RESEARCH ARTICLE

A Cross-Setting Study of Chimpanzee (*Pan troglodytes*) Personality Structure and Development: Zoological Parks and Yerkes National Primate Research Center

ALEXANDER WEISS[1]*, JAMES E. KING[2], AND WILLIAM D. HOPKINS[3]
[1]*Department of Psychology, The University of Edinburgh, Edinburgh, United Kingdom*
[2]*Department of Psychology, University of Arizona, Tucson, Arizona*
[3]*Department of Psychology, Berry College, Mount Berry, Georgia*

This study addressed whether personality ratings using a 43 adjective questionnaire based on the Five-Factor Model generalized from a sample of 202 zoo-housed chimpanzees to a sample of 175 chimpanzees housed in Yerkes National Primate Research Center. Mean interrater reliabilities of adjectival ratings were lower for the chimpanzees housed in Yerkes. In addition, rank order of the interrater reliabilities of items differed between settings. To compare factor structure, we first examined whether we could replicate the original six factor structure found in an earlier study of 100 zoo chimpanzees using principal factors analysis in the Yerkes sample and 102 new zoo chimpanzees. The dominance, extraversion, conscientiousness, and agreeableness factors were clearly replicated in the Yerkes sample and the 102 new zoo chimpanzees. The Neuroticism and Openness factors did not replicate in the Yerkes sample, but they also did not replicate in the new zoo chimpanzees. These findings suggest the need to sample more adjectives representing neuroticism and openness in future versions of the questionnaire. We next sought to determine whether factor structure, as determined by principal components analysis, remained invariant across the two settings. This analysis revealed dominance, extraversion, conscientiousness, and agreeableness factors in both settings and a high level of congruence between the zoo and Yerkes samples for these factors. Finally, we tested whether factor scores in the two samples were similarly related to age and sex. With the exception of differences in age effects for dominance and agreeableness, age, and sex effects were consistent across samples. These findings suggest that, whereas there may be differences in the ease with which ratings are made, personality structure, and development are largely consistent across widely differing settings. Am. J. Primatol. 69:1264–1277, 2007. © 2007 Wiley-Liss, Inc.

Key words: Chimpanzee; Five-factor model; chimpanzee; sex; age; laboratory; zoo

*Correspondence to: Alexander Weiss, Department of Psychology, The University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK. E-mail: alex.weiss@ed.ac.uk

Received 21 July 2006; revised 26 January 2007; revision accepted 29 January 2007

DOI 10.1002/ajp.20428
Published online 30 March 2007 in Wiley InterScience (www.interscience.wiley.com).

INTRODUCTION

The study of great ape personality had its historical antecedents in the writings of early researchers who frequently attributed personality traits to their subjects [Köhler, 1925; Yerkes, 1925, 1943; Yerkes & Learned, 1925; Yerkes & Yerkes, 1929]. Recently, to understand the structure of personality in different species including great apes, researchers have turned to judgments by knowledgeable informants about specific personality attributes. The validity of nonhuman primate personality ratings is shown by their heritability [Weiss et al., 2000] and relationships to other variables, such as behavior in rhesus macaques [Capitanio, 1999] and chimpanzees [Pederson et al., 2005], subjective well-being in chimpanzees [King & Landau, 2003], immune function in rhesus macaques [Maninger et al., 2003], and cortisol levels in chimpanzees [Anestis, 2005].

Human personality is most commonly conceptualized using the Five-Factor Model, [FFM; Digman, 1990; McCrae & Costa, 2003] which posits that five dimensions or domains—neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness—describe individual differences in human personality. The evolutionary bases of these domains is supported by their occurrence in over 50 cultures [McCrae et al., 2005]. There are also consistent sex differences across cultures, with females scoring higher in Neuroticism and Agreeableness than males [McCrae et al., 2005]. Although these domains are mostly stable in adulthood, there is evidence for modest age-related declines in neuroticism, extraversion, openness to experience, and increases in agreeableness and conscientiousness [McCrae et al., 2005; Roberts et al., 2006; Srivastava et al., 2003].

In their review of the animal personality literature, Gosling and John [1999] used the FFM framework to show that neuroticism, extraversion, agreeableness, and, to a lesser extent, Openness often emerged in studies of animal personality. They also noted that the personality factor structures of several species, including nonhuman primates include a dimension related to dominance that is not found in humans. Considering the importance of social rank in primate societies [Bernstein, 1981], the emergence of dominance-like personality dimensions in nonhuman primates is not surprising. Gosling and John [1999] noted that the only evidence of a factor similar to human conscientiousness was a factor reflecting predictability and low aggression identified in a study of zoo-housed chimpanzees that revealed analogues of all five human dimensions and a broad factor related to dominance [King & Figueredo, 1997].

These findings suggest that dominance, neuroticism, extraversion, agreeableness, and possibly openness are ancestral and were present in the ancestors of modern primates. In contrast, the absence of a dominance-like factor among humans and the presence of conscientiousness domains in humans and chimpanzees may be derived characteristics.

In a recent study of chimpanzees, King et al. [2005] found that the dominance, extraversion, agreeableness, and conscientiousness were present in zoos and a naturalistic sanctuary in the Republic of the Congo (Brazzaville), ruling out the possibility that the conscientiousness domain was an artifact of zoo settings. On the other hand, neuroticism and openness did not generalize to the more naturalistic setting. However, the fact that neuroticism and openness also did not generalize across two independent samples of zoo chimpanzees suggests that the lack of factor replicability probably resulted from there being only three markers of neuroticism and two markers of openness in the original version of the questionnaire.

A study on orangutan personality also suggested that conscientiousness is a recently evolved personality dimension [Weiss et al., 2006]. This study used a questionnaire that included the same items King and Figueredo [1997] used to rate chimpanzees, three additional Neuroticism, and two additional Openness items. Factor analysis revealed extraversion, dominance, neuroticism, and agreeableness domains, but no evidence for distinct Openness or conscientiousness factors. Instead, Weiss et al. [2006] found a factor they named ''Intellect'' that combined aspects of openness and conscientiousness.

In this study we wish to determine whether personality structure as well as sex and age differences generalize from zoo-housed chimpanzees to laboratory chimpanzees. This study addresses several questions. Do differences in the relationships between chimpanzees and their caregivers or environmental affordances lead to setting-based differences in the interrater reliabilities of items or factors? Moreover, does the laboratory environment lead to different personality dimensions or differences in sex and age effects? Two types of environmental effects may influence between-setting differences. First, chimpanzees in laboratory settings are housed in smaller groups and are more likely to come into direct contact with humans than zoo-housed chimpanzees. Second, zoos have a longer history of providing chimpanzees with environmental enrichment than laboratories.

METHODS

Subjects

Participating zoos in the United States were accredited by the Association of Zoos and Aquariums. The participating zoo in Australia was accredited by the Australasian Regional Association of Zoological Parks and Aquaria. Yerkes National Primate Research Center is fully accredited by the American Association for Accreditation of Laboratory Animal Care.

The total sample included 377 chimpanzees (*Pan troglodytes*) who were rated between March 6, 1993 and February 15, 2004. Subjects were between .8 and 57.6 years old ($M = 18.4$, $SD = 12.4$). The zoo sample included 78 males and 124 females housed in 17 zoos in the United States and 1 zoo in Australia. Ages ranged from .8 to 55.2 years ($M = 16.5$, $SD = 12.2$). Males ranged in age from .8 to 50.2 years ($M = 14.8$, $SD = 11.7$); females ranged in age from 2.0 to 55.2 years ($M = 17.6$, $SD = 12.5$). The zoo sample included 41 males ranging in age from 2.4 to 50.2 years ($M = 18.5$, $SD = 12.5$) and 59 females ranging in age from 2.8 to 55.2 years ($M = 19.0$, $SD = 11.6$) who were subjects in King and Figueredo's [1997] study. The remaining 37 males ranging in age from .8 to 45.5 years ($M = 10.7$, $SD = 9.1$) and 65 females ranging in age from 2.0 to 46.3 years ($M = 16.3$, $SD = 13.1$) were rated after the original study.

The zoo chimpanzees had varied rearing histories: 126 were captive-born, 48 were wild-born, and 28 had unknown origins. Three zoo chimpanzees were singly housed and the remaining chimpanzees were members of groups ranging from 4 to 25 individuals ($M = 9.2$, $Mdn = 8$, $SD = 7.7$).

The Yerkes sample contained 68 males and 107 females housed at Yerkes National Primate Research Center in Atlanta, GA. Ages ranged from 1.6 to 57.6 years ($M = 20.5$, $SD = 12.4$); males ranged in age from 1.6 to 40.5 years ($M = 17.2$, $SD = 9.4$) and females ranged in age from 1.9 to 57.6 years ($M = 22.7$, $SD = 13.6$).

Yerkes chimpanzees resided in a variety of housing conditions, including being singly housed, pair-housed, housed in smaller social groups of 3–7 individuals, and housed in larger social groups of 8–22 individuals ($M = 4.5$, *Mdn* = 3, *SD* = 3.1). All Yerkes chimpanzees had access to indoor–outdoor caging with visual and auditory access to chimpanzees in adjacent cages.

The Yerkes chimpanzees had a variety of rearing histories, including conspecific mother-reared ($n = 62$), human-reared ($n = 83$), and wild-caught ($n = 30$). Mother-reared chimpanzees were those reared by their biological mother for more than the first 30 days of life. Nursery-reared subjects were those brought to the Yerkes nursery before 31 days of life. The standard protocol for hand rearing chimpanzees has been described in detail elsewhere [Bard & Gardner, 1996]. Wild-caught chimpanzees were older animals that had been captured from the wild and imported to the United States before 1973.

Raters

The 202 zoo chimpanzees were rated by 90 zoo employees or volunteers of the ChimpanZoo program of the Jane Goodall Institute who had conducted behavioral observations as part of another study. The length of experience raters had with chimpanzees was available for 43 raters who rated 141 of the chimpanzees ($M = 5.9$ years, $SD = 4.2$). Raters completed rating forms for between 1 and 37 chimpanzees, each ($M = 8.7$ chimpanzees, $SD = 6.50$).

The 175 Yerkes chimpanzees were rated by 16 raters who were faculty, staff, and students familiar with the chimpanzees. Raters at Yerkes had, on average, less experience with the chimpanzees than those in zoos ($M = 4.1$ years, $SD = 4.1$). Two raters rated 75 of the chimpanzees and a single rater rated the remaining 100 chimpanzees. Raters completed rating forms for between 3 and 66 chimpanzees, each ($M = 15.6$, $SD = 18.7$).

Instrument

The Chimpanzee Personality Questionnaire [CPQ; King & Figueredo, 1997] instructs raters to base ratings on overall impressions and not estimated frequencies of particular behaviors. The CPQ also instructed raters to not discuss their ratings with other raters.

The CPQ contains 43 adjectives followed by one to three sentences that define the adjective within the context of chimpanzee behavior and are consistent with dictionary definitions. Clarifying sentences also served to reduce ambiguity, which some suggest is a potential problem when only adjectives are used [Briggs, 1992, p 258]. Forty-one of the adjectives were sampled from the 75 subscales based on a factor analysis of students' self-evaluations on 1,431 trait descriptive terms of the Big Five [Goldberg, 1990, Table 3]. Two additional items, *clumsy* and *autistic*, were added by King and Figueredo [1997]. Each adjective was rated on a seven-point Likert scale (1 = ''Displays either total absence or negligible amounts of the trait''; 7 = ''Displays extremely large amounts of the trait'').

Although most raters of both samples answered all of the questionnaire items, there were a few cases in which raters did not complete some of the questions. In all cases a score of 4 (''Displays about average amounts of the trait'') was substituted for the missing value. This approach is commonly used in human personality studies [Costa & McCrae, 1992].

RESULTS

Interrater Reliabilities of Adjectives

Two types of intraclass correlations were computed to assess interrater reliabilities of individual adjectives [Shrout & Fleiss, 1979]. The first, ICC(3,1), assessed the reliability of individual ratings and is comparable across studies with different numbers of raters. The second, ICC(3,*k*), assessed the reliability of mean ratings averaged across all raters.

To compute intraclass correlation coefficients for zoo chimpanzees, the mean squares for chimpanzee(zoo) and Rater × Chimpanzee(Zoo) were obtained using a GLM with Type III Sums of Squares [PROC GLM; SAS Institute, 1999]. Interrater reliabilities for individual ratings ranged from .14 (*sensitive*) to .64 (*active*) and had a mean of .35. Interrater reliabilities for mean ratings ranged from .38 (*sensitive*) to .88 (*active*) and had a mean of .66.

To compute intraclass correlation coefficients for Yerkes chimpanzees, the mean squares for chimpanzee and Rater × Chimpanzee were obtained using a GLM with Type III Sums of Squares [PROC GLM; SAS Institute, 1999]. One adjective in the Yerkes sample (*disorganized)* was not reliable as it had negative interrater reliability estimates. For the ratings that had positive interrater reliabilities, individual ratings ranged from .06 (*imitative*) to .44 (*submissive*) and had a mean of .26 indicating that interrater reliabilities for the Yerkes chimpanzees were slightly lower than for the zoo chimpanzees. Interrater reliabilities for the mean ratings of adjectives with positive reliabilities ranged from .11 (*imitative*) to .62 (*submissive*) and had a mean of .40.

After converting interrater reliabilities into Fisher's z' scores, we compared the reliabilities of individual ratings between settings to determine whether the rank orders of adjective reliabilities in the different subsamples were comparable. Rank orders of the interrater reliabilities were preserved between ratings of the original 100 chimpanzees and the 95 new zoo chimpanzees ($r = .87$, $P < .0001$). The correlation of interrater reliabilities of ratings made on the 75 Yerkes chimpanzees and the pooled sample of 195 zoo chimpanzees ($r = .41$, $P < .01$) was significantly lower ($P < .05$, McNemar, 1955, formula 45). Thus, rank orders of reliabilities differed between zoo ratings and those made at Yerkes. In other words, adjectival descriptors with relatively high or low interrater reliabilities among zoo chimpanzees did not preserve their position relative to other adjectival descriptors in the Yerkes data.

Data Reduction

We first examined whether the six factors derived in the original study of 100 zoo chimpanzees [King & Figueredo, 1997] generalized to the Yerkes chimpanzees. The original factor structure was derived via principal factors analysis (PFA), weighting adjectival ratings by number of raters, and varimax rotation [King & Figueredo, 1997]. We used this same procedure to extract six factors from the Yerkes ratings. We compared factor structures using a targeted orthogonal Procrustes rotation, which determines the degree of similarity between two factor structures by testing how congruent they are after rotating the axes of one factor structure as closely as possible to those of a specified target [McCrae et al., 1996]. Congruence coefficients greater than .90 indicate that loadings clearly generalized across samples, though some [Haven & ten Berge, 1977] suggest that congruence coefficients greater than .85 indicate replicability.

Rotating the factor structure of the Yerkes ratings to those of the zoo ratings[1] indicated an overall congruence of .90. The congruences for the first four factors,

[1]The congruence coefficients will be unaffected by which of the two factor structures is chosen as the target.

dominance (D_{CH}[2]), extraversion (E_{CH}), conscientiousness (C_{CH}), and agreeableness (A_{CH}) were .96, .95, .95, and .94, respectively, indicating that they were replicated in the Yerkes sample. The congruences for neuroticism (N_{CH}) and openness (O_{CH}) were .68 and .53, respectively, indicating that they were not replicated in the Yerkes sample. The congruences for all but six items (*dominant*, *imitative*, *excitable*, *disorganized*, *predictable*, and *sensitive*) were greater than or equal to .85.

One possible reason why N_{CH} and O_{CH} did not replicate is that the environment offered at Yerkes did not support behaviors that covary in a way consistent with these factors. We tested this possibility by seeing whether these factors are found in another sample of zoo chimpanzees. The replication of N_{CH} and O_{CH} in another sample of zoo chimpanzees would indicate that this difference reflects differences between zoo and Yerkes chimpanzees. If these factors do not replicate, we would conclude that other reasons, including shortcomings of the CPQ, were responsible.

We rotated the factor structure of the 102 zoo chimpanzees that were not part of the original study to the factor structure of the original 100 chimpanzees. The congruences were .96, .93, .92, .95, .53, and .66 for D_{CH}, E_{CH}, C_{CH}, A_{CH}, N_{CH}, and O_{CH}, respectively, ruling out the possibility that N_{CH} and O_{CH} did not generalize because of differences between Yerkes and zoos.

To determine whether a simpler data reduction approach would yield a more replicable structure, we used principal components analysis (PCA) with no weighting of observations. PCA differs from PFA in that it assumes that all variance is true-score variance and has been criticized on these grounds [see Nunnally & Bernstein, 1994, pp 522–525 for a detailed discussion]. We used parallel analysis [Horn, 1965; O'Connor, 2000] to determine the number of factors to extract. Parallel analysis uses successive iterations of randomly generated data sets to determine the sampling distribution of eigenvalues. If the eigenvalue of a factor derived from actual data exceeded the 95th percentile of the sampling distribution, it was accepted as significant.

Since the first four eigenvalues for zoo chimpanzee ratings (9.63, 7.98, 7.06, and 2.55) exceeded chance levels, we extracted four factors and subjected them to varimax rotation. The four factors (see first four columns in Table I) were recognizable and similar to the four zoo factors that were replicated in a naturalistic sanctuary [King et al., 2005].

The first factor was primarily characterized by two items clearly related to dominance: *dominant* and (not) *submissive*. Additional items that loaded on this factor included markers of low neuroticism such as (not) *fearful*; markers of low agreeableness such as *manipulative*, *defiant*, and (not) *gentle*; and one conscientiousness marker, *persistent*. We labeled this factor dominance (D_{CH}). Items with salient loadings on the second factor included markers of the temperamental aspects of extraversion such as *active* and (not) *lazy* and interpersonal aspects of extraversion such as *social* and (not) *solitary*. We therefore labeled this factor extraversion (E_{CH}). Adjectives with salient loadings on the third factor included conscientiousness markers such as (not) *disorganized*, *predictable*, (not) *reckless*, and (not) *clumsy*. Other adjectives that loaded on this factor are typically related to neuroticism such as (not) *excitable* and (not) *jealous*. Given that this factor was predominantly made up of conscientiousness markers, we named this factor conscientiousness (C_{CH}). The final factor was

[2]We have adopted the convention of labeling factors using the first letter of the factor and a subscript that indicates species (CH = chimpanzee, OR = orangutan, HU = human).

TABLE I. Structures of Zoo, Yerkes, and Pooled Samples Derived via Principal Components Analysis

	Zoo[a]				Yerkes[b]				Pooled[a]			
Item	D_{CH}	E_{CH}	C_{CH}[c]	A_{CH}	D_{CH}	E_{CH}	C_{CH}[c]	A_{CH}	D_{CH}[c]	E_{CH}	C_{CH}[c]	A_{CH}
Dominant	**.83**	−.24	.03	.13	**.83**	−.04	−.15	.02	**.81**	−.10	−.25	.08
Submissive	**−.81**	.09	−.21	.02	**−.84**	−.04	−.13	.08	**−.85**	−.05	.05	.05
Dependent	**−.79**	.31	−.21	−.06	**−.77**	.13	−.24	.11	**−.83**	.18	−.03	.03
Timid	**−.78**	−.16	−.27	−.06	**−.70**	−.33	−.36	.08	**−.78**	−.33	−.12	.00
Persistent	**.75**	.24	−.13	.11	**.62**	.26	−.28	.02	**.58**	.24	**−.40**	.09
Bullying	**.74**	−.08	−.33	−.27	**.74**	.05	−.33	−.24	**.63**	−.01	**−.52**	−.23
Independent	**.73**	−.20	.21	.13	**.73**	−.16	.14	.04	**.77**	−.14	.02	.08
Decisive	**.72**	−.01	.22	.26	**.66**	.00	.11	.16	**.69**	.06	−.02	.23
Aggressive	**.72**	−.09	−.36	−.34	**.67**	.04	**−.46**	−.30	**.57**	.00	**−.60**	−.27
Fearful	**−.71**	−.06	**−.47**	.04	**−.59**	.11	−.36	.10	**−.72**	.02	−.20	.11
Cautious	**−.70**	−.14	−.13	.25	**−.65**	−.06	.02	**.41**	**−.66**	−.13	.16	.33
Defiant	**.59**	.17	**−.43**	−.25	**.60**	.05	**−.49**	−.09	**.44**	.09	**−.63**	−.12
Stingy	**.59**	−.09	**−.40**	−.04	**.62**	−.10	−.38	−.14	**.49**	−.10	**−.54**	−.09
Gentle	**−.57**	.01	**.43**	**.51**	**−.50**	.06	.34	**.58**	**−.42**	.05	**.53**	**.52**
Manipulative	**.53**	.13	−.32	.17	**.55**	.06	−.34	.29	**.42**	.04	**−.45**	.23
Active	−.05	**.90**	−.10	−.17	.11	**.80**	−.20	−.22	−.08	**.85**	−.21	−.12
Playful	−.14	**.88**	−.10	−.09	.05	**.86**	.03	.03	−.13	**.88**	−.05	.04
Lazy	.09	**−.85**	−.03	.15	−.12	**−.82**	.05	.18	.04	**−.86**	.07	.08
Solitary	−.06	**−.78**	−.12	−.01	−.10	**−.78**	−.22	−.04	−.06	**−.79**	−.12	−.06
Sociable	.01	**.76**	.18	.32	.05	**.61**	.27	**.47**	.02	**.70**	.22	**.42**
Inquisitive	.21	**.76**	−.07	.22	.17	**.77**	.06	.16	.12	**.76**	−.07	.26
Imitative	−.16	**.74**	−.16	−.06	−.25	.33	−.34	.32	−.33	**.51**	−.21	.19
Friendly	−.26	**.66**	.26	**.47**	−.25	**.42**	.34	**.53**	−.21	**.54**	.36	**.50**
Depressed	−.09	**−.67**	−.37	−.12	−.24	**−.69**	**−.41**	−.05	−.21	**−.74**	−.31	−.12
Inventive	.31	**.60**	−.01	.30	.35	**.57**	−.07	.27	.24	**.59**	−.14	.35
Affectionate	−.10	**.58**	.27	**.50**	−.14	**.50**	.36	**.54**	−.09	**.56**	.35	**.53**
Unemotional	.08	**−.53**	.35	−.01	−.32	**−.51**	.12	.12	−.05	**−.56**	.27	−.01
Erratic	.06	−.06	**−.73**	−.27	.18	−.18	**−.76**	−.26	−.05	−.20	**−.76**	−.25
Impulsive	.23	.27	**−.71**	−.20	.34	.17	**−.65**	−.25	.09	.23	**−.75**	−.14
Excitable	−.06	−.05	**−.70**	−.16	.21	.19	**−.74**	−.15	−.09	.10	**−.73**	−.06
Stable	.26	.02	**.62**	.29	−.02	−.17	**.60**	**.51**	.24	−.06	**.60**	.38
Jealous	**.46**	.11	**−.58**	−.17	**.51**	.18	**−.43**	−.11	.33	.10	**−.62**	−.12
Disorganized	−.20	−.02	**−.57**	−.19	−.26	−.31	**−.56**	−.21	−.32	−.33	**−.47**	−.25
Predictable	−.13	−.31	**.56**	.27	.00	−.13	**.40**	**.40**	.09	−.21	**.48**	.27
Autistic	−.14	−.11	**−.53**	.16	−.10	−.23	**−.57**	.11	−.24	−.27	**−.50**	.14
Irritable	**.48**	−.38	**−.52**	−.19	.38	−.26	**−.64**	−.30	.28	−.35	**−.67**	−.23
Reckless	**.41**	.31	**−.48**	**−.44**	**.56**	.18	**−.46**	−.37	.34	.26	**−.61**	−.37
Clumsy	−.17	−.13	**−.46**	.05	−.27	−.36	**−.55**	.08	−.30	**−.40**	**−.41**	−.01
Sympathetic	−.14	.03	.14	**.86**	−.30	.15	.25	**.71**	−.17	.09	.30	**.78**
Sensitive	.03	.02	.08	**.79**	−.25	.29	.19	**.55**	−.08	.12	.21	**.69**
Helpful	−.12	.31	.17	**.75**	−.06	.20	.03	**.72**	−.09	.25	.14	**.74**
Protective	.37	−.06	.14	**.70**	.10	.07	−.09	**.78**	.22	.01	.02	**.77**
Intelligent	**.42**	.11	.32	**.53**	.39	.33	.21	**.41**	**.41**	.30	.16	**.51**

Note: Salient loadings are in boldface.
[a]Loadings after varimax rotation of factor loadings.
[b]Loadings after orthogonal targeted Procrustes rotation to factor structure of 202 zoo chimpanzees.
[c]Loadings have been reflected.

primarily defined by adjectives such as *sympathetic* and *sensitive*, markers of the positive pole of agreeableness. It was also defined by *intelligent*, a marker of openness. We labeled the factor agreeableness (A_{CH}).

Parallel analysis revealed that the first four eigenvalues derived from PCA of the Yerkes sample (11.01, 8.11, 4.21, and 2.73) were significant. We extracted four factors and used a targeted orthogonal Procrustes rotation [McCrae et al., 1996] to compare the factor structure of the Yerkes sample to that of the zoo sample (see second four columns in Table I). Rotating the factor structure of the Yerkes sample ratings to those of the zoo sample ratings indicated a high degree of overall congruence (.95). The congruences for D_{CH}, E_{CH}, C_{CH}, and A_{CH} were .96, .93, .96, and .94, respectively, indicating that the factors replicated across samples. In addition, congruence coefficients for the loadings of all but three items (*imitative*, *unemotional*, and *sensitive*) exceeded .85.

Based on these findings, we pooled the zoo and Yerkes samples and extracted four factors (see third four columns in Table I). We then created unit-weighted factor scores based on the definitions of these four factors; items with loadings greater than or equal to .40 were given a weight of +1, items with loadings less than or equal to −.40 were given a weight of −1, and remaining items were given a weight of 0. When comparing samples, unit-weighting is desirable because factor scores that are weighted by exact loadings may differ substantially from sample to sample [Gorsuch, 1983, p 269]. For ease of interpretation, unit-weighted factor scores were converted into *T*-scores, which have a mean of 50 and standard deviation of 10.

Interrater Reliabilities and Internal Consistencies of Factor Scores

Table II shows the interrater reliabilities of individual and mean ratings and internal consistencies of the four factors derived from the zoo and the Yerkes samples. Because error variance is reduced in aggregate scores, the interrater reliabilities of ratings for factors, which were defined by multiple items, will necessarily be higher than reliabilities of individual items. Interrater reliabilities of single ratings, *ICC*(3,1), showed that ratings for D_{CH}, E_{CH}, and A_{CH}, but not C_{CH}, were slightly more consistent across raters in zoos than in Yerkes. Also, in both settings, the reliabilities of D_{CH} and E_{CH} were higher than those of C_{CH} and A_{CH}.

Mean Difference for Across Settings

We first tested mean differences across settings using a general linear model with Type III Sums of Squares. The Yerkes and zoo chimpanzees did not differ in

TABLE II. Interrater Reliabilities and Internal Consistencies of Zoo and Yerkes Chimpanzees

	Zoo			Yerkes		
Factor	*ICC*(3,1)[a]	*ICC*(3,*k*)[b]	α	*ICC*(3,1)[a]	*ICC*(3,*k*)[b]	α
D_{CH}	.63	.87	.93	.48	.65	.91
E_{CH}	.65	.88	.93	.38	.55	.89
C_{CH}	.46	.77	.89	.38	.55	.90
A_{CH}	.38	.71	.84	.22	.36	.77

[a]Reliability of individual ratings [Shrout and Fleiss, 1979].
[b]Reliability of mean ratings [Shrout and Fleiss, 1979].

D_{CH} ($d = .14$, $F(1,375) = 1.74$, $P > .05$), but Yerkes chimpanzees scored lower than zoo chimpanzees on E_{CH} ($d = .76$, $F(1,375) = 51.99$, $P < .0001$), C_{CH} ($d = .40$, $F(1,375) = 15.23$, $P < .0001$), and A_{CH} ($d = .63$, $F(1,375) = 37.35$, $P < .0001$). We ran a second analysis in which we compared settings after controlling for sex, age, and the Age × Sex interaction; the results did not differ.

Sex Differences Across Settings

To test whether sex differences differed across settings, we conducted a general linear model with Type III Sums of Squares in which each of the four personality *T*-scores were predicted by sex, setting, and the Sex × Setting interaction. A significant interaction effect would indicate that mean level sex differences did not generalize across settings.

Male chimpanzees were higher than females in D_{CH} ($d = .47$, $F(1,373) = 20.00$, $P < .0001$) and E_{CH} ($d = .21$, $F(1,373) = 4.94$, $P < .05$), but were lower than females in C_{CH} ($d = .42$, $F(1,373) = 17.37$, $P < .0001$) and A_{CH} ($d = .37$, $F(1,373) = 13.02$, $P < .001$). The Sex × Setting interaction was not a significant predictor of any of the four personality domains (all $P > .05$), indicating similar sex differences in zoo and Yerkes chimpanzees.

Age-Related Differences Across Settings

To compare age-related differences across settings for the four replicable personality factors, we used a series of general linear models with Type I Sums of Squares. In the first step, we determined whether there was a linear age effect after controlling for setting and sex. Because the personality factors were converted into *T*-scores, the linear regression coefficient (b) indicates the change in standard deviation units that can be expected over 10 years time. We then tested whether including quadratic and cubic age effects, respectively, improved model fit after controlling for setting, sex, and all lower level age effects. We then tested whether interactions of setting with any of the significant age effects explained additional variance. If any interaction between age effects and setting were significant, this indicated that the age effects did not generalize across samples.

D_{CH} increased linearly as a function of age ($F(1,373) = 15.56$, $P < .0001$), but the size of the effect was modest ($b = .16$). Sequential inclusion of quadratic ($\Delta R^2 = .05$, $F(1,372) = 23.08$, $P < .0001$) improved model fit over a model that only included setting, sex, and linear age effects, but including cubic age effects ($\Delta R^2 = .01$, $F(1,371) = 2.72$, $P > .05$) did not. The Age × Setting interaction for D_{CH} explained a significant additional proportion of variance ($\Delta R^2 = .01$, $F(1,371) = 5.30$, $P < .05$), indicating that the linear component of age effects did not generalize across settings. Including the Age^2 × Setting effect in the model did not explain a significant amount of additional variance ($\Delta R^2 = .01$, $F(1,370) = 3.60$, $P > .05$), indicating that the quadratic component of age differences generalized across settings (see Fig. 1a).

E_{CH} decreased as a function of age ($F(1,373) = 186.80$, $P < .0001$) and the effect size ($b = -.44$) was larger than the relationship between age and D_{CH}. Adding terms for the quadratic ($\Delta R^2 = .06$, $F(1,372) = 46.89$, $P < .0001$) and cubic ($\Delta R^2 = .01$, $F(1,371) = 8.98$, $P < .01$) effects of age significantly improved model fit. Neither age × setting ($\Delta R^2 < .01$, $F(1,370) = 1.04$, $P > .05$), Age^2 × Setting ($\Delta R^2 < .01$, $F(1,369) = 1.00$, $P > .05$), nor Age^3 × Setting ($\Delta R^2 < .01$, $F(1,368) = 1.46$, $P > .05$) increased the proportion of variance accounted for, suggesting that age differences in E_{CH} generalized across settings (see Fig. 1b).

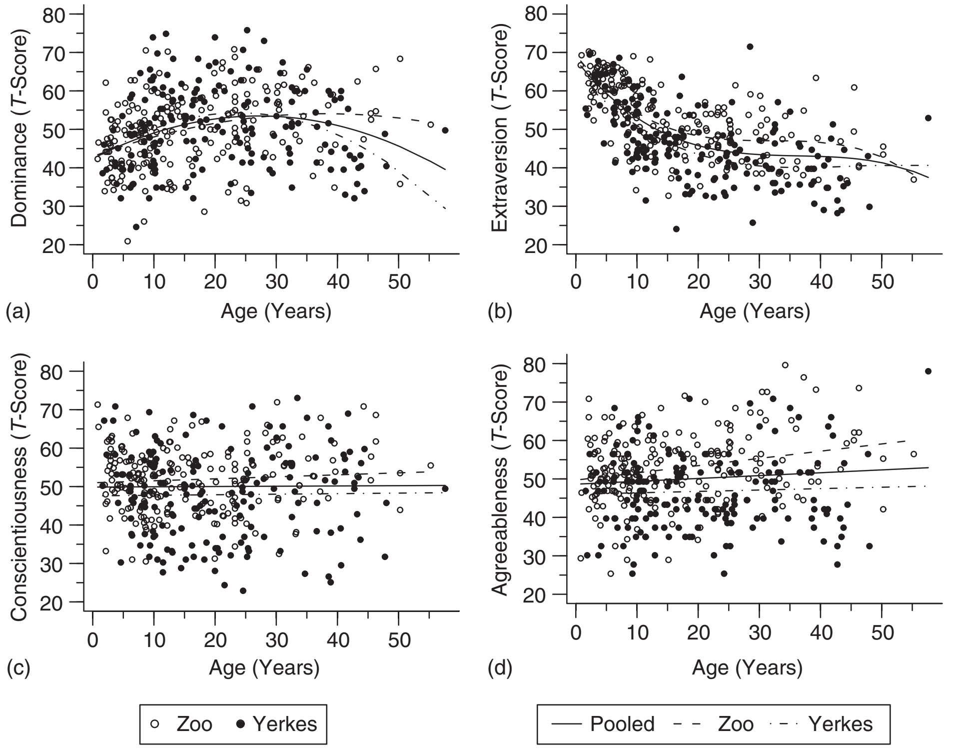

Fig. 1. Predicted personality domain *T*-scores as a function of age for the pooled, zoo, and Yerkes samples.

Age was a negligible ($b = .01$) and nonsignificant ($F(1,373) = .03$, $P > .05$) predictor of C_{CH}. The Age × Setting interaction was also not significant ($F(1,372) = .45$, $P > .05$), indicating that these age effects generalized across settings (see Fig. 1c).

There was a positive but small ($b = .10$) association between age and A_{CH} ($F(1,373) = 5.83$, $P < .05$). Including a quadratic term for age did not significantly model fit ($\Delta R^2 < .01$, $F(1,372) = .23$, $P > .05$). The Age × Setting interaction was a significant predictor of A_{CH} ($\Delta R^2 = .01$, $F(1,372) = 4.47$, $P < .05$) indicating that age effects differed between settings (see Fig. 1d).

DISCUSSION

We found that four chimpanzee personality factors, D_{CH}, E_{CH}, C_{CH}, and A_{CH}, generalized from zoos to a laboratory setting. Earlier, King et al. [2005] found that these factors generalized from zoo settings to that of a large, naturalistic sanctuary. This finding is also largely consistent with cross-cultural findings in human personality research showing that factor structure is largely invariant across multiple differing cultures [McCrae et al., 2005].

The fact that N_{CH} and O_{CH} did not generalize to chimpanzees living in a naturalistic sanctuary [King et al., 2005] or a laboratory does not indicate the non-existence of these factors in chimpanzees. Instead, given that these dimensions are found across a wide range of species [Gosling, 2001] but were also absent in another zoo sample, it is more likely that this failure to generalize reflects a problem with the CPQ. In a future study, it would be useful to reassess these chimpanzees on a more recent version of the CPQ that includes more markers of these domains.

There were other similarities across the two settings. Namely, the interrater reliabilities for D_{CH} and E_{CH} were higher than those for A_{CH} and C_{CH}. This is broadly consistent with findings in humans and nonhuman animals [see Gosling, 2001 for a review], indicating, perhaps, that A_{CH} and C_{CH} are manifested less visibly than D_{CH} and E_{CH}. Tests for sex differences revealed that, in both settings, males were higher in D_{CH} and E_{CH}, but lower in A_{CH} and C_{CH} than females. Similar sex differences have been found in humans [McCrae et al., 2005].

There were linear and quadratic age effects for D_{CH}. In zoo and Yerkes chimpanzees, age-related differences suggested that D_{CH} increased until approximately age 30 after which it declined, with this decline being more rapid in the Yerkes chimpanzees. There were linear, quadratic, and cubic age effects for and E_{CH}. These age-related differences in E_{CH} suggested that in the zoo and Yerkes sample there was a steep E_{CH} decline until approximately 20 years of age followed by a second, less steep decline until approximately 45 years of age, and a third, sharper decline. There were no age differences in C_{CH} in either sample. Linear age effects indicated that older chimpanzees in zoos and Yerkes had higher A_{CH} scores; this effect was greater among zoo chimpanzees.

A possible explanation for differences in age effects on D_{CH} and A_{CH} is that smaller group sizes are less effective at maintaining or supporting behaviors related to these domains. Future studies may wish to investigate this possibility by examining age differences and changes in D_{CH}, A_{CH}, or related behaviors in groups that differ in the number of members.

The direction of the linear age effects on E_{CH} and A_{CH} mirrored human findings, though the age effects on E_{CH} were larger. The lack of age differences in C_{CH} did not resemble human findings [McCrae et al., 2005; Roberts et al., 2006; Srivastava et al., 2003]. These findings suggest that, with the exception of age-related differences in D_{CH} and A_{CH}, the influences of age and sex on the four robust personality domains were mostly invariant across settings. This invariance is strong evidence that the processes underlying sex and age effects were probably biologically based as they were independent of rearing and present housing conditions.

In addition to the different age effects on D_{CH} and A_{CH}, there were other differences between settings. The mean interrater reliabilities for the adjectives and the interrater reliabilities of D_{CH}, E_{CH}, and A_{CH} were higher in zoos than in Yerkes. There are at least four possible reasons for why the reliability of ratings may differ between settings.

First, personality in adulthood is more stable over time than personality in adolescence [Roberts & DelVecchio, 2000; Terracciano et al., 2006]. As such, reliabilities of ratings might differ between two settings because of age differences of the subjects in these settings. However, since the mean ages of chimpanzees at zoos and Yerkes were comparable, we can rule out this possibility.

Second, personality ratings of nonhuman primates are probably more reliably rated by individuals who are familiar with the subjects [see, e.g., Martau et al., 1985]. Although the raters at Yerkes were less familiar with the chimpanzees than those in zoos, raters in both settings worked with the subjects for a longer period of time than did raters in Martau et al.'s study. In addition, comparing reliabilities of items between two zoo samples and the zoo and Yerkes sample indicated that the reliabilities also differed in terms of the rank order of the reliabilities of individual adjectives. These facts lead us to reject the possibility that degree of familiarity resulted in interrater reliability differences between the zoo and Yerkes ratings.

Two remaining possibilities are that the difference in reliabilities resulted from differences in the nature of the relationship between chimpanzees and the raters or some aspect of the environment, such as the size of social groups or degree of enrichment [Podberscek & Gosling, 2000]. In both cases, interrater reliabilities would differ as some behaviors may be more clearly expressed by chimpanzees or easily observed in some settings than in others. While we cannot distinguish between these two possibilities without further research, both are consistent with the finding that the rank order of the reliabilities of adjectives differed between settings.

There were differences in mean personality factor scores between the zoo and the Yerkes chimpanzees. Yerkes chimpanzees were rated as being lower in E_{CH}, C_{CH}, and A_{CH}. Although these differences may be attributed to environmental differences, it is also possible that these differences reflected the quality of the rater-chimpanzee relationships between settings or founder effects.

Taxonomy was a crucial first step in the biological sciences. Similarly, the development of human or animal personality research requires identifying a robust taxonomy of personality domains. In human personality research, the well-known FFM [Digman, 1990; McCrae & Costa, 2003] has been widely accepted as accurately describing human personality. A critical finding that led to the wide adoption of the FFM for research in a variety of fields was its constancy across different cultures [McCrae & Terracciano, 2005]. There is strong, direct evidence that four chimpanzee personality factors, D_{CH}, E_{CH}, A_{CH}, and C_{CH}, generalize across settings ranging from zoos to natural forest to laboratory and indirect evidence that a more comprehensive measure would also lead to generalizable N_{CH} and O_{CH} dimensions.

Much like the need for ratings to be consistent across raters, evidence that personality structure generalizes across settings is as critical for research on nonhuman animals as it is for humans. A chimpanzee personality structure that varied as a function of setting would have only limited usefulness as a focus of research that sought to elucidate the relationship between chimpanzee personality and behavior or other outcomes such as health and life-history variables. The consistency of personality within species is also crucial as it indicates that personality structure is a consequence of evolutionary adaptation to the physical and social environment in which the species evolved and not a result of present environmental circumstances [Sih et al., 2004]. As such it is a good foundation for comparative primate personality studies that could yield clues to the selective pressures that shaped differences in the characters of apes and men.

ACKNOWLEDGMENTS

This research was supported by NIH grants NS-36605, NS-42867, and RR 00165 to the Yerkes National Primate Research Center. The Yerkes Center is fully accredited by the American Association for Accreditation of Laboratory Animal Care. American Psychological Association guidelines for the ethical treatment of animals were adhered to during all aspects of this study.

We thank the zoos, Yerkes National Primate Center, and the raters at these institutions for completing the questionnaires. We also thank Mike Allerhand for his assistance in producing the figures and Melissa S. Gerald for her helpful comments and suggestions.

REFERENCES

Anestis SF. 2005. Behavioral style, dominance rank, and urinary cortisol in young chimpanzees (*Pan troglodytes*). Behaviour 142: 1245–1268.

Bard KA, Gardner KH. 1996. Influences on development in infant chimpanzees: enculturation, temperament, and cognition. In: Russon AE, Bard KA, editors. Reaching into thought: the minds of the great apes. Cambridge, England: Cambridge University Press. p 235–256.

Bernstein IS. 1981. Dominance relationships and ranks—explanations, correlations and empirical challenges. Behav Brain Sci 4: 449–453.

Briggs SR. 1992. Assessing the five-factor model of personality description. J Pers 60: 253–293.

Capitanio JP. 1999. Personality dimensions in adult male rhesus macaques: prediction of behaviors across time and situation. Am J Primatol 47:299–320.

Costa PT Jr, McCrae RR. 1992. Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI) professional manual. Odessa, FL: Psychological Assessment Resources.

Digman JM. 1990. Personality structure: emergence of the five-factor model. Annu Rev Psychol 41:417–440.

Goldberg LR. 1990. An alternative ''description of personality'': the big-five factor structure. J Pers Soc Psychol 59:1216–1229.

Gorsuch RL. 1983. Factor analysis (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gosling SD. 2001. From mice to men: what can we learn about personality from animal research? Psychol Bull 127:45–86.

Gosling SD, John OP. 1999. Personality dimensions in nonhuman animals: a cross-species review. Curr Dir Psychol Sci 8:69–75.

Haven S, ten Berge JMF. 1977. Tucker's coefficient of congruence as a measure of factorial invariance: an empirical study (Heymans Bulletin 290 EX). University of Groningen.

Horn JL. 1965. A rationale and test for the number of factors in factor analysis. Psychometrika 30:179–185.

King JE, Figueredo AJ. 1997. The five-factor model plus dominance in chimpanzee personality. J Res Pers 31:257–271.

King JE, Landau VI. 2003. Can chimpanzee (*Pan troglodytes*) happiness be estimated by human raters? J Res Pers 37:1–15.

King JE, Weiss A, Farmer KH. 2005. A chimpanzee (*Pan troglodytes*) analogue of cross-national generalization of personality structure: zoological parks and an African sanctuary. J Pers 73:389–410.

Köhler W. 1925. The mentality of apes. New York, NY: Harcourt, Brace and Company.

Maninger N, Capitanio JP, Mendoza SP, Mason WA. 2003. Personality influences tetanus-specific antibody response in adult male rhesus Macaques after removal from natal group and housing relocation. Am J Primatol 61:73–83.

Martau PA, Caine NG, Candland DG. 1985. Reliability of the Emotions Profile Index, primate form, with *Papio hamadryas*, *Macaca fuscata*, and two *Saimiri* species. Primates 26:501–505.

McCrae RR, Costa PT Jr. 2003. Personality in adulthood: a five-factor theory perspective. New York, NY: Guilford Press.

McCrae RR, Terracciano A, et al. 2005. Universal features of personality traits from the observer's perspective: data from 50 cultures. J Pers Soc Psychol 88: 547–561.

McCrae RR, Zonderman AB, Bond MH, Costa PT Jr, Paunonen SV. 1996. Evaluating replicability of factors in the revised NEO personality inventory: Confirmatory factor analysis versus Procrustes rotation. J Pers Soc Psychol 70:552–566.

McNemar Q. 1955. Psychological statistics (2nd ed.). New York: Wiley.

Nunnally, JC, Bernstein IH. 1994. Psychometric theory (3rd ed.). New York: McGraw-Hill.

O'Connor BP. 2000. SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. Behav Res Meth Ins C 32:396–402.

Pederson AK, King JE, Landau VI. 2005. Chimpanzee (*Pan troglodytes*) personality predicts behavior. J Res Pers 39:534–549.

Podberscek AL, Gosling SD. 2000. Personality research on pets and their owners: conceptual issues and review. In: Podberscek AL, Serpell JA, editors. Companion animals and us. Cambridge, England: Cambridge University Press. p 143–167.

Roberts BW, DelVecchio WF. 2000. The rank-order consistency of personality from childhood to old age: a quantitative review of longitudinal studies. Psychol Bull 126: 3–25.

Roberts BW, Walton KE, Viechtbauer W. 2006. Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. Psychol Bull 132:1–25.

SAS Institute. 1999. SAS/STAT user's guide, version 8. Cary, NC: Author.

Shrout PE, Fleiss JL. 1979. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 86:420–428.

Sih A, Bell AM, Johnson JC, Ziemba RE. 2004. Behavioral syndromes: an integrative overview. Q Rev Biol 79:241–277.

Srivastava S, John OP, Gosling SD, Potter J. 2003. Development of personality in early and middle adulthood: Set like plaster or persistent change? J Pers Soc Psychol 84: 1041–1053.

Terracciano A, McCrae RR, Costa PT Jr. 2006. Personality plasticity after age 30. Pers Soc Psychol Bull 32: 999–1009.

Weiss A, King JE, Figueredo AJ. 2000. The heritability of personality factors in chimpanzees (*Pan troglodytes*). Behav Genet 30: 213–221.

Weiss A, King JE, Perkins L. 2006. Personality and subjective well-being in orangutans (*Pongo pygmaeus* and *Pongo abelii*). J Pers Soc Psychol 90:501–511.

Yerkes RM. 1925. Almost human. New York, NY: The Century Company.

Yerkes RM. 1943. Chimpanzees: a laboratory colony. New Haven, CT: Yale University Press.

Yerkes RM, Learned BW. 1925. Chimpanzee intelligence and its vocal expression. Baltimore, MD: Williams and Wilkins.

Yerkes RM, Yerkes AW. 1929. The great apes: a study of anthropoid life. New Haven, CT: Yale University Press.