Predictive Validity of Personality Types Versus Personality Dimensions From Early Childhood to Adulthood: Implications for the Distinction Between Core and Surface Traits

Jens B. Asendorpf, *Humboldt University of Berlin*

Jaap J. A. Denissen, *Humboldt University and Max Planck Institute for Human Development, Berlin*

This study compared the long-term predictive validity of person-centered personality types and variable-centered personality dimensions assessed between ages 4–6 years in a population sample of 154 children. Results indicated that the predictive power of both approaches was remarkably robust between age 17 and 22, and even increased in the case of aggressiveness. At age 22 the long-term predictive ability of both types and dimensions was about equal, both in terms of the multivariate percentage of explained variance and the number of significant longitudinal correlates. This pattern is consistent with the notion that personality types and variables represent core personality domains that predict a wide range of longitudinal outcomes. However, the predictive ability of both approaches was larger for personality traits and intelligence than for social relationships outcomes. Implications for the distinction between core and surface traits are discussed.

For a long time, personality psychologists have stressed the importance of considering the within-person organization of personality traits. In the United States, Allport's now-famous definition states that "personality is the dynamic organization *within the individual* of those psychophysical traits that determine his unique adjustments to his environment" (1937, p. 48; emphasis added). A theoretical rationale for such an approach is that complex systems such as the human psyche are best described by the

Jens B. Asendorpf and Jaap J. A. Denissen, Institute for Psychologie.
Correspondence should be directed to the first author at Humboldt University of Berlin, Institute for Psychology, Rudower Chaussee 18, Berlin 10099, Germany. E-mail: jens.asendorpf @online.de.

Merrill-Palmer Quarterly, July 2006, Vol. 52, No. 3, pp. 486–513.

configuration or patterning of their respective subsystems (Bergman, 2001). In theory there can be an unlimited number of possible configurations, but in practice there are typically one or more configurations that are more frequently observed (Magnusson, 1998). Such typical configurations are considered *types*. Accordingly, research on personality types belongs to the person-centered tradition in personality research. Because types capture unique information about the way personality dimensions are organized at the individual level, they are also closest to Allport's definition. As the current study will argue, this within-person organization captured in personality types may represent a particularly stable, "core" feature of personality. To test this notion, we investigated how personality types are related to long-term personality and social relationships outcomes.

One of the most common typologies is the classification system proposed by J. H. Block and J. Block (1980), which is based on a theory of ego-control and ego-resiliency. In this dimensional model of personality in childhood, *ego-resiliency* refers to the tendency to respond flexibly rather than rigidly to changing situational demands, particularly stressful situations. *Ego-control* refers to the tendency to contain (overcontrol) rather than express (undercontrol) emotional and motivational impulses. Subsequent research (e.g., Asendorpf, Borkenau, Ostendorf, & van Aken, 2001; Asendorpf & van Aken, 1999; Robins, John, Caspi, Moffitt, & Stouthamer-Loeber, 1996) has shown that ego-resiliency and ego-control are related in an inverted U-shaped fashion. That is, high resiliency is related to intermediate scores in control, whereas both high and low control are related to low resiliency. Consequently, three rather than four types are expected: resilients (with intermediate scores on ego-control), undercontrollers (high on undercontrol), and overcontrollers (high on ego-control). These types (referred to as RUO types throughout this article) were confirmed through replicated Q-factor or cluster analysis for both children and adults (e.g., Asendorpf & van Aken, 1999; Asendorpf et al., 2001; Barbaranelli, 2002; Boehm, Asendorpf, & Avia, 2002; Costa, Herbst, McCrae, Samuels, & Ozer, 2002; Rammstedt, Riemann, Angleitner, & Borkenau, 2004; Robins et al., 1996; Schnabel, Asendorpf, & Ostendorf, 2002). However, because not all studies have succeeded in replicating the RUO types (Asendorpf, Caspi, & Hofstee, 2002), they may not generalize across all samples and observers.

Despite empirical and theoretical arguments for the person-centered approach, most research in personality psychology has been carried out in the variable-centered tradition (note that in the current study the term *dimensions* is often used instead of *variables* because personality types can also be summarized with categorical variables). In the variable-centered approach, associations among two or more dimensions in a specific *population* are the focus

of investigation. One of the most influential developments in this area has been the establishment of the Five-Factor Model of personality description (John & Srivastava, 1999), which maps most trait labels onto five more or less independent factors, the Big Five: neuroticism, extraversion, agreeableness, conscientiousness, and culture (openness to experience). Although most of the research leading to the establishment of this model rested on self-ratings of adults, this model has also been increasingly applied to parental and teacher judgments of children and adolescents (e.g., Digman, 1989; John, Caspi, Robins, Moffitt, & Stouthamer-Loeber, 1994; Kohnstamm, Halverson, Mervielde, & Havill, 1998).

Work in the mainstream variable-centered tradition has mostly ignored the issue of personality types, and little research has attempted to compare—let alone integrate—both perspectives in a broader theoretical framework (Hart, Atkins, & Fegley, 2003; Robins & Tracy, 2003). Some studies have compared the concurrent predictive validity of types versus dimensions, but almost all have relied on cross-sectional data. Such comparisons have generally favored the variable-centered approach (Costa et al., 2002; van Leeuwen, de Fruyt, and Mervielde, 2004), though one study reported mixed results (Huey & Weisz, 1997). According to Hart et al. (2003), however, the utility of both approaches may depend on the type of problem that is being studied. Specifically, they hypothesized that "when the interest is on the movement of individuals through time, the person-centered approach may be preferable" (p. 89). The current study tests this notion by comparing the long-term predictive validity of both approaches. In this article we argue that the results of this comparison have implications for the theoretical status of types versus dimensions in terms of the distinction between core and surface traits (Asendorpf & van Aken, 2003a). To understand the logic behind this claim, this distinction is briefly explained in the following section.

Theoretical Status of Types and Dimensions as Representing Core Versus Surface Personality Traits

A number of authors have hypothesized that there exist multiple levels of personality that differ in their level of generality (McAdams, 2001). Specifically, it is hypothesized that there exist highly stable "core" traits, as well as "surface" traits that are more easily influenced by random fluctuations (Asendorpf & van Aken, 2003a; McCrae et al., 2000). Hart et al. (2003, p. 90) argued for a similar distinction between "global personality traits and structures that influence thinking and behavior in many contexts" (i.e., core traits) and "well-defined patterns of behavior, thought and

emotion that are expressed in particular contexts" (i.e., surface traits). Of course, the distinction between core and surface traits is more a matter of degree: all personality factors are hypothesized to be influenced by a different mixture of both stable and fluctuating factors. In Figure 1 this is depicted by a vertical double-headed arrow that represents a continuum rather than discrete categories.

For some constructs, evidence exists regarding their relative location on the stability continuum. For example, according to McCrae et al. (2000) the Big Five represent "core" traits, whereas acquired habits, roles, and relationships are "surface" characteristics. Accordingly, the Big Five should be better able to predict changes in social relationships than vice versa. This prediction has been confirmed by empirical research (Asendorpf & van Aken, 2003a; Asendorpf & Wilpers, 1998). In Figure 1 the Big Five (B5) are therefore located closer to the "core" end of the continuum, whereas social relationships (Re) are closer to the "surface" end. Less is known about the relative position of other constructs. Because IQ has been shown to be moderately heritable (McGue, Bouchard, & Iacono, 1993) and highly stable (Deary, Whalley, Lemmon, Crawford, & Starr, 2000), this factor may be placed even higher than the Big Five, yet empirical support for this notion is lacking (depicted in Figure 1 by the overlap between IQ and the Big Five). Finally, temperament is regarded by many researchers as traits that are rooted in neurobiology, appear early in life, and are the basis of later personality (Shiner, 1998). Conceptually at least, temperament should be located closer to the "core" end of the continuum, yet once again, empirical research is sparse.

Regarding the relative place of types (T) versus dimensions (D) on the stability continuum, three possibilities exist (see Figure 1, right panel). First, it is possible that types are more basic than dimensions (option 1). In the Block and Block (1980) model, ego-resiliency and ego-control are fundamental personality characteristics that govern interactions with the outside world. In addition, types are derived from the relative *patterning* of dimensions, so they are not affected by individual-specific response tendencies biasing the mean level of personality scales (Asendorpf, 2003). To the extent that such response biases are environmental (e.g., mood), types may capture a more stable component of personality. The second possibility (option 2) is that types and dimensions are located at same level of the stability continuum. Because types summarize individuals' response patterns on multiple dimensional scales, they may not add any information over and above the dimensions on which they were based. This second option would imply that both types and dimensions are located at exactly the same level of the stability continuum. Finally, the third possibility

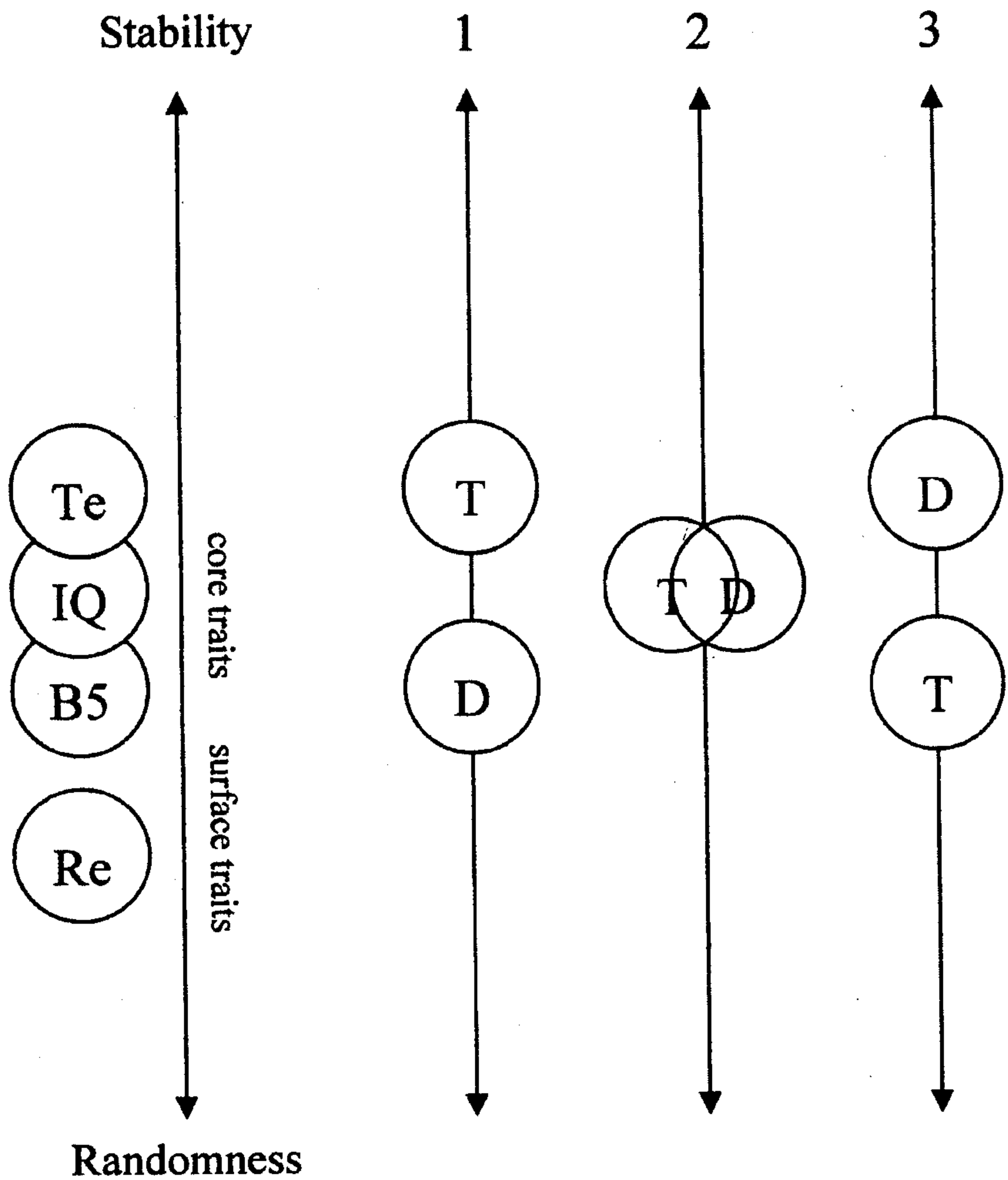

Figure 1. Distinction between core and surface personality traits and the relative position of types and dimensions on this dimension. Meaning of signs: T = types, D = dimensions, Te = temperament, IQ = intelligence, B5 = Big Five, Re = social relationships. The vertical axis between stability and randomness represents a continuum that includes highly stable "core" and more variable "surface" traits. Because social relationships have been identified by previous research as more representative of surface traits, they have been placed lower on the continuum, whereas the lack of research on the relative position of other constructs is represented by the overlap between circles. In addition, three possibilities with varying relative position of types versus dimensions on this dimension are depicted.

(option 3) is that the dimensions represent more basic and stable constructs than the types. Types are based on the relative similarity of individuals' personality dimensions with some ideal or empirical prototype. On the basis of this information, individuals are then assigned to different type categories on an all-or-nothing basis (alternatively, see Asendorpf, 2006). Especially individuals with ambiguous constellations may be misclassified because of relatively minor fluctuations of dimensional scores. If this is the case, types may be located lower on the stability continuum than dimensions.

Implications of the Theoretical Status of Traits Versus Dimensions for Their Long-Term Predictive Validity

According to a model of personality development proposed by Fraley and Roberts (2005), the stability of personality differences shows a negative exponential decrease across time until a stability asymptote well above zero is reached. The decrease is due to differential personality change as a result of person-environment (PE) transactions, whereas the positive asymptote is due to a constant influence of a "hidden" (latent) third factor on personality differences (e.g., genetic differences or effects of "frozen" early gene-environment transactions). Fraley and Roberts's (2005) model and the corresponding stability continuum have two implications for long-term personality predictions: across repeated PE transactions in time, (1) core personality traits are predicted to be more stable than surface traits (due to their stable "core"), and (2) the higher stability of personality differences makes it more likely that core personality traits have an impact on environmentally influenced surface traits (assuming cumulative effects of PE transactions) in terms of both the amount of explained variance and the number of effects. This is why, in Figure 2, the arrow between P_1 and P_2 is thicker than the arrow between E_1 and E_2, and why the transactional arrow between P_1 and E_2 is thicker than the arrow between E_1 and P_2.

To date, research that compares the longitudinal predictive validity of the types versus dimensions in a head-to-head fashion is mostly lacking. However, results from studies assessing the predictive validity of one or the other approach separately suggest that both are significantly related to long-term outcomes. To date, perhaps the most impressive evidence for the predictive validity of the RUO types comes from the Dunedin Longitudinal Study, which followed a large, representative sample of New Zealand children from age 3 into adulthood. Results showed that at ages 18, 21, and 26, the inhibited (overcontrolled) children showed internalizing tendencies and the undercontrolled children showed externalizing tendencies (Caspi, 2000;

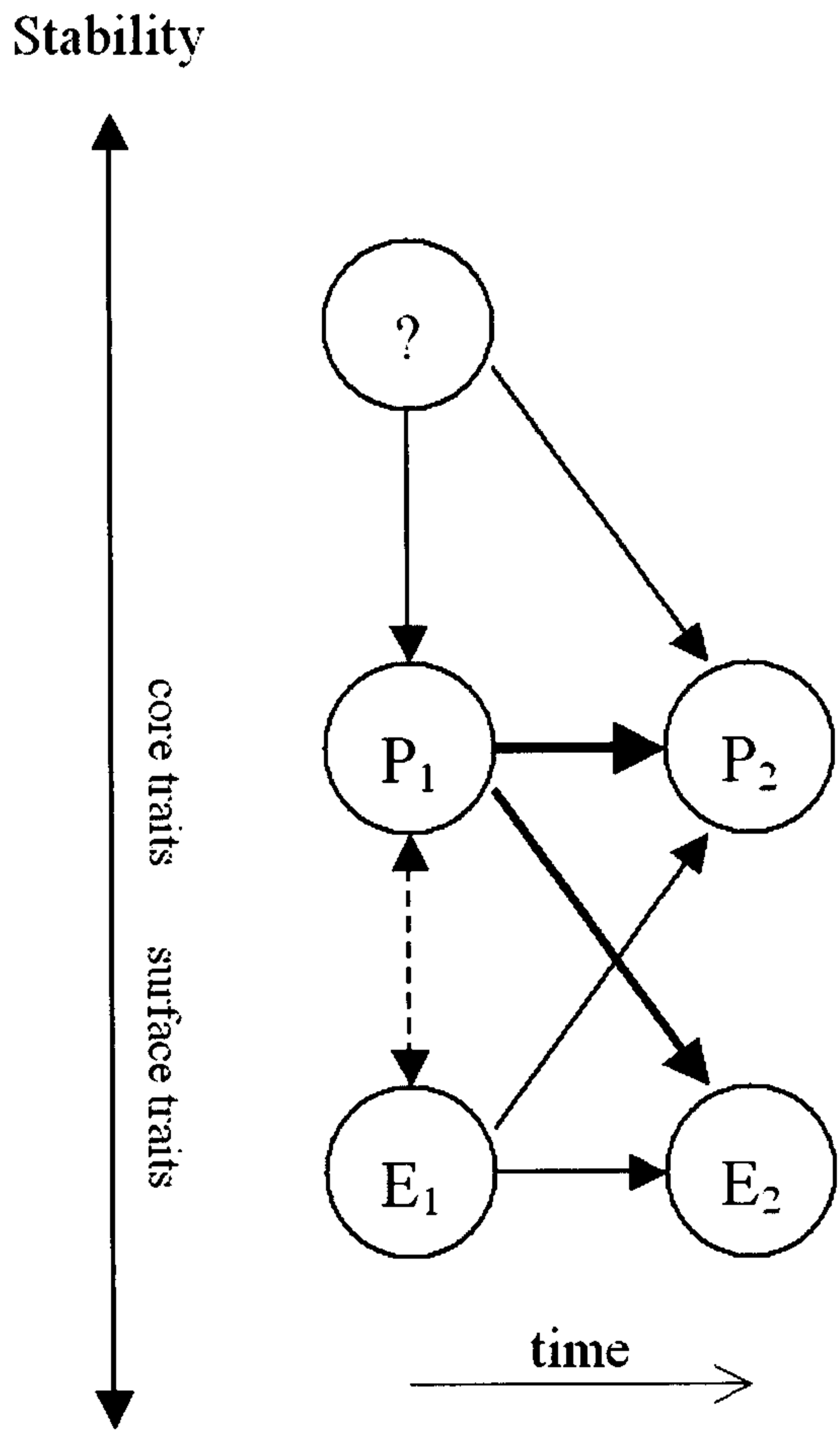

Figure 2. Model of personality stability by Fraley and Roberts (2005). Meaning of signs: ? = unknown stabilizing factor (e.g., genes, "frozen" PE transactions), P = personality, E = environment. According to this model, personality is influenced by a "hidden" stable factor that prevents long-term stability coefficients from reaching zero levels. In contrast, environmental factors are hypothesized to be more fluctuating. Because personality variables have a higher a priori stability, they are more likely to influence environmental factors than vice versa (the thickness of the arrows represents the strength of influence; the dotted double-headed arrow represents associations between P and E at t = 0). The distinction between core and surface traits made in the current study can be applied to the Fraley and Roberts (2005) model, which gives rise to the prediction that constructs that are directly influenced by the stability factor are representative of core traits, whereas more fluctuating, environmental constructs are more representative of surface traits.

Caspi, Harrington, Milne, Amell, Theodore, & Moffitt, 2003; Caspi & Silva, 1995; see Asendorpf & van Aken, 1999, and Pulkkinen, 1996, for similar evidence). The Big Five scales have also been shown to have substantial validity in the prediction of longitudinal outcomes (e.g., Abe, 2005; Asendorpf & van Aken, 2003b; Markey, Markey, & Tinsley, 2003). These findings suggest that both types and dimensions are located high on the stability continuum. However, they say comparatively little about the relative position of the two approaches with regard to each other.

In an earlier part of the current study, Asendorpf and van Aken (1991) found that the stability of the types between ages 4–6 was not smaller than between ages 6–10 in spite of a doubling of the retest interval, which supports the notion of types as highly stable core constructs. However, because the latest waves of the current study did not include a reassessment of the types, it was not possible to evaluate the first implication of the distinction between core and surface traits (i.e., higher stability for core traits than for surface traits). Therefore, the current analysis focuses on the second implication, that more stable, "core" traits have greater potential for longitudinal predictions than more "surface" traits.

To our knowledge, the only study that compared the longitudinal predictive ability of types versus dimensions is a study by Asendorpf (2003), who studied 99 German adolescents (part of the same sample as the one used in the current study) between ages 4–6 and 17. Results showed that the variable approach accounted for most of the concurrent, cross-sectional variance. However, both approaches fared almost equally well when they were applied to the prediction of longitudinal data. In fact, both types and dimensions added incremental variance to the prediction of the age 17 outcomes when they were entered as a second block into the regression (i.e., after accounting for the variance explained by the other set of predictors). From this evidence it could be concluded that both types and dimensions are located at a comparable level of the stability continuum (i.e., option 2).

To arrive at a comprehensive picture of the predictive ability of both the type approach and the variable approach, the current outcome measures included a broad range of socially important personality domains, such as social relationships, intellectual ability, and temperamental traits. In addition, as explained above, these outcome domains can be expected to encompass a relatively broad range of the stability continuum. Since the previous assessment at age 17, participants have gone through a number of very significant transformations, including puberty, change of schools, and (for many) leaving home. This makes it difficult to argue that any existing long-term consistencies between childhood personality criteria and adulthood

outcomes are environmentally mediated. Instead, if either types or dimensions were able to achieve such long-term predictions, this would be a strong indication for their status as "core" traits. To test this notion, the current study investigated the ability of the childhood types/dimensions to predict variance in outcomes assessed at age 22 and compared these findings to those reported by Asendorpf (2003) for age 17.

Summary and Hypotheses

Relatively little is known about the theoretical status of types versus dimensions. The current study investigates this issue from the perspective of the distinction between core and surface traits. Although the relative status of types versus dimensions on this continuum is unclear at present, the Big Five factors of personality description have been hypothesized to be highly representative of core traits (McCrae et al., 2000). In addition, Hart et al. (2003, p. 90) concluded from their study on long-term predictive outcomes of the types that "personality types . . . are nearer to the global pole of the dimension that to the specific, context-dependent end." Indeed, evidence reported by Asendorpf (2003) suggests that whereas variables outperform the types in cross-sectional predictions, both approaches perform equally well in longitudinal predictions. This can be interpreted as evidence for the position that types, just like dimensions, represent "core" dimensions of personality. Accordingly, it was hypothesized that both types and dimensions are able to predict a broad range of outcomes across a wide time period. From this overarching hypothesis, three hypotheses were derived: (1) types and dimensions predict an equal amount of variance in outcomes at age 22; (2) types and dimensions predict an equally broad range of outcomes at age 22; (3) the predictive ability of types and dimensions remains constant between ages 17 and 22.

Method

Participants

The target sample for the present study consisted of 154 participants who were assessed at ages 17 and 22. These participants were part of the Munich Longitudinal Study on the Genesis of Individual Competencies (LOGIC; Weinert & Schneider, 1999). The LOGIC sample originally consisted of 204 children whose first language was German and who started to attend 20 preschools in the Munich area in the fall of 1984 when they were 3–4 years old. After 1 year, another 26 participants of the same birth cohort were added to the sample. This initial sample of 230 children (119 males) was

rather unbiased because the schools were selected from a broad spectrum of neighborhoods, and more than 90% of the parents who were asked for permission gave their consent for their child's participation. After age 12, the LOGIC sample was reassessed at age 17 and 22.

Measures

Age 4–6. The 54-item short version of the California Child Q-Set (Block & Block, 1980) was adapted to German by bilingual parents (Göttert & Asendorpf, 1989). This instrument provides personality descriptors (sample items: "creative," "likes to compete," "fearful and anxious") that have to be sorted according to nine salience categories (1 = extremely uncharacteristic, 9 = extremely characteristic). Judges are instructed to sort exactly 6 items into each of the 9 categories (forced equal distribution).

Age 17. At age 17 the LOGIC sample included 174 participants (93 males), which amounts to 76% of the original sample. Participants' parents provided ratings of the child on age-appropriate bipolar adjective pairs developed by Ostendorf (1990) that measure the Big Five factors of personality (extraversion, neuroticism, agreeableness, conscientiousness, and culture) with 8 items using a 5-point response scale (median $\alpha = .87$, range .84–.93). Temperamental factors were assessed with a 5-item scale measuring sociability developed by Asendorpf and Wilpers (1998 (sample item: "likes to be around a large number of people"), and 2 scales of 4 items each measuring shyness (sample item: "feels shy in the presence of other people") and aggressiveness (reverse coded sample item: "seldom reacts aggressively") developed by Asendorpf and van Aken (1999). These scales had good reliabilities (median $\alpha = .83$, range .78–.84). In total, parental ratings were available for 149 children. Twenty-one of these cases consisted of mothers' judgments, 3 of fathers' judgments, and the remaining 125 consisted of the average rating by both parents. Intelligence was assessed with the vocabulary subtest of the German version of the Wechsler scales for adults (HAWIE-R; Tewes, 1991) and the German version of the Culture Fair Intelligence Test (CFT-20; Weiß, 1987). Finally, participants judged the quality of social relationships with mothers, same-sex friends, and opposite-sex friends (excluding classmates) using a German adaptation of the Network of Relationships Inventory (NRI; Furman & Buhrmester, 1985). This instrument assesses the amount of instrumental help, intimacy, esteem enhancement, and reliability in specific relationships. For the current analyses, these facets were averaged to create a composite scale of relationship quality ($\alpha > .70$).

Age 22. At age 22, 154 participants were still in the LOGIC sample (81 males; retention rate of 89% compared to age 17). The Big Five factors

were assessed with the German version of the NEO-Five Factor Inventory (NEO-FFI, framed in the third person; Borkenau & Ostendorf, 1993), which consists of 60 items (responses scale 1–5). Scale α's were mostly satisfactory (for mother ratings: range .70–.84, average .77; for father ratings: range .68–.85, average .76). An inspection of the correlations between mother and father ratings yielded a high level of inter-rater agreement (range .60–.75, average $r = .66$).

To assess temperamental aspects of personality, parents filled out the same scales as at age 17 to measure aggressiveness, shyness, and sociability. Alpha reliabilities were satisfactory (for mother ratings: range .70–.76, average .76; for father ratings, range .68–.82, average .77). and cross-rater agreement was high (range .56–.67, average $r = .62$). In total, parental ratings were available for 120 children. Twenty-two of these cases consisted of mothers' judgments, 3 of fathers' judgments, and the remaining 95 consisted of the average rating by both parents.

Psychometric intelligence was assessed with the nonverbal Series, Classifications, Matrices, and Topologies scales of the Culture Fair Intelligence Test (CFT-20; Weiß, 1987; $n = 145$, $\alpha = .81$) and the vocabulary subtest of the German revised Wechsler Intelligence Test for Adults (HAWIE-R; $n = 151$, $\alpha = .81$). A composite IQ measure was created by z-standardizing and averaging both measures. For the 12 participants with missing data for one of the two IQ measures, the standardized score on the other, completed test was used instead, resulting in a total of 154 usable responses.

Social relationships were assessed with an ego-centered social network questionnaire (see Asendorpf & Wilpers, 1998; Neyer, 1997), which differed from the NRI used at age 17. This instrument requires participants to list important people they had interacted with during the past month. A list of 14 relationship categories (e.g., parents, friends, classmates) was provided to guide the selection of contact persons. Participants listed an average of 10.0 ($SD = 4.6$) network partners (51% female) with an average age of 31.5 ($SD = 4.8$). Participants were free to name any person who was important to them, but some relationship categories were named especially frequently, with an average of 0.97 mothers, 0.80 fathers, 4.18 friends, and 0.56 partners (values less than one indicate that not everybody listed a person in the corresponding category). After listing the contact persons, they rated each individual relationship on five Likert scales: personal closeness, exchange of personal thoughts and feelings, felt acceptance, conflict, and emotional support (response scale 1–5). Exploratory factor analysis of these ratings using Varimax rotation resulted in a clear two-factor solution explaining 73% of the variance. Closeness, exchange, acceptance, and sup-

port all loaded very highly on the first factor (primary loadings between .72 and .85, secondary loadings between -.42 and .11), whereas conflict dominated the second factor (loading = .96). The items of the first factor were averaged to form a composite index of the quality of social relationships (α = .82), whereas the conflict rating was retained as a separate variable. Scores for both relationship quality and conflict covered the entire scale range (i.e., between 1 and 5).

Procedure

Ages 4–6. All LOGIC participants attended a preschool or kindergarten from ages 4 through 6. At the end of each school year, the child's main teacher provided a Q-sort description of the child. Each teacher was assisted by a trained examiner, who outlined the Q-sort procedure in detail and answered any questions about the procedure. To increase the reliability of the judgments, the three Q-sorts at ages 4–6 were averaged for every item individually. These averaged items were subjected to a Q-factor analysis with three factors that were interpreted as Q-sort prototypes of resilient, overcontrolled, and undercontrolled children (Asendorpf & van Aken, 1999). Children were then classified according to their factor loadings on these factors (which are identical to the Q-correlation between their average Q-sort and the factor) according to a procedure developed by Robins et al. (1996). Of the maximum number of 154 participants in the current sample (i.e., subjects who provided IQ data at age 22), 103 could be classified according to their childhood personality type, with 54% resilient, 18% overcontrolled, and 27% undercontrolled individuals. The findings of the current study are based on this subsample of classified individuals.

Following the procedure of John et al. (1994), the Big Five personality factors were assessed by Q-sort indices consisting of four to six Q-sort items for each factor. The reliabilities of the five indices were satisfactory (median α = .78, range .72–.89), and the indices showed substantial concurrent external validities (see Asendorpf & van Aken, 2003b, for details).

Ages 17–22. When participants were 17 and 22 years old they visited the Munich Max Planck Institute for an extensive assessment session that lasted about 5 hours. Participants completed the social relationships questionnaires and the IQ tests, in addition to measures unrelated to the current study. For their participation in each wave of the LOGIC study, they received 70 Euros. Around the time of participants' laboratory visit, their parents received a package of questionnaires, including the current Big Five and temperamental scales. Parents were asked to rate the personality of their children and return the questionnaire by mail.

Attrition effects on predictor and criterion measures. As stated above, not all individuals who were assessed at age 22 could be Q-typed at age 4–6. Attrition analyses were conducted to see how these participants differed from the remaining sample in terms of childhood IQ and SES and adult IQ and personality. Out of a total of 20 analyses, only one difference was significant: individuals who could not be Q-typed had a significantly higher adult IQ level, $F(1, 109) = 3.94$, $p = .05$. Additionally, it was tested whether individuals included in the current analyses differed from the original 230 LOGIC participants. First, a dropout dummy was created for subjects who had missing IQ and Q-type data. In a total of 9 attrition analyses, IQ-dropouts did not differ from the original participants in terms of gender, SES, and childhood IQ (ps > .30), but they were rated by their teachers are more neurotic, $F(1, 149) = 4.65$, $p = .03$. A second dropout variable was created to identify subjects who had to be excluded from the current analyses because of missing age 4–6 Q-type and age 22 parent-judged personality data. Out of a total number of 9 attrition analyses, the only statistical difference was for personality-dropouts to be judged by their teachers as significantly more neurotic, $F(1, 149) = 4.02$, $p = .05$.

Analysis Strategy

The goal of the present study was to compare the long-term predictive validity of the type and variable approaches. For this purpose, hierarchical multiple regressions were carried out to predict age 17 and 22 outcome variables (i.e., three temperamental variables, the composite IQ score, the Big Five factors, and the six social relationship indices). The predictive validity of the dimensional approach was tested with the five Q-sort dimensions as predictors. For the type approach, dummy variables were created to specify participants of the two non-resilient types. Specifically, one dummy variable coded overcontrolled individuals with a value of 1 and coded undercontrolled and resilient individuals with 0. The other dummy variables coded undercontrolled individuals with 1 and coded the others as 0. Two regression analyses including both types and variables as predictors were carried out for each criterion variable. Predictors were entered as blocks to avoid problems due to possible colinearity between the predictors (see Huey & Weisz, 1997, p. 409). In one regression, dummy-coded type indicators were entered as a first block of predictors, followed by a second block of dimensional predictors. In the other regression, the order of the two blocks of predictors was reversed. On a multivariate level, the R^2 of the first block of criteria is indicative of the overall predictive power, whereas the ΔR^2 of the second block can be treated as an index of each type of predictor's ability to explain unique variance.

Results

Correlations Between Age 22 Outcome Variables

As can be seen in Table 1, correlations between the Big Five factors were low to moderate (range -.40–.42). By comparison, the associations between the Big Five factor extraversion and temperamental sociability and (low) shyness were quite high (.68–.77), just like the correlation between agreeableness and aggression (-.51). IQ did not show many associations with other outcome variables, with the exception of small correlations with culture (.26) and sociability (-.21). Similar associations have been reported in the literature (for culture and IQ see Ackerman & Heggestad, 1997, and Ashton, Lee, Vernon, & Lang, 2000; for sociability and crystallized intelligence see Rolfhus & Ackerman, 1999). Finally, social relationships were not generally associated with personality variables. The only exceptions were relationship quality with mothers, which was positively related to conscientiousness, and conflict with mothers, which was negatively related to conscientiousness and positively to aggressiveness. The association between conscientiousness and parental relationships has been reported in the literature (Asendorpf & Wilpers, 1998), whereas the relation between aggressiveness and conflict can be explained by the fact that the mother was in almost all cases the informant of the aggressiveness ratings. Finally, there were moderate associations between different social relationships categories (e.g., between mothers and peers and between other-sex and same-sex peers).

Comparison Between Types and Dimensions in Predicting Long-Term Outcomes

The current study addressed three hypotheses regarding the comparative long-term predictive validity of personality types versus dimensions. According to Hypothesis 1, types and dimensions should predict an equal amount of variance in long-term outcomes at age 22.[1] Second, Hypothesis

1. According to Costa et al. (2002), the predictive power of personality types derives in part from their ability to capture interactions between single dimensions. Accordingly, it could be argued that the predictive validity of the type approach should be pitted against the combined additive and interactive effects of the dimensional approach. However, such an approach would be biased in favor of the dimensional approach because of the higher number of predictor variables that are needed to represent main and interaction effects (Asendorpf, 2003). To illustrate this point, we compared the incremental validity of a third regression block consisting of all possible two-way interactions of the five variables (i.e., 10 in number) to that of five sets of 10 randomly created variables. Results indicated that the average incremental validity of the "real" interaction block amounted to a ΔR^2 of .12, which may seem very promising at first. However, this value proved to be almost equal to the average ΔR^2 of .13 (range .10–.15), which was found when the predictive

Table 1. *Concurrent Correlations Between Age 22 Criterion Variables: Parent-Rated Personality, Psychometric IQ, and Self-Rated Social Relationships*

	Parent-rated temperament				Parent-rated Big Five					Self-rated relationships					
	shy	soc	agg	iq	neu	ext	cul	agr	con	mqu	mco	squ	sco	oqu	oco
-shy	1.00	–.45**	.14	–.05	.40**	–.68**	–.04	–.32**	–.28**	–.04	–.03	–.10	.02	–.13	–.08
-soc		1.00	–.21*	–.21*	–.17	.77**	.12	.42**	.26**	–.04	.06	.05	.03	.01	.09
-agg			1.00	–.06	.37**	–.24**	–.03	–.51**	–.27**	–.12	.23*	–.01	.12	.07	.06
iq				1.00	–.18*	–.08	.26**	–.11	–.04	.01	–.07	–.11	–.14	–.12	–.12
-neu					1.00	–.38**	.07	–.32**	–.40**	–.03	.05	.01	.10	.00	.08
-ext						1.00	.19*	.37**	.32**	.04	–.01	.12	–.01	.13	.09
-cul							1.00	.11	.13	.14	–.14	.01	–.07	.19	.02
-agr								1.00	.42**	.10	–.05	.06	–.17	–.11	–.10
-con									1.00	.26**	–.29**	.01	–.22*	–.13	–.04
-mqu										1.00	–.41**	.44**	–.28**	.40**	–.14
-mco											1.00	–.08	.15	–.15	.16
-squ												1.00	–.28**	.54**	–.28**
-sco													1.00	–.15	.55**
-oqu														1.00	–.17
-oco															1.00

Note. neu = neuroticism, ext = extraversion, cul = culture, agr = agreeableness, con = conscientiousness, shy = shyness, soc = sociability, agg = aggressiveness, mqu = quality of relationship with mother, mco = conflict in relationship with mother, squ = quality of same-sex peer relationships, sco = conflict in same-sex peer relationships, oqu = quality of other-sex peer relationships, oco = conflict in other-sex peer relationships. *n* for parent-rated personality = 120; for parent-rated personality × IQ = 120; for parent-rated personality × self-rated relationships = 94–118; for self-rated relationships = 112–147.

* $p < .05$; ** $p < .01$

2 states that types and dimensions contribute to an equal number of significant effects at age 22. Third, Hypothesis 3 predicts that the predictive ability of types and dimensions remains constant across time. Table 2 shows the *R*-square values of both approaches in predicting psychometric intelligence, parent-reported personality, and self-reported social relationship outcomes at ages 17 and 22. In the following, this information is summarized according to its relevance to the three hypotheses formulated in the introduction.

Temperamental traits. At age 22 both types and dimensions fared equally well in the prediction of temperamental traits. Specifically, types explained an average of 10% of the variance (range 4%–18%) when inserted as a first block and 6% (range 2%–9%) when inserted as a second block in the multiple regression analysis, whereas the dimensions explained an average of 9% (range 4%–14%) and 4% (range 2%–5%) of the variance as first and second blocks, respectively. Second, both approaches added a comparative number of significant effects when entered as a first block, with types contributing to the prediction of age 22 shyness and both approaches contributing to the prediction of aggressiveness. However, the type approach added significant variance over and above the dimension approach in the prediction of aggressiveness. Third, the predictive power of both approaches showed no decrease with age.

IQ. Regarding the prediction of IQ, the dimensions were somewhat superior to the types. At age 22 the dimensions explained a higher percentage of explained variance than the types when entered as a first block of the regression (13% vs. 7%). Second, at age 22 both types and dimensions contributed to the prediction of variance in intelligence as a first block of the regression, so the number of significant effects was the same. Third, between age 17 and 22 the predictive power of the dimensions was reduced from 18% to 13%, whereas the predictive power of the types remained constant at 7%. Accordingly, the (lower) predictive ability of the types was better conserved than the (higher) predictive ability of the dimensions in predicting adult intelligence.

Big Five factors. Regarding the prediction of adult Big Five personality traits, the verdict seemed slightly in favor of the dimensions, but the difference was not large. Entered as a first or second block in the regression, the

---

ability of the random variables was averaged across all five Monte Carlo simulations. Apparently, the use of a large number of predictors can lead to an accumulation of small chance effects that do not capture anything significant. From this the conclusion can be drawn that the comparison in the current study is already unfair to the type approach, since it includes only two (dummy) predictors, whereas the variable approach is allowed five. To tip the balance even more in favor of the variables by including more predictors would lead to biased conclusions that capitalize on chance.

Table 2. Percentage of Initially and Incrementally Explained Variance of Hierarchical Multiple Regression Analyses of Age 4–6 Types/Dimensions Predicting Age 22 Temperament, Big Five factors, IQ, and Social Relationships

	Age 22				Age 17 [a]			
	Model 1		Model 2		Model 1		Model 2	
Criterion	Step 1 R^2 2 dummy types	Step 2 ΔR^2 5 Q–Sort dimensions	Step 1 R^2 5 Q–Sort dimensions	Step 2 ΔR^2 2 dummy types	Step 1 R^2 2 dummy types	Step 2 ΔR^2 5 Q–Sort dimensions	Step 1 R^2 5 Q–Sort dimensions	Step 2 ΔR^2 2 dummy types
Temperament								
-shy	.08*	.02	.08	.02	.08*	.04	.11*	.01
-soc	.04	.05	.04	.06	.03	.03	.06	.00
-agg	.18**	.05	.14*	.09*	.12**	.06	.09	.09*
Average	.10	.04	.09	.06	.08	.04	.09	.03
iq	.07*	.06	.13*	.00	.07*	.13*	.18**	.01
Big Five								
-neu	.06	.08	.14*	.00	.17**	.05	.21**	.01
-ext	.05	.02	.05	.01	.08*	.05	.13*	.01
-cul	.05	.08	.12	.01	.08*	.08	.07	.08*
-agr	.10*	.04	.10	.04	.09**	.07	.16**	.01
-con	.11**	.04	.13	.02	.11**	.05	.16**	.00
Average	.07	.05	.11	.02	.11	.06	.15	.02

Social relationships								
-mqu	.01	.10	.07	.04	.03	.08	.10*	.00
-mco	.04	.04	.07	.01	.00	.10*	.09	.01
-squ	.00	.05	.02	.03	.01	.03	.03	.00
-sco	.04	.02	.03	.02	.00	.02	.01	.01
-oqu	.01	.13	.09	.04	.03	.04	.03	.03
-oco	.00	.02	.01	.01	.04	.06	.07	.03
Average	.01	.05	.05	.02	.02	.05	.06	.02

Note. neu = neuroticism, ext = extraversion, cul = culture, agr = agreeableness, con = conscientiousness, shy = shyness, soc = sociability, agg = aggressiveness, mqu = quality of relationship with mother, mco = conflict in relationship with mother, squ = quality of same-sex peer relationships, sco = conflict in same-sex peer relationships, oqu = quality of other-sex peer relationships, oco = Conflict in other-sex peer relationships. *N* for parent-rated personality = 120; for parent-rated personality × IQ = 120; for parent-rated personality × self–rated relationships = 94–118; for self-rated relationships = 112–147. The outcomes of the regression analyses regarding the age 17 outcomes were already reported in Asendorpf (2003, his Table 5), with the exception of age 17 shyness and social relationships.

* $p < .05$; ** $p < .01$

dimensions explained between 3% and 4% more variance than the types at age 22. Second, both approaches were comparable in terms of the number of significant effects, with the types contributing to two effects and the dimensions to one effect at age 22. Third, both approaches suffered from a slight reduction in their predictive ability between ages 17 and 22. Specifically, the reduction of the percentage of variance explained by the first regression block was 4% for both approaches.

Quality of social relationships. At age 22 neither the types nor the dimensions accounted for a significant portion of explained variance in social relationship outcomes. Accordingly, both approaches can be considered equivalent in their long-term (in)ability to predict the quality of social relationships. In every case, the amount of variance the dimensions could explain was very low, except in the case of quality and conflicts in relationships with mothers and quality in relationships with other-sex peers. Second, types and variables did not contribute to any significant effects at age 22. Third, the predictive ability of both approaches was not substantially reduced with age, with differences between ages 17 and 22 not exceeding 1% of explained variance.

Summary of Findings

First, it was hypothesized that both types and dimensions have about equal long-term predictive validity at age 22. This hypothesis was fully confirmed when the temperamental variables were used as outcomes, but there were some slight differences in the case of the other outcome variables. Because the difference in these cases only ranged between 4% (Big Five) and 6% (intelligence), support for Hypothesis 1 can be considered moderate. Second, Hypothesis 2 stated that types and dimensions predict an equally broad range of long-term outcomes. Results provided partial support for this prediction, with types predicting five outcomes at age 22 (shyness, aggressiveness, IQ, agreeableness, and conscientiousness) and dimensions predicted three outcomes (aggressiveness, IQ, and neuroticism). Third, it was expected that the predictive ability of types and dimensions would remain constant between ages 17 and 22. This hypothesis was fully confirmed for both social relationships and temperamental outcomes, partly confirmed for intelligence (when the types served as predictors), and partly disconfirmed for Big Five outcomes. Because the latter reductions were not large (e.g., from 11% to 7% in the case of the types and from 15% from 11% in the case of the variables), support for Hypothesis 3 can be considered moderate, especially for the types.

Discussion

Long-Term Predictive Validity of Personality Characteristics

Starting from the assumption that both the type approach and the dimensional approach are located at a comparatively high level of the stability continuum, the current study compared their ability to predict long-term outcomes. As dependent variables, intelligence, parent-rated personality traits, and self-rated social relationships were assessed at ages 17 and 22. Overall, neither the type approach nor the variable approach showed much reduction in terms of the overall percentage of explained variance between ages 17 and 22. For example, the type approach explained between 8% and 10% of the variance in parent-rated temperamental traits at both ages. This remarkable stability for both approaches also emerged for self-rated social relationships, whereas the reductions in predictive validity for psychometric intelligence and Big Five personality traits were generally small. This is quite surprising, given that the period between 17 and 22 is characterized by a number of significant life transitions, such as leaving the parental home, getting a job or going to university, and so forth. Of course, one could argue that both types and dimensions had a somewhat limited predictive validity to begin with. However, it should be noted that these validity coefficients corresponded to multiple correlations up to .42 (parent-reported aggressiveness predicted by the RUO types), which is quite substantial given the time gap (18 years) and the change in assessment sources (teachers vs. psychometric tests, parental ratings, and self- ratings). The fact that both the types and the dimensions consolidated such a performance after five additional, potentially turbulent years is remarkable indeed.

Multiple reviews (e.g., Olweus, 1979; Roberts & DelVecchio, 2000) have demonstrated that the stability of personality traits decreases as the time between two assessments increases. However, these reviews are based on studies that typically calculated single-validity coefficients using only two time points. In this context it is not surprising that a study of personality between ages 10 and 15 achieves a higher stability coefficient than a study that assesses personality at ages 10 and 35. However, the current study started with personality measurements at a very young age (4–6 years), with regular reassessments until age 22. From a theoretical perspective, it makes sense that the decrease in predictive validity fades out at some point. Indeed, Caspi et al. (2003) reported similar findings, which they explained with increasing genome-environment covariance across time. Indeed, the Fraley and Roberts (2005) model infers a source of stable influence on personality traits (e.g., genes) to explain the higher-than-zero asymptotic stability of

personality variables. Applied to the issue of long-term predictive ability, this means there should be clear limits to the decrease in predictive validity of childhood personality indicators, just as found in the present study.

Unexpectedly, the ability of both types and dimensions to predict variance in aggressiveness was higher at age 22 than at age 17. This finding needs to be explained because it violates the rule of decreasing predictive power with increasing prediction interval. One possibility is that the aggressiveness ratings at age 17 were somehow influenced by environmental "error" that was no longer present at age 22. For example, Moffitt's (1993) taxonomy of antisocial behavior distinguishes between life-course-persistent individuals and a larger group of adolescent-limited antisocial individuals, who display aggressive behavior in adolescence under the influence of normative social-environmental factors. If such factors also influenced the aggressiveness ratings in the current study, this could explain the limited ability of childhood personality to predict aggressiveness during adolescence (i.e., around age 17) compared to later ages (i.e., around 22).

In the introduction we postulated a preliminary categorization of the current outcome variables. Specifically, we proposed that social relationships are characteristic of surface personality traits, whereas variables in the three broad categories—temperament, the Big Five factors, and psychometric intelligence—are more characteristic of core traits. The current finding that only a very limited percentage of variance in social relationships can be predicted by childhood personality is consistent with their conceptualization as surface personality traits. Empirical evidence is very limited with regard to the relative status of the three hypothesized core domains (i.e., temperament, Big Five, intelligence), but results from the present study suggest that they are located at approximately the same level of the stability continuum. Averaging the contributions of types and dimensions for each outcome category at age 22 (across regression models 1 and 2) resulted in 9.5%, 10.0%, and 9.0% of initially explained variance in temperamental factors, psychometric intelligence, and Big Five factors, respectively. In addition, across both approaches the average amount of incrementally explained variance in these variables was 5.0%, 3.0%, and 3.5%.

On a more conceptual level, the results suggest that the categories used to classify personality variables (such as the distinction between temperament, Big Five, intelligence, and social relationships used in the current study) may not be optimal. Specifically, the only adult traits that could be significantly predicted by the childhood predictors were shyness, aggressiveness, intelligence, neuroticism, agreeableness, and conscientiousness. Although the current categorization regarded sociability as a temperamen-

tal trait, the percentage of explained variance for these variables was much smaller than for age 22 aggressiveness. In addition, the percentage of explained variance in extraversion was not much higher than the percentage of explained variance in social relationships. It seems that additional research is needed to develop more sophisticated categories that group variables on the stability continuum.

Theoretical Status of Types and Dimensions as Core or Surface Characteristics

One of the goals of the present study was to investigate the relative status of types and dimensions as core or surface characteristics. Specifically, three options were considered: (1) types are more characteristic of core traits than dimensions are, (2) dimensions are more characteristic of core traits than types are, and (3) both approaches represent core traits. The present results point most strongly toward the third option. By comparing the long-term predictive ability of these approaches, the current study has made a contribution to the literature that cross-sectional studies are not able to make. In such cross-sectional designs, dimensions have typically outperformed the types, yet longitudinal designs such as those used in the present study and by Asendorpf (2003) yield a much more balanced picture, with both approaches contributing to long-term predictions and the type approach even outperforming the variable approach in terms of the number of significant effects. Specifically, entered as a first block, the types added to the prediction of age 22 shyness, aggressiveness, IQ, agreeableness, and conscientiousness; entered as a second block, they added to the prediction of aggressiveness over and above the contribution of the dimensions (i.e., 6 effects). Entered as a first block, the dimensions contributed to the prediction of aggressiveness, IQ, and neuroticism, but no effects were significant when entered as a second block (i.e., 3 effects).

The fact that the type approach was relatively on par with the variable approach is quite remarkable. After all, there were five dimensions and only three personality types, which should work in favor of the dimensions. Moreover, because of their dimensional character, dimensions allow a finer-grained level of between-person differentiation, whereas types only allow for crude yes-or-no categorizations. Against this background, the empirical draw between types and dimensions is impressive. This result is consistent with Asendorpf's (2003) finding that the type approach did better when predicting longitudinal compared to concurrent outcomes. As stated above, one of the reasons for this could lie in the fact that types capture the within-person organization of traits that is (when Q-sorts are used) independent of

mean-level tendencies that might fluctuate across time (e.g., because of mood). Accordingly, they may be highly representative of core (as opposed to surface) traits.

A final note needs to be made regarding the fact that at age 22, both types and dimensions explained significant portions of the variance when entered as a first block in the regressions but failed to add incremental variance when entered as a second block (with the exception of the incremental validity of the types in predicting age 22 aggressiveness). This implies that the "working ingredients" behind most effects of either types or dimensions are not uniquely contained in either approach. Instead, childhood types and dimensions seem to share considerable variance regarding their ability to predict adult outcomes. Again, this pattern is most consistent with the notion that both approaches are located at approximately the same level of the stability continuum (i.e., option 2).

Implications

In the light of the current findings, it seems too early to issue a definite verdict on the relative superiority of either the type approach or the variable approach. From a practical perspective, then, the answer to the question of what approach future studies should use when predicting long-term outcomes is likely to be differentiated, with the best type of predictors depending on the type of outcome measure and the time lag between the assessment of predictors and outcomes (Hart et al., 2003). For example, in the current study the types did not do a very good job at predicting social relationships outcomes, which were better explained by the variable approach. In contrast, the types were better suited to predict more temperamental outcomes like aggressiveness. Moreover, due to the fact that they tap into more reliable intra-individual personality organization, types might do a better job when the time lag between predictor and outcome is especially high.

An additional reason for using the type approach exists: compared to dimensions, types are much easier to communicate (Asendorpf, 2003; Hart et al., 2003). In fact, much of lay psychology seems to revolve around types, with popular magazines offering tests to "find out your love/communication/humor type," and so forth. The popular appeal of the type approach might partly be explained by its ability to summarize personality information under one label. For example, in terms of their Big Five profile, undercontrolled individuals are characterized by low agreeableness and conscientiousness (see Asendorpf et al., 2001). Summarizing this information by means of the comprehensive label "undercontrol" seems a good compromise between information overload and oversimplification, which

may be especially useful when cognitive resources are scarce. This feature is especially appealing because the associations between dimensional variables are often complex and non-linear (Magnusson, 1998, p. 60) and may thus be difficult to understand. Accordingly, the type approach may be preferred over including complex statistical interaction terms between dimensional dimensions, especially when psychologists try to communicate results to policy makers or a popular audience (Hart et al., 2003).

Limitations

Of course, the current study suffers from some restrictions that limit the generalizability of its findings. Although the number of attrition effects was small compared to the number of assessed dimensions, some systematic attrition effects occurred. First of all, children who could not be assigned to one of the RUO types had higher IQ levels, whereas participants who provided IQ or personality data at age 22 were rated by their teachers at age 4–6 as less neurotic. It should be noted, however, that this sample restriction should have made it more difficult to find significant results for intelligence and neuroticism. In this light, it is quite impressive that the childhood personality was still able to predict a significant portion of the variance in both variables.

A second limitation is the somewhat limited size of the current sample. Because some individuals typically cannot be categorized to one of the existing types, this problem is even further exacerbated when applying the type approach. Indeed, the distribution of the RUO types has been frequently found to be uneven, and the number of non-resilient individuals may be limited in small samples. Thus, future research should use larger samples in trying to replicate the current findings. Note, however, that the small sample size limited the current study's ability to find significant results. Under these circumstances, the fact that the type approach nevertheless produced six significant effects in predicting the age 22 parent-rated trait outcomes can be considered impressive. Also, the current study spanned a longitudinal period of almost 20 years, using a fairly representative sample that was assessed during multiple time points. Faced with limited resources, such an extensive investment in data quality often comes at the cost of sample size. However, as Bergman (2001) correctly points out, precisely these kinds of data are needed to study personality as a dynamic and complex system. This way, future research may be able to develop more sophisticated theories about the relations between core and surface personality traits and how this distinction can further a better understanding of personality stability and person x environment transactions.

References

Abe, J. A. A. (2005). The predictive validity of the Five-Factor Model of personality with preschool age children: A nine year follow-up study. *Journal of Research in Personality, 39,* 423–442.

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin, 121,* 219–245.

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

Asendorpf, J. B. (2003). Head-to-head comparison of the predictive validity of personality types and dimensions. *European Journal of Personality, 17,* 327–346.

Asendorpf, J. B. (2006). Typeness of personality profiles: A continuous person-centered approach to personality data. *European Journal of Personality, 20,* 83–106.

Asendorpf, J. B., Borkenau, P., Ostendorf, F., & van Aken, M. A. G. (2001). Carving personality description at its joints: Confirmation of three replicable personality prototypes for both children and adults. *European Journal of Personality, 15,* 169–198.

Asendorpf, J. B., Caspi, A., & Hofstee, W. B. K. (Eds.). (2002). The puzzle of personality types. *European Journal of Personality, 16,* S1–S5.

Asendorpf, J. B., & van Aken, M. A. G. (1991). Correlates of the temporal consistency of personality patterns in childhood. *Journal of Personality, 59,* 689–703.

Asendorpf, J. B., & van Aken, M. A. G. (1999). Resilient, overcontrolled, and undercontrolled personality prototypes in childhood: Replicability, predictive power, and the trait-type issue. *Journal of Personality and Social Psychology, 77,* 815–832.

Asendorpf, J. B., & van Aken, M. A. G. (2003a). Personality-relationship transaction in adolescence: Core versus surface personality characteristics. *Journal of Personality, 71,* 629–666.

Asendorpf, J. B., & van Aken, M. A. G. (2003b). Validity of Big Five personality judgments in childhood: A 9-year longitudinal study. *European Journal of Personality, 17,* 1–17.

Asendorpf, J. B., & Wilpers, S. (1998). Personality effects on social relationships. *Journal of Personality and Social Psychology, 74,* 1531–1544.

Ashton, M. C., Lee, K., Vernon, P. A., & Lang, K. L. (2000). Fluid intelligence, crystallized intelligence, and the Openness/Intellect factor. *Journal of Research in Personality, 34,* 198–207.

Barbaranelli, C. (2002). Evaluating cluster analysis solutions: an application to the Italian NEO personality inventory. *European Journal of Personality, 16,* S43–S55.

Bergman, L. R. (2001). A person approach in research on adolescence: Some methodological challenges. *Journal of Adolescent Research, 16,* 28–53.

Block, J. H., & Block, J. (1980). The role of ego-control and ego-resiliency in the organization of behavior. In W. A. Collins (Ed.), *Minnesota Symposium on Child Psychology* (Vol. 13, pp. 39–101). Hillsdale: Erlbaum.

Boehm, B., Asendorpf, J. B., & Avia, M. D. (2002). Replicable types and subtypes of personality: Spanish NEO-PI samples. *European Journal of Personality, 16,* S25–S41.

Borkenau, P., & Ostendorf, F. (1993). *NEO-Fünf-Faktoren-Inventar (NEO-FFI) nach Costa und McCrae* [NEO-Five-Factor Inventory according to Costa and McCrae]. Göttingen: Hogrefe.

Caspi, A. (2000). The child is father of the man: Personality continuities from childhood to adulthood. *Journal of Personality and Social Psychology, 78,* 158–172.

Caspi, A., Harrington, H., Milne, B., Amell, J. W., Theodore, R. F., & Moffitt, T. E. (2003). Children's behavioral styles at age 3 are linked to their adult personality traits at age 26. *Journal of Personality, 71,* 495–513.

Caspi, A., & Silva, P. A. (1995). Temperamental qualities at age three predict personality traits in young adulthood: Longitudinal evidence from a birth cohort. *Child Development, 66,* 486–498.

Costa, P. T., Jr., Herbst, J. H., McCrae, R. R., Samuels, J., & Ozer, D. J. (2002). The replicability and utility of three personality types. *European Journal of Personality, 16,* S73–S87.

Deary, I. J., Whalley, L. J., Lemmon, H., Crawford, J., & Starr, J. M. (2000). The stability of individual differences in mental ability from childhood to old age: Follow-up of the 1932 Scottish Mental Survey. *Intelligence, 28,* 49–55.

Digman, J. M., 1989. Five robust trait dimensions: Development, stability, and utility. *Journal of Personality, 57,* 195–214.

Fraley, C. R., & Roberts, B. W. (2005). Patterns of continuity: A dynamic model for conceptualizing the stability of individual differences in psychological constructs across the life course. *Psychological Review, 112,* 60–74.

Furman, W., & Buhrmester, D. (1985). Children's perceptions of the personal relationships in their social networks. *Developmental Psychology, 21,* 1016–1022.

Göttert, R., & Asendorpf, J. B. (1989). Eine deutsche Version des California-Child-Q-Sort, Kurzform [A German short version of the California Child Q-Sort]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 21,* 70–82.

Hart, D., Atkins, R., & Fegley, S. (2003). Personality and development in childhood: A person-centered approach. *Monographs of the Society for Research in Child Development, 68.*

Huey, S. J., Jr., & Weisz, J. R. (1997). Ego control, ego resiliency, and the five-factor model as predictors of behavioral and emotional problems in clinic-referred children and adolescents. *Journal of Abnormal Psychology, 106,* 404–415.

John, O. P., Caspi, A., Robins, R. W., Moffitt, T. E., & Stouthamer-Loeber, M. (1994). The "Little Five": Exploring the nomological network of the five-factor model of personality in adolescent boys. *Child Development, 65,* 160–178.

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

Kohnstamm, G. A., Halverson, C. F., Jr., Mervielde, I., & Havill, V. (Eds.). (1998). *Parental descriptions of child psychology: Developmental antecedents of the Big Five?* Mahwah: Erlbaum.

Magnusson, D. (1998). The logic and implications of a person-oriented approach. In R. B. Cairns, L. R. Bergman, & J. Kagan (Eds.), *Methods and models for studying the individual* (pp. 33–64). Thousand Oaks, CA: Sage.

Markey, C. N., Markey, P. M., & Tinsley, B. J. (2003). Personality, puberty, and preadolescent girls' risky behaviors: Examining the predictive value of the five-factor model of personality. *Journal of Research in Personality, 37,* 405–419.

McAdams, D. P. (2001). The psychology of life stories. *Review of General Psychology, 5,* 100–122.

McCrae, R. R., Costa, P. T. J., Ostendorf, F., Angleitner, A., Hrebickova, M., Avia, M. D., et al. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology, 78,* 173–186.

McGue, M., Bouchard, T. J., Jr., & Iacono, W. G. (1993). Behavior genetics of cognitive ability: A life-span perspective. In R. Plomin & G. E. McClearn (Eds.), *Nature, nurture, and psychology* (pp. 59–76). Washington, DC: American Psychological Association.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100,* 674–701.

Neyer, F. J. (1997). Free recall or recognition in collecting egocentered networks: The role of survey techniques. *Journal of Social and Personal Relationships, 14,* 305–316.

Olweus, D. (1979). Stability of aggressive reaction patterns in males: A review. *Psychological Bulletin, 86,* 852–875.

Ostendorf, F. (1990). *Sprache und Persönlichkeitsstruktur: Zur Validität des Fünf-Faktoren-Modells der Persönlichkeit* [Language and personality structure: Validity of the five-factor model of personality]. Regensburg: Roderer.

Pulkkinen, L. (1996). Female and male personality styles: A typological and developmental analysis. *Journal of Personality and Social Psychology,* 70, 1288–1306.

Rammstedt, B., Riemann, R., Angleitner, A., & Borkenau, P. (2004). Resilients,

overcontrollers, and undercontrollers: The replicability of the three personality prototypes across informants. *European Journal of Personality, 18,* 1–14.

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126,* 3–25.

Robins, R. W., John, O. P., Caspi, A., Moffitt, T. E., & Stouthamer-Loeber, M. (1996). Resilient, overcontrolled, and undercontrolled boys: Three replicable personality types. *Journal of Personality and Social Psychology, 70,* 157–171.

Robins, R. W., & Tracy, J. L. (2003). Setting an agenda for a person-centered approach to personality development: Commentary. *Monographs of the Society for Research in Child Development, 68,* 110–122.

Rolfhus, E. L., & Ackerman, P. L. (1999). Assessing individual differences in knowledge: Knowledge, intelligence, and related traits. *Journal of Educational Psychology, 91,* 511–526.

Schnabel, K., Asendorpf, J. B., & Ostendorf, F. (2002). Replicable types and subtypes of personality: German NEO-PI-R versus NEO-FFI. *European Journal of Personality, 16,* S7–S24.

Shiner, R. L. (1998). How shall we speak of children's personalities in middle childhood? A preliminary taxonomy. *Psychological Bulletin, 124,* 308–332.

Tewes, U. (1991). *Hamburg-Wechsler Intelligenztest für Erwachsene, Revision 1991* [Hamburg-Wechsler intelligence test for adults, revision 1991]. Bern: Huber.

van Leeuwen, K., de Fruyt, F., & Mervielde, I. (2004). A longitudinal study of the utility of the resilient, overcontrolled, and undercontrolled personality types as predictors of children's and adolescents' problem behavior. *International Journal of Behavioral Development, 28,* 210–220.

Weinert, F. E., & Schneider, W. (Eds.). (1999). *Individual development from 3 to 12: Findings from a longitudinal study.* Cambridge: Cambridge University Press.

Weiß, R. (1987). Grundintelligenztest Skala 2 (CFT 20) [Culture Fair Intelligence Test] (3rd ed.). Göttingen: Hogrefe.