European Journal of Personality
Eur. J. Pers. **17**: 327–346 (2003)
Published online 24 September 2003 in Wiley InterScience (www.interscience.wiley.com). **DOI**: 10.1002/per.492

Head-to-head Comparison of the Predictive Validity of Personality Types and Dimensions

JENS B. ASENDORPF*
Humboldt-Universität zu Berlin, Germany

Abstract

The utility of a configural type approach for predictions from personality is currently controversial. Configural types predict important personality correlates, but continuous dimensions based on the same data often fare much better in cross-sectional head-to-head comparisons. However, many such comparisons can be considered unfair to the type approach, confound diverse differences between type and dimensional approaches, and rely only on cross-sectional data. A sequence of analyses is reported that include fairer comparisons and deconfound differences due to the number of predictors, categorical versus dimensional predictors, dichotomized dimensions versus configural types, dimensional versus type criterion variables, and cross-sectional versus longitudinal predictions. The results suggest incremental validity of configural types over dimensions only in a very few cases. Copyright © 2003 John Wiley & Sons, Ltd.

INTRODUCTION

Personality can be defined as 'the dynamic organization *within the individual* of those psychophysical systems that determine his unique adjustments to his environment' (Allport, 1937, p. 48, italics added). Thus, personality psychology is concerned with the description, prediction, and explanation of this within-person organization, or personality structure. An individual's personality structure can be studied within a trait perspective on personality (Allport, 1937; Funder, 1991): psychologically meaningful characteristics on which individuals reliably differ (traits) are isolated, and an individual's personality is described by a configuration of traits. This article explores the extent to which the idiographic information in the individual trait configurations can be utilized in nomothetic predictions of personality correlates by grouping the individual configurations into personality types and predicting the external correlates by differences between these types.

*Correspondence to: Jens B. Asendorpf, Institut für Psychologie, Humboldt-Universität zu Berlin, Oranienburger Strasse 18, 10178 Berlin, Germany. E-mail: jens.asendorpf@rz.hu-berlin.de

Received 24 October 2002
Accepted 14 March 2003

Two approaches to deriving personality types

It is obvious from the outset that the predictive power of such a type approach depends on how much valid information is retained in the between-type differences, and how much is lost because it is contained in the within-type differences. Therefore, methods are needed that maximize the between-type variance relative to the within-type variance. Two different approaches have been frequently used: Q-factor analysis of individual Q-sort profiles, and cluster analysis of individual profiles in multiple personality variables.

Block (1961, 1971) developed the Q-sort approach to personality types. Each individual is described by a knowledgeable informant who sorts trait descriptions according to how well they fit the individual's personality (Q-sort). The individual Q-sort profiles of many respondents are then factor analysed by Q-factor analysis, that is, the profiles are reduced to a few uncorrelated prototypic profiles (the Q-factors). Finally, the individuals are classified into personality types according to the best-fitting Q-factor. As correlations between personality profiles ignore between-person differences in overall level and scatter, this method should only be applied to Q-sort measures such as the California Adult Q-Sort (Block, 1971) and the California Child Q-Sort (Block & Block, 1980) that force the judges to produce Q-sorts with a prescribed mean and variance.

Most studies have used the California child or adult Q-sets. Their items were originally selected by Block and colleagues to cover a wide range of age-appropriate traits. John, Caspi, Robins, Moffitt, and Stouthamer-Loeber (1994) showed that the Big Five factors of personality description are well represented by the items of the child Q-set; the same applies to the adult Q-set (McCrae, Costa, & Busch, 1986). The number of Q-factors can be determined by studying the replicability of the Q-factors when the sample is split into two random halves. This criterion was first applied by York and John (1992), and has been used in all later Q-type studies. Using this criterion, three, but not more, Q-factors were replicable in studies of children and were labeled resilient, overcontrolled, and undercontrolled, as originally suggested by Block and Block (1980) (Asendorpf & van Aken, 1999; Hart, Hofmann, Edelstein, & Keller, 1997; Robins, John, Caspi, Moffitt, & Stouthamer-Loeber, 1996).

The second approach to personality types, the cluster analysis approach, derives personality types through the cluster analysis of individual profiles in multiple trait dimensions. For example, the Big Five dimensions of the five-factor model of personality description (John & Srivastava, 1999) can be clustered. Each individual is described by a profile of five scores. These profiles are grouped by cluster analysis into relatively homogeneous clusters. Each cluster represents a personality type, and the average profile of the cluster members describes a personality prototype. Profiles that differ with regard to overall level and scatter can be treated as different if an appropriate measure of profile similarity is used (e.g. the Euclidean distance of the profiles).

Caspi and Silva (1995) clustered children's scores in three factors obtained from behavioural ratings by examiners who observed the children in various testing situations. The resulting five clusters have been repeatedly shown to predict important long-term outcomes over a period of more than 20 years (Caspi, 2000). Van Lieshout, Haselager, Riksen-Walraven, and van Aken (1995) clustered Big Five scales that were derived from the California Child Q-set. In a study of female adults, Pulkkinen (1996) clustered composites that comprised a total of 62 personality variables.

In an informal review of Q-factor and clustering studies, Caspi (1998) concluded that the resilient, overcontrolled, and undercontrolled types are not only regularly detected in

Q-sort studies but also in clustering studies, although the number of clusters varied across the clustering studies. Asendorpf, Borkenau, Ostendorf, and van Aken (2001) tested this hypothesis empirically by applying a within-study replication criterion similar to the criterion used in Q-sort studies. In two studies of Big Five self-ratings in adulthood and one study of Big Five parental ratings of their children, Asendorpf et al. (2001) found that only three clusters were replicable. These types showed substantial consistency, not only across the three clustering studies but also with the Q-sort types derived by Asendorpf and van Aken (1999), and could once again be interpreted as resilient, overcontrolled, and undercontrolled.

In an attempt to study the generality of these three personality types, Asendorpf, Caspi, and Hofstee (2002) asked colleagues from Germany, Spain, Italy, Belgium, and the United States to apply the approach of Asendorpf et al. (2001) to various Big Five data sets that referred to nonclinical samples in adulthood and adolescence. The results were mixed. The three types were found in the majority of the seven studies, but only six out of 21 cross-study agreement coefficients surpassed the conventional criterion of $kappa = 0.60$ for sufficient agreement, and the cluster centres in two of the seven samples were clearly not consistent with the resilient, overcontrolled, and undercontrolled prototypes. Thus, these prototypes represent frequently found personality types but they are not necessarily the prototypes that best describe personality differences.

Utility of the type approach for predictions from personality

Whereas personality description through a configuration of traits is a routine procedure for describing individuals, particularly in applied settings, the utility of configural types for the prediction of important personality concomitants and long-term outcomes is currently controversial. The routine procedure here is multiple regression, that weighs scores on multiple trait dimensions for all individuals in the same way such that the weighed sum maximally predicts a criterion variable (e.g. emotional or social adjustment, academic or job achievement). This exclusively variable-centred approach ignores the individuals' personality structure.

Therefore, it has been repeatedly suggested to make better use of the 'idiographic' information on personality structure in 'nomothetic' predictions of personality correlates (Asendorpf & van Aken, 1999; Asendorpf et al., 2001, 2002; Block, 1971; Caspi, 1998; York & John, 1992). One approach is to use information on differences between personality types for prediction. If only two personality types are distinguished, one dummy-coded variable (score 1 for membership in one type, score 0 otherwise) captures all information on differences between the two types, and can be used for prediction through simple regression. If $k > 2$ types are distinguished, $k - 1$ dummy-coded variables capture all information on differences between the types, and can be used for prediction through multiple regression (see e.g. Costa, Herbst, McCrae, Samuels, & Ozer, 2002).

The main advantage of this approach is that information on individuals' personality structure is preserved at least in part in the definition of the types, and can be used for prediction. The main disadvantage of this approach is that data on interindividual differences is lost in the transition from individual personality structure to personality types. The higher the interindividual variation within the types, the more data are lost for prediction. Thus, there is a trade-off between preserving some information on personality structure and losing information on within-type variation. Therefore, it is an empirical question to what extent and under which conditions a type approach can compete with, or even outperform, a dimensional approach.

This trade-off is not always acknowledged in discussions of the relative merits of a type approach to personality. Sometimes the type approach is dismissed with the statistical truism that reducing one continuous trait dimension to discrete types means losing information for prediction. However, this argument for the univariate case does not necessarily apply to the multivariate case of two or more dimensions. In the multivariate case, intraindividual between-trait differences are a new source of information that can be utilized by type approaches but not by dimensional approaches.

The predictive power of personality types can be empirically compared with the predictive power of personality dimensions by deriving types and dimensions from the same multivariate personality data, and predicting personality correlates by two hierarchical multiple regressions. In one regression, dummy-coded type indicators are 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 is reversed. This approach allows one to evaluate not only the overall predictive power of multiple types and multiple dimensions but also the incremental predictive validity of the dimensions over the types, and vice versa.[1]

Costa et al. (2002) applied this approach to two data sets, contrasting the predictive power of five Big Five scales (NEO-PI-R; Costa & McCrae, 1992) with two dummy-coded indicators for three personality types that were empirically derived from the multivariate data through a clustering procedure proposed by Asendorpf et al. (2001). For numerous continuous cross-sectional personality correlates, they found substantial predictive power of the types and the dimensions, and significant incremental validity of the dimensions over the types but not vice versa. Thus, dimensions fared clearly better than types in these analyses. When I applied their approach to various available data sets that included both Big Five dimensions and external correlates, I found virtually the same result. These findings question the utility of a three type approach to predictions from personality, relative to the Big Five.

Factors that may influence the predictive power of types versus dimensions

However, this cannot be the last word on predictions by types versus dimensions in general because the outcome of such a head-to-head comparison depends on numerous factors that are either confounded in Costa et al.'s (2002) particular study, or are not sufficiently varied. Below I briefly discuss some important factors that may influence the outcome of head-to-head comparisons between types and dimensions.

Number and intercorrelations of the predictors

A simple reason that may explain most of the superiority of the dimensions in Costa et al.'s (2002) approach is that there were 5 dimensional predictors but only 2 predictors that reflected differences between the types. Everything else being equal, 5 predictors yield stronger predictions than 2 predictors unless the 5 predictors show much higher intercorrelations than the 2 predictors. Costa et al. (2002) used NEO-PI-R factor scores that maximized the predictive strength of the dimensions because they are uncorrelated by

[1]One reviewer suggested including interaction terms between the predictors in the regression equation as a dimensional equivalent for including configural information. For dichotomized scores, this approach comes close to a type approach. Compared with empirically derived types, such *a priori* classifications share with multivariate extreme group approaches the problem that (i) the borders between the resulting classes of persons are not empirically determined and hence somewhat arbitrary and nonoptimal in many cases, and (ii) the resulting number of possible configurations used for prediction exponentially increases with the number of (dichotomized) dimensions, which leads to serious problems of inflated predictions in the case of the Big Five.

definition; in contrast, dummy-coded type indicators are correlated because membership for the non-type overlaps across types.

Cross-sectional versus long-term prediction
If it is true that types describe basic enduring intraindividual differences, they may have a particular advantage in long-term predictions. The individual configuration of traits may remain similar even if the individual trait scores fluctuate a lot. Consider the example of changes due to situation-dependent affective biases in personality judgments (e.g. traits are judged more positively after success than after failure; Funder, 1999). These biases do not affect intraindividual differences between traits that share a similar level of positivity. Interestingly, long-term longitudinal studies provide today perhaps the most successful examples of applications of a type approach (see e.g. Asendorpf & van Aken, 1999; Block, 1971; Caspi, 2000). Therefore, contrasting types and dimensions only in cross-sectional predictions may underestimate the predictive power of types.

What is predicted: dimensions versus types
Because any configural type approach capitalizes on intraindividual differences between personality traits, it seems *a priori* more likely that a type approach better predicts criteria that are also based on such intraindividual differences such as the same types judged by different informants or at a later age, or another type classification. Therefore, contrasting types and dimensions only in the prediction of dimensions may underestimate the predictive power of types.

Variable- versus person-centred personality assessment
In variable-centred personality assessments, each trait is judged independently from other traits with regard to an explicitly described or implicitly used reference group. Therefore, this approach is expected to be particularly sensitive to interindividual differences. In contrast, advocates of a person-centred approach to personality, such as Block (1971, 1993), claim that person-centred assessment methods that explicitly ask the judges to contrast different traits within one individual, such as in a Q-sort, are more sensitive to intraindividual differences and may therefore result in more reliable and valid configural types. If this is true, types based on Q-sorts may show a higher predictive power than types that are based on variable-centred assessments such as Big Five judgments.

Fairness of comparisons between dimensions and types

Thus, studies such as that of Costa et al. (2002) that play three types obtained by variable-centred assessments against five uncorrelated traits in the cross-sectional prediction of continuous outcome dimensions can be considered unfair for studies of the general predictive utility of type versus dimensional approaches because every factor that potentially moderates the relative advantage of dimensions over types was chosen in favour of dimensions. Studies that vary these moderating factors within or between samples are needed to draw a more balanced picture of the relative (dis)advantages of a type approach to prediction. The present study provides important pieces of such a picture and can be thus regarded as a first step towards a fairer comparison.

Losses and gains of the type approach

Predictions from personality types are handicapped by the fact that the predictors are dichotomous variables whereas personality dimensions are mostly continuous (e.g. the Big

Five). Advocates of a type approach claim that the information about individual trait configurations provided by the types may balance or even outweigh this handicap if this configural information can be utilized for the prediction. A rough and simple way of estimating these losses and gains of a type approach is to contrast three different predictions with one another: predictions by types, predictions by continuous dimensions, and predictions by the same dimensions that are dichotomized through median split.

Predictions from dichotomized dimensions constitute something like a baseline expectation for predictions from types because they share the 'dichotomization handicap' but convey no information on individual trait configurations. The predictive advantage of types over dichotomized dimensions is a rough estimate of how much valid configural information is extracted by the types. Similarly, the advantage of continuous over dichotomized dimensions is a rough estimate of how much valid information is lost if the continuous multivariate personality space is reduced to a few discrete personality types.

Although the exact losses and gains are expected to vary with the investigated constructs, it can be assumed that both continuous dimensions and types predict personality correlates better than the same dichotomized dimensions.

Hypotheses

From the preceding discussion of the factors contributing to relative (dis)advantages of types over dimensions in predictions from personality, I derived the following testable hypotheses.

Hypothesis 1. Everything else being equal, (i) an increasing number of types increases the relative advantage of a type approach to prediction, (ii) an increasing number of dimensions increases the relative advantage of a dimensional approach to prediction, and (iii) continuous dimensions predict personality correlates better than the same dimensions dichotomized through median split.

This hypothesis describes statistical truisms. Therefore, not a significant confirmation of the three subhypotheses is interesting, but instead their effect sizes. These effect sizes may be small particularly in cases (i) and (ii) if types or dimensions of low reliability are included.

Hypothesis 2. Everything else being equal, types predict personality correlates better than dimensions that are dichotomized through median split.

Hypothesis 3. Everything else being equal, types predict long-term outcomes of personality better than concurrent correlates.

Hypothesis 4. Everything else being equal, types predict personality correlates that are based on trait configurations better than personality correlates that are not based on such information.

These hypotheses were tested in the following two studies.

BERLIN RELATIONSHIP STUDY

Data from the first assessment of the Berlin Relationship Study, a short-term longitudinal study on personality and social relationships, were reanalysed for the purpose of the

present study. Because most methods have been described in detail by Asendorpf and Wilpers (1998) and Asendorpf et al. (2001), they are only briefly summarized here.

Method

Participants

When students of Humboldt University, Berlin, enrolled a few weeks before their first term started, they were asked to participate in a longitudinal study on students' social relationships. Only students below 23 years of age were included. During the second week of their first term, 173 females (age 18–22 years, $M = 20.0$) and 64 males (age 18–22 years, $M = 20.4$) participated in the first session. They constituted the sample of the study by Asendorpf and Wilpers (1998). Because of the smaller male sample, we repeated the study one year later with a second sample of 75 males (age 19–24 years, $M = 20.8$). Because the results for the two male samples were virtually identical with regard to all major variables, we pooled the two male samples, resulting in a more sex-balanced sample ($N = 312$; 173 females, 139 males).

Measures of personality, self-esteem, and loneliness

The Big Five factors of personality were assessed by the German version of the NEO-FFI by Costa and McCrae (1992) (Borkenau & Ostendorf, 1993). In addition, the subfactors *shyness* and *sociability* were assessed by five-item scales (e.g. 'I feel uneasy at parties and in large groups'; 'I find people more stimulating than everything else'). The items of these seven scales were mixed and responded to on five-point scales. Reliabilities were highly similar to those reported by Asendorpf and Wilpers (1998).

Global self-esteem as well as *self-esteem toward same- and opposite-sex peers* were assessed by German short versions of the SDQIII by Marsh and O'Neill (1984). Sample items are 'Overall, I have a lot of respect for myself', 'I am popular with members of the same sex', and 'I make friends easily with members of the opposite sex'. The six items with the highest corrected item–scale correlations in the original questionnaire were selected and translated into German. The resulting six-item scales showed satisfactory reliabilities ($\alpha > 0.79$). *Loneliness* was assessed by a short version of the UCLA loneliness scale (sample item 'I feel lonely'). A German version of the UCLA scale (Döring & Bortz, 1993) was reduced to the five highest-loading items on the first factor, representing feelings of loneliness, and the 5 highest-loading items on the second factor, representing feelings of social isolation, in a factor analysis of all 20 items in a representative sample of the general German population ($N = 592$). The resulting ten-item scale was highly reliable ($\alpha = 0.91$). The items of these four scales were mixed and responded to on five-point scales.

Measures of social relationships

The participants listed all persons who were currently personally important to them, indicated their sex and age and the duration of the relationship with them, and rated the quality of the relationship during the last 3 months on eight Likert scales (see Asendorpf & Wilpers, 1998, for more details). In the analysis by Asendorpf et al. (2001) of the data, five scales produced significant between-type differences: *contact frequency* (six-point scale 0–5, 'less than once a month'–'daily'), *closeness* of the relationship (five-point scale 1–5, 'very distant'–'very close'), *available support* (five-point scale 1–5, 'If I have problems, I would turn to this person to talk about my problems', 'never'–'always'), *frequency of conflict* (five-point scale 1–5, 'never'–'nearly at every encounter'), and *in love* (five-point scale 1–5, 'not at all'–'very much so').

Results

Selectivity of the sample

The representativeness of the sample ($N = 312$) was tested by comparing the means and standard deviations of the NEO-FFI scales with the norms that are based on 2112 subjects of the German population (mean age 29 years; Borkenau & Ostendorf, 1993). The initial sample had significantly higher scores in openness to new experience, $t(2422) = 6.28$, $p < 0.001$, and in agreeableness, $t(2422) = 5.72$, $p < 0.001$, than the normative sample, but the effect sizes of these differences were modest (d below 0.26 in both cases). The variances of the scales were unrestricted (mean SD 0.60 versus 0.58 in the normative sample). Thus, there was no evidence for attenuated correlations due to biased sampling.

Derivation of personality types

Ward's hierarchical clustering procedure was applied to the raw NEO-FFI scale scores, followed by the nonhierarchical K-means procedure (see Asendorpf et al., 2001, for the rationale of this approach and a more detailed description). Because earlier studies showed a particularly high replicability within and across studies for a three-cluster solution (see Asendorpf et al., 2001, 2002), the three-cluster solution was considered in the present study. The consistency between the resulting cluster solution and an independently derived three-cluster solution for a larger sample of 730 young German adults (reported by Asendorpf et al., 2001) was evaluated by a method outlined in detail by Asendorpf et al. (2001). Basically, each participant was assigned to the best-fitting cluster centre derived from the Berlin Relationship Study and from the other, larger study, and the agreement between these two classifications was computed by using Cohen's *kappa*. The resulting *kappa* was 0.85, which indicates a high similarity of the prototypic Big Five patterns for the three clusters between the two studies. As in the earlier studies by Asendorpf and colleagues, the clusters were therefore labelled resilient, overcontrolled, and undercontrolled.

For a fairer comparison with dimensional predictors, a six-cluster solution was additionally computed because differences between six types can be represented by five dummy-coded type indicators, and the predictions from the types were compared with predictions from five dimensions.

Prediction of personality dimensions from types versus dimensions

Predicted through hierarchical multiple regression were those personality dimensions that were significantly related to the three types in the study by Asendorpf et al. (2001): shyness, sociability, loneliness, and self-esteem (separately for global, with same-sex peers, and with opposite-sex peers). To test hypothesis 1(i), these personality dimensions were predicted by both the three- and the six-type solution. The three-type solution provided a test as parallel as possible to the Costa et al. (2002) analysis (two dummy-coded variables are needed to represent all type differences) whilst the six-type solution was fairer to the type approach with regard to the number of predictors (five dummy-coded variables are needed to represent all type differences).

To test hypothesis 2, both the continuous and dichotomized (median split) Big Five scores were used. This approach results in 2 (two versus five dummy-coded type indicators) $\times$ 2 (continuous versus dichotomized scores) $\times$ 2 (order in which types versus dimensions were entered) = 8 analyses (see Tables 1 and 2).

Numerous observations can be made in Table 1. First, the two dummy-coded type indicators significantly predicted all dependent variables, explaining 14% variance on average. Their significant predictive power is not at all surprising because they were

Table 1. Head-to-head comparison of two dummy-coded types and five (dichotomized) NEO-FFI scales in the concurrent prediction of personality

	Models 1&2	Model 1	Model 2	Model 3		Model 4	
	Step 1 R^2 2 dummy variables	Step 2 ΔR^2 5 scales		Step 1 R^2 5 scales Continuous	Step 2 ΔR^2 2 dummy variables	Step 1 R^2 5 scales Dichotomized	Step 2 ΔR^2 2 dummy variables
		Continuous	Dichotomized				
Shyness	0.11***	0.29***	0.10***	0.39***	0.01	0.21***	0.00
Sociability	0.12***	0.43***	0.31***	0.55***	0.00	0.42*	0.01
Loneliness	0.18***	0.13***	0.07***	0.30***	0.00	0.22*	0.03**
Self-esteem							
Global	0.30***	0.20***	0.07***	0.50***	0.00	0.32*	0.05***
With same-sex peers	0.07***	0.16***	0.10***	0.23***	0.00	0.16*	0.01
With opposite-sex peers	0.04***	0.14***	0.08***	0.17***	0.01	0.12***	0.00
Average	0.14	0.23	0.12	0.36	0.00	0.24	0.02

$N = 312$. The third dummy-coded type is redundant with the others. The NEO-FFI scales were dichotomized using median split.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

selected for this property in the first place. Second, the five dichotomized Big Five indicators explained consistently more variance than the two type indicators (24 versus 14%), which can be mainly attributed to the larger number of predictors.

Third, hypothesis 1(iii) was confirmed. As Table 1 indicates, the dichotomized Big Five scores showed a lower incremental predictive validity than the continuous scores, no matter whether the scores were entered before or after the type indicators (contrast

Table 2. Head-to-head comparison of five dummy-coded types and five (dichotomized) NEO-FFI scales in the concurrent prediction of personality

	Models 1&2	Model 1	Model 2	Model 3		Model 4	
	Step 1 R^2 5 dummy variables	Step 2 ΔR^2 5 scales		Step 1 R^2 5 scales Continuous	Step 2 ΔR^2 5 dummy variables	Step 1 R^2 5 scales Dichotomized	Step 2 ΔR^2 5 dummy variables
		Continuous	Dichotomized				
Shyness	0.22***	0.19***	0.05***	0.39***	0.01	0.21***	0.06***
Sociability	0.28***	0.27***	0.18***	0.55***	0.01	0.42*	0.04***
Loneliness	0.19***	0.13***	0.07***	0.30***	0.02	0.22*	0.04**
Self-esteem							
Global	0.32***	0.20***	0.08***	0.50***	0.01	0.32*	0.07***
With same-sex peers	0.15***	0.09***	0.05**	0.23***	0.01	0.16*	0.03*
With opposite-sex peers	0.11***	0.08***	0.04**	0.17***	0.01	0.12***	0.03
Average	0.21	0.16	0.08	0.36	0.01	0.24	0.05

$N = 312$. The sixth dummy-coded type is redundant with the others. The NEO-FFI scales were dichotomized using median split.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

columns 3 with 4, and 5 with 7); average incremental validity of the dichotomized scores was 12%. However, contrary to hypothesis 2, the types explained less variance than the dichotomized predictors (14 versus 24%), which can be attributed to the larger number of dichotomized predictors.

Fourth, when the type indicators were entered after the continuous scores, the incremental validity of the types was virtually zero for each dependent variable, replicating the main findings of Costa et al. (2002). Fifth, when the two type dummy-coded indicators were entered after the five dichotomized scores, they did explain significant incremental variance for loneliness and global self-esteem, but the additionally explained variance on average was small (2%).

In the fairer analysis with five dummy-coded type indicators, the pattern changed considerably in favour of the type approach (contrast Table 2 with Table 1), confirming hypothesis 1(i). The type approach explained now 50% more of the variance (21%), more than the incremental validity of the continuous dimensions, but the six types explained considerably less variance than the continuous variables (21 versus 36%), and showed again zero incremental validity. Contrary to hypothesis 2, the types fared somewhat worse than the dichotomized dimensions (21 versus 24% explained variance on average), although they now showed clearer incremental validity (5% additionally explained variance). Thus, in these fairer comparisons, the types fared less well than the continuous dimensions and even less well than the dichotomized dimensions.

Prediction of social relationships from types versus dimensions

One objection against this analysis concerns the fairly high correlations between the Big Five scales and the continuous dependent variables, which rely in part on intrinsic correlations between constructs. Sociability and shyness can be considered subfactors of extraversion, and did indeed correlate 0.71 and -0.57 with extraversion, respectively. Similarly, global self-esteem was closely linked to neuroticism, $r = -0.67$. A replication of the main predictive patterns for less intrinsically related variables was desirable.

Therefore, exactly the same procedure was applied to the prediction of those social relationship variables that were significantly related to the three types in the study by Asendorpf et al. (2001). As Tables 3 and 4 show, the same main pattern was found. The pattern was however less consistent across the different variables, which can be attributed to a floor effect, and the explained variances were smaller and therefore showed more chance variation across methods. Again, the three or six types showed no incremental validity over the continuous scores, but in one case significant incremental validity over the dichotomized scores. However, this one significant finding can be attributed to chance due to the large number of statistical tests.

Discussion

The results of the prediction of continuous criteria in this cross-sectional study supported hypotheses 1(i) and 1(iii), but not hypothesis 2. The findings suggest that empirically derived configural types derived through cluster analysis from Big Five dimensions predict important continuous personality and social relationship correlates, but show weaker predictions than continuous and even dichotomized Big Five scales in fairer comparisons where the number of predictors is the same. Furthermore, there was no evidence that the types display incremental validity over continuous dimensions, and little evidence even for dichotomized dimensions.

Table 3. Head-to-head comparison of two dummy-coded types and five (dichotomized) NEO-FFI scales in the concurrent prediction of social relationships

	Models 1&2	Model 1	Model 2	Model 3		Model 4	
	Step 1 R^2 2 dummy variables	Step 2 ΔR^2 5 scales		Step 1 R^2 5 scales Continuous	Step 2 ΔR^2 2 dummy variables	Step 1 R^2 5 scales Dichotomized	Step 2 ΔR^2 2 dummy variables
		Continuous	Dichotomized				
No. social relationships	0.03**	0.06**	0.04**	0.09***	0.00	0.06**	0.01
No. peer relationships	0.04**	0.08***	0.07***	0.12***	0.00	0.11***	0.00
Age of relationship partners	0.03**	0.10***	0.09***	0.12***	0.01	0.12***	0.01
Duration of relationships	0.04**	0.03	0.04*	0.07***	0.00	0.07***	0.00
Frequency of conflict	0.05***	0.06***	0.03	0.11***	0.00	0.06***	0.02*
Closeness to mother	0.03**	0.03	0.01	0.06***	0.00	0.03	0.02
Available support from father	0.03**	0.04*	0.04*	0.07***	0.00	0.07**	0.00
Contact frequency with peers	0.05***	0.05**	0.03*	0.10***	0.00	0.07***	0.01
In love with peers	0.02*	0.00	0.02	0.01	0.01	0.02	0.02
Average	0.04	0.05	0.04	0.08	0.00	0.07	0.01

$N = 312$ (308 for relationship with mother, 293 for relationship with father). The third dummy-coded type is redundant with the others. The NEO-FFI scales were dichotomized using median split.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Table 4. Head-to-head comparison of five dummy-coded types and five (dichotomized) NEO-FFI scales in the concurrent prediction of social relationships

	Models 1&2	Model 1	Model 2	Model 3		Model 4	
	Step 1 R^2 5 dummy variables	Step 2 ΔR^2 5 scales		Step 1 R^2 5 scales Continuous	Step 2 ΔR^2 5 dummy variables	Step 1 R^2 5 scales Dichotomized	Step 2 ΔR^2 5 dummy variables
		Continuous	Dichotomized				
No. social relationships	0.07***	0.03	0.02	0.09***	0.02	0.06**	0.03
No. peer relationships	0.09***	0.03*	0.03	0.12***	0.01	0.11***	0.02
Age of relationship partners	0.05**	0.09***	0.08***	0.12***	0.02	0.12***	0.01
Duration of relationships	0.03	0.05*	0.05**	0.07***	0.01	0.07***	0.00
Frequency of conflict	0.09***	0.04*	0.02	0.11***	0.01	0.06***	0.05**
Closeness to mother	0.05**	0.02	0.00	0.06***	0.01	0.03	0.02
Available support from father	0.06**	0.03	0.03	0.07***	0.01	0.07**	0.02
Contact frequency with peers	0.06**	0.05**	0.04*	0.10***	0.02	0.07***	0.03
In love with peers	0.01	0.02	0.02	0.01	0.01	0.02	0.01
Average	0.06	0.04	0.03	0.08	0.01	0.07	0.02

$N = 312$ (308 for relationship with mother, 293 for relationship with father). The sixth dummy-coded type is redundant with the others. The NEO-FFI scales were dichotomized using median split.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

To test the remaining hypotheses, data from a longitudinal study that allowed long-term predictions were analysed.

THE LOGIC STUDY

Data from the LOGIC study (Weinert & Schneider, 1999) were analysed for the purpose of the present study. Included are teacher Q-sort judgments of the participants' personality at ages 4–6, and personality and cognitive/achievement outcomes at age 17 for the 99 participants who had non-missing data at all assessments. Because the methods used up to age 12 are described in detail by Asendorpf and van Aken (1999, 2003), and the same methods were used for age 17, they are only briefly summarized here.

Method

Participants

The participants were part of the Munich Longitudinal Study on the Genesis of Individual Competencies (LOGIC). The LOGIC sample originally consisted of 230 children (119 boys, 111 girls) who were studied every year from their first or second year in preschool until age 12. The sample was fairly unbiased because the schools were selected from a broad spectrum of neighbourhoods, more than 90% of the parents who were asked for permission gave their consent for their child's participation, and attrition until age 12 was low (19% over 8 years) and unsystematic (see Weinert & Schneider, 1999, for this initial part of the study). After age 12, the LOGIC sample was reassessed once more at age 17. Attrition was again low (6% over 5 years), resulting in 174 participants at age 17. A comparison with the 56 drop-outs did not reveal significant differences in terms of socioeconomic status of the family, or with ego-control or ego-resiliency that were assessed in the first year through a teacher Q-sort (see below). The present study includes those 99 participants whose personality was judged by their teachers in preschool and kindergarten at ages 4, 5, and 6 years, and by their mother or father at age 17.

Assessments and Measures

Teacher Q-sort. The 54-item short version of the California Child Q-Set (CCQ; Block & Block, 1980) was adapted into German (Göttert & Asendorpf, 1989). All LOGIC participants attended a preschool, or kindergarten, from age 4 through to 6. At the end of each school year, the child's main teacher provided a Q-sort description of the child according to a fixed, nine-point distribution, ranging from 'extremely uncharacteristic' to 'extremely characteristic'. The teacher was instructed to sort exactly six items into each of the nine categories (forced equal distribution). To increase the reliability of the judgments, the three Q-sorts at ages 4–6 were averaged itemwise.

Parental personality judgments. At age 17, the child's mother or father judged the child on 40 age-appropriate bipolar adjective pairs that assessed each factor of the five-factor model of personality with eight items (*extraversion*, *emotional stability*, *conscientiousness*, *agreeableness*, *culture*). The items were derived from the highest-loading items on the first five factors of a pool of 179 bipolar adjective pairs by Ostendorf (1990) in a multi-step procedure (see Asendorpf & van Aken, 1999, for details). The items were answered on a five-point response scale (with labels *very*, *somewhat*, *neither/nor*, *somewhat*, and *very*). The reliabilities of the scales were satisfactory (median $\alpha = 0.87$, range 0.84–0.93).

The parents also answered a questionnaire that included four-item scales assessing *aggressiveness* and *inhibition*. The items of the two scales were randomly mixed with other items and answered on a seven-point response scale. The reliabilities of the scales were satisfactory (median $\alpha = 0.83$, range 0.78–0.84).

Of the 99 children, 81 were judged by both parents, 16 only by the mother, and two only by the father. Because the between-parent correlations were substantial, median 0.61, range 0.49–0.72, the two parental judgments were averaged.

Intelligence tests. Verbal intelligence was assessed with the vocabulary subtest of the German version of the Wechsler scales for adults (HAWIE-R; Tewes, 1991). Nonverbal intelligence was assessed with the German version of the Culture Fair Intelligence Test (CFT-20; Weiß, 1987). Because the available test norms referred to different birth cohorts, total IQ scores were computed with regard to the full LOGIC sample at age 17. The internal consistencies of the verbal and nonverbal IQ variables were satisfactory ($\alpha > 0.82$). The two IQ-scores correlated 0.38, and were averaged, resulting in a total IQ score.

Deviation from expected school grade. At age 17, an expected school grade was determined on the basis of the child's age and the Bavarian rules of age-appropriate schooling. Deviations from expected grade were computed by substracting children's actual school grade at age 17 from their expected grade at this age. Thus, except for a few gifted children who attended higher grades than expected, these deviation scores are zero or negative, and more negative scores indicate cognitive and/or social–emotional problems in school.

Results

Derivation of personality types and personality dimensions for ages 4–6

Because Asendorpf and van Aken (1999, 2003) provided a detailed description of the procedures, they are only briefly summarized here.

Personality types. The averaged Q-sort items for ages 4–6 were subjected to a Q-factor analysis with three factors that were interpreted as Q-sort prototypes of resilient, overcontrolled, and undercontrolled children. 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). In essence, they were classified as resilient/overcontrolled/undercontrolled if their average Q-sort was more similar to the resilient/overcontrolled/undercontrolled prototype than to the other two prototypes. Of the 99 children, 52% were classified as resilients, 20% as overcontrollers, and 27% as undercontrollers.

Big Five Q-sort indices. Following the procedure of John et al. (1994), the Big Five personality factors (extraversion, neuroticism, agreeableness, conscientiousness, and openness to experience) were assessed by Q-sort indices consisting of four to six Q-sort items for each factor (see Asendorpf & van Aken, 2003, for details). The reliabilities of the five indices were satisfactory, median $\alpha = 0.78$, range 0.72–0.89, and the indices showed substantial concurrent external validities (see Asendorpf & van Aken, 2003, for details).

Ego-control/resiliency Q-sort dimensions. For each of the three Q-sort assessments, the Q-sort dimensions ego-control and ego-resiliency were derived as suggested by Block and Block (1980). Each child's Q-sort profile was correlated with prototypic Q-sort profiles obtained from experts for ego-control and ego-resiliency (see Block & Block, 1980). The

resulting scores were then averaged across ages, resulting in one ego-control score and one ego-resiliency score for each child (see also Asendorpf & van Aken, 1999).

Derivation of personality types for age 17

Personality types were derived from the parental Big Five scores at age 17 with exactly the same procedure that was used for deriving personality types in the LOGIC study at age 12 (Child Adjective Study in Asendorpf et al., 2001), and in the Berlin Relationship Study as described above. Thus, Ward's hierarchical clustering procedure was applied to the raw Big Five scores, followed by the nonhierarchical K-means procedure; only the three-cluster solution was considered. The consistencies between the resulting cluster solution and the solutions derived within the same longitudinal sample at age 12 and the Berlin Relationship Study was evaluated by the method described above in the section on the Berlin Relationship Study. The consistency with the age 12 clusters was $kappa = 0.73$. Because of the sufficient consistency, the clusters were once more labelled resilient, overcontrolled, and undercontrolled. The consistency with the Berlin Relationship Study clusters was $kappa = 0.64$, despite the different samples, ages, judges, and Big Five instruments.

Prediction of personality dimensions from types versus dimensions

A first head-to-head comparison contrasted the longitudinal predictive power of the two dummy-coded type indicators with the five Big Five Q-sort indices (see Table 5). Predicted at age 17 were personality dimensions as well as two key indicators of cognitive ability and achievement, IQ and grade deviation. The same models as in the analyses of the Berlin Relationship Study were contrasted with one another. The analyses are a conceptual replication of the analyses summarized in Table 1 except that they refer to long-term

Table 5. Head-to-head comparison of two dummy-coded Q-types and five (dichotomized) Q-sort indices at ages 4–6 in prediction of outcomes at age 17

	Predictors at ages 4–6						
	Models 1&2	Model 1	Model 2	Model 3		Model 4	
	Step 1 R^2 2 dummy variables	Step 2 ΔR^2 5 Q-sort indices		Step 1 R^2 5 Q-sort indices Continuous	Step 2 ΔR^2 2 dummy variables	Step 1 R^2 5 Q-sort indices Dichotomized	Step 2 ΔR^2 2 dummy variables
		Continuous	Dichotomized				
Big Five							
Emotional stability	0.17***	0.05	0.02	0.21***	0.01	0.14*	0.04
Extraversion	0.08*	0.05	0.03	0.13*	0.01	0.09	0.02
Openness	0.09**	0.07	0.03	0.16**	0.01	0.10	0.03
Agreeableness	0.08*	0.08	0.05	0.07	0.08*	0.04	0.08*
Conscientiousness	0.11**	0.05	0.07	0.16**	0.00	0.18**	0.01
Shyness	0.08*	0.04	0.02	0.11*	0.01	0.06	0.03
Aggressiveness	0.12**	0.06	0.05	0.09	0.09*	0.07	0.10**
IQ	0.07*	0.13*	0.07	0.18**	0.01	0.11*	0.03
Deviation from grade	0.08*	0.08	0.03	0.16*	0.00	0.08	0.04
Average	0.10	0.07	0.04	0.14	0.02	0.10	0.04

$N = 99$. The third dummy-coded type is redundant with the others. The Q-sort indices were dichotomized using median split.

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

predictions, not to cross-sectional predictions. The explained variances were therefore expected to be much smaller.

Indeed, as a comparison of Tables 1 and 5 shows, the two dummy-coded type indicators plus the five continuous dimensions explained on average 37% of the variance of the concurrent concomitants, but only 17% of the long-term outcomes. Interestingly, this was mainly due to a decreased incremental validity of the continuous predictors (reduction from 23 to 7% explained variance), whereas the predictive power of the types was less strongly reduced (reduction from 14 to 10%).

Also, the types showed this time significant incremental longitudinal validity even over the five continuous Q-dimensions. This result cannot be attributed to chance because four out of 18 incremental validities were significant and the pattern was the same for both the continuous and the dichotomized Big Five indices and for different operationalizations of the conceptually similar traits (non)agreeableness and aggressiveness. More detailed *post hoc* analyses showed that the incremental validity of the types was not due to a particular weakness of the agreeableness variable at ages 4–6. The teacher Q-sort scale for agreeableness at ages 4–6 was reliable, $\alpha = 0.88$, and correlated -0.35, $p < 0.001$, with parental aggressiveness judgments at ages 4–6, and -0.38, $p < 0.001$, with observed aggressiveness in the preschool and kindergarten peer group at ages 4–6 (see Asendorpf & van Aken, 2003).

The weak long-term predictive power of the Big Five indices for agreeableness seems to be due to the fact that nonagreeableness and aggressiveness at age 17 are related not only to low ego-control but also to low ego-resiliency. When the Big Five and the dummy-coded undercontrolled type at ages 4–6 were entered as a first block into a hierarchical multiple regression, the dummy-coded resilient type alone accounted for an additional 5.4% of the variance in agreeableness at age 17 ($p < 0.05$). It seems that being not a resilient child is an important moderator of the long-term outcome of early undercontrol, a moderator that is not sufficiently captured in linear combinations of the Big Five dimensions.

Not surprisingly, the continuous Q-dimensions were better predictors than the dichotomized Q-dimensions, confirming hypothesis 1(iii). Again, hypothesis 2 was not confirmed. However, the two dummy-coded type indicators explained as much variance as the five dichotomized Big Five dimensions (10%), despite the fact that this was an unfair comparison for the types.

To test hypotheses 1(ii) and 3, the next head-to-head comparison was identical to the preceding one except that the two Q-dimensions ego-resiliency and ego-control were used for the predictions instead of the five Big Five Q-sort indices. Because there were now two predictors for both types and dimensions, this analysis provides a fairer comparison of the predictive power of types versus dimensions (see Table 6).

Hypothesis 1(ii) was confirmed because the two Q-sort dimensions explained less variance at age 17 than the five Big Five indices. This time hypothesis 2 was confirmed because the types explained somewhat more variance (10 versus 8%) than the dichotomized Q-sort dimensions. Hypothesis 3, that types are relatively better predictors in long-term studies, was also confirmed, because the types explained in the long-term predictions as much variance as the continuous dimensions (10% in both cases), whereas they fared clearly worse in the cross-sectional predictions (see Tables 1–4). Also, the incremental validity of the Q-sort dimensions (dichotomized or not) over the Q-sort types was not significant in 17 out of 18 cases; the one significant finding may be due to chance. However, the type indicators also did not show significant incremental validity, either over the continuous or over the dichotomized Q-sort dimensions.

Table 6. Head-to-head comparison of two dummy-coded Q-types and two (dichotomized) Q-dimensions at ages 4–6 in prediction of outcomes at age 17

	Predictors at ages 4–6						
	Models 1&2	Model 1	Model 2	Model 3		Model 4	
	Step 1 R^2 2 dummy variables	Step 2 ΔR^2 2 Q-dimensions		Step 1 R^2 2 Q-dimensions	Step 2 ΔR^2 2 dummy variables	Step 1 R^2 2 Q-dimensions	Step 2 ΔR^2 2 dummy variables
		Continuous	Dichotomized	Continuous		Dichotomized	
Big Five							
Emotional stability	0.17***	0.03	0.02	0.18***	0.02	0.15***	0.04
Extraversion	0.08*	0.01	0.00	0.07*	0.02	0.05	0.04
Openness	0.09**	0.04	0.02	0.10**	0.03	0.07*	0.04
Agreeableness	0.08*	0.03	0.00	0.06	0.05	0.03	0.05
Conscientiousness	0.11**	0.02	0.02	0.13**	0.01	0.09*	0.05
Shyness	0.08*	0.02	0.01	0.08*	0.02	0.04	0.04
Aggressiveness	0.12**	0.00	0.02	0.09*	0.04	0.13**	0.02
IQ	0.07*	0.01	0.03	0.07*	0.01	0.07*	0.02
Deviation from grade	0.08*	0.08*	0.06	0.15***	0.01	0.13**	0.01
Average	0.10	0.02	0.02	0.10	0.02	0.08	0.03

$N = 99$. The third dummy-coded type is redundant with the others. The Q-sort dimensions were dichotomized using median split.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Prediction of personality types from personality types and dimensions

Finally, hypothesis 4 was tested by predicting the three personality types at age 17 instead of the continuous outcome variables (see Table 7). Thus, these are once more fairer comparisons between types and dimensions. Because the dependent variables were dichotomies, logistic regressions were used.

Hypothesis 4 that types are relatively better predictors for configural criteria was not confirmed. On the contrary, the types at ages 4–6 explained somewhat less variance of the type differences at age 17 than the dimensions, whereas types and dimensions predicted continuous outcomes equally well (see Table 6). Interestingly, the overcontrolled type at age 17 could not be significantly predicted at all. The type predictors showed no significant incremental validity over continuous or dichotomized dimensions, but again it should be noted that it was hard to find incremental validity because the overall predictive power in these long-term predictions was limited. This also becomes obvious when the long-term stability of the three types is considered; it was highly significant but low, $kappa = 0.23$, $p < 0.001$.

Discussion

Compared to the cross-sectional predictions in the Berlin Relationship Study, the types clearly fared better in the long-term predictions of the LOGIC study. Indeed, it was the surprisingly strong long-term predictive power of the simple distinction between only three personality types that had originally attracted my attention to the type approach when I analysed long-term outcomes of the types at age 12 (Asendorpf & van Aken, 1999). It is noteworthy that every single Big Five dimension at age 17, judged by the parents, as well as important cognitive outcomes such as IQ and deviations from expected school grade, could be significantly predicted from the three types at ages 4–6, judged by preschool and

Table 7. Head-to-head comparison of two dummy-coded Q-types and two (dichotomized) ego-control/resiliency Q-sort dimensions at ages 4–6 in prediction of Q-types at age 17

	Predictors at ages 4–6						
	Models 1&2	Model 1	Model 2	Model 3		Model 4	
	Step 1 R^2 2 dummy variables	Step 2 ΔR^2 2 Q-dimensions		Step 1 R^2 2 Q-dimensions Continuous	Step 2 ΔR^2 2 dummy variables	Step 1 R^2 2 Q-dimensions Dichotomized	Step 2 ΔR^2 2 dummy variables
		Continuous	Dichotomized				
Resilient	0.11**	0.01	0.01	0.11**	0.01	0.09*	0.03
Overcontrolled	0.00	0.04	0.03	0.02	0.02	0.03	0.00
Undercontrolled	0.11**	0.05	0.01	0.15***	0.01	0.09**	0.03
Average	0.07	0.03	0.02	0.09	0.01	0.07	0.02

$N = 99$. The third dummy-coded type predictor is redundant with the others. The Q-sort dimensions were dichotomized using median split. Reported are results for logistic regressions (Cox–Snell R^2).
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

kindergarten teachers. The Big Five dimensions at ages 4–6 added significant incremental predictive power over the three types only in the case of IQ. Furthermore, the three types even outperformed the Big Five dimensions in long-term predictions of disagreeableness and aggressiveness.

More detailed analyses showed that the early assessed Big Five did not sufficiently capture children's resiliency, which turned out to show significant incremental predictive power over both ego-undercontrol and all Big Five dimensions in predictions of later nonagreeableness and agressiveness. Thus, there was one instance where the type approach was more predictive than the Big Five dimensions.

However, this advantage was lost when the three types were contrasted with the two Q-sort dimensions ego-control and ego-resiliency. In this fairer comparison to the type approach, the types explained as much variance as the Q-sort dimensions, and more than the dichotomized Q-sort dimensions, but showed no incremental validity over these dimensions, even over the dichotomized scores.

In addition, the types failed to show a particular advantage in predicting configural outcomes, and they even predicted them less well than continuous dimensions. As a note of caution it should be added that the sample size of the LOGIC longitudinal sample ($N = 99$) was relatively small for these multivariate analyses. On balance, the long-term predictions showed a relative advantage of a type approach when compared to cross-sectional predictions, but incremental validity over continuous dimensions was found only in one particular case.

GENERAL DISCUSSION

The results from these two studies provide a more balanced picture of the (dis)advantages of a type approach relative to a continuous dimension approach than the particular comparison by Costa et al. (2002), which was unfair for a general comparison of type versus dimensional approaches in many respects. When various factors that were expected to moderate the outcome of such comparisons were systematically varied, the type approach fared better in fairer comparisons where the number of predictors were the same

for both approaches. However, in cross-sectional predictions, the type approach was always less predictive than a continuous dimension approach even in such fairer comparisons. Only in long-term predictions was the type approach equally predictive, and in one instance even more predictive than continuous dimensions that were based on the same data. Furthermore, contrasting types with dichotomized dimensions showed only in a few cases an advantage of the configural information inherent in the types. Finally, contrary to what many might expect, the type approach was found to be less predictive of configural outcomes than of continuous outcomes.

This last surprising finding raises questions about the interpretation of the differences between the cross-sectional and the long-term predictions. Can differences in predictive power be attributed only to the cross-sectional versus longitudinal nature of the data, or are there other differences between the Berlin Relationship and the LOGIC study that can account for the results?

An important difference has been already pointed out in the introduction, namely the method of assessing personality profiles. Weaker predictions involved types derived through cluster analysis from Big Five scales, either on the predictor side in the Berlin Relationship Study, or on the outcome side in the prediction of such types from earlier Q-sort based types and dimensions in the LOGIC study. Stronger predictions involved Q-sort based types on the predictor side. Thus, differences between cross-sectional and long-term predictions, and differences between predictions of continuous versus configural outcomes, are confounded with this methodological variation in the present studies. Future longitudinal studies are needed that vary the assessment method (e.g. Q-sort versus Big Five scales) in the same sample of participants independently from cross-sectional versus long-term predictions.

There are good arguments that person-centred assessments may be more sensitive to within-person differences between traits than variable-centred assessments, and there are good arguments that these intraindividual differences may be particularly important for long-term predictions. One outcome of such studies may be that person-centred assessments have a particular advantage for long-term predictions. Such a finding would bring us back full circle to the origins of person-centred assessments in Blocks's (1971) longitudinal study, and would confirm Block's early intuitions about how we should study personality the long way (see also Block, 1993).

Whatever the outcome of such a test will be, the results of the present study suggest that, from the perspective of predictive power, nearly all variations in the relative predictive power of types will be on the disadvantage side. Thus, types have little utility for predictions from personality; the costs of losing within-type information are (with few exceptions) higher, and sometimes much higher, than the gains from the use of type-typical configural information.

However, a clear advantage of types is that results on type differences are much easier to communicate to a wider audience than correlations or regression coefficients. The concept of a personality type as a group of people who share a similar personality profile is intuitively more appealing to the public than a continuous personality dimension, and clinicians and public health authorities frequently use categorical classifications and are trained to frame questions and answers in terms of increases and decreases of risk for groups of people. Therefore, correlational findings are often illustrated with differences between high- versus low-scoring groups on one or two dimensions. If the important differences for prediction cut across three or more dimensions, this approach is difficult to accomplish. Types offer an alternative in such cases and have the particular advantage that

the notion of personality as an individual configuration of traits is preserved in the definition and description of the types.

My conclusion from the present study is, therefore, that configural types are not a serious alternative to dimensions in psychological predictions from personality, but that they are a serious option for describing personality differences, presenting results on personality correlates to the general public, or introducing students to personality psychology (see also Costa et al., 2002).

Last but not least, the limited utility of configural types for prediction should not be confused with a limited utility of a person-centred approach to personality. Person-centred approaches can treat similarities and differences between the profiles of two individuals, or between an individual's profile and a prototypical profile as a continuous dimension (see already Block, 1961, 1971), and more recent multi-level analysis methods such as hierarchical linear modelling make it possible to simultaneously study within-person and between-person differences without reducing the between-person information to categories (Bryk & Raudenbush, 1992). To study personality seriously as the individual within-person organization of behaviour we need to incorporate rich information both on within-person and on between-person differences.

ACKNOWLEDGEMENTS

Part of the data was presented at the 11th European Conference on Personality in Jena, July 2002. I wish to thank Marcel A. G. van Aken and the rest of the LOGIC crew for many years of cooperation, and Iain Glen for comments and stylistic corrections.

REFERENCES

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

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), S1–S96.

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. (2003). 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.

Block, J. (1961). *The Q-sort method in personality assessment and psychiatric research*. Springfield, IL: Thomas.

Block, J. (1971). *Lives through time*. Berkeley, CA: Bancroft.

Block, J. (1993). Studying personality the long way. In D. Funder, R. D. Parke, C. Tomlinson-Keasey, & K. Widaman (Eds.), *Studying lives through time: Personality and development* (pp. 9–41). Washington, DC: American Psychological Association.

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, NJ: Erlbaum.

Borkenau, P., & Ostendorf, F. (1993). *NEO-Fünf-Faktoren Inventar (NEO-FFI) NEO-Five-Factor Inventory.* Göttingen: Psychologie.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models*. Newbury Park, CA: Sage.
Caspi, A. (1998). Personality development across the life course. In N. Eisenberg (Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 311–388). New York: Wiley.
Caspi, A. (2000). The child is the father of the man: Personality continuities from childhood to adulthood. *Journal of Personality and Social Psychology, 78*, 158–172.
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., 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*(S1), S73–S87.
Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory NEO-PI-R and NEO Five-Factor Inventory NEO-FFI professional manual*. Odessa, FL: Psychological Assessment Resources.
Döring, N., & Bortz, J. (1993). Psychometrische Einsamkeitsforschung: Deutsche Neukonstruktion der UCLA Loneliness Scale [Psychometric loneliness research: A new German version of the UCLA loneliness scale]. *Diagnostica, 39*, 224–239.
Funder, D. C. (1991). Global traits: A neo-Allportian approach to personality. *Psychological Science, 2*, 31–39.
Funder, D. (1999). *Personality judgment: A realistic approach to person perception*. San Diego, CA: Academic.
Göttert, R., & Asendorpf, J. (1989). Eine deutsche Version des California-Child-Q-Sort, Kurzform [A German short version of the California Child Q-Set]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 21*, 70–82.
Hart, D., Hofmann, V., Edelstein, W., & Keller, M. (1997). The relation of childhood personality types to adolescent behavior and development: A longitudinal study of Icelandic children. *Developmental Psychology, 33*, 195–205.
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.
Marsh, H. W., & O'Neill, R. (1984). Self-description questionnaire III: The construct validity of multidimensional self-concept ratings by late adolescents. *Journal of Educational Measurement, 21*, 153–174.
McCrae, R. R., Costa, P. T., & Busch, C. M. (1986). Evaluating comprehensiveness in personality systems: The California Q-set and the five-factor model. *Journal of Personality, 54*, 430–446.
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.
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.
Tewes, U. (1991). *Hamburg–Wechsler Intelligenztest für Erwachsene, Revision 1991 [Hamburg–Wechsler intelligence test for adults, revision 1991]*. Bern, Switzerland: Huber.
Van Lieshout, C. F. M., Haselager, G. J. T., Riksen-Walraven, J. M., & van Aken, M. A. G. (1995). *Personality development in middle childhood*. Paper presented at the Society for Research in Child Development, Indianapolis, IN, 1995.
Weinert, F. E., & Schneider, W. (Eds.). (1999). *Individual development from 3 to 12: Findings from the Munich Longitudinal Study*. Cambridge: Cambridge University Press.
Weiß, R. (1987). *Grundintelligenztest Skala 2 (CFT 20) [Culture Fair Intelligence Test]* (3rd ed.). Göttingen: Hogrefe.
York, K. L., & John, O. P. (1992). The four faces of Eve: A typological analysis of women's personality at midlife. *Journal of Personality and Social Psychology, 63*, 494–508.

