wider array of cardiovascular outcomes, but type ARC constructs provide comparable predictive power for some outcomes more parsimoniously. The classification accuracy for some outcomes approached predictive models for similar outcomes in epidemiology, which used physiological measurements, variable selection techniques, and shorter follow-up periods. These considerations lead to the following recommendations:

A. Clustering methods based on split-sample and/or resampling replicability using oblique Big Five scores appear to be a useful approach for recovering ARC types.

B. Investigators should continue to examine a range of clustering solutions, as prior work has shown some variability in the number and nature of prototypes found.

C. Individuals should be characterized not just in terms of discrete category membership but in terms of dimensional similarity to each prototype profile.

D. In predictive analyses, typological and trait approaches should both be considered. They may provide complementary information about a given outcome. In modeling situations where degrees of freedom need to be preserved, types or prototypicality may provide a more parsimonious predictor set.

Strengths, Limitations, and Future Directions

Our findings must be interpreted in the light of a balanced consideration of the strengths and limitations of our study. First, we did not address the developmental **antecedents** giving rise to ARC constructs among the study participants. This should be an important topic for further research. Second, our midlife health outcomes were self-reported, rather than based on medical examinations. To the extent that self-reporting errors are random with respect to teacher ratings of personality, this would likely attenuate our statistical power. In order for this to artificially induce spurious associations, children assessed by their teachers as Undercontrolled or Overcontrolled would have to systematically over-report the incidence of bad-health outcomes over the 40-year follow-up period. Although not inconceivable, we find this unlikely given the consistency, and the magnitude, of the childhood-adulthood associations. (Note that potentially under-reporting of midlife health outcomes among Undercontrolled and Overcontrolled children could occur with comparable likelihood, in which case the predictive power of the childhood ARC types would be systematically underestimated.)

A third qualification is that the adult follow-up is ongoing, so not all of the child cohort could be included in our adult predictive models. Some children were unlocatable for follow-up, and others are still being contacted and recruited. This is inevitable in large cohort studies, particularly those with time frames of this length, so later analyses involving

additional follow-up data will be useful. Especially important will be analyses of the data from the adult medical/physical examinations, which are now ongoing.

A fourth caveat is that we examined only one set of outcomes, motivated by general theories linking personality and health and by the high public health relevance of cardiovascular conditions. Future work might examine other outcomes, including socioeconomic status and career success, interpersonal relationships, and other important life factors. In addition, future analyses should examine the stability of ARC constructs because prior reports indicated considerable personality development between childhood and midlife in this sample (Hampson & Goldberg, 2006).

Finally, the Hawaii sample reflects a unique demographic population, birth cohort, and set of childhood experiences in a particular locale during a particular era of history. Such considerations must be appropriately weighed in determining the extent to which our findings might extend to populations that are very different, such as modern urban youth. This underscores the need for examination of ARC constructs in demographically different populations.

Contrasted with these qualifications are several study strengths. First, our research includes the largest cohort and most extensive analysis of ARC constructs of which we are aware. Second, we deployed nuanced methodological procedures based on resampling and dimensionalizing similarity to categories. Third, we were able to examine outcomes of great individual and public health consequence, in the form of general health and cardiovascular conditions. Fourth, a few of these outcomes could be predicted with non-trivial accuracy relative to specialized risk models used in epidemiology and medicine. This is surprising given the simple specifications of our models, the use of only childhood personality data (which shares no common method variance with the outcomes), and a follow-up duration of roughly four decades; the development of more optimized predictive models, possibly using algorithmic modeling techniques (Breiman, 2001), is an important topic for future investigation. Fifth, other than the classic Terman cohort, we are unaware of any other work on childhood personality in such a large cohort over such a long follow-up period, with a comparable breadth of personality information and follow-up time.

In the final analysis, the ARC constructs are an interesting person-centered approach to personality, with a rich theoretical history (Block, 1971). Although quantitative tools for type research are less commonly utilized than factoring methods for Big-Five trait-based approaches, they too have a rich history (e.g., Cronbach & Gleser, 1953; Gibson, 1959). We believe that ARC and Big Five approaches can fruitfully co-exist, and perhaps be implemented in a complementary manner in personality research. Certainly in many cases classic trait approaches will be more useful, but we encourage open consideration of ARC constructs, or at least continuing efforts to understand them.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Allport, GW. Personality: A psychological interpretation. Holt; New York: 1937.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders-IV-TR. American Psychiatric Association; Arlington VA: 2000.

Anderson KM, Wilson PWF, Odell PM, Kannel WB. An updated coronary risk profile: a statement for health professionals. Circulation. 1991; 83:357–363.

Asendorpf JB. Editorial: The puzzle of personality types. European Journal of Personality. 2002; 16:S1–S5.

Asendorpf JB. Head-to-head comparison of the predictive validity of personality types and dimensions. European Journal of Personality. 2003; 17:327–346.

Asendorpf JB. On artifacts and meaning in person-centered analyses: Comment on McCrae, Terracciano, Costa, and Ozer (2005). European Journal of Personality. 2006a; 20:45–47.

Asendorpf JB. Typeness of personality profiles: A continuous person-centered approach to personality data. European Journal of Personality. 2006b; 20:83–106.

Asendorpf JB, Borkenau P, Ostendorf F, Van Aken MAG. Carving personality description at its joints: Confirmation of three replicable personality prototypes for both children and adults. European Journal of Personality. 2001; 15:169–198.

Asendorpf JB, Denissen JJA. Predictive validity of personality types versus personality dimensions from early childhood to adulthood: Implications for the distinction between core and surface traits. Merrill-Palmer Quarterly. 2006; 52:486–513.

Asendorpf JB, Van Aken MAG. Resilient, overcontrolled, and undercontrolled personality prototypes in childhood: Replicability, predictive power, and the trait-type issue. Journal of Personality and Social Psychology. 1999; 77:815–832. [PubMed: 10531673]

Barbaranelli C. Evaluating cluster analysis solutions: An application to the Italian NEO personality inventory. European Journal of Personality. 2002; 16:S43–S55. SpecIssue.

Barrett, P. Person-target profiling.. In: Beauducel, A.; Biehl, B.; Bosnjak, M.; Conrad, W.; Schonberger, G.; Wagener, D., editors. Multivariate research strategies: A festschrift for Werner Wittman. Shaker-Verlag; Aachen: 2005. p. 63-118.

Benjamini Y, Idler EL. Community studies reporting associations between self-rated health and mortality – additional studies, 1995 to 1998. Research on Aging. 1999; 21:392–401.

Bergman LR, Trost K. The person-oriented versus the variable-oriented approach: Are they complementary, opposites, or exploring different worlds? Merrill-Palmer Quarterly. 2006; 52:601–632.

Block, J. Lives through time. Bancroft Books; Berkeley, CA: 1971.

Block, JH.; Block, J. The role of ego control and ego resiliency in the organization of behavior.. In: Collins, WA., editor. Development of cognition, affect, and social relations: The Minnesota symposium on child psychology. Erlbaum; Hillsdale, NJ: 1980. p. 39-101.

Boehm B, Asendorpf JB, Avia MD. Replicable types and subtypes of personality: Spanish NEO-PI samples. European Journal of Personality. 2002; 16:S25–S41.

Breckenridge JN. Replicating cluster analysis: Method, consistency, and validity. Multivariate Behavioral Research. 1989; 24:147–161.

Breckenridge JN. Validating cluster analysis: Consistent replication and symmetry. Multivariate Behavioral Research. 2000; 35:261–285.

Breiman L. Statistical modeling: The two cultures. Statistical Science. 2001; 16:199–231.

Caspi, A. Personality development across the life course.. In: Eisenberg, N., editor. Handbook of child psychology. 5th ed.. Wiley; New York: 1998. p. 311-388.

Caspi A, Silva PA. Temperamental qualities at age-3 predict personality-traits in young adulthood - longitudinal evidence from a birth cohort. Child Development. 1995; 66:486–498. [PubMed: 7750379]

Centers for Disease Control. Summary health statistics for US Adults: National Health Interview Survey, 2008. Us Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics; Hyattsvile, MA: 2009. DHHS Publication No. (PHS) 2010-1570

Chapman BP, Duberstein P, Lyness JL. The distressed personality type: Replicability and general health associations. European Journal of Personality. 2007; 21:911–929.

Chapman BP, Fiscella KF, Duberstein PR, Coletta M, Kawachi I. Can the influence of childhood socioeconomic status on men's and women's adult body mass be explained by adult socioeconomic status or personality? Findings from a national sample. Health Psychology. 2009; 28:419–427. [PubMed: 19594266]

Chapman BP, Fiscella KF, Kawachi I, Duberstein PR. Socioeconomic status, personality, and all-cause mortality in the US. American Journal of Epidemiology. in press.

Chernoff H. The use of faces to represent points in k-dimensional space graphically. Journal of the American Statistical Association. 1973; 68:361–368.

Costa JT, Herbst JH, McCrae RR, Samuels J, Ozer PJ. Replicability and utility of three personality types. European Journal of Personality. 2002; 16:573–587.

Costa, PT., Jr.; Piedmont, RL. Interpretation of Revised NEO Personality Inventory profiles of Madeline G.: Self, partner, and an integrated perspective.. In: Wiggins, JS., editor. Paradigms of personality assessment. Guilford Press; New York: 2003.

Costa, PT.; Widiger, TA. Personality disorders and the five-factor model of personality. 2nd ed.. American Psychological Association; Washington, DC: 2001.

Cronbach LJ, Gleser GC. Assessing similarity between profiles. Psychological Bulletin. 1953; 50:456–473. [PubMed: 13112334]

De Fruyt F, Denollet J. Type D personality: A Five Factor Model perspective. Psychology and Health. 2002; 17:671–683.

De Fruyt F, Mervielde I, Van Leeuwen K. The consistency of personality type classification across samples and five-factor measures. European Journal of Personality. 2002; 16:S57–S72.

Denollet J, Van Heck GL. Psychological risk factors in heart disease - what type D personality is (not) about. Journal of Psychosomatic Research. 2001; 51:465–468. [PubMed: 11602215]

Denollet J, Pedersen SS, Vrints CJ, Conraads VM. Usefulness of type D personality in predicting five-year cardiac events above and beyond concurrent symptoms of stress in patients with coronary heart disease. American Journal of Cardiology. 2006; 97:970–973. [PubMed: 16563897]

Denollet J, Schiffer AA, Kwaijtaal M, Hooijkaas H, Hendriks EH, Widdershoven JW, Kupper N. Usefulness of type D personality and kidney dysfunction as predictors of interpatient variability in inflammatory activation in chronic heart failure. The American Journal of Cardiology. 2009; 103:399–404. [PubMed: 19166697]

Digman JM. Five robust trait dimensions: Development, stability, and utility. Journal of Personality. 1989; 57:195–214. [PubMed: 2671337]

Digman, JM. Child personality and temperament: Does the five-factor model embrace both domains.. In: Halverson, CF., Jr.; Kohnstamm, GA.; Martin, RP., editors. The developing structure of temperament and personality from infancy to adulthood. Erlbaum; Hillsdale, NJ: 1994. p. 323-338.

Everitt, B.; Landau, S.; Leese, W. Cluster analysis. 4th ed.. A Hodder Arnold; London: 2001.

Eysenck HJ. Personality, stress, and disease: An interactionist perspective. Psychological Inquiry. 1991a; 2:221–232.

Eysenck HJ. Reply to criticisms of the Grossarth-Maticek studies. Psychological Inquiry. 1991b; 2:297–323.

Friedman HS, Booth-Kewley S. Personality, type A behavior, and coronary heart disease: The role of emotional expression. Journal of Personality and Social Psychology. 1987; 53:783–792. [PubMed: 3681651]

Friedman HS. Long-term relations of personality and health: Dynamism, mechanisms, and tropisms. Journal of Personality. 2000; 68:1089–1107. [PubMed: 11130733]

Friedman HS, Tucker JS, Tomlinson-Keasey C, Scwartz JE, Wingard DL, Criqui MH. Does childhood personality predict longevity? Journal of Personality and Social Pyshcology. 1993; 65:176–185.

Funder DC. Global traits: A neo-Allportian approach to personality. Psychological Science. 1991; 2:31–39.

Gibson WA. Three multivariate models: Factor analysis, latent structure analysis, and latent profile analysis. Psychometrika. 1959; 24:229–252.

Goldberg LR. The structure of phenotypic personality traits. American Psychologist. 1993; 48:26–34. [PubMed: 8427480]

Goldberg LR. Analyses of Digman's child-personality data: Derivation of Big-Five factor scores from each of six samples. Journal of Personality. 2001; 69:709–743. [PubMed: 11575511]

Goodwin RD, Friedman HS. Health status and the five-factor personality traits in a nationally representative sample. Journal of Health Psychology. 2006; 11:643–654. [PubMed: 16908463]

Grossarth-Maticek R, Eysenck HJ, Vetter H. Personality type, smoking habit and their interaction as predictors of cancer and coronary heart disease. Personality and Individual Differences. 1988; 9:479–495.

Hampson SE, Dubanoski JP, Hamada W, Marsella AJ, Matsukawa J, Suarez E, Goldberg LR. 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. 2001; 35:375–387.

Hampson SE, Goldberg LG. A first large cohort study of personality trait stability over the 40 years between elementary school and midlife. Journal of Personality and Social Psychology. 2006; 91:473–779.

Hampson SE, Goldberg LG, Vogt TM, Dubanoski JP. Forty years on: Teacher's assessments of children's personality traits predict self-reported health behaviors and outcomes at midlife. Health Psychology. 2006; 25:57–64. [PubMed: 16448298]

Hart D, Atkins R, Fegley S. Personality and development in childhood: A person-centered approach. Monographs of the Society of Research in Child Devlopment. 2003; 68:vii–109.

Hart D, Atkins R, Matsuba MK. The association of neighborhood poverty with personality change in adulthood. Journal of Personality and Social Psychology. 2008; 94:1048–61. [PubMed: 18505317]

Hart D, Burock D, London B, Atkins R, Bonilla-Santiago G. The relation of personality types to physiological, behavioural, and cognitive processes. European Journal of Personality. 2005; 19:391–407.

Hart D, Hofmann V, Edelstein W, Keller M. The relation of childhood personality types to adolescent behavior and development: A longitudinal study of Icelandic children. Developmental Psychology. 1997; 33:195–205. [PubMed: 9147829]

Herzberg PY, Hoyer J. Personality prototypes in adult offenders. Criminal Justice and Behavior. 2009; 36:259–274.

Herzberg PT, Roth M. Beyond resilients, undercontrollers, and overcontrollers? An extension of personality prototype research. European Journal of Psychological Assessment. 2006; 22:139–148.

Hofstee WKB. Types and variables: Towards a congenial procedure for handling personality data. European Journal of Personality. 2002; 16:S89–S96. SpecIssue.

Hofstee WKB, Barelds DPH, Ten Berge JMF. Structuring assessments of psychopathology. Journal of Individual Differences. 2007; 27:87–92.

Idler EL, Benjamini Y. Self-rated health and mortality: a review of twenty-seven community studies. Journal of Health and Social Behavior. 1997; 38:21–37. [PubMed: 9097506]

Isselbacher, KJ.; Braunwald, E.; Wilson, JD.; Martin, JB.; Fauci, AS.; Kasper, DL. Harison's principles of internal medicine. 13th ed.. McGraw-Hill; New York: 1994.

Jemal A, Ward E, Hao Y, Thun M. Trends in the leading causes of death in the United States, 1970-2002. JAMA-Journal of the American Medical Association. 2005; 294:1255–1259.

John, OP.; Srivastava, S. The Big Five trait taxonomy: History, measurement, and theoretical perspectives.. In: Pervin, LA.; John, OP., editors. Handbook of personality: Theory and research. 2nd ed.. Guilford; New York: 1999. p. 102-138.

Kass R, Rafferty AE. Bayes factors. Journal of the American Statistical Association. 1995; 90:773–795.

Laursen B, Hoff E. Person-centered and variable-centered approaches to longitudinal data. Merrill-Palmer Quarterly. 2006; 52:377–389.

Magnusson, D. Holistic interactionism: A perspective for research on personality development.. In: Pervin, LA.; John, OP., editors. Handbook of personality: Theory and research. 2nd ed.. Guilford; New York: 1999. p. 219-249.

McCrae RR. Agreement of personality profiles across observers. Multivariate Behavioral Research. 1993; 28:25–40.

McCrae RR, Costa PT. Reinterpreting the Myers-Briggs Type Indicator from the perspective of the five-factor model of personality. Journal of Personality. 1989; 57:17–40. [PubMed: 2709300]

McCrae RR, Costa PT Jr. Personality trait structure as a human universal. American Psychologist. 1997; 52:509–516. [PubMed: 9145021]

McCrae RR, Terracciano A, Costa PT, Ozer DJ. From types to typological thinking: A reply to Asendorpf. European Journal of Personality. 2006; 20:49–51.

Morizot J, Le Blanc M. Searching for a developmental typology of personality and its relations to antisocial behavior: A longitudinal study of an adjudicated men sample. Criminal Behaviour and Mental Health. 2003; 13:241–277. [PubMed: 14654863]

Morizot J, Le Blanc M. Searching for a developmental typology of personality and its relations to antisocial behavior: A longitudinal study of a representative sample of men. Journal of Personality. 2005; 73:139–182. [PubMed: 15660676]

Myers, IB.; McCaulley, MH. MBTI Manual: A guide to the development and use of the Myers-Briggs Type Indicator. Consulting Psychologist Press; Mountain View, CA: 1985.

Pulkkinen L. Female and male personality styles: A typological and developmental analysis. Journal of Personality and Social Psychology. 1996; 70:1288–1306. [PubMed: 8667169]

Raciborski R. Graphical representation of multivariate data using Chernoff faces. The Stata Journal. in press.

Rammstedt B, Riemann R, Angleitner A, Borkenau P. Resilients, overcontrollers, and undercontrollers: The replicability of the three personality prototypes across informants. European Journal of Personality. 2004; 18:1–14.

Roberts BW, Kuncel NR, Shiner R, Caspi A, Goldberg L. The power of personality: A comparative analysis of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. Perspectives in Psychological Science. 2007; 4:313–336.

Robins RW, John OP, Caspi A, Moffitt TE, Stouthamer-Loeber M. Resilient, overcontrolled, and undercontrolled boys: Three replicable personality types. Journal of Personality and Social Psychology. 1996; 70:157–171. [PubMed: 8558407]

Roth M, von Collani G. A head-to-head comparison of big-five types and traits in the prediction of social attitudes: Further evidence for a five-cluster typology. Journal of Individual Differences. 2007; 28:138–149.

Rothman, KJ.; Greenland, S.; Lasch, T. Modern epidemiology. 3rd ed.. Lippincott, Williams, & Wilkins; Philadelphia, PA: 2008.

Rusten, J.; Cunningham, IC. Theophrastus: Characters; Herodas: Mimes; Sophron and other mime fragments. 3rd ed.. Harvard University Press; Cambridge, MA: 2002.

Schnabel K, Asendorpf JB, Ostendorf F. Replicable types and subtypes of personality: German NEO-PI-R versus NEO-FFI. European Journal of Personality. 2002; 16:S7–S24.

Smith, TW.; Ruiz, JM. Psychosocial influences on the development and course of coronary heart disease: Current status and implications for research and practice. 2002. p. 548-568.

Smith TW, Spiro A. Personality, health, and aging: Prolegomenon for the next generation. Journal of Research in Personality. 2002; 36:363–394.

Stern, WI. Die differentielle psychologie in ihren methodischen grundlagen [Methodological foundations of differential psychology]. Johann Ambrosius Barth; Leipzig, Germany: 1911.

Temoshok L. Personality, coping style, emotion and cancer: Toward an integrative model. Cancer Survival. 1987; 6:545–567.

Tufte, ER. The visual display of quantitative information. Graphics Press; Cheshire, CT: 2001.

von Eye A, Bogat GA. Person-oriented and variable-oriented research: Concepts, results, and development. Merrill-Palmer Quarterly. 2006; 52:390–420.

Ware JE, Sherbourne CD. The MOS 36-item short-form health survey. Medical Care. 1992; 20:473–480. [PubMed: 1593914]

Widiger TA, Trull TJ. Plate tectonics in the classification of personality disorder: Shifting to a dimensional model. American Psychologist. 2007; 62:71–83. [PubMed: 17324033]

Wilson PWF, D'Agnostino RB, Levy D, Belanger AM, Sibershatz A, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation. 1998; 97:1837–1847. [PubMed: 9603539]
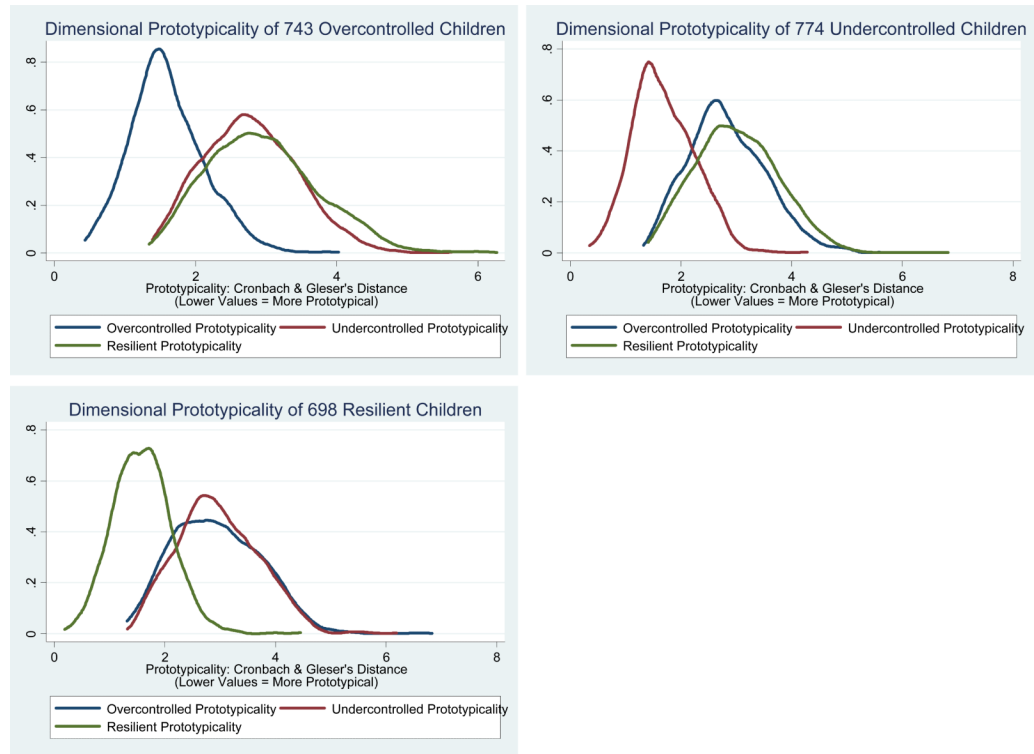
Figure 1.
Dimensional Assessment of Protoypicality by Type
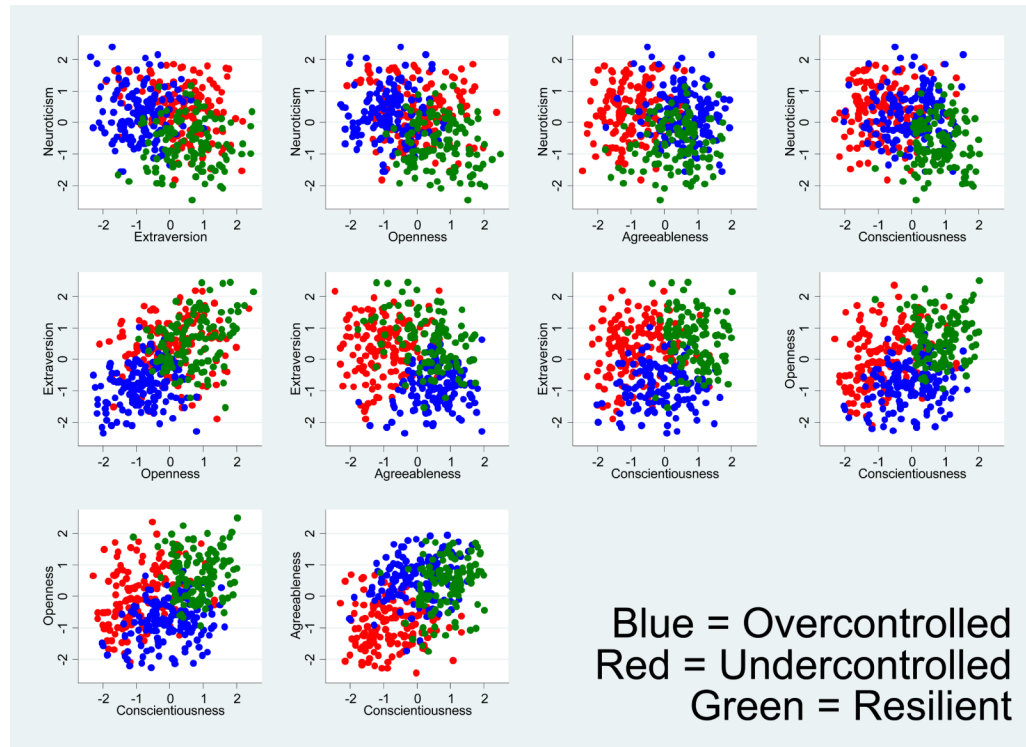
Figure 2.
Location of Types in Big 5 Space

Table 1

Characteristics of Empirical Kappa Distributions From Bootstrapped Split-Sample Cross-Validation Analyses

Number of Clusters	Mean	Median	Standard Deviation	Interquartile Range	Minimum / Maximum	Skewness	Kurtosis
2	.68	.76	.26	.50, .90	-.12, .99	-.93	2.90
3	**.83**	**.89**	**.16**	**.78, .94**	**.01, .98**	**-1.89**	**6.69**
4	.68	.73	.21	.54, .86	.18, .97	-.69	2.34
5	.61	.61	.16	.49, .73	.18, .96	-.03	2.26
6	.57	.55	.14	.47, .66	.17, .93	.17	2.65
7	.54	.54	.12	.46, .62	.24, .89	.14	2.81
8	.53	.53	.11	.46, .61	.18, .83	-.06	2.88
9	.53	.53	.09	.47, .60	.27, .81	.05	2.73

Note. **Child Sample Oahu-Kauai(n = 2,215).** Measures of central tendency, dispersion, and shape of empirical Kappa distributions resulting from 1,000 bootstrap split-sample cross-validations for solutions ranging from 2-9 clusters. All distributions differ significantly from by one-another by Komolgorov Smirnov non-parametric test of equality of distribution functions at p < Bonferoni corrected critical value of .00179, except distributions for 7 & 8 clusters and 8 & 9 clusters. The most replicable cluster solution is bolded.

Table 2

Big 5 Factor Scores by ARC Type

	Neuroticism M (SD)	Extraversion M (SD)	Open-Ness M (SD)	Agree-Ableness M (SD)	Concscientiousness M (SD)
Overcontrolled (n = 743)	.22 (.9)	-.87 (.69)	-.81 (.67)	.65 (.65)	0 (.83)
Undercontrolled (n = 774)	.5 (.86)	.29 (.82)	.06 (.88)	-.94 (.68)	-.77 (.71)
Resilient (n = 698)	-.77 (.75)	.61 (.79)	.78 (.71)	.35 (.79)	.85 (.7)

Note. Factor score means = 0; SD = 1

Table 3

Associations Between Childhood ARC Type and Midlife Health Outcomes

Categorical ARC Types

	SF-1		Hypertension		Hypercholesterolemia		Diabetes Melitus		Cardiac Arrest		Stroke		Cardiovascular Disease	
	b/se	P	b/se	p	b/se	p	b/se	P	b/se	p	b/se	p	b/se	P
Male	-.08 .06	.214	.21 .08	.007	.33 .08	<.001	-.06 .11	.564	.43 .2	.03	-.14 .2	.478	-.06 .12	.62
Resilent	Reference		Reference		Reference		Reference		Reference		Reference		Reference	
Overcontrolled	-.18 .08	.015	.25 .1	.009	-.06 .1	.511	0 .14	.974	.12 .22	.584	-.24 .31	.438	.19 .14	.196
Undercontrolled	-.18 .08	.015	.14 .1	.168	-.03 .1	.795	.09 .13	.498	.01 .22	.951	.47 .23	.042	.01 .15	.97

Dimensional ARC Prototypicality Scores

	SF 1		Hypertension		Hypercholesterolemia		Diabetes Melitus		Cardiac Arrest		Stroke		Cardiovascular Disease	
	b/se	p	b/se	p	b/se	p	b/se	P	b/se	p	b/se	p	b/se	P
Male	-.07 .06	.234	.22 .08	.007	.35 .08	<.001	-.08 .11	.499	.43 .2	.029	-.16 .2	.429	-.08 .12	.533
Resilient	.09 .03	.001	-.09 .03	.01	-.01 .03	.688	-.05 .05	.292	-.07 .08	.397	-.11 .09	.203	-.13 .05	.016
Undercontrolled	-.01 .03	.83	-.05 .03	.129	-.04 .03	.204	.02 .05	.718	-.07 .08	.36	.17 .09	.052	-.05 .05	.312
Overcontrolled	-.02 .03	.369	.02 .03	.626	-.02 .03	.605	-.05 .05	.291	.03 .08	.722	-.11 .09	.212	.02 .05	.655

Note. Results from ordinal (SF-1) and binary (others) probit models. Parameter estimates indicate the standard deviation change in the outcome associated with predictor. Dimensional prototypicality scores are scaled so that 1 unit indicates a 25% greater resemblance to ARC prototype. SF-1 = SF-1 General Health Item.

Table 4

Incremental Validities Evaluated With Bayesian Information Criteria (BIC)

Predictor Set	SF-1 BIC	Hypertension (27.2%) BIC	Hypertension AUC 95%CI	Hypercholesterolemia (26.8%) BIC	Hypercholesterolemia AUC 95% CI	Diabetes Melitus (7.2%) BIC	Diabetes AUC 95%CI	Cardiac Arrest (1.7%) BIC	Cardiac Arrest AUC 95%CI	Stroke (1.5%) BIC	Stroke AUC 95%CI	Cardiovascular Disease (5.6%) BIC	CVD AUC 95%CI
Types	3300	1393	.55 .51,.58	1388	.52 .48,.55	637	.52 .46,.58	232	.53 .41,.65	199	.68 .57,.79	522	.55 .48,.62
Prototypicality	3303	1398	.56 .52,.60	1394	.53 .49,.56	643	.55 .49,.61	238	.57 .44,.71	207	.68 .53,.82	524	.60 .53,.67
Traits	3326	1417	.57 .53,.60	1410	.55 .51,.59	649	.59 .53,.65	250	.62 .50,.74	220	.68 .54,.82	531	.64 .58,.71
Types + Traits	3327	1420	.57 .54,.61	1417	.55 .52,.59	662	.59 .53,.65	263	.63 .51,.75	230	.72 .60,.84	544	.65 .58,.71
Prototypicality + Traits	3333	1426	.57 .54,.61	1424	.55 .52,.59	668	.60 .54,.66	269	.66 .56,.76	240	.70 .56,.83	550	.66 .60,.72
Types + Prototypicality	3317	1410	.56 .53,.60	1407	.53 .49,.56	567	.55 .49,.61	251	.59 .47,.72	219	.71 .59,.83	536	.61 .54,.68
Types + Traits + Prototypicality	3353	1437	.58 .54,.62	1436	.56 .52,.59	682	.60 .54,.66	283	.66 .55,.77	252	.72 .60,.84	562	.67 .61,.73

Note. 40-year incidence rates of each cardiovascular condition in parentheses. BICs from probit models, adjusted for gender. *Lower* BICs indicate better model fit, penalized for number of parameters. Differences of 8 indicate "strong evidence" and differences of 10 or more indicate "very strong evidence" favoring the lower over the higher model (Kass & Rafferty, 1995. AUC = Area Under the Receiving Operator Curve of outcome classification produced by the model. This is the proportion of true (vs. false) positive across all possible cut-points for classifying observations based on the predicted probability of model. Higher values are better, and random guesses about outcome category produce an AUC of .5. 95% CI = 95% Confidence Interval for AUC; models without overlapping confidence intervals are significantly different, p < .05. SF-1 = SF-1 General Health Item