Replicable Types and Subtypes of Personality: German NEO-PI-R versus NEO-FFI

KONRAD SCHNABEL[1]*, JENS B. ASENDORPF[1]
and FRITZ OSTENDORF[2]

[1]*Humboldt-Universität zu Berlin, Germany*
[2]*Universität Bielefeld, Germany*

Abstract

In two German samples we investigated whether three major personality prototypes (resilient, overcontrolled, and undercontrolled) can be replicated across different Big Five instruments (NEO-FFI and NEO-PI-R), and whether they can be differentiated into replicable subtypes. For prototypes and subtypes, replicability within and consistency across samples was quantitatively measured. Three replicable prototypes were confirmed, but only two replicable resilient subtypes (well adjusted and assertive) were found. It is concluded that the search for a hierarchical taxonomy of people will require more investigation. Copyright © 2002 John Wiley & Sons, Ltd.

INTRODUCTION

When people describe an object they rarely say, for example, 'It has four legs, it is 1.25 m long, and 61 cm high'. Instead, depending on the situation they prefer to use, for example, the word 'sheep' or 'table'. The perceiver immediately relates the word to a fairly complex mental image that includes the four legs, the length, and the height, but also many more features. In the case of a sheep one could assume a woolly coat, and in the case of a table a smooth surface. People have particular terms for certain configurations of attributes that represent what is *typical* for a group of objects. Personality researchers, however, have not, on the whole, tended to do this, and instead of classifying people into types they prefer to concentrate their efforts on classifying variables. Thus, they have tended to develop methods to systemize the enormous number of characteristics in which people reliably differ (i.e. traits).

The five-factor, or Big Five, model of personality is currently the most common classification for personality traits. This model proposes that at the broadest level of

*Correspondence to: Konrad Schnabel, Institut für Psychologie, Humboldt-Universität zu Berlin, Oranienburger Strasse 18, 10178 Berlin, Germany. E-mail: konrad.schnabel@psychologie.hu-berlin.de

description, five basic personality dimensions can be distinguished: neuroticism, extraversion, openness to experience (or culture), agreeableness, and conscientiousness (see, e.g. John & Srivastava, 1999). Thus, the Big Five represents a *dimensional*, or variable-centred, approach to personality, which is both broad (i.e. including a maximum spectrum of different traits) and efficient (i.e. managing this with a minimum set of dimensions).

In contrast, the *typological*, or person-centred, approach aims at developing a taxonomy of personality types. Whereas the dimensional approach usually investigates personality dimensions in isolation, and is silent about the overall structure of personality *within* individuals, the typological approach classifies people on the basis of their *individual* personality structure. For a long time there was little consensus on a taxonomy of people (Block, 1971). Recently, researchers have stressed the advantages of the typological approach, and progress has been made toward the identification of a replicable and generalizable personality typology.

AN EMPIRICALLY DERIVED PERSONALITY TYPOLOGY: RESILIENTS, OVERCONTROLLERS, AND UNDERCONTROLLERS

Across numerous recent studies, three major personality prototypes have consistently emerged, despite the fact that these studies differed considerably in sets of personality traits (e.g. Big Five scales, Q-sort patterns), judgment (self versus other), sample characteristics (age, gender, geographic location), and methods of deriving types (Q-factor versus cluster analysis; Asendorpf & van Aken, 1999; Asendorpf, Borkenau, Ostendorf, & van Aken, 2001; Caspi, 1998; Robins, John, & Caspi, 1998; Robins, John, Caspi, Moffitt, & Stouthamer-Loeber, 1996). Although the labels that were given to the personality types differed slightly across these studies, the prototypes could always be identified as resilient, overcontrolled, and undercontrolled. These labels refer to the two dimensions of ego-control and ego-resiliency proposed by the personality theory of J. H. Block and J. Block (1980). Ego-resiliency describes the ability to respond flexibly rather than rigidly to changing situational demands, such as stress, conflict, or uncertainty, and is related to the constructs of competence, social intelligence, and coping. Ego-control refers to the tendency to contain versus express motivational and emotional impulses, and is related to the constructs of introversion–extraversion, reflection–impulsivity, internalizing–externalizing, and delay of gratification.

These studies adopted one of two different approaches to empirically deriving personality types in the multivariate trait space: Q-factor analysis and cluster analysis. Q-factor analysis starts with Q-sort profiles, i.e. trait descriptions that are sorted according to how well they fit the individual. Judges are asked to produce Q-sorts with equal mean and variance such that the Pearson correlation between the profiles is an adequate measure of profile similarity. Intercorrelations among these profiles are then factor-analysed by Q-factor analysis. The resulting Q-factors represent prototypes, and individual factor loadings are an index of the similarity of a person with the prototype. Individuals are classified into types according to their best-fitting Q-factor. Cluster analysis, on the other hand, starts with multiple variables (e.g. questionnaire scales). Individual profiles are grouped into relatively homogenous clusters, resulting in maximized similarity of persons within clusters and minimized similarity between clusters. The mean profile of all cluster

members represents the prototype, and a person's Euclidean distance to a prototype is an index of the person's prototypicality for the type.

Regardless of whether prototypes were directly identified by cluster analysis of Big Five scales (Asendorpf et al., 2001; van Lieshout, Haselager, Riksen-Walraven, & van Aken, 1995) or by Q-factor analysis of Q-sort patterns and subsequently described in terms of the Big Five (Asendorpf et al., 2001; Robins et al., 1996, 1998), the Big Five profiles of the three prototypes showed considerable consistency across studies. Resilients had a generally well adjusted profile, with below average neuroticism and above average or intermediate scores in the remaining four dimensions. Overcontrollers scored high in neuroticism and low in extraversion, and undercontrollers were distinguished by their particularly low scores in conscientiousness and agreeableness.

THE SEARCH FOR PERSONALITY SUBTYPES

At the prototype level, the application of methodologically rigorous standards (i.e. replicability within random halves of the sample, generalizability across gender, developmental period, and ethnicity) resulted in the identification of the described prototypes that are assumed to constitute a core set of types for any generalizable personality typology (Caspi, 1998). In contrast, little progress has been made that would allow for a multi-level taxonomy of personality types. One approach is trying to identify groups of individuals with different personality profiles *within* each personality prototype. Through such a hierarchical taxonomy, more specific and subtle individual differences that remain undetected in the noise of broader variance between prototypes could be determined. However, up to now there exist only two studies along these lines.

Pulkkinen (1996) used Ward's hierarchical cluster procedure to analyse extensive self-report data of young adults from Finland. In a first step, Pulkkinen extracted two major clusters for both men and women. The *adjusted* prototype represented a positive adaptation to life, whereas the *conflicted* prototype represented internalizing problems for women and externalizing problems for men. In a second step, Pulkkinen gradually increased the number of inspected clusters from two to seven. For women, consistency with the results of other studies was low already at the prototype level, so only the male subtypes were reported.

The adjusted men were subdivided into resilients and introverts. These types correspond closely to the resilient and overcontrolled prototypes described earlier. The introverts were further subdivided into passive adjusters and overcontrollers. Compared with overcontrollers, passive adjusters were characterized by lower intellectual interest and lower neuroticism. Overcontrollers were described as neurotic and lacking self-confidence. The resilient men were subdivided into constructives and extraverts. Both groups were extraverted, socially integrated, and had positive life attitudes, although the extraverts scored even higher in extraversion than the constructives and had intermediate rather than below average scores in neuroticism. The conflicted men were further subdivided into undercontrollers and anomic revellers. Anomic revellers had very low scores in conscientiousness and were not content with their present life situation. Undercontrollers were further divided into antisocials and neurotics. Antisocials had acute problems in social integration, whereas neurotics scored especially high in neuroticism.

Unfortunately Pulkkinen (1996) did not examine the replicability of the identified personality types nor did she optimize the hierarchically derived cluster solution by a

non-hierarchical clustering procedure. These limitations do not apply to the study by Robins et al. (1998), who performed Q-factor analyses within prototypes based on California Child Q-Set data of adolescent boys. These authors started with the three prototypes described earlier (identified on the basis of Q-factor analysis of Q-sort patterns) and computed additional Q-factor analyses within each prototype. They identified two replicable subtypes for both resilients and undercontrollers. The subtypes of the resilient prototype were labelled as agentic and communal. Agentics scored high in extraversion and low in neuroticism, whereas communals scored high in agreeableness and conscientiousness. Both subtypes were high in ego resiliency, but whereas agentic resilients were somewhat undercontrolled, communals were somewhat overcontrolled. The subtypes of the undercontrolled prototype were labelled as impulsive and antisocial. Whereas the impulsives scored intermediate in agreeableness and conscientiousness, the antisocials were characterized by extremely low scores in these two Big Five dimensions. Both subtypes scored low in ego resiliency and high in undercontrol, with the antisocials having even more extreme scores in both scales than the impulsives.

THE PRESENT STUDY

In summary, there have been only two studies that have investigated the replicability of personality subtypes: one using Q-factor analysis and one study using a less than optimal clustering procedure without considering type replicability. Therefore, the main purpose of this study was to identify replicable subtypes based on cluster analysis of Big Five profiles. It was hypothesized first that the three prototypes Caspi (1998) considered to represent a generalizable typology would be replicable both within and between our studies. Second, we expected that within these prototypes, at least the subtypes that were reported by Robins et al. (1998), could be identified. We therefore (i) applied in two fairly large samples a combination of both hierarchical and non-hierarchical clustering procedures to two different, but related Big Five inventories, (ii) evaluated the replicability of the derived types within random splits of the samples using a quantitative coefficient of consistency, and (iii) studied the consistency of the replicable types across studies. To maximize the consistency of our results with earlier findings by Asendorpf et al. (2001), we used exactly the same clustering procedures as these authors.

METHOD

This section starts with a description of the two studies followed by the method of how we derived personality prototypes and subtypes.

The German NEO-PI-R study

Participants were part of a larger sample of 1200 subjects aged 16–83 years who were recruited through advertisements in local newspapers, by psychology students who asked acquaintances for their participation, or in psychology courses. To ensure comparability of the sample with the NEO-FFI study (see below) we excluded all participants who were younger than 20 or older than 30 years of age. Subsequently, the sample was balanced for sex ($N = 786$, 393 men and 393 women; mean age 23.9 years, SD = 2.93 years) by random

exclusion of 11 men. All participants completed the German NEO-PI-R (Ostendorf & Angleitner, unpublished), which is a translation of the Revised NEO Personality Inventory developed by Costa and McCrae (1992). This inventory assesses the Big Five factors Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness as well as six facets for each of these dimensions. Facet scales consist of eight items, resulting in a sum of 48 items for each Big Five scale. To make our results comparable to earlier studies, we analysed personality types on the basis of the five Big Five scales rather than the 30 facets. The internal consistencies of the five scales were high (in each case, $\alpha > 0.85$).

The German NEO-FFI study

Participants were part of the German NEO-FFI normative sample that comprised a total of 2112 subjects. The German NEO-FFI (Borkenau & Ostendorf, 1993) is a translation of the NEO Five-Factor-Inventory of Costa and McCrae (1989), which assesses each of the factors Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness with 12 items. The NEO-FFI is a subtest of the NEO-PI-R because each NEO-FFI item is also represented in the NEO-PI-R. Of the 2112 participants, 195 were excluded because they had missing values. To ensure comparability with the NEO-PI-R study, only participants within an age range from 18 to 24 years were considered for the analysis. Of the resulting 858 subjects, we further excluded 128 randomly selected females to equalize the sex ratio of our sample. This led to a final sample of 730 participants (365 women and 365 men; mean age 21.6 years, SD = 1.69). The internal consistencies of the five scales ranged from 0.70 for Openness to experience to 0.86 for Conscientiousness. This study is identical to the adult NEO-FFI study reported by Asendorpf et al. (2001).

Derivation of personality prototypes and subtypes

The derivation of personality prototypes exactly followed the method used by Asendorpf et al. (2001). In a first step, Ward's hierarchical clustering procedure was applied. The resulting three-, four-, and five-cluster solutions were subsequently optimized through the nonhierarchical *k*-means procedure. Only those solutions that showed a median replicability (Cohen's κ) of at least 0.60 across ten random splits of the full sample were considered acceptable.

The consistency of the prototypes across studies was evaluated by assigning the *z*-standardized individual Big Five patterns of one study to (i) the best-fitting original *z*-standardized cluster centres that were derived on the basis of raw scores from this study, and (ii) the best-fitting *z*-standardized cluster centres of the other study. This resulted in two classifications per subject which were then cross-classified. Because this procedure could be applied to both studies, the resulting κs were averaged. Again, a mean κ-value of at least 0.60 was considered acceptable. As explained in the Editorial in a footnote, this procedure deviated slightly and for good reasons from the procedure of Asendorpf et al. (2001) of computing cross-study consistencies.

For the derivation of subtypes the same procedure (Ward, followed by *k*-means, evaluation of cluster replicability and cross-study consistency) was repeated *within* each prototype. Because of the smaller *n* for the subtype analyses, only the two- and three-cluster solutions were considered for analysis.

Table 1.   Replicability of the personality prototypes within studies

| Study | N | Median replicability for prototypes | | |
		3	4	5
NEO-PI-R	786	0.72	0.58	0.40
NEO-FFI	730	0.69	0.54	0.51

Note. Median replicability refers to ten random splits within sex and age.

RESULTS

Replicability and cross-study consistency of prototypes

The median replicabilities of the three-, four-, and five-prototype solutions in the NEO-PI-R and NEO-FFI studies are presented in Table 1. In both studies, only the three-cluster solution reached sufficient replicability. For the three-prototype solution, median κ-values were 0.72 (NEO-PI-R) and 0.69 (NEO-FFI), although there was considerable variance between the ten random splits within each sample, with κ-values ranging from 0.24 to 0.91 (NEO-PI-R), and from 0.25 to 0.87 (NEO-FFI).

In the NEO-PI-R study, even the four-prototype solution reached almost acceptable replicability, yet it was noticeably lower than for the three-prototype solution. Accordingly, the three-prototype solution was considered the most suitable division of the whole sample within each study. To compare the cluster solutions of the two studies, their consistency was calculated across studies. The mean κs were 0.72, 0.52, and 0.46 for the three-, four-, and five-prototype solution, respectively. Thus, only the three-prototype solution of both studies showed acceptable consistency. All in all, the three-prototype solution proved both acceptable replicability within and acceptable consistency across the two studies.

Description of the prototypes

Figure 1 describes the mean Big Five patterns of several prototype solutions that were found in the two studies. For better comparability, the Big Five scores were first standardized (i.e., z-transformed) and then averaged within each prototype. As can be seen in panels A and B, the Big Five patterns of the three-prototype solution could be unambiguously identified as resilient, overcontrolled, and undercontrolled in both studies, and correspond to those found in previous studies (Asendorpf et al., 2001; Robins et al., 1998).

The *resilient prototype* is characterized by low scores in Neuroticism and high scores in Conscientiousness. Resilient individuals also have above average scores in Extraversion, and average scores in Openness and Agreeableness. The most specific characteristic of the *overcontrolled prototype* is its high Neuroticism and low Extraversion, whereas scores were average in the other Big Five scales. However, the overcontrolled pattern based on the NEO-PI-R is additionally characterized by low Openness. The *undercontrolled prototype* is characterized by low Conscientiousness in both studies. Unlike the results for other studies (Asendorpf et al., 2001; Asendorpf & van Aken, 1999), neither the under-controlled individuals in the NEO-PI-R study nor those in the NEO-FFI study were characterized by low scores in Agreeableness. In the NEO-FFI study, they scored average
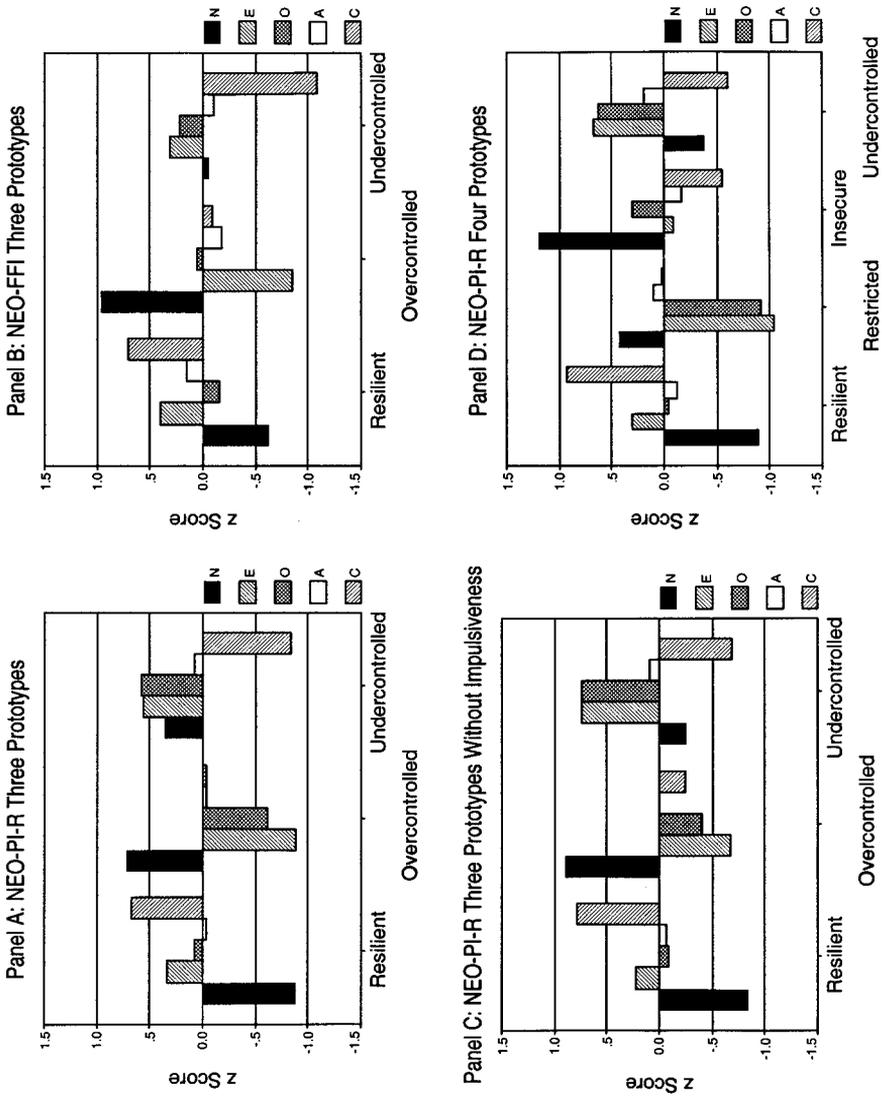
Figure 1. Three major personality prototypes characterized by their Big Five patterns in two studies (panels A and B); three prototypes that resulted when the impulsiveness facet was excluded from the NEO-PI-R study (panel C); and the four-prototype solution based on the NEO-PI-R (panel D). N, Neuroticism; E, Extraversion; O, Openness; A, Agreeableness; C, Conscientiousness.

in Neuroticism, Extraversion, and Openness, and in the NEO-PI-R study they scored above average in Extraversion and Openness. In sum, the specific characteristics of the three prototypes in both studies were that resilients scored low in Neuroticism and high in Conscientiousness, overcontrollers were neurotic and introverted, and undercontrollers were unconscientious.

Because in the NEO-PI-R each Big Five scale is hierarchically subdivided into six facets, the three prototypes based on the NEO-PI-R could be additionally described in terms of these facets. Table 2 lists the mean facet scores within each prototype. For better comparability z scores are reported. Two main observations can be made from Table 2. First, prototypes that were characterized by exceptionally high or low scores in any Big Five scale (see Figure 1) had high or low scores in the subordinated facets. Second, there is however one main exception: Overcontrollers scored high in all neuroticism facets except in impulsiveness, in which they had average scores. In contrast, undercontrollers

Table 2. NEO-PI-R facet profiles of the three replicable personality prototypes

Scale	Facet	Prototype		
		Resilient	Overcontrolled	Undercontrolled
Neuroticism	Anxiety	−0.74	0.64	0.26
	Angry Hostility	−0.65	0.46	0.34
	Depression	−0.76	0.71	0.19
	Self-Consciousness	−0.65	0.75	0.01
	Impulsiveness	−0.46	−0.06	0.66
	Vulnerability	−0.78	0.68	0.25
Extraversion	Warmth	0.24	−0.67	0.44
	Gregariousness	0.16	−0.60	0.46
	Assertiveness	0.42	−0.64	0.17
	Activity	0.16	−0.47	0.32
	Excitement-Seeking	0.06	−0.41	0.38
	Positive Emotions	0.31	−0.83	0.53
Openness	Fantasy	−0.07	−0.37	0.51
	Aesthetics	0.01	−0.34	0.36
	Feelings	0.02	−0.45	0.48
	Actions	0.17	−0.63	0.48
	Ideas	0.22	−0.37	0.14
	Values	−0.03	−0.28	0.34
Agreeableness	Trust	0.19	−0.39	0.19
	Straightforwardness	0.01	0.22	−0.25
	Altruism	0.14	−0.31	0.17
	Compliance	0.00	0.08	−0.10
	Modesty	−0.29	0.34	0.00
	Tender-Mindedness	−0.20	−0.06	0.33
Conscientiousness	Competence	0.64	−0.45	−0.33
	Order	0.41	0.13	−0.68
	Dutifulness	0.48	0.04	−0.67
	Achievement Striving	0.44	−0.03	−0.55
	Self-Discipline	0.69	−0.20	−0.67
	Deliberation	0.34	0.28	−0.76

Note. Reported are z scores.

were characterized by exceptionally high scores in impulsiveness, whereas they were average or slightly above average in the remaining neuroticism facets. These findings from a person-centered approach suggested that the impulsiveness facet might not fit very well into the neuroticism scale.

Because of this problem with the impulsiveness facet, we calculated an additional cluster analysis based on NEO-PI-R scale means that were computed without the items of the impulsiveness facet. Results are depicted in Figure 1, panel C. As can be seen in the figure, the Big Five patterns of the three prototypes did not change very much. Compared with panel A, overcontrollers of panel C were slightly more neurotic, whereas the opposite was true for undercontrollers. Nevertheless, the Neuroticism scores of the three prototypes became more similar to the NEO-FFI, which represents a short form of the NEO-PI-R, but whose Neuroticism scale does not contain any impulsiveness items.

In sum, inclusion of impulsiveness slightly decreased differences between over-controllers and undercontrollers in Neuroticism and reduced correspondence with the NEO-FFI study. However, exclusion of impulsiveness had almost no effect on the consistency with the NEO-FFI cluster solution at the scale level ($\kappa = 0.71$ versus 0.72 for all facets), and the within-study replicability even decreased somewhat (median $\kappa = 0.62$ versus 0.69 for all facets). Therefore, the following analyses of the NEO-PI-R were based on the full Neuroticism scale scores.

Comparisons of the group sizes of the three prototypes revealed that in both studies resilients were the largest group, followed by overcontrollers and undercontrollers. This finding is consistent with earlier studies (Asendorpf et al., 2001; Asendorpf & van Aken, 1999; Robins et al., 1996). The mean proportions across both studies were 41% resilients, 31% overcontrollers, and 28% undercontrollers. There were significant sex differences in the distribution of the three prototypes in the NEO-PI-R, $\chi^2(n = 786, df = 2) = 18.3$, $p < 0.001$, but not in the NEO-FFI, $\chi^2(n = 730, df = 2) = 4.58$, $p > 0.10$. To determine which specific prototypes showed sex differences in the NEO-PI-R study, *a posteriori* group comparisons were calculated contrasting the sex proportion of each prototype with that of the two pooled remaining prototypes. Whereas there were significantly more men (58%) than women (42%) within the group of resilients, $\chi^2(n = 299, df = 1) = 11.9$, $p < 0.01$, the opposite was true for the undercontrollers, $\chi^2(n = 230, df = 1) = 15.4$, $p < 0.001$, 61% women, 39% men. Within the overcontrolled type, the ratio of men and women was virtually identical (49% women, 51% men).

The four-prototype solution based on the NEO-PI-R is also presented in Figure 1, panel D. It should be noted that these prototypes were neither replicable within random splits of the sample nor consistent with the NEO-FFI. Therefore, they may not represent a reliable prototype solution within the NEO-PI-R study. They are nevertheless reported here because this may facilitate comparability with the results of other studies. The resilient and the undercontrolled prototype from the four-prototype solution correspond to those from the three-prototype solution. However, the resilients from the four-prototype solution scored somewhat higher in Conscientiousness than the resilients from the three-prototype solution, whereas the opposite was true for the undercontrollers. The overcontrollers seem to have split into *restricted* subjects with extremely low scores in Extraversion and openness and into *insecures* with extremely high scores in Neuroticism.

When these two prototypes were pooled, and the consistency of the resulting three clusters with the three-prototype solution was calculated, Cohen's κ was 0.68. This result illustrates that the four-prototype solution can be regarded as a variant of the three-prototype solution with the overcontrolled prototype subdivided into two types. Inspection

Table 3.   Within- and between-study consistency of the personality subtypes

Study	Resilient		Overcontrolled		Undercontrolled	
	2	3	2	3	2	3
	Within-study replicability of the subtypes					
NEO-PI-R	$n = 299$		$n = 257$		$n = 230$	
	0.68	0.51	0.36	0.37	0.22	0.48
NEO-FFI	$n = 323$		$n = 218$		$n = 189$	
	0.26	0.40	0.19	0.34	0.24	0.37
	Between-study consistency of the subtypes					
	0.44	0.25	0.19	0.29	0.11	0.30

Note. Median replicability refers to ten random splits within sex.

of group sizes indicated that resilients (29%) and undercontrollers (26%) were slightly more frequent than restricted subjects (23%) and insecures (22%).

Replicability and consistency of subtypes

In a next step the prototypes were divided into subtypes using the same clustering procedures as for the derivation of types. Median replicabilities within the two studies and cross-study consistencies of the two- and three-subtype solutions within each prototype are presented in Table 3. Only the two-subtype solution of the resilient prototype within the NEO-PI-R study reached sufficient replicability. In neither study did subtypes for the overcontrolled or for the undercontrolled prototype reach acceptable replicability. Furthermore, the two resilient subtypes within the NEO-PI-R study were not replicable within the NEO-FFI study. When the two resilient subtypes of both studies were compared, they showed low consistency ($\kappa = 0.44$). Likewise, none of the other subtype solutions showed acceptable consistency between the NEO-FFI and NEO-PI-R studies. Therefore, only the two-subtype solution within the NEO-PI-R study was considered as an acceptable division of the resilient prototype.

Description of the two resilient subtypes

Figure 2 describes the two-subtype solutions that resulted from the clustering procedure within the NEO-PI-R. Again, the Big Five scores were first standardized and then averaged within each subtype. As can be seen in panel A, the two resilient subtypes could be identified as well adjusted and assertive. The *well adjusted* subtype was characterized by average or slightly below average scores in Extraversion, Openness, and Agreeableness. Like all resilients, the well adjusted subjects scored low in Neuroticism and high in Conscientiousness. The most specific characteristics of the *assertive* subtype were its high Extraversion and Openness, which clearly distinguished it from the well adjusted subtype. In contrast, the two resilient subtypes based on the NEO-FFI, which were not replicable and are therefore not depicted in Figure 2, differed only in their Openness.

The two-subtype solutions based on the NEO-PI-R are also reported for the overcontrolled and the undercontrolled prototype, although they did not reach the replicability criterion. The overcontrolled subtypes closely corresponded to the two
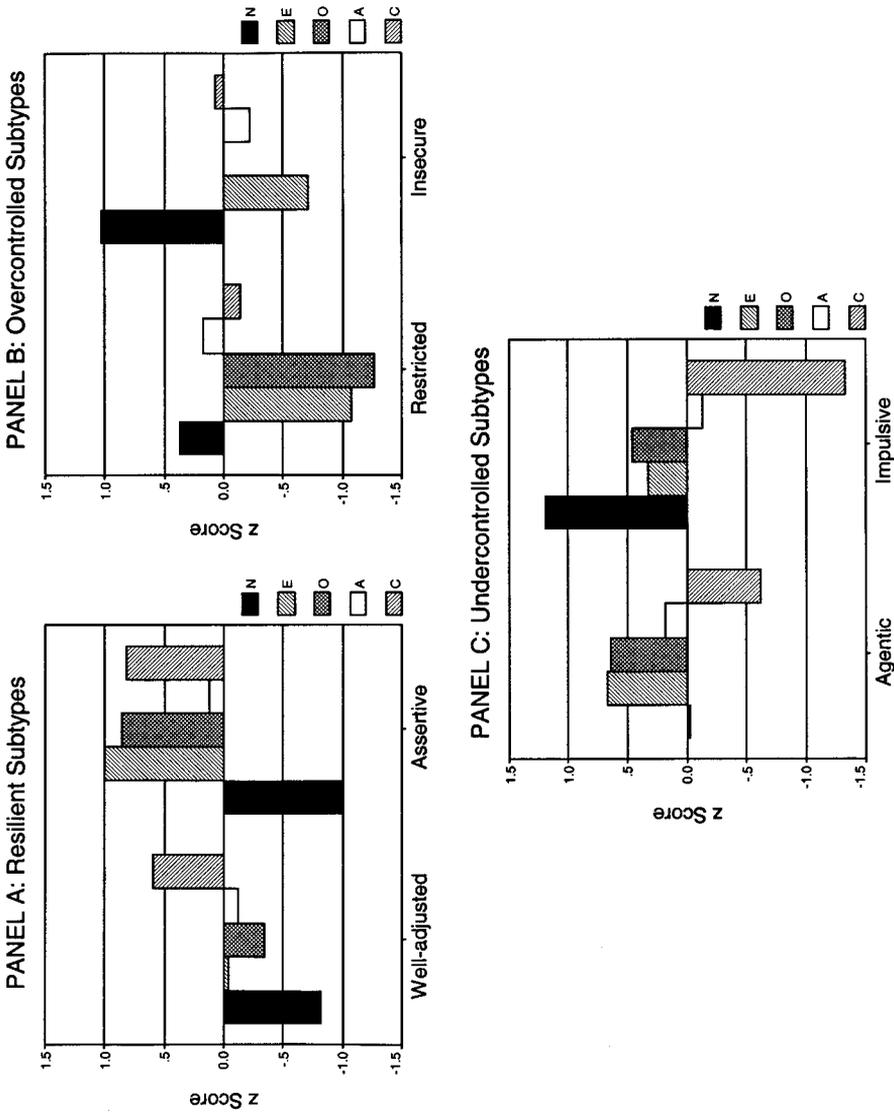
Figure 2.   Two-subtype solutions within the prototypes of the NEO-PI-R study characterized by their Big Five patterns. N, Neuroticism; E, Extraversion; O, Openness; A, Agreableness; C, Conscientiousness.

prototypes that were already described within the four-prototype solution. However, in the subtype solution, restricted subjects had somewhat lower scores in Openness than in the prototype solution, and insecures were not unconscientious. When the overcontrollers of the three-prototype solution were divided into the two subtypes and consistency with the four-prototype solution was calculated, κ was 0.62. This illustrates once more that the four-prototype solution resulted in a split of the overcontrollers.

As with the three prototypes, the two resilient subtypes were also described in terms of the NEO-PI-R facet subscales. The mean z-standardized facet scores within each subtype are reported in Table 4. Again, most subtypes that had high or low scores in any of the Big Five scales had also high or low scores in the subordinated facets. However, assertives scored low in all neuroticism scales ($z < -0.65$) except in impulsiveness ($z = -0.25$). Inspection of the highest absolute z scores for both subtypes indicated that well-adjusted subjects described themselves as not depressed ($z = -0.70$), not anxious ($z = -0.67$), and not vulnerable ($z = -0.65$), all of which are facets of the Neuroticism scale, whereas assertives characterized themselves as not vulnerable ($z = -0.99$), assertive ($z = 0.92$), or competent ($z = 0.92$), which are facets of Neuroticism, Extraversion, and Conscientiousness,

Table 4.   NEO-PI-R facet profiles of the two resilient subtypes

Scale	Facet	Well adjusted	Assertive
Neuroticism	Anxiety	−0.67	−0.88
	Angry Hostility	−0.65	−0.65
	Depression	−0.70	−0.87
	Self-Consciousness	−0.52	−0.88
	Impulsiveness	−0.58	−0.25
	Vulnerability	−0.65	−0.99
Extraversion	Warmth	−0.07	0.80
	Gregariousness	−0.07	0.58
	Assertiveness	0.14	0.92
	Activity	−0.12	0.66
	Excitement-Seeking	−0.10	0.34
	Positive Emotions	0.03	0.82
Openness	Fantasy	−0.33	0.41
	Aesthetics	−0.33	0.63
	Feelings	−0.34	0.67
	Actions	−0.08	0.63
	Ideas	0.01	0.60
	Values	−0.31	0.50
Agreeableness	Trust	0.04	0.44
	Straightforwardness	0.04	−0.06
	Altruism	−0.01	0.40
	Compliance	0.06	−0.10
	Modesty	−0.21	−0.44
	Tender-Mindedness	−0.46	0.25
Conscientiousness	Competence	0.48	0.92
	Order	0.37	0.49
	Dutifulness	0.45	0.54
	Achievement Striving	0.26	0.79
	Self-Discipline	0.62	0.81
	Deliberation	0.44	0.17

Note. Reported are z scores.

Table 5. Mean prototypicality scores of the two resilient subtypes

	Subtype				
	Well adjusted			Assertive	
Prototype	*M*	*SD*		*M*	*SD*
Resilient	−0.63	0.05		−0.84	0.06
Overcontrolled	−1.44	0.04		−2.42	0.06
Undercontrolled	−1.63	0.05		−1.66	0.07

respectively. In sum, main characteristics of the well adjusted subjects are low scores in all Neuroticism facets whereas the assertives are additionally characterized by facets that reflect self-assurance.

Inspection of group sizes indicated that the well adjusted subtype was more frequent (65%) than the assertive subtype (35%). There were no significant sex differences in the distribution of the two subtypes, $\chi^2(n = 299, df = 1) = 1.13, p > 0.28$.

To examine the similarity of the two subtypes with the three prototypes, individual prototypicality scores were analysed. As in the Asendorpf et al. (2001) study, prototypicality was defined as unity minus the Euclidean distance between an individual Big Five pattern and a prototype's Big Five pattern. This measure was unity if an individual pattern corresponded perfectly with a prototype, and declined with the dissimilarity of the individual pattern to the prototype. Table 5 reports the mean prototypicality scores of the well adjusted and assertive subtypes.

Differences between the two subtypes in their prototypicality for the three prototypes were tested by a 2×3 ANOVA with subtype as a between-subject factor and prototype as a within-subject factor. Significant effects were found for subtype, $F(1, 297) = 30.69$, $p < 0.001$, prototype, $F(2, 594) = 926.37$, $p < 0.001$, and the subtype $\times$ prototype interaction, $F(2, 594) = 149.77$, $p < 0.001$. *A posteriori* comparisons *within* subtypes revealed that both subtypes had higher prototypicality scores for the resilient prototype than for both the overcontrolled and the undercontrolled prototype (in each of the four cases, $t > 15.9, p < 0.001$). Well adjusted subjects were more prototypical for the overcontrolled than for the undercontrolled prototype, $t(192) = 5.59, p < 0.001$, whereas the opposite was true for assertives, $t(105) = 18.0, p < 0.001$. *A posteriori* comparisons *between* the two subtypes showed that well adjusted subjects had higher prototypicality scores than assertives for both the resilient, $t(297) = 2.65, p < 0.001$, and the overcontrolled prototypes, $t(297) = 13.0, p < 0.001$. However, the two subtypes did not show significant differences in their prototypicality scores for the undercontrolled prototype, $t < 1$. In sum, as the subtypes were both members of the resilients, they had highest prototypicality scores for the resilient prototype. In addition, well adjusted subjects were more overcontrolled than undercontrolled, whereas assertives were more undercontrolled than overcontrolled.

DISCUSSION

The results of the present study (i) confirmed our hypothesis of three main prototypes for the NEO-PI-R, (ii) confirmed earlier findings by Asendorpf et al. (2001) that only these three prototypes were replicable within and consistent across studies, and (iii) showed a

replicable division of the resilient prototype into a well adjusted and an assertive subtype. However, these two subtypes were replicable only within the NEO-PI-R study and not within the NEO-FFI study, and they were not consistent across these two studies. Furthermore, no replicable subtypes were found for the overcontrolled and the under-controlled prototype.

Thus, at the prototype level the broad picture was the expected one, but at the subtype level the clustering strategy was only partly successful in detecting subtypes. In addition, closer inspection of the results revealed numerous differences to findings of earlier studies. We begin our discussion with these differences, and then move to a broader discussion of the problems of and future perspectives for the person-centred study of personality.

Differences to findings of other studies

Prototype level
In the NEO-PI-R study there were significant sex differences in the distribution of the three prototypes. Men were more frequently resilients than women, and *less* frequently undercontrollers. These sex differences did not occur in the NEO-FFI study and are not consistent with previous research (Asendorpf et al., 2001). In particular, Asendorpf and van Aken (1999) found exactly the opposite sex distributions for children. Girls were more frequently resilients than boys and *less* frequently undercontrollers. Perhaps these unexpected sex differences may be due to the fact that the resilients based on the NEO-PI-R were particularly low in Neuroticism, whereas the undercontrollers were slightly neurotic. Women normally score higher in Neuroticism than men (see, e.g., Borkenau & Ostendorf, 1993), and this was also the case for the women in the NEO-PI-R study, $t(784) = -5.75$, $p < 0.001$. Therefore, women might have been less frequent within particularly non-neurotic resilients, and more frequent within the slightly neurotic undercontrollers.

In contrast to the results of other studies (Asendorpf et al., 2001; Robins et al., 1996, 1998), the undercontrollers based on the NEO-PI-R and the NEO-FFI were not disagreeable. A reason for this may be that, in contrast to most of the cited studies, the subjects of our studies were adults rather than adolescents and were not judged by others but judged themselves. It might be typical for undercontrollers to consider themselves as more agreeable than others do, particularly in adulthood. However, Asendorpf et al. (2001) found for another Big Five self-rating instrument that adult undercontrollers rated themselves as disagreeable.

Subtype level
It was hypothesized that the subtypes found by Robins et al. (1998) would be replicable in our studies. Concerning the resilient subtypes, the assertives of our study scored high in extraversion and low in neuroticism like the agentics in the study of Robins et al. However, in contrast to the communals of Robins et al., the well adjusted subjects of our study neither were agreeable nor extremely conscientious. As with the communals and the agentics in the study by Robins et al. (1998), the well adjusted subjects were somewhat overcontrolled, and the assertives were somewhat undercontrolled. In contrast to the results of Robins et al. (1998), no replicable undercontrolled subtypes were found. It should be noted that Robins et al. used Q-factor analysis for Q-sort patterns of adolescent boys who were judged by their primary caregivers, whereas we used cluster analyses of male and female adults' self-rated Big Five profiles; thus, there were numerous differences between the Robins et al. and our study that may have contributed to the different results at the subtype level.

When our subtypes were compared to those reported by Pulkkinen (1996), even less consistency emerged. Differences may be attributed to the fact that she (i) used only a hierarchical clustering procedure for (ii) different self-report data (e.g. Eysenck Personality Questionnaire, Eysenck, & Eysenck, 1975; Sensation Seeking Scale, Zuckerman, 1979), and (iii) analysed women and men separately. As there was for her female types only small consistency already at the prototype level, our subtypes were only compared with her male subtypes. Concerning the resilient subtypes, the high extraversion of the assertives in our study seemed to correspond with the extraverts of Pulkkinen's study, but the assertives in our study additionally had very low scores in Neuroticism and high scores in Conscientiousness. The low Neuroticism of the well adjusted subjects in our study is comparable to that of the constructives in Pulkkinen's study, but in contrast to the constructives of Pulkkinen's study the well adjusted subjects were not extraverted, not agreeable, and more conscientious.

In sum, the two resilient subtypes that were replicable in our study were consistent with the two resilient subtypes of the other studies to the extent that in all studies the two subtypes predominantly differed in their extraversion.

The problem of scale unreliability

The two resilient subtypes were replicable within the NEO-PI-R but not within the NEO-FFI study. A probable reason is that the Big Five scales of the NEO-FFI contain only one-quarter of the items of the NEO-PI-R and are therefore less reliable. For example, Cronbach's αs for the Extraversion and Openness scales of the NEO-PI-R were 0.89 and 0.88, respectively, whereas for the NEO-FFI, they were 0.81 and 0.70. It should be noted that the two resilient subtypes within the NEO-PI-R predominantly differed in their extraversion and openness. Furthermore, when the reliabilities of the two scales were inspected only for the resilient prototype, αs were 0.86 for both the extraversion and openness scales of the NEO-PI-R but only 0.74 and 0.72 for the NEO-FFI, respectively. Thus, the reliability of the extraversion scale showed a stronger decline in the NEO-FFI than in the NEO-PI-R study when the reliability only for resilients was considered. All in all, the scale reliabilities of the NEO-FFI may have been not sufficiently high for a reliable distinction of the resilient subtypes.

The problem with impulsiveness

In our study, overcontrollers scored high in all neuroticism facets except in impulsiveness, whereas the opposite was true for undercontrollers. When this facet was excluded from analysis, overcontrollers were more and undercontrollers less neurotic. Conceptually, impulsiveness clearly relates to undercontrol and not to overcontrol, and our data confirmed this expectation. Thus, from our person-centred perspective, inclusion of impulsiveness items in the neuroticism scale somewhat blurred the distinction between over- and undercontrollers.

There were problems with impulsiveness from a variable-centred perspective, too. Inspection of the factor loadings of the NEO-PI-R items in a varimax-rotated principal component analysis with the number of factors fixed to five revealed that every single impulsiveness item had higher loadings on either the Extraversion, the Openness, or the Conscientiousness factors than on the Neuroticism factor. Additional factor analyses of facet scores showed that the impulsiveness subscale loaded higher on Extraversion ($r = 0.55$) than on Neuroticism ($r = 0.36$). Ostendorf and Angleitner (2000) reported

comparable loadings for a larger sample of German speaking people ($N = 10\,748$). Therefore, the impulsiveness subscale may not represent an adequate facet of Neuroticism.

In sum, the clustering procedure seems to be quite sensitive to even slight deficiencies in both scale reliability and facet structure. Therefore, care should be taken to select traits that serve as an adequate basis for cluster analysis.

The problem of sample size

Perhaps the most serious problem for finding replicable subtypes is the smaller sample size for the replication analyses. Whereas the full sample is split into halves in within-study replications of prototypes, the prototypes themselves are further split into halves in within-study replications of subtypes. For example, the within-study replication of the NEO-PI-R prototypes was based on comparisons of 393 subjects whereas the within-study replication of the undercontrolled subtypes was based on comparisons of only 94 or 95 subjects. Samples four times as large as the present ones are needed to yield subtype replications that are based on the same number of data as the prototype replications in the present study. From this view point, it is not surprising that our search for replicable subtypes was not really successful.

Perhaps it is possible to find replicable subtypes in samples much larger than the present ones. In addition, alternative methods of replication that make better use of the data might also increase the chance to find replicable subtypes (e.g. bootstrapping procedures; see Barbaranello, this Special Issue). Finally, cross-study replications that do not rely on sample splits also increase this chance (see particularly the attempts of replicating our subtypes in Spanish samples by Boehme et al., this Special Issue).

The problem of adequate variables

A more fundamental problem for person-centred studies of personality types is that a few broad personality dimensions such as the Big Five may be not sufficient for distinguishing personality types that matter. Referring to our initial metaphor, objects with legs such as sheep and tables can be described on a few general-purpose dimensions such as number of legs, length, and height, but these general dimensions are insufficient to distinguish sheep from tables. If one is interested in sheep or tables, it is important not to confuse sheep with tables, or tables with sheep. Variables such as living versus nonliving or soft versus hard surface easily do the distinction.

Similarly, the general-purpose Big Five dimensions of personality description may be insufficient to carve out differences between personality types with sufficient precision. We do not want to imply that personality types are as distinct at the construct level as sheep and tables, but we do suggest that profiles based on the Big Five are not necessarily the optimal approach for distinguishing personality types. One obvious alternative is to study profiles based on Big Five facets rather than the Big Five themselves. We did this, using the NEO-PI-R facets, but the results were highly similar to those reported in this article, probably because the intraindividual differences between the Big Five are more marked than the intraindividual differences between facets of the same Big Five scale (see also Table 2).

Another alternative is to question the Big Five as the optimal starting point for person-centred analyses of personality differences. It should be noted however that methods tailored to the person-centred approach such as Q-factor analysis of Q-sorts have resulted

in similar replicable personality types as Big Five based clustering methods (Asendorpf et al., 2001; Robins et al., 1996).

CONCLUSION

This study replicated the result of several other studies that personality can be described in terms of three major prototypes: resilient, overcontrolled, and undercontrolled. However, the search for a more differentiated classification in terms of subtypes had only limited success. The resilient prototype could be reliably divided only into a well adjusted and an assertive subtype. Inspection of their Big Five profiles showed that these subtypes were only partially consistent with other studies. Clearly, 'the study of personality types is still in its infancy' (Robins et al., 1998, p. 136). Future studies may try to find replicable subtypes with Big Five profiles that are based on (i) larger samples, (ii) superior methods of within-study replicability such as bootstrapping, and (iii) even more reliable scales that (iv) exhibit a simpler factor structure, and explore alternative variables and methods for the description of personality differences.

REFERENCES

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., & 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.

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

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.

Caspi, A. (1998). Personality development across the life course. In W. Damon (Ed.), *Handbook of child psychology*, 5th ed. N. Eisenberg (Vol. Ed.), *Social, emotional, and personality development* (Vol. 3, pp. 311–388). New York: Wiley.

Costa, P. T., Jr., & McCrae, R. R. (1989). *The NEO PI/FFI manual supplement*. Odessa, FL: Psychological Assessment Resources.

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

Eysenck, H. J., & Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire*. London: Hodder and Stoughton.

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* (pp. 102–138). New York: Guilford.

Ostendorf, F., & Angleitner A. (2000). Personality differences between Austrians, Swiss, East and West Germans as measured by the Revised NEO Personality Inventory. In R. R. McCrae, & F. Ostendorf (Chairs), *Personality structure*. Symposium conducted at the 10th European Conference on Personality, Cracow, Poland.

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. (1998). The typological approach to studying personality. In R. B. Cairns, L. R. Bergman, & J. Kagan (Eds.), *Methods and models for studying the individual* (pp. 135–160). Beverly Hills, CA: Sage.

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.

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, USA, 1995.

Zuckermann, M. (1979). *Sensation seeking: Beyond the optimal level of arousal.* Hillsdale, NJ: Erlbaum.