Journal of Comparative Psychology

Integrative Personality Assessment in Wild Assamese Macaques (Macaca assamensis)

Anja Ebenau, Christoph von Borell, Lars Penke, Julia Ostner, and Oliver Schülke

Online First Publication, July 18, 2019. http://dx.doi.org/10.1037/com0000190

American Psychological Association

Journal of Comparative Psychology

0735-7036/19/$12.00

http://dx.doi.org/10.1037/com0000190

Integrative Personality Assessment in Wild Assamese Macaques (*Macaca assamensis*)

Anja Ebenau
University of Goettingen; Leibniz Institute for Primate Research, Goettingen, Germany; and Leibniz ScienceCampus Primate Cognition, Goettingen, Germany

Christoph von Borell and Lars Penke
Leibniz ScienceCampus Primate Cognition, Goettingen, Germany, and University of Goettingen

Julia Ostner and Oliver Schülke
University of Goettingen; Leibniz Institute for Primate Research, Goettingen, Germany; and Leibniz ScienceCampus Primate Cognition, Goettingen, Germany

In nonhuman animals, individuals of the same sex and age differ in their behavior patterns consistently across time, comparable with human personality differences. To draw conclusions about the adaptive value of behavior traits, it is essential to study them in the wild where animals are subject to the ecological pressures that promoted the evolution of behavior strategies. This study was conducted in the Phu Khieo Wildlife Sanctuary, Thailand, on 4 groups of habituated wild Assamese macaques by observers who had familiarized themselves with the subjects over the course of an ongoing long-term study. We used a multimethod approach enabling the most comprehensive understanding of variation in stable interindividual differences in a species-typical ecological setting. We combined trait ratings (TRs), assessed with observer-report questionnaires (54-item Hominoid Personality Questionnaire) of 107 individuals of diverse age-sex classes, with behavior codings (BCs) of 24 adult males. We found male and female personality constructs to be congruent and examined reliability and construct validity. Combining trait rating and behavioral coding, we found two solutions with five factors to best describe the personality structure of the males: one structure comprised the dimensions $\text{Gregariousness}_{BC}$, $\text{Aggressiveness}_{BC}$, Sociability_{BC} and Vigilance_{BC}, complemented by a Confidence_{TR} domain and the other structure Opportunism_{TR}, Confidence_{TR}, Friendliness_{TR}, Activity_{TR} complemented with Vigiance_{BC}. We discuss our findings with regard to the importance of construct validity and reproducibility in the context of method development and standardization in nonhuman animal personality research.

Keywords: personality, wild macaques, behavior coding, trait rating, integrative assessment

Supplemental materials: http://dx.doi.org/10.1037/com0000190.supp

"Personality in the broadest sense is the internal organization of behavior that is stable over considerable time periods in the individual yet varies among the individuals of a population on latent dimensions" (Uher, 2008a, p. 476). Personality has a moderately heritable component and is systematically associated with differences in fitness parameters, such as survival, reproductive success,

Anja Ebenau, Department of Behavioral Ecology, Johann-Friedrich-Blumenbach Institute for Zoology and Anthropology, University of Goettingen; Research Group Social Evolution in Primates, German Primate Center, Leibniz Institute for Primate Research, Goettingen, Germany; and Leibniz ScienceCampus Primate Cognition, Goettingen, Germany. Christoph von Borell and Lars Penke, Leibniz ScienceCampus Primate Cognition, Goettingen, Germany, and Biological Personality Psychology, Georg Elias Mueller Institute of Psychology, University of Goettingen; Julia Ostner and Oliver Schülke, Department of Behavioral Ecology, Johann-Friedrich-Blumenbach Institute for Zoology and Anthropology, University of Goettingen; Research Group Social Evolution in Primates, German Primate Center, Leibniz Institute for Primate Research, Goettingen, Germany; and Leibniz ScienceCampus Primate Cognition, Goettingen, Germany.

We thank the National Research Council of Thailand (NRCT) and the Department of National Parks, Wildlife and Plant Conservation (DNP) for permission to conduct this study and for all the support granted (permit 0002/2424). We are thankful to the three anonymous reviewers for their constructive and helpful comments on the manuscript. We are grateful to I. Chatchawarn, J. Prabnasuk, K. Nitaya T. Wongsnak, M. Pongjantarasatien, K. Kreetiyutanont, M. Kumsuk and W. Saenphala (PKWS) for their cooperation over the years and permission to carry out this study. We thank A. Koenig and C. Borries, who developed the field site. Our special thanks goes to S. Jumrudwong, W. Nueorngshiyos, N. Juntuch, J. Wanart, R. Intalo, T. Kilawit, N. Pongangan, B. Klaewklar, N. Bualeng, D. Gutleb, P. Saisawatdikul, K. Srithorn, M. Swagemakers and T. Wisate for their excellent help in the field. This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project 254142454/GRK 2070.

Correspondence concerning this article should be addressed to Oliver Schülke, Department of Behavioral Ecology, Johann-Friedrich-Blumenbach Institute for Zoology and Anthropology, University of Goettingen, Kellnerweg 6, 37077 Goettingen, Germany. E-mail: oschuel@gwdg.de

and health in both animals and humans (Deary, Weiss, & Batty, 2010; Penke & Jokela, 2016; Wolf & Weissing, 2012). Animal personality became an expanding topic of behavioral ecology and comparative psychology in the last 2 decades (Carere & Maestripieri, 2013; Dingemanse, Kazem, Réale, & Wright, 2010; Mehta & Gosling, 2008; Uher, 2008b). Yet, there are still no consistent definitions for personality dimensions, or standardized assessment methods, and differences in species-specific behavior repertoires make cross-species comparisons challenging (Gosling, 2001; Lilley, Kuczaj, & Yeater, 2017).

The two most common assessment methods in animal personality research are trait ratings (TRs) and behavior coding (BC) from naturalistic observations or experimental data (Réale, Reader, Sol, McDougall, & Dingemanse, 2007; Vazire, Gosling, Dickey, & Schapiro, 2007). In human psychology, personality is mainly assessed via self-reported statement ratings on personality inventories (Koski, 2011a; Uher, 2011) and expressed in broad bipolar or monopolar dimensions, often referred to as the so-called "Big Five": Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience (Digman, 1990; John & Srivastava, 1999). In a top-down approach, personality questionnaires have been modified for animals to allow TRs by human observers familiar with the individuals (e.g., Hominoid Personality Questionnaire [HPQ]; Freeman & Gosling, 2010; Weiss et al., 2009). To avoid issues with validity due to possible anthropomorphic projections in TRs (Freeman et al., 2013; Uher, 2018), naturalistic observational approaches use previously established ethograms defining behaviors and situations at the species level to quantify single behaviors or behaviors in a specific situation (Massen, Antonides, Arnold, Bionda, & Koski, 2013; Uher, 2015). Experimental observational approaches create situations to tap into one (Dammhahn, 2012) or several personality dimensions at a time (Carter, Marshall, Heinsohn, & Cowlishaw, 2012; Koski & Burkart, 2015; Massen et al., 2013; Réale et al., 2007; Vazire et al., 2007).

The choice of method typically depends on time and feasibility. TRs by familiar observers are less time-consuming and can cover a broader range of individual traits compared with BCs in experimental or natural settings. However, TR require that several raters are familiar with the individuals because different observers know an individual from different situations and may have unique observer biases (Freeman, Gosling, & Schapiro, 2011). Experimentation is better able to assess nonsocial personality dimensions, such as boldness and exploration, because relevant situations occur rarely and unpredictably in naturalistic observations (Massen et al., 2013), but experimentation is not always feasible, especially in wild populations (Neumann, Agil, Widdig, & Engelhardt, 2013; Tkaczynski et al., 2018). When several methods have been applied to assess personality of the same subjects, convergent validity across methods was often low (Freeman et al., 2011; John & Soto, 2007; Uher & Visalberghi, 2016).

Such disparity may result from the disadvantages of BCs, which may struggle with reliably detecting individual differences in behaviors that occur at low frequencies (Freeman et al., 2013) and behaviors that change from day to day, across seasons, or with changing social and physical environment (Brommer & Class, 2017; Uher, 2011). Aggregation may overcome these shortcomings in behavior measurements, as a higher aggregation level inevitably leads to higher reliability scores (for discussion about the *principle of aggregation*, see Uher, 2011). In TRs, however, variability in behavior due to random variance is partly cancelled out because observers implicitly aggregate an animal's behavior over time when they form an image of a subject's personality and subject it to their own memory. This aggregation comes at a cost though because such images may be biased toward specific, more memorable events, and the assessment of subjects further may be influenced by discussions among raters and may be modified through recalled memories (which are reshaped every time they are recalled). The advantage of BC studies is that behaviors are directly perceivable and measurable. Hence, they are not susceptible to most of the biases that can influence TRs (Freeman et al., 2011; van Aken & Asendorpf, 2018). The abovementioned problems can be overcome by repeated experimentation and long-term observations when even rare behaviors occur often enough to reliably assess individual variation and all individuals are assessed across a variety of social and ecological contexts.

The strengths and weaknesses of TR and BC have been repeatedly discussed (Freeman et al., 2011; Koski, 2014; Uher, 2018; Uher, Addessi, & Visalberghi, 2013; Uher & Visalberghi, 2016; Weiss, 2017). It has been argued that the most comprehensive understanding of variation in stable interindividual differences may come from studies combining approaches to compensate for the weaknesses that each technique individually has (Koski, 2014). Multimethod approaches in field and lab work use item selection based on reliability measures (e.g., test–retest) and analyses of construct validity (Eckardt et al., 2015; Garai, Weiss, Arnaud, & Furuichi, 2016) and may generate integrative/complemented personality structures to provide a more complete picture (e.g., BC and experiments: Massen & Koski, 2014; Neumann et al., 2013; Uher, Addessi, et al., 2013; BC and TRs: Iwanicki & Lehmann, 2015; Uher, Werner, & Gosselt, 2013; all three methods: Tkaczynski et al., 2018). As can be expected from the differences in strength and weaknesses of the methods, personality structures derived from TRs and BCs do not always clearly correspond to each other. Aggressiveness-related trait-rating dimensions are mainly well-supported with aggressive behavior (Pritchard, Sheeran, Gabriel, Li, & Wagner, 2014; Uher, Werner, et al., 2013; Vazire et al., 2007), yet aggressive behavior may also correlate with a Dominance dimension (Freeman et al., 2013; Uher, Werner, et al., 2013; Uher & Visalberghi, 2016) or Excitability (Tkaczynski et al., 2018). Sociability mostly displays sociable or affiliative facets of behavior, yet grooming also correlates with Confident (Capitanio, 1999) or Irritability (Garai et al., 2016). The behaviors active and playful are often correlated (Uher & Visalberghi, 2016; Vazire et al., 2007), but playful also correlates with anxious and scratch (Iwanicki & Lehmann, 2015). Further, being physically active (i.e., not resting) does not always correlate with social activity (Koski, 2011a). Thus, multimethod approaches may reveal more subtle and complete personality structures because each method captures aspects of personality that are not fully grasped by the other (Garai et al., 2016; Tkaczynski et al., 2018; Uher & Visalberghi, 2016).

Different species of the genus *Macaca* vary in their social style, that is, in aspects of affiliation, aggression, dominance, nepotism, maternal behavior, and socialization (Thierry, Singh, & Kaumanns, 2004). Adams and colleagues (2015) showed that similarities in personality dimensions capturing aggression and social competence across species of macaques are related to similarities in social styles (ranging from despotic via intermediate to egalitarian) and that the personality

structure of the intermediately tolerant Assamese macaques (*Macaca assamensis*) studied here clustered with that of egalitarian crested macaques (*Macaca nigra*). In contrast to many other male primates, male Assamese macaques form differentiated nonkin social relationships with other males (Schülke, Bhagavatula, Vigilant, & Ostner, 2010), with females (Haunhorst, Schülke, & Ostner, 2016), and with immatures (Minge, Berghänel, Schülke, & Ostner, 2016) and engage in frequent coalition formation and different affiliative behaviors, such as grooming, gentle touch, and male–infant–male interactions (Kalbitz, Schülke, & Ostner, 2017; Ostner & Schülke, 2014) while maintaining a strictly linear dominance hierarchy at the same time (Ostner, Heistermann, & Schülke, 2008).

The aim of this study was to explore the personality structure of male Assamese macaques in their natural habitat to provide the baseline for future studies on the socioecology of interindividual differences. We used BC to emphasize prosocial personality traits that received less attention in previous studies (Koski, 2014) but may influence social partner choice (Massen & Koski, 2014) and individual fitness (Seyfarth & Cheney, 2013). The BC was combined with TR to develop a more comprehensive personality model that integrates the strengths of both methods. Our discussion will contribute to the ongoing debate about personality assessment in nonhuman animals (Iwanicki & Lehmann, 2015; Tkaczynski et al., 2018; Uher & Visalberghi, 2016).

Method

Study Site and Subjects

Fieldwork took place in the Phu Khieo Wildlife Sanctuary (PKWS: 16°5′–35′N, 101°20′–55′E), which is part of the roughly 6,500 km^2 interconnected and well-protected Western Isaan Forest Complex in Northeastern Thailand (Borries, Larney, Kreetiyutanont, & Koenig, 2002). The study area is covered by hill evergreen forest and harbors a diverse community of large mammals and predators (Borries et al., 2002). Data for this study were collected on four multimale–multifemale groups from April 2014 (groups ASM and AOM) or October 2014 (groups ASS and AOS) to March 2016. Group sizes at the beginning of behavior data collection are shown in Table S1 in the online supplemental materials.

Data Collection

Trait ratings. All adult females and males of the four groups were rated with the 54-item HPQ (King & Figueredo, 1997; Weiss et al., 2009), which had previously been used to assess personality in different macaque species, including Assamese macaques (Adams et al., 2015). Each adjective item is defined within the context of general behaviors common to primates. For example, "fearful" is defined as "Subject reacts excessively to real or imagined threats by displaying behaviors such as screaming, grimacing, running away or other signs of anxiety or distress." Adjective items were rated on a 7-point scale, from 1 (*displays either total absence or negligible amounts of the trait*) to 7 (*displays extremely large amounts of the trait*; Weiss et al., 2009). Ratings were done in an office away from the animals, and observers were instructed to base their ratings on overall impressions of the individual's behavior and to not discuss the questionnaires with each other. The Thai project members used a Thai language version of the questionnaire that was back-translated to ensure accuracy. Within the study period from April 2014 to March 2016, four Thai field assistants, Anja Ebenau and another doctoral-level student, all familiar with the subjects for 6 months to 7 years in April 2014, rated the animals twice. Eighty-one individuals were rated in the first year, all adult males in March 2015, and all adult females and subadult males in September 2015. Of these 81, 74 individuals were rated again in March 2016. Animals were rated by the same observers who collected the behavior data within the study period.

One of the groups (ASM) had been rated earlier (2009) by eight observers (two of them rated again in this study) familiar with the animals from focal animal data collection (60 adult and juvenile males and females; Adams et al., 2015). Of these 60 individuals, 26 had emigrated into a nonstudy group or died by the beginning of this study and were included here to extend our data set to ensure a higher observations-to-variables ratio.

Behavior codings. We collected 4,628 hr of focal animal observations (Altmann, 1974) from 24 adult males (mean per subject = 193 hr; range = 86–284 hr) of the four study groups. Focal animals were included in the study if they were present longer than 3 months within 1 year of the 2-year study period. Individuals were followed for 40 min with continuous recording of (a) all approaches and departures within 1.5m of the focal animal, (b) all affiliative and agonistic social interactions with directionality and the identities of interaction partners, and (c) several solitary behaviors (e.g., scratching, yawning, display; Supplement Table S2 in the online supplemental materials). Activity of the focal animal was recorded instantaneously at 2-min intervals. Every 10 min, we recorded the identities of all individuals within a 5-m sphere around the focal animal. An effort was made to distribute observation time equally across individuals and time of the day. This study includes an extended data set of socially motivated behaviors, such as affiliative triadic male–infant–male interactions (Kalbitz et al., 2017; Paul, Kuester, & Arnemann, 1996).

Data Analysis

All statistical analyses were run with R (R Core Team, 2017).

Trait ratings. We reduced the data to only those raters who used the entire 7-point scale, those items that were rated by more than half of the remaining reliable observers, and those items that showed normal distribution. We calculated interrater reliability, a measure of consistency across raters, and test–retest reliability, examining consistency over time, to assess reliable adjective items that would then be included in a factor analysis.

For each year (2009, 2015, and 2016), data were visually inspected via scatterplots, histograms, and Q-Q plots to examine the rating performance of observers and the distribution of each adjective item. Three observers had to be excluded from further analyses due to incorrect use of the coding protocol. Specifically, in 2009, three of eight observers and in 2015 and 2016 one of six observers (one of the two, who rated in all 3 years) did not use the complete 7-point scale, and used 1 (*displays either total absence or negligible amounts of the trait*) instead of 4 as mean rating value. The item "autistic" was excluded because eight of the remaining reliable nine observers did not rate it, leaving 53 items for further analysis.

Interrater reliability was measured as intraclass correlation coefficients (ICC; Shrout & Fleiss, 1979). ICCs assess reliability by comparing item variability across all observers. ICC quantifies the similarity of single TRs of each adjective for every individual among all observers (3,1), or mean TRs of each adjective for every individual among all observers (3,k). A good level of agreement for primate studies has been set at ICC(3,k) = 0.48 (Freeman & Gosling, 2010). A meaningful item selection is based on positive values when confidence intervals do not include zero, and often a cutoff criterion of ICC(3,k) > 0.4 is applied (Cicchetti, 1994).

Temporal stability of TRs was assessed with Pearson correlations (corrected for multiple testing by the false discovery rate) of mean adjective items, averaged across observers, for each individual from one rating to the next. High retest reliability measures for the TRs of 74 individuals present in 2015 and 2016 ($N_{\text{adjective items}} = 53$; $M = 0.66$; range: 0.3–0.92) support data aggregation. Thus, we assessed the overall personality structure from TRs of 107 subjects, a combination of mean ratings of 74 individuals present in 2016, plus seven individuals present in 2015 (who emigrated or died later), and additional 26 individuals that were present in the study group only in 2009. Only temporally stable adjective items, that is, significant positive correlations > 0, were subjected to a factor analysis.

After data reduction based on reliability analyses, two measures of sampling adequacy were applied to check for moderate intercorrelations and hence factorability of variables. The Kaiser–Mayer–Olkin (KMO) index compares values of correlations between items and those of partial correlations to check if at least two of three variables correlate with each other ("KMO" function in "psych" package; Revelle, 2018). The Bartlett's test of sphericity compares the observed correlation matrix to the identity matrix with no correlation and is significant when it deviates from identity ("bartlett.test" function in "REdaS" package; Maier, 2015).

The number of factors to be retained for factor analysis was determined using the "fa.parallel" function ("psych" package) and the "paran" function ("paran" package; Dinno, 2012) to perform a scree test (Cattell, 1966) with parallel analysis (Horn, 1965). Exploratory factor analysis was done with the "fa" function of the "psych" package. Principle axis factoring with promax rotation was applied to attain an optimal simple structure that maximize high loadings on one factor for each variable. Oblique factor rotation allows for intercorrelations of factors, which is more adequate for TRs by observers, as well as BCs (Uher & Visalberghi, 2016). For interpretation, salient factor loadings were considered to be $\geq |0.4|$, and items with salient cross loadings were included in the factor on which they had the highest loading (Field, Miles, & Field, 2012). A clean factor structure is achieved by item loadings above $|0.30|$ with preferably no cross loadings, and no factors with less than three items (Costello & Osborne, 2005; Freeman et al., 2013). Behavioral personality research, however, does often accept dimensions with less than three salient items (Koski, 2011b; Manson & Perry, 2013; Pritchard et al., 2014; Seyfarth, Silk, & Cheney, 2012; Sussman, Ha, Bentson, & Crockett, 2013). Internal consistency for each retained personality dimension was calculated as Cronbach's α on mean ratings of adjective items across raters for all 107 subjects ("alpha" function in "psych" package). Alpha ranges from zero to one, with higher values indicating greater internal consistency.

For further analyses, factor scores for each dimension were extracted with the "factor.scores" function ("psych" package) using the regression ("Thurstone") method. Rank order stability of the personality structures from one year to the next (March 2015 and 2016) was assessed with the "factor.congruence" function ("psych" package), comparing the factor loadings of respective dimensions. Confidence intervals (CI) were calculated with the "boot.data" function. For those calculations, we applied 1,000 iterations and sampled 60 out of 74 individuals with no replacement. Further, to demonstrate that the convergent structure relies on individual stability, individual factor scores were correlated for every dimension applying false discovery rate corrections for multiple testing. Due to a rather small sample size, which exaggerates the effect of single individuals on the analysis output, the same variables and individuals were submitted to factor analysis for both years.

The data sets for TR and BC were imbalanced. TR were conducted for male and female infants, juveniles, and adults to increase the number of observations above that of items in the factor analysis. Observational data for BC were collected only on adult males though. To assess whether this imbalance affected comparisons of constructs derived from both methods, we used a parceling approach for item reduction in the TR data set (Little, Cunningham, Shahar, & Widaman, 2002). We ran separate factor analyses for males and females and compared the resulting constructs (further details in online supplementary material). The resulting separate four factor solutions for males ($n = 59$) and females ($n = 48$) had high factor congruence of factor loadings ($M = 0.93$, range: 0.90–0.98), suggesting they were similar or equal (Lorenzo-Seva & ten Berge, 2006). We ran further tests to assess whether the trait-rated factors from the males only ($n = 24$) structure differed from the structure derived from the full data set on males and females ($n = 107$) and found high factor congruence in factor loadings again ($M = 0.98$, range: 0.96–0.99). All further analyses were run on the TR construct built from the full data set.

Behavior codings. Variables for behavioral coding were derived from a long ethogram provided in the supplement (Table S2 in the online supplemental materials). To arrive at our final set of variables for behavioral coding, we first omitted variables that could not be measured reliably either because the behavior was shown very rarely by the adult male subjects ("yawning," "treeshake," "submission," "reconciliation," "rejection," "dominance," "interference," "ignorance," and "playface") or because of problems with data collection; the variable "fidgeting," which should have captured how often an individual changed between different activities, was omitted because we were unable to record these changes in the continuous protocol, and our instantaneous recording interval was too long to capture its effects. In an effort to reduce the overall number of variables, we collapsed those variables that had been defined per age-sex class into just one variable for all partners. We then separately assessed for each study period whether those behaviors that comprised an aggregated variable were positively correlated among each other and retained only those that were. For "friendly behavior," this meant that male-infant-male interaction ("MIMI") was added, but "peer," "play," "kiss," "genital touch," "mount," and "present" were omitted from the aggregate. We had planned to aggregate aggressive behaviors into "threats" and "aggression"; we tested the validity of these constructs by principal component analysis and found three instead of two principle components that we named "mild aggression," "overt aggression," and "physical aggression." The aggressive

behaviors "open mouth," "point," "pretend grab," "bite," and "chase" were omitted because they were not consistently correlated to behaviors from one of the three aggregated aggression variables across years. New aggregated variables were built from correlated behaviors that individually occurred at rather low rates: "auto-grooming" and "scratching" into "self-directed," "body-contact" and "grooming" into "contact time" and "contact diversity," "involvement," "policing" and "agonistic support" aggregated into "agonistic involvement"; the low frequency variables mentioned earlier were omitted because they did not fit into any aggregate variable. The aggregate variable "vocalization" was dropped because components were moved into other aggregates; "growl"' was included in "overt aggression." One variable, "being left," was omitted to reduce the number of variables further.

After this first data reduction, 20 behavior variables were defined and extracted from the focal animal observations (Table S7 in the online supplemental materials). Most of the behavior variables were calculated as rate per hour, corrected by the individual observation time. Other variables were calculated as proportions, across all partners in social measures. To assess the personality construct, variables were aggregated from the overall data collection period to maximize construct stability, which is compromised by low frequencies (Freeman et al., 2013). Reliability was assessed with Pearson correlations of mean values from one year to the next, and only the 18 stable out of the 20 variables were used (Table S7 in the online supplemental materials). The 2 years of data collection were split in half, to have 1 year of BCs before the TRs in March 2015 and 2016, respectively. In the first period, continuous data collection started half a year later for six adult males due to age-class change in October 2014. All behavior variables were visually inspected via histograms and Q-Q plots to examine their distribution pattern. "Display" was log transformed for the single years and square root transformed for the both years together. Variables were z-transformed before submitted to further analyses.

After data reduction, behavior variables were treated as described for the TRs (mentioned earlier) to evaluate their suitability for factor analysis, except that further methods were considered to determine the number of factors to retain for factor analysis (besides scree test with parallel analysis using "fa.parallel" and "paran" function), to search for similar results across several criteria to overcome data insecurities (Freeman et al., 2013; Preacher & MacCallum, 2003). Additional methods were the very simple structure criterion (Revelle & Rocklin, 1979) and Velicer's minimum average partial test (Velicer, 1976), both using the "vss" function ("psych" package), as well as Kaiser's rule (Kaiser, 1960) and scree tests with parallel analysis using the "nScree" function ("nFactors" package; Raiche, 2010). Exploratory factor analysis and further steps were implemented as described earlier for the rating analysis, except that maximum likelihood method with oblimin rotation was applied and that the CIs for factor congruence were calculated based on a sample of 19 out of 22 individuals (bootstrapping required subsampling).

Construct validity. Convergent construct validity examines the correlation between different measures of the same construct (John & Soto, 2007). Convergent construct validity between personality constructs, derived from TRs and BCs, was assessed in two ways. First, individual factor scores of 24 adult males were correlated (Pearson's *r*) for every dimension (Freeman et al., 2013; Garai et al., 2016). Second, percentage bend correlation (Wilcox, 1994) was used to measure the relation between the behavior data from 24 adult males and the overall personality scores from questionnaire ratings from 107 individuals (Iwanicki & Lehmann, 2015; Morton et al., 2013; Tkaczynski et al., 2018). Z-transformed behavior variables were correlated via the "pbcor" function ("WRS2" package; Mair, Schoenbrodt, & Wilcox, 2017) with factor scores of every TR personality dimension.

Results

Trait Ratings

Eight adjective items were excluded because of low interrater reliability, leaving 45 items, with $M_{ICC} = 0.6$; range: 0.4–0.87 (Table S8 in the online supplemental materials). The remaining 45 adjective items were submitted to measures of sampling adequacy and met the criteria. The overall KMO index was higher than 0.5 (KMO = 0.87; range: 0.57–0.92), and the Bartlett's test for homogeneity of variances was significant, Bartlett's K^2 (44) = 360.17; $p > 2.2e\text{-}16$.

The visual inspection of the scree plot with parallel analysis suggested retaining four factors to explain most of the variability in the data. Two items (thoughtless and unperceptive) had loadings $< |0.4|$ on all factors and were excluded. Factor analysis with the remaining 43 items yielded a similar construct. Four dimensions were extracted: Opportunism_{TR}, Confidence_{TR}, Activity_{TR}, and Friendliness_{TR} (Table 1). Opportunism_{TR} comprises items such as aggressive, bullying, irritable, impulsive, not gentle, and not stable. Confidence_{TR} includes dominant, not a follower, not vulnerable, not timid, and not submissive. Activity_{TR} is described by being curious, active, playful, not depressed, and not lazy. Friendliness_{TR} is characterized with affectionate, sociable, friendly, and not solitary. All four factors together explain 72% of the item variance (Table 1). The communalities are the sum of the squared factor loadings for a given variable and may be interpreted as a reliability indicator (Field, 2000). No item loadings were below |0.4|, there were relatively few cross loadings, and mean item communality h^2 was rather high ($M = 0.72$; range: 0.39–0.93), with only three items below 0.5 (clumsy, individualistic, and intelligent), indicating robust personality dimensions (Costello & Osborne, 2005; de Winter, Dodou, & Wieringa, 2009).

Correlations among personality dimensions, as computed with the "fa" function ($M = |0.16|$; range: |0.01–0.44|), were generally weak, apart from a moderate correlation between Opportunism_{TR} and Activity_{TR}. Internal consistencies for the factor scores ($M_{\text{Cronbach's } \alpha} = 0.92$; range: 0.88–0.98) were high.

Rank order stability of personality dimensions from March 2015 to March 2016 was confirmed, with similar results for the congruence of factor loadings Φ ($M = 0.95$; range: 0.91–0.98) and the correlations of factor scores *r* ($M = 0.81$; range: 0.72–0.87; Table 2).

Behavior codings

Test–retest reliabilities *r* of behavior variables were rather low ($N_{\text{behavior variables}} = 20$; $M = 0.39$; range: −0.16–0.69; Table S7 in the online supplemental materials). Two variables, grooming symmetry and overt aggression, did not meet the criteria for temporal stability, that is, positive correlations >0, and were excluded from further analyses.

Table 1
Personality Structure After Promax Rotation With Factor Loadings and Item Communalities (H^2) Derived from Trait Ratings

Item	$Opportunism_{TR}$	$Confidence_{TR}$	$Activity_{TR}$	$Friendliness_{TR}$	h^2
Bullying	**.96**	.21	−.12	−.02	.92
Aggressive	**.93**	.28	−.1	−.01	.92
Irritable	**.93**	0	.01	−.09	.88
Greedy	**.92**	.21	−.38	.15	.77
Jealous	**.92**	.01	−.14	.10	.75
Manipulative	**.87**	.24	−.17	.09	.77
Defiant	**.81**	.09	.12	.02	.81
Excitable	**.78**	−.31	.15	−.05	.77
Impulsive	**.72**	.12	.27	−.01	.82
Erratic	**.71**	−.17	.25	−.21	.76
Distractable	**.55**	**(−.49)**	.16	−.05	.58
Persistent	**.49**	**(.44)**	.20	0	.63
Cool	**−.46**	**(.42)**	−.36	.13	.63
Unemotional	**−.62**	.12	−.39	−.03	.76
Stable	**−.66**	.16	−.36	.15	.79
Gentle	**−.77**	−.17	−.06	.41	.80
Dominant	.30	**.82**	−.19	.05	.81
Decisive	.20	**.79**	−.16	.03	.70
Independent	−.07	**.68**	−.32	**(−.42)**	.65
Protective	.17	**.61**	−.21	**(.47)**	.75
Intelligent	.10	**.60**	−.13	.05	.39
Individualistic	0	**.53**	−.08	**(−.49)**	.41
Clumsy	.36	**−.53**	−.52	.22	.44
Quitting	.18	**−.71**	−.37	.01	.60
Anxious	.18	**−.74**	.05	−.21	.66
Timid	−.24	**−.78**	.11	−.15	.77
Fearful	.25	**−.83**	−.04	.01	.68
Submissive	−.23	**−.86**	−.04	.10	.82
Vulnerable	−.11	**−.88**	−.1	.03	.83
Follower	.03	**−.91**	.01	.30	.78
Playful	−.17	−.17	**.84**	.09	.65
Active	.09	.01	**.79**	−.06	.68
Curious	.32	−.06	**.61**	.15	.73
Inquisitive	.31	−.04	**.61**	.08	.68
Depressed	−.07	**(−.44)**	**−.48**	−.29	.71
Cautious	.19	**(−.42)**	**−.58**	−.09	.50
Lazy	−.16	.13	**−.73**	−.03	.68
Affectionate	−.10	−.10	.01	**.91**	.81
Sympathetic	−.31	−.11	−.02	**.81**	.71
Friendly	**(−.46)**	−.14	−.05	**.77**	.77
Sociable	.22	.11	.14	**.76**	.79
Helpful	.12	.30	.16	**.66**	.74
Solitary	−.15	−.24	−.22	**−.64**	.72
Variance explained	27%	22%	11%	12%	

Note. TR = trait ratings. Salient factor loadings ≥ |.4| are shown in bold. Parentheses are used to highlight double loadings for bold values.

The remaining 18 behavior variables were submitted to measures of sampling adequacy. The overall KMO index for the total observation period was slightly lower than 0.5 (KMO = 0.46; range: 0.20–0.76), and the Bartlett's test for homogeneity of variances was significant, Bartlett's $K^2(17) = 1014.8$; $p < 2.2e\text{-}16$. Low KMO values might result from the rather small observations-to-variables ratio, and these variables were not rejected.

The visual inspection of the scree plot and parallel analysis ("psych" and "paran" package) suggested to retain three factors to explain most of the variability in the data. The very simple structure criterion suggested four and the Velicer's minimum average partial test suggested five factors to retain ("psych" package). The "nFactors" package revealed for Kaiser's rule five factors and for scree test and parallel analysis four factors to retain.

Due to these inconsistencies, we ran two factor analyses, extracting three and four factors, respectively. The four-factor solution, with the dimensions $Gregariousness_{BC}$, $Aggressiveness_{BC}$, $Sociabiliy_{BC}$, and $Vigilance_{BC}$, (Table 3) yielded more robust factors, as described in the following text (Costello & Osborne, 2005), though less robust than in the TR analysis, with seven of 18 variables showing item communalities below 0.5. The four-factor solution had higher mean communality values, four-factor solution: mean (h^2) = 0.62; range: 0.3–1; three-factor solution: mean (h^2) = 0.55; range: 0.2–0.95, as well as a higher proportion of variance explained, four-factor solution: 62%; three-factor solution: 55%. The variable "display" had no reliable loading (<|0.4|) in the four-factor solution and cross loadings of |0.4| (in two dimensions) in the three-factor solution, reflecting its

Table 2
Rank-Order Stability of Personality Dimensions from Trait Ratings in 2015 and 2016 With Factor Congruence of Factor Loadings (First Number in Cell) and Pearson Correlations of Extracted Factor Scores (Second Number in Cell)

	2016			
2015	$Opportunism_{TR}$	$Confidence_{TR}$	$Friendliness_{TR}$	$Activity_{TR}$
$Opportunism_{TR}$	**.97/.87**	−.04/−.20	−.03/−.04	−.18/**−.39**
$Confidence_{TR}$	−.03/−.12	**.98/.79**	−.17/**−.29**	.05/.07
$Friendliness_{TR}$	−.02/−.07	−.11/−.17	**.94/.72**	−.02/**−.29**
$Activity_{TR}$	−.20/**−.46**	.01/.14	−.01/−.13	.91/**.84**

Note. TR = trait rating. $N_{adjectives}$ = 41. Values with 90% confidence intervals excluding zero are shown in bold. $N_{individuals}$ = 74. Significant values after correcting for false discovery rate are shown in bold.

instability. In addition, the fourth factor was mainly marked by the variable "vigilant," with a very high factor loading (0.99) and communality ($h^2 = 1$), but the lowest communality ($h^2 = 0.2$) and a weak factor loading (−0.43) in the three-factor solution. This emphasizes that "vigilant" represents a separate factor. This factor, however, is rather unstable, with only one additional variable "not active" loading on it, which is probably, due to the lack of more nonsocial variables in the data. In other personality studies, vigilance also groups with negatively loaded activity behaviors, or other behaviors, not captured in our analysis, for example, provisioning (Iwanicki & Lehmann, 2015).

As with the rating analysis, there were also moderate intercorrelations *r* among dimensions ($M = |0.16|$; range: |0.02–0.35|), speaking against varimax factor rotation. Internal consistency of the personality dimensions was fair to excellent ($M_{Cronbach's\ \alpha} = 0.77$; range: 0.62–0.92). Alpha for $Vigilance_{BC}$ was rather low at 0.62, but comparable with other studies where personality dimensions include less than three variables (Manson & Perry, 2013).

To assess rank order stability, the three-factor solution was applied for the single years, even though there was a Heywood case (i.e., factor loading >1.0) for "quitting" in the second year. There was no clear result for factor determination, considering different methods and the criteria for construct robustness, as mentioned earlier. This ambiguity underlines the need to aggregate the data (2-year observation period) to overcome the shortcoming of low frequencies in BCs, to enable interpretation of the personality structure. Rank-order stability of personality dimensions from one year to the next was moderate and revealed some differences between the congruence of factor loadings Φ ($M = 0.69$; range: 0.37–0.92) and the correlations of factor scores *r* ($M = 0.40$; range: 0.24 – 0.57; Table 4).

Construct Validity Trait Rating With Behavior Coding

Significant positive correlations of individual factor scores were found between personality dimensions $Friendliness_{TR}$ and

Table 3
Personality Structure Derived from Behavior Codings After Oblimin Rotation With Factor Loadings and Item Communalities (h^2)

Variable	$Gregariousness_{BC}$	$Aggressiveness_{BC}$	$Sociability_{BC}$	$Vigilance_{BC}$	h^2
Neighbor diversity	**.91**	.08	−.14	−.06	.95
Tolerance	**.88**	.02	.13	.12	.78
Friendly approach	**.84**	.10	−.10	.01	.80
Active	**−.52**	.08	−.04	**(−.46)**	.46
Alone	**−.99**	.06	−.05	.06	.94
Quitting	.04	**.85**	.10	−.12	.77
Mild aggression	−.14	**.73**	.13	.16	.49
Agonistic involvement	−.02	**.67**	.03	−.08	.45
Physical aggression	.27	**.60**	−.14	.16	.61
Leaving	.21	**.48**	−.24	.01	.43
Friendly behavior	−.01	.14	**.99**	−.02	1.00
Contact diversity	.03	−.21	**.56**	.11	.34
Peripheral	−.34	−.24	**.49**	−.06	.55
Initiation	−.17	−.05	**.42**	−.28	.38
Contact time	.38	−.19	**.42**	−.10	.30
Self-directed	−.35	.01	**−.44**	.24	.38
Vigilant	−.03	.00	−.03	**.99**	1.00
Display	.39	.38	−.06	−.24	.46
Variance explained	25%	15%	13%	9%	

Note. BC = behavior coding. Salient factor loadings >|.4| are shown in bold, salient double loadings in parentheses.

Table 4
Rank-Order Stability of Personality Dimensions From Behavior Codings in April 2014 to March 2015 and April 2015 to March 2016 With Factor Congruence of Loadings (First Number in Cell) and Pearson Correlation of Extracted Factor Scores (Second Number in Cell)

	2015 to 2016		
2014 to 2015	$Gregariousness_{BC}$	$Sociability_{BC}$	$Quitting_{BC}$[a]
$Gregariousness_{BC}$	**.92/.57**	.01/−.27	.38/.04
$Sociability_{BC}$	−.01/.01	*.78*/*.40*	.02/.25
Mild $Aggression_{BC}$[a]	.18/−.01	−.25/−.24	.37/.24

Note. BC = behavior coding. $N_{behaviors}$ = 18. Significant values, i.e., 95% confidence interval excluding zero, are shown in bold. $N_{individuals}$ = 22. Significant values after controlling for false discovery rate are shown in bold, and trends are in italics.
[a] Third factor named after behavior variables with highest loading on that factor, resembling, but not equivalent to, $Aggressiveness_{BC}$ in the four-factor solution of the aggregated data.

$Gregariousness_{BC}$ (r = .69; $p < .001$), as well as between $Sociability_{BC}$ and $Activity_{TR}$ (r = .63; p = .001), and a low negative correlation was found between $Friendliness_{TR}$ and $Sociability_{BC}$ (r = −0.41; p = .046; Table 5; Figure 1). Regarding the relations between behavior variables, many of which were integrated from several behaviors and personality dimensions based on TRs (Table 6), "mild aggression" correlated with $Opportunism_{TR}$ (ρ percentage bent [ρ_{pb}] = 0.41; p = .049). "Display" (ρ_{pb} = 0.59; p = .003), "vigilant" (ρ_{pb} = −0.43; p = .035), and "self-directed" (ρ_{pb} = −0.56; p = .005) correlated with $Confidence_{TR}$. "Friendly behavior" (ρ_{pb} = 0.55; p = .006) and "peripheral" (ρ_{pb} = 0.43; p = .037) correlated with $Activity_{TR}$. Nine out of 18 variables correlated with $Friendliness_{TR}$: "active" (ρ_{pb} = −0.44; p = .031), "alone" (ρ_{pb} = −0.66; p = .001), "display" (ρ_{pb} = 0.44; p = .033), "friendly approach" (ρ_{pb} = 0.60; p = .002), "leaving" (ρ_{pb} = 0.63; p = .001), "neighbor diversity" (ρ_{pb} = 0.65; p = .001), "physical aggression" (ρ_{pb} = 0.50; p = .012), "peripheral" (ρ_{pb} = −0.45; p = .026), and "tolerance" (ρ_{pb} = 0.60; p = .002).

Discussion

This study was not designed as a validation study comparing personality structures derived with different methods or with the goal to compare results with the published literature. An elaborate method comparison is hampered by the imbalance in our data sets for the two methods and also would have required deriving a priori predictions about which factors from each method would be congruent (Koski et al., 2017; Tkaczynski et al., 2018). A method comparison would also have benefitted from including rating of adjectives that had been derived from the species' behavior repertoire instead of applying constructs derived from lexical accounts in humans (Uher, Asendorpf, & Call, 2008). A validation study would assess which constructs derived by one method are congruent with a construct derived by the second method and thereby arrive at a possibly reduced set of constructs that have been validated. Accepting that to some extent each method will grasp slightly different aspects of personality, the aim of this study was to build a comprehensive personality structure by retaining dimensions that were not congruent across methods and thereby extending the structure space beyond what could be achieved by a single method. Before we describe this structure, we will relate the personality structures and dimensions developed here with published work on the same and other macaque species and other nonhuman primates to establish how robust and reproducible they are across studies using slightly to vastly different methods.

Reproducibility of Trait Ratings and Construct Validity

A previous study on Assamese macaque personality (Adams et al., 2015) did not use item selection and extracted five personality dimensions from HPQ ratings (labeled as $Confidence_{TR}$, $Activity_{TR}$, $Openness_{TR}$, $Friendliness_{TR}$, and $Opportunism_{TR}$), whereas our rating analysis revealed four dimensions: $Opportunism_{TR}$, $Confidence_{TR}$, $Activity_{TR}$, and $Friendliness_{TR}$, which we named the same when replicated. Our $Opportunism_{TR}$ domain comprised all adjectives representing $Opportunism_{TR}$, as well as four out of 10 items of $Openness_{TR}$ in Adams et al. BCs from adult males also resulted in a four-factor structure, with $Gregariousness_{BC}$, $Aggressiveness_{BC}$, $Sociability_{BC}$, and $Vigilance_{BC}$, which has considerable overlap with the personality structure from TR and does reflect previous findings in other macaque studies (Neumann et al., 2013). $Friendliness_{TR}$ was convergent with $Gregariousness_{BC}$ and inverse $Sociability_{BC}$, whereas $Activity_{TR}$ correlated with $Sociability_{BC}$. In addition, $Opportunism_{TR}$ was weakly associated with $Aggressiveness_{BC}$. $Confidence_{TR}$ replicated very well but showed no association with the behavior personality dimensions.

The TR dimension $Friendliness_{TR}$ confirmed previous findings (Adams et al., 2015; Konečná et al., 2008; Sussman et al., 2013;

Table 5
Construct Validity as Pearson Correlations of Individual Factor Scores Derived from Trait Ratings and Behavior Codings

	Behavior coding			
Trait rating	$Gregariousness_{BC}$	$Aggressiveness_{BC}$	$Sociability_{BC}$	$Vigilance_{BC}$
$Opportunism_{TR}$	−.26	*.36*	.06	−.08
$Confidence_{TR}$	.30	.29	.07	−*.36*
$Friendliness_{TR}$	**.69**	*.35*	**−.41**	.20
$Activity_{TR}$	−.13	.14	**.63**	−.05

Note. TR = trait rating; BC = behavior coding. Bold type indicates values significant after controlling for false discovery rate, and italicized values signify statistical trends.

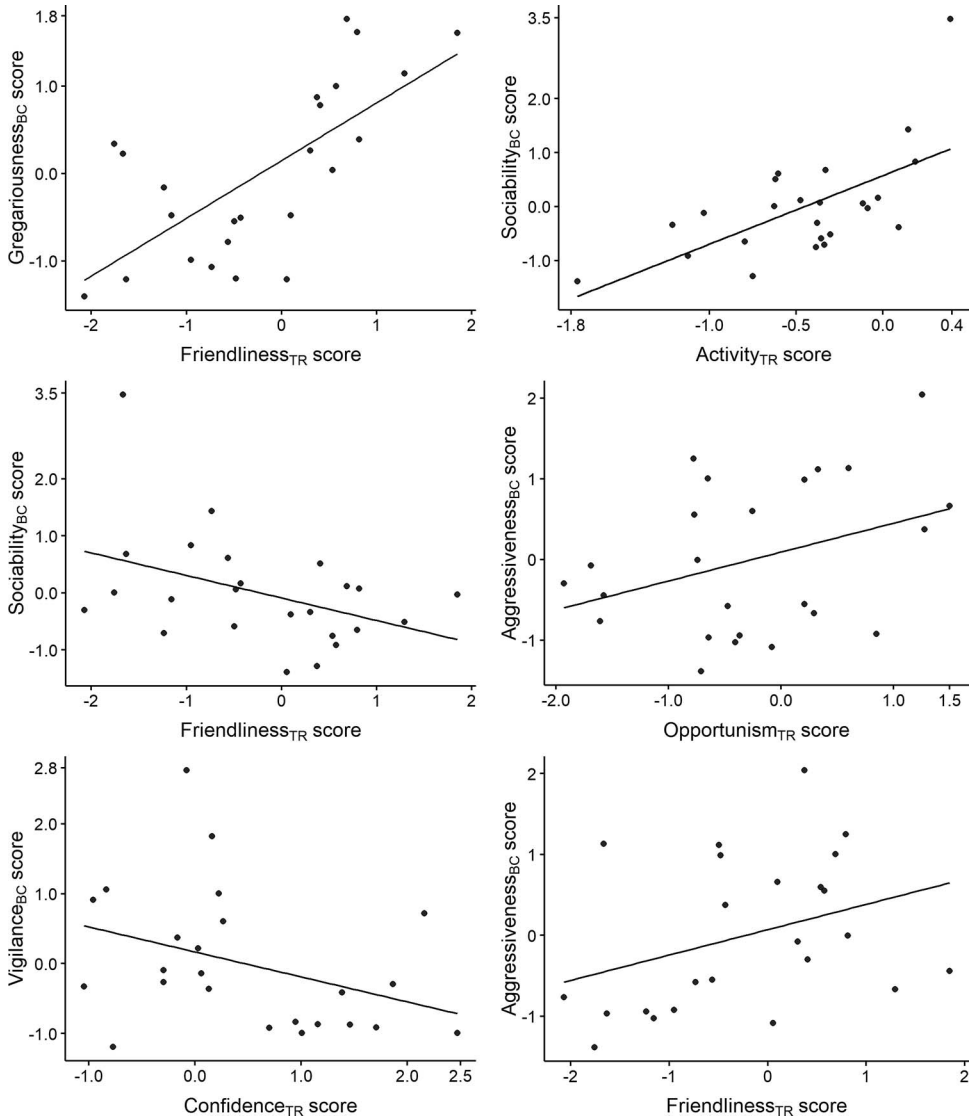

Figure 1. Congruence in factor scores of 24 males between constructs from behavior coding and trait rating. For statistics, see Table 5.

Weiss, Adams, Widdig, & Gerald, 2011): The rather strong link between Friendliness$_{TR}$ and Gregariousness$_{BC}$ is explained by all behavior variables belonging to Gregariousness$_{BC}$ having significant correlations with Friendliness$_{TR}$. More specifically, adjectives, such as "sympathetic" (kind toward others), "sociable" (seeks and enjoys company), and "not solitary" (spends considerable time alone), correlate with behaviors like "friendly approach" or "tolerance," "neighbor diversity," and "not alone." Gregariousness$_{BC}$ represents spatial proximity aspects that have been suggested to measure social integration, that is, close proximity to and higher rates of affiliation with diverse partners. The evidence for a Gregariousness$_{BC}$ dimension is mixed for macaques. There are similar domains with the focus on close and distant proximity in Barbary, crested, and Rhesus macaques (Capitanio, 1999; Neumann et al., 2013; Tkaczynski et al., 2018) or domains with a mixture of behaviors belonging to our Gregariousness$_{BC}$ and Sociability$_{BC}$ in Barbary, long-tailed, lion-tail, and Tibetan macaques (Pritchard et al., 2014; Rouff, Sussman, & Strube, 2005; Tkaczynski et al., 2018; Uher, Werner, et al., 2013).

The negative correlation between Friendliness$_{TR}$ and Sociability$_{BC}$, while being weak, still seems unexpected. Friendliness$_{TR}$ does not seem to cover the behavior personality dimension Sociability$_{BC}$, as four out of the five variables with positive factor loadings on Sociability$_{BC}$ correlated negatively with Friendliness$_{TR}$. Sociability$_{BC}$ seems to have two facets: First, it was primarily described by high rates of short-term affiliation (such as "embrace"), with a factor loading of 0.99 (mean loadings of other variables <0.5) and a negative correlation with Friendliness$_{TR}$, and thus may reflect an aspect of social integration, not included in Friendliness$_{TR}$. Second, Sociability$_{BC}$ seems to express variation in social bonding, that is, the tendency to focus affiliation on a few close partners (Ostner & Schülke, 2014). Individuals with stronger bonds have a higher "contact time" (time in body contact and grooming), with fewer social

Table 6
Correlation Coefficients (ρ_{Pb}) of Behavior Variables With Personality Factor Scores from Trait Ratings (TR)

Behavior variable	Personality Dimension$_{TR}$							
	Opportunism$_{TR}$	*p*	Confidence$_{TR}$	*p*	Activity$_{TR}$	*p*	Friendliness$_{TR}$	*p*
Active	.34	.11	.02	.92	.13	.06	**−.44**	.03
Agonistic involvement	.07	.75	.32	.13	−.15	.48	.14	.52
Alone	.32	.13	−.34	.10	.24	.27	**−.66**	.00
Contact diversity	−.21	.32	.01	.96	.10	.63	−.40	.05
Contact time	−.40	.06	.29	.17	−.01	.95	−.03	.88
Display	−.23	.28	**.59**	.00	−.18	.40	**.44**	.03
Friendly approach	−.02	.92	.08	.71	−.25	.24	**.60**	.00
Friendly behavior	.06	.79	.01	.96	**.55**	.01	−.34	.11
Initiation	.27	.21	−.24	.25	.19	.37	−.34	.11
Leaving	.15	.49	.22	.31	.14	.52	**.63**	.00
Mild aggression	**.41**	.05	−.11	.61	.08	.72	.10	.64
Neighbor diversity	−.20	.35	.22	.29	−.33	.12	**.65**	.00
Physical aggression	.05	.82	.17	.43	−.19	.39	**.50**	.01
Quitting	.34	.10	.23	.28	.17	.42	.32	.13
Self-directed	.09	.68	**−.56**	.01	−.20	.36	−.04	.85
Peripheral	.15	.48	−.10	.66	.43	.04	**−.45**	.03
Tolerance	−.34	.11	.15	.49	−.32	.13	**.60**	.00
Vigilant	−.04	.85	**−.43**	.04	−.13	.55	.06	.79

Note. TR = trait rating. Significant values after false discovery rate correction shown in bold.

partners leading to a negative correlation of "contact diversity" with Friendliness$_{TR}$. Also, "contact time," as a measure for "affectionate" (closeness and frequent grooming with others), was not correlated with Friendliness$_{TR}$, as high scores may result from a few long or from many short periods of contact. In addition, rates of "friendly behavior" do not need to be high for strongly bonded individuals, as they might engage in prolonged but less frequent social interactions with their bonded partners. Crested and rhesus macaques exhibit a similar behavior Sociability$_{BC}$ dimension, with high grooming and affiliation rates and diverse grooming and affiliation (Neumann et al., 2013; von Borell, Kulik, & Widdig, 2016). Interestingly, "peripheral" (staying outside the group center) included in Sociability$_{BC}$ underlines that there is no need for individuals to stay close to the core of the group to be socially integrated. In Barbary and crested macaques being "central" loads on a dimension most similar to this study's Gregariousness$_{BC}$ (Neumann et al., 2013; Tkaczynski et al., 2018).

Opportunism$_{TR}$ completely comprises adjectives referring to Adams and colleagues' Opportunism$_{TR}$, as well as four out of 10 adjectives belonging to their Openness$_{TR}$ dimension, that is, "excitable," "impulsive," "erratic," and "distractible" (Adams et al., 2015). The remaining five 'Openness$_{TR}$ adjectives' from Adams (e.g., "thoughtless," "innovative") were excluded from our analysis and "individualistic" loaded on Confidence$_{TR}$. It was unexpected that Opportunism$_{TR}$ and Aggressiveness$_{BC}$ were not strongly correlated. This weak relationship was likely driven by the behaviors "mild aggression" and "quitting" because observers may have perceived more frequent behaviors as more important ("mild aggression" was five times more frequent than "physical aggression"). Similarly, in chimpanzees, aggression-related TRs converged strongly with "threat" but only weakly with the lower frequency "attack" behavior measures (Vazire et al., 2007). "Physical aggression" and "leaving" correlated positively with Friendliness$_{TR}$, which is similar to a positive association of "contact aggression" and the sociopositive Extraversion$_{TR}$ found in chimpanzees (Freeman et al., 2013). As "physical aggression" is often expressed in more serious conflicts, which in chimpanzees and Assamese macaques often turn polyadic, an individual acting aggressively may at the same time be supportive and helpful in a coalitionary context. "Helpful" is included in Friendliness$_{TR}$, possibly causing the positive correlation between "physical aggression" and Friendliness$_{TR}$. In addition, there is a general link between affiliative and aggressive interactions among partners due to the increasing probability of conflict with increasing time spent in close proximity to others (Silk et al., 2010; Streich, Widdig, Bercovitch, Nürnberg, & Krawczak, 2002). In line with this, in macaque personality studies, affiliative and aggressive behaviors are frequently correlated as are social behavior and aggressiveness-related rating dimensions (Capitanio, 1999; Rouff et al., 2005; Tkaczynski et al., 2018; Uher, Werner, et al., 2013). So far, rating Opportunism$_{TR}$ domains are described for Assamese (this study), crested (Adams et al., 2015), pigtail, long-tailed (Sussman et al., 2013; Uher, Werner, et al., 2013), and Tibetan macaques (Pritchard et al., 2014), whereas Confidence$_{TR}$ or Dominance$_{TR}$ dimensions are found in rhesus (Adams et al., 2015; Capitanio, 1999; Weiss et al., 2011) and Confidence$_{TR}$ and Opportunism$_{TR}$ or Excitability$_{TR}$ in Barbary macaques (Adams et al., 2015; Konečná, Weiss, Lhota, & Wallner, 2012; Tkaczynski et al., 2018).

Confidence$_{TR}$ is almost equivalent to the dimension found previously (Adams et al., 2015). Confidence$_{TR}$ was weakly related to Vigilance$_{BC}$ as a statistical trend, given the negative correlation of "vigilant" behavior with Confidence$_{TR}$. However, the two dimensions seem not convergent, as other candidate behaviors for Confidence$_{TR}$ did not correlate with "vigilant." For instance, "display," presumably a dominance behavior (Freeman et al., 2013; Uher, Werner, et al., 2013; Tkaczynski et al., 2018), showed no reliable loading on any behavior personality dimension, but a prominent positive correlation with Confidence$_{TR}$. In addition,

"self-directed" behavior correlated negatively with $Confidence_{TR}$, suggesting it is a measure of anxiety (Brent et al., 2014; Iwanicki & Lehmann, 2015; Koski, 2011b; Neumann et al., 2013; Tkaczynski et al., 2018), but here it only loaded negatively, although weakly, on $Sociability_{BC}$. There was a lack of behavior variables supporting $Confidence_{TR}$-adjectives like "dominant" and "submissive," because adult Assamese macaques rarely show unprovoked submissive behavior toward other individuals. Similarly, a study on wild Barbary macaques did not find behavior variables to be correlated with their $Confidence_{TR}$ dimension, which is a mixture of items belonging to our $Confidence_{TR}$ and $Opportunism_{TR}$ (Tkaczynski et al., 2018). $Confidence_{TR}$ was correlated with dominance rank though in Hanuman langurs and with rank stability in Barbary macaques (Konečná et al., 2008; Konečná et al., 2012). Taken together, a fifth behavior dimension replicating $Confidence_{TR}$ may be possible.

In contrast to the reverse correlation of "vigilant" and "active" in this study, a positive correlation of "vigilance" with "playfulness" and "physical activity" was found in long-tailed macaques (Uher, Addessi, et al., 2013) and vigilance increased with increasing play behavior in brown capuchins (Morton et al., 2013). In these studies, vigilance was measured as social attentiveness, monitoring the activities of other group members. Further, "vigilant" loaded on the Anxiety dimension derived from playback experiments in wild Barbary macaques (Tkaczynski et al., 2018). However, it was only used as playback reaction measure and not as a neutral variable in the everyday BC. A behavior study with free-ranging rhesus macaques (von Borell et al., 2016) described a $Fearfulness_{BC}$ dimension, resembling some of the behavior variables of this study loading on $Gregariousness_{BC}$ (i.e., friendly approach, proximity, and resting) grouped together with "submissive." If our "vigilance" represented an anxiety measure, it would most likely group with other potential anxiety behaviors, but instead it represented its own dimension. Thus, the "vigilance" variable in this study reflected a general attentiveness to the surrounding environment, social or ecological.

$Activity_{TR}$ shared the most descriptive adjectives with highest loadings, for example, "active" (considerable time moving or engaging in energetic behavior) and "not lazy" and "playful," with earlier findings (Adams et al., 2015). In addition, our $Activity_{TR}$ included "curious," as in other studies of macaques (Barbary: Konečná, et al., 2012; long-tailed: Uher, Werner, et al., 2013; Tibetan: Pritchard et al., 2014), but the $Activity_{TR}$ dimension was not found across all rating studies of macaques. $Activity_{TR}$ was not correlated with "active" behavior, but with "friendly behavior" and "peripheral" instead, leading to a strong correlation between $Activity_{TR}$ and $Sociability_{BC}$. $Activity_{TR}$ may thus be mainly perceived as socially active and spending a greater amount of time outside the group center, yet not necessarily alone. However, there was no convergence between $Activity_{TR}$ and $Sociability_{BC}$ because the correlation was only driven by one social behavior (out of six behaviors loading on $Sociability_{BC}$); it was not correlated with "active" behavior, despite the high loadings of "active" and "not lazy" on $Activity_{TR}$. Generally, rating and behavior Activity dimensions are related to social behaviors in primates (Konečná et al., 2008; Tkaczynski et al., 2018; Uher, Werner, et al., 2013; Vazire et al., 2007), but further analyses are needed to better understand these links.

The inconsistencies identified in this descriptive comparison between our and previous findings for Assamese macaques (Adams et al., 2015) could be due to the sensitivity of the personality structure to rather small sample sizes, leading to small observations-to-variables ratios for the factor analyses (TRs with 107 individuals and 43 adjective items). If we assume moderately plastic personality, the construct might change due to new animals being rated in 2015 and 2016 and 19 of the 34 animals already rated in 2009 having matured into adulthood (Dingemanse & Wolf, 2013). Inconsistencies may additionally be due to differences in method; the original study (Adams et al., 2015) neither excluded unreliable raters nor items with low interrater reliability, whereas we excluded half of the items included in the original $Openness_{TR}$ dimension, which was found in four out of five species studied. Some studies found similar Openness dimensions (Konečná et al., 2012; Uher, Werner, et al., 2013), but others did not (Capitanio, 1999; Tkaczynski et al., 2018). Differences in method may thus explain the main discrepancies between ours and the previous study on the same population. Another primate species where personality has been studied with different methods is the common marmoset (Iwanicki & Lehmann, 2015; Koski et al., 2017). Discrepancies between the personality structures derived in the two studies, including the number of personality dimensions derived, were explained by differences in methods used (Koski et al., 2017).

Comparison With Other Multimethod Primate Studies

In the following text, we provide a brief review of seven studies of macaques and 11 studies of other genera of primates that used multimethod approaches (macaques: Capitanio, 1999; Pritchard et al., 2014; Rouff et al., 2005; Sussman et al., 2013; Tkaczynski et al., 2018; Uher, Werner, et al., 2013; this study; other nonhuman primates: Carter et al., 2012; Eckardt et al., 2015; Freeman et al., 2013; Garai et al., 2016; Iwanicki & Lehmann, 2015; Konečná et al., 2008; Morton et al., 2013; Pederson, King, & Landau, 2005; Uher & Asendorpf, 2008; Uher & Visalberghi, 2016; Vazire et al., 2007). To date, TR studies identified a unique macaque $Friendliness_{TR}$ dimension described with the same adjectives in all species, whereas other primate studies often define two discriminant social domains, mostly named $Extraversion_{TR}$ and $Agreeableness_{TR}$. In nearly all of the reviewed multimethod studies, social rating domains were best validated, revealing positive correlations with social behaviors or behavior domains, generally resembling a mixed pattern of behaviors reflecting social integration (e.g., proximity) and bonding (e.g., grooming skew). $Opportunism_{TR}$- and $Dominance_{TR}$-related dimensions (or $Confidence_{TR}$ as in most macaques) were validated in half of the studies. $Dominance_{TR}$ and $Opportunism_{TR}$ are not seen as interchangeable because $Dominance_{TR}$, described with dominant and confident adjectives, is indeed often correlated with dominant and aggressive behaviors. In contrast, $Opportunism_{TR}$ is correlated with dominance behavior only, if the two aspects are integrated in one Dominant–Competitive–Aggressive dimension. Most studies describe either $Opportunism_{TR}$ (macaques) or $Dominance_{TR}$ (other species), and one third made both domains part of the personality profile. Other dimensions, as $Excitability_{TR + BC}$, $Emotionality_{TR + BC}$ or $Boldness_{TR + BC}$, are less reliably measured and validated. Half of the studies in nonmacaque species, and one in

macaques, define an $\text{Openness}_{\text{TR + BC}}$ domain, consistently loaded with "curious," which in all cases correlated with "playful" behavior.

Disparities between methods measuring the same construct, as found in the majority of studies, demand multimethod approaches to check for validity in every species under study to further ensure reproducibility. We should not assume validity if only shown once, as most personality dimensions have not been reliably validated in the nonhuman literature (but see discussion for nonhuman primates in the study by Koski et al., 2017). Particularly, studies on wild primates with low sample sizes and limited observational data (partly due to lower behavior frequencies compared with captivity) need to be replicated. Replication studies on the same subjects, groups, and populations could build upon each other enabling longer observation periods, and hence larger sample sizes, which may allow including behaviors with low frequencies. Eventually, male, female, and juvenile behavior data could be integrated. Subsequent studies could focus on a single dimension, which might require experimental settings (e.g., boldness), which are difficult to conduct in the wild (e.g., playback experiments in demanding habitats). Altogether, this may lead to more complex and fine-grained personality structures of nonhuman animals. In addition, mixed findings in personality profiles between species also call for replication studies to understand the socioecological relevance of similar behaviors in different species.

A Comprehensive Personality Structure for Male Assamese Macaques

We propose two solutions for how to combine the eight dimensions from TR and BC. Congruence between $\text{Friendliness}_{\text{TR}}$ and $\text{Gregariousness}_{\text{BC}}$ and between $\text{Activity}_{\text{TR}}$ and $\text{Sociability}_{\text{BC}}$ make them largely redundant. Based on the considerations laid out above about the role of behavior frequency, we also consider $\text{Opportunism}_{\text{TR}}$ and $\text{Aggressiveness}_{\text{BC}}$ to reflect a similar dimension of interindividual variation in behavior. Despite the correlation between factor loadings on $\text{Confidence}_{\text{TR}}$ and $\text{Vigilance}_{\text{BC}}$ being the same as between $\text{Opportunism}_{\text{TR}}$ and $\text{Aggressiveness}_{\text{BC}}$ dimensions, we considered congruence to be low between the former because $\text{Confidence}_{\text{TR}}$ was a much broader construct with several items that were not related statistically to $\text{Vigilant}_{\text{BC}}$. It follows that the full five-factor structure is either built from all TR dimensions plus $\text{Vigilance}_{\text{BC}}$ or from all BC dimensions plus $\text{Confidence}_{\text{TR}}$. Both solutions have their strengths and weaknesses.

We demonstrated the importance of examining of rating performances of observers, as well as reliability analyses for item selection. However, even if a strong agreement among observers can ensure that TRs are not purely idiosyncratic interpretations, all observers may be biased in the same way (Freeman et al., 2011; Koski, 2011a; Uher, 2008; Uher & Asendorpf, 2008). This aspect of TRs deserves continued attention. For instance, an individual that is mainly staying in the periphery of the group and spotted less often will more likely be rated as unsocial, although it is extremely social when joining the group. In statistical terms, TRs have to be preferred over BC results here because the former were much more robust.

Advantages of BC include that behaviors are directly perceivable and measurable and therefore more objective (Freeman et al., 2011) so that more subtle variation in sociability can be detected. So far, TR studies in macaques identified a single Friendliness dimension (Adams et al., 2015), whereas BC studies frequently found two distinct prosocial personality domains (Capitanio, 1999; Neumann et al., 2013; Tkaczynski et al., 2018; this study). Researchers in animal and human personality studies stress that social personality dimensions can only be identified in the context of social relationships, especially in species with complex societies (Cooper, 2002; Koski, 2011b; Reis, Collins, & Berscheid, 2000). This is highlighted by the fact that affinitive behavior loaded on a different factor than actual affiliative interaction variables because individuals may tolerate and be tolerated sometimes without implications for friendly contact behaviors.

One disadvantage of the BC data is that measurement reliability for individual differences may be low for rare behaviors that may, however, be biologically very important such as support in agonistic conflicts. Observers may be better able to represent variation in rare behaviors in their ratings.

We conclude that two solutions with five factors best describe the personality structure of the male Assamese macaques studied here. One five-factor structure is built from all TR dimensions plus $\text{Vigilance}_{\text{BC}}$ and the other from all BC dimensions plus $\text{Confidence}_{\text{TR}}$. Both solutions have their strengths and weaknesses. We encourage the use of integrative approaches including TRs, BCs, and experiments embracing the transdisciplinary philosophy of science paradigm (Uher, 2018), which aims to broaden the horizon in personality research in a transdisciplinary way.

References

Adams, M. J., Majolo, B., Ostner, J., Schülke, O., De Marco, A., Thierry, B., . . . Weiss, A. (2015). Personality structure and social style in macaques. *Journal of Personality and Social Psychology, 109,* 338–353. http://dx.doi.org/10.1037/pspp0000041

Altmann, J. (1974). Observational study of behavior: Sampling methods. *Behaviour, 49,* 227–266. http://dx.doi.org/10.1163/156853974X00534

Borries, C., Larney, E., Kreetiyutanont, K., & Koenig, A. (2002). The diurnal primate community in a dry evergreen forest in Phu Khieo Wildlife Sanctuary. *Northeast Thailand, 50,* 75–88.

Brent, L. J. N., Semple, S., Maclarnon, A., Ruiz-Lambides, A., Gonzalez-Martinez, J., & Platt, M. L. (2014). Personality Traits in Rhesus Macaques (*Macaca mulatta*) are heritable but do not predict reproductive Output. *International Journal of Primatology, 35,* 188–209. http://dx.doi.org/10.1007/s10764-013-9724-6

Brommer, J. E., & Class, S. A. (2017). Personality from the perspective of behavioral ecology. In J. Vonk, A. Weiss, & S. A. Kuczaj (Eds.), *Personality in nonhuman animals* (pp. 73–107). New York, NY: Springer. http://dx.doi.org/10.1007/978-3-319-59300-5_5

Capitanio, J. P. (1999). Personality dimensions in adult male rhesus macaques: Prediction of behaviors across time and situation. *American Journal of Primatology, 47,* 299–320. http://dx.doi.org/10.1002/(SICI)1098-2345(1999)47:4<299::AID-AJP3>3.0.CO;2-P

Carere, C., & Maestripieri, D. (2013). *Animal personalities: Behavior, physiology, and evolution*. Chicago, IL: The University of Chicago Press. http://dx.doi.org/10.7208/chicago/9780226922065.001.0001

Carter, A. J., Marshall, H. H., Heinsohn, R., & Cowlishaw, G. (2012). Evaluating animal personalities: Do observer assessments and experimental tests measure the same thing? *Behavioral Ecology and Sociobiology, 66,* 153–160. http://dx.doi.org/10.1007/s00265-011-1263-6

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1,* 245–276. http://dx.doi.org/10.1207/s15327906mbr0102_10

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6,* 284–290. http://dx.doi.org/10.1037/1040-3590.6.4.284

Cooper, M. L. (2002). Personality and close relationships: Embedding people in important social contexts. *Journal of Personality, 70,* 757–782. http://dx.doi.org/10.1111/1467-6494.05023

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation, 10,* 173–178. http://pareonline.net/getvn.asp?v=10&n=7

Dammhahn, M. (2012). Are personality differences in a small iteroparous mammal maintained by a life-history trade-off? *Proceedings of the Royal Society B: Biological Sciences, 279,* 1–7. http://dx.doi.org/10.1098/rspb.2012.0212

Deary, I. J., Weiss, A., & Batty, G. D. (2010). Intelligence and personality as predictors of illness and death: How researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychological Science in the Public Interest, 11,* 53–79. http://dx.doi.org/10.1177/1529100610387081

de Winter, J. C. F., Dodou, D., & Wieringa, P. A. (2009). Exploratory factor analysis with small sample sizes. *Multivariate Behavioral Research, 44,* 147–181. http://dx.doi.org/10.1080/00273170902794206

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41,* 417–440. http://dx.doi.org/10.1146/annurev.ps.41.020190.002221

Dingemanse, N. J., Kazem, A. J. N., Réale, D., & Wright, J. (2010). Behavioural reaction norms: Animal personality meets individual plasticity. *Trends in Ecology and Evolution, 25,* 81–89. http://dx.doi.org/10.1016/j.tree.2009.07.013

Dingemanse, N. J., & Wolf, M. (2013). Between-individual differences in behavioural plasticity within populations: Causes and consequences. *Animal Behaviour, 85,* 1031–1039. http://dx.doi.org/10.1016/j.anbehav.2012.12.032

Dinno, A. (2012). *paran: Horn's test of principal components/factors*. Retrieved from https://CRAN.R-project.org/package=paran

Eckardt, W., Steklis, H. D., Steklis, N. G., Fletcher, A. W., Stoinski, T. S., & Weiss, A. (2015). Personality dimensions and their behavioral correlates in wild Virunga mountain gorillas (Gorilla beringei beringei). *Journal of Comparative Psychology, 129,* 26–41. http://dx.doi.org/10.1037/a0038370

Field, A. (2000). *Discovering statistics using SPSS for Windows: Advanced techniques for beginners (Introducing Statistical Methods series*). London, United Kingdom: Sage publications.

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. London, United Kingdom: Sage publications.

Freeman, H. D., Brosnan, S. F., Hopper, L. M., Lambeth, S. P., Schapiro, S. J., & Gosling, S. D. (2013). Developing a comprehensive and comparative questionnaire for measuring personality in chimpanzees using a simultaneous top-down/bottom-up design. *American Journal of Primatology, 75,* 1042–1053. http://dx.doi.org/10.1002/ajp.22168

Freeman, H. D., & Gosling, S. D. (2010). Personality in nonhuman primates: A review and evaluation of past research. *American Journal of Primatology, 72,* 653–671. http://dx.doi.org/10.1002/ajp.20833

Freeman, H. D., Gosling, S. D., & Schapiro, S. J. (2011). Comparison of methods for assessing personality in nonhuman primates. In A. Weiss, J. E. King, & L. Murray (Eds.), *Personality and temperament in nonhuman primates* (pp. 17–40). New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4614-0176-6_2

Garai, C., Weiss, A., Arnaud, C., & Furuichi, T. (2016). Personality in wild bonobos (Pan paniscus). *American Journal of Primatology, 78,* 1178–1189. http://dx.doi.org/10.1002/ajp.22573

Gosling, S. D. (2001). From mice to men: What can we learn about personality from animal research? *Psychological Bulletin, 127,* 45–86. http://dx.doi.org/10.1037/0033-2909.127.1.45

Haunhorst, C. B., Schülke, O., & Ostner, J. (2016). Opposite-sex social bonding in wild Assamese macaques. *American Journal of Primatology, 78,* 872–882. http://dx.doi.org/10.1002/ajp.22554

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30,* 179–185. http://dx.doi.org/10.1007/BF02289447

Iwanicki, S., & Lehmann, J. (2015). Behavioral and trait rating assessments of personality in common marmosets (*Callithrix* jacchus). *Journal of Comparative Psychology, 129,* 205–217. http://dx.doi.org/10.1037/a0039318

John, O. P., & Soto, C. J. (2007). The importance of being valid: Reliability and the process of construct validation. In R. W. Robins, C. R. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 461–494). New York, NY: Guilford Press.

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* (Vol. 2, pp. 102–138). New York, NY: Guilford Press.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20,* 141–151. http://dx.doi.org/10.1177/001316446002000116

Kalbitz, J., Schülke, O., & Ostner, J. (2017). Triadic male-infant-male interaction serves in bond maintenance in male Assamese macaques. *PLoS ONE, 12,* e0183981. http://dx.doi.org/10.1371/journal.pone.0183981

King, J. E., & Figueredo, A. J. (1997). The Five-Factor Model plus Dominance in chimpanzee personality. *Journal of Research in Personality, 31,* 257–271. http://dx.doi.org/10.1006/jrpe.1997.2179

Konečná, M., Lhota, S., Weiss, A., Urbánek, T., Adamová, T., & Pluhácek, J. (2008). Personality in free-ranging Hanuman langur (Semnopithecus entellus) males: Subjective ratings and recorded behavior. *Journal of Comparative Psychology, 122,* 379–389. http://dx.doi.org/10.1037/a0012625

Konečná, M., Weiss, A., Lhota, S., & Wallner, B. (2012). Personality in Barbary macaques (*Macaca* sylvanus): Temporal stability and social rank. *Journal of Research in Personality, 46,* 581–590. http://dx.doi.org/10.1016/j.jrp.2012.06.004

Koski, S. E. (2011a). How to measure animal personality and why does it matter? Integrating the psychological and biological approaches to animal personality. In M. Inoue-Murayama, S. Kawamura, & A. Weiss (Eds.), *From genes to animal behavior* (pp. 115–136). Tokyo, Japan: Springer. http://dx.doi.org/10.1007/978-4-431-53892-9_5

Koski, S. E. (2011b). Social personality traits in chimpanzees: Temporal stability and structure of behaviourally assessed personality traits in three captive populations. *Behavioral Ecology and Sociobiology, 65,* 2161–2174. http://dx.doi.org/10.1007/s00265-011-1224-0

Koski, S. E. (2014). Broader horizons for animal personality research. *Ecology and Evolution, 2,* 1–6.

Koski, S. E., Buchanan-Smith, H. M., Ash, H., Burkart, J. M., Bugnyar, T., & Weiss, A. (2017). Common marmoset (*Callithrix jacchus*) personality. *Journal of Comparative Psychology, 131,* 326–336. http://dx.doi.org/10.1037/com0000089

Koski, S. E., & Burkart, J. M. (2015). Common marmosets show social plasticity and group-level similarity in personality. *Scientifc Reports, 5,* 8878. http://dx.doi.org/10.1038/srep08878

Lilley, M. K., Kuczaj, S. A., & Yeater, D. B. (2017). Individual differences in nonhuman animals: Examining boredom, curiosity, and creativity. In J. Vonk, A. Weiss, & S. A. Kuczaj (Eds.), *Personality in nonhuman animals* (pp. 257–275). New York, NY: Springer. http://dx.doi.org/10.1007/978-3-319-59300-5_13

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9,* 151–173. http://dx.doi.org/10.1207/S15328007SEM0902_1

Lorenzo-Seva, U., & ten Berge, J. M. F. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology, 2,* 57–64. http://dx.doi.org/10.1027/1614-2241.2.2.57

Maier, M. J. (2015). *REdaS: Companion package to the book "R: Einführung durch angewandte Statistik"*. Retrieved from https://CRAN.R-project.org/package=REdaS

Mair, P., Schoenbrodt, F., & Wilcox. (2017). *WRS2: Wilcox robust estimation and testing*. Retrieved from https://CRAN.R-project.org/package=WRS2

Manson, J. H., & Perry, S. (2013). Personality structure, sex differences, and temporal change and stability in wild white-faced capuchins (*Cebus capucinus*). *Journal of Comparative Psychology, 127,* 299–311. http://dx.doi.org/10.1037/a0031316

Massen, J. J. M., Antonides, A., Arnold, A.-M. K., Bionda, T., & Koski, S. E. (2013). A behavioral view on chimpanzee personality: Exploration tendency, persistence, boldness, and tool-orientation measured with group experiments. *American Journal of Primatology, 75,* 947–958. http://dx.doi.org/10.1002/ajp.22159

Massen, J. J. M., & Koski, S. E. (2014). Chimps of a feather sit together: Chimpanzee friendships are based on homophily in personality. *Evolution and Human Behavior, 35,* 1–8. http://dx.doi.org/10.1016/j.evolhumbehav.2013.08.008

Mehta, P. H., & Gosling, S. D. (2008). Bridging human and animal research: A comparative approach to studies of personality and health. *Brain, Behavior, and Immunity, 22,* 651–661. http://dx.doi.org/10.1016/j.bbi.2008.01.008

Minge, C., Berghänel, A., Schülke, O., & Ostner, J. (2016). Patterns and consequences of male–infant relationships in wild Assamese macaques (*Macaca assamensis*). *International Journal of Primatology, 37,* 350–370. http://dx.doi.org/10.1007/s10764-016-9904-2

Morton, F. B., Lee, P. C., Buchanan-Smith, H. M., Brosnan, S. F., Thierry, B., Paukner, A., . . . Weiss, A. (2013). Personality structure in brown capuchin monkeys (*Sapajus apella*): Comparisons with chimpanzees (*Pan troglodytes*), orangutans (*Pongo spp.), and rhesus macaques (Macaca mulatta*). *Journal of Comparative Psychology, 127,* 282–298. http://dx.doi.org/10.1037/a0031723

Neumann, C., Agil, M., Widdig, A., & Engelhardt, A. (2013). Personality of wild male crested macaques (*Macaca* nigra). *PLoS ONE, 8,* e69383. http://dx.doi.org/10.1371/journal.pone.0069383

Ostner, J., Heistermann, M., & Schülke, O. (2008). Dominance, aggression and physiological stress in wild male Assamese macaques (*Macaca* assamensis). *Hormones and Behavior, 54,* 613–619. http://dx.doi.org/10.1016/j.yhbeh.2008.05.020

Ostner, J., & Schülke, O. (2014). The evolution of social bonds in primate males. *Behaviour, 151,* 871–906. http://dx.doi.org/10.1163/1568539X-00003191

Paul, A., Kuester, J., & Arnemann, J. (1996). The sociobiology of male-infant interactions in Barbary macaques, *Macaca sylvanus*. *Animal Behaviour, 51,* 155–170. http://dx.doi.org/10.1006/anbe.1996.0013

Pederson, A. K., King, J. E., & Landau, V. I. (2005). Chimpanzee (*Pan troglodytes*) personality predicts behavior. *Journal of Research in Personality, 39,* 534–549. http://dx.doi.org/10.1016/j.jrp.2004.07.002

Penke, L., & Jokela, M. (2016). The evolutionary genetics of personality revisited. *Current Opinion in Psychology, 7,* 104–109.

Preacher, K. J., & MacCallum, R. C. (2003). Repairing tom swift's electric factor analysis machine. *Understanding Statistics, 2,* 13–43. http://dx.doi.org/10.1207/S15328031US0201_02

Pritchard, A. J., Sheeran, L. K., Gabriel, K. I., Li, J.-H., & Wagner, R. S. (2014). Behaviors that predict personality components in adult free-ranging Tibetan macaques *Macaca* thibetana. *Current Zoology, 60,* 362–372. http://dx.doi.org/10.1093/czoolo/60.3.362

Raiche, G. (2010). *nFactors: An R package for parallel analysis and non graphical solutions to the Cattell scree test*. Retrieved from https://CRAN.R-project.org/package=nFactors

R Core Team. (2017). R: A language and environment for statistical computing (Version 3.4.2.). Retrieved from https://cran.r-project.org/mirrors.html

Réale, D., Reader, S. M., Sol, D., McDougall, P. T., & Dingemanse, N. J. (2007). Integrating animal temperament within ecology and evolution. *Biological Reviews of the Cambridge Philosophical Society, 82,* 291–318. http://dx.doi.org/10.1111/j.1469-185X.2007.00010.x

Reis, H. T., Collins, W. A., & Berscheid, E. (2000). The relationship context of human behavior and development. *Psychological Bulletin, 126,* 844–872. http://dx.doi.org/10.1037/0033-2909.126.6.844

Revelle, W. (2018). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Retrieved from https://CRAN.R-project.org/package=psych

Revelle, W., & Rocklin, T. (1979). Very simple structure: An alternative procedure for estimating the optimal number of interpretable factors. *Multivariate Behavioral Research, 14,* 403–414. http://dx.doi.org/10.1207/s15327906mbr1404_2

Rouff, J. H., Sussman, R. W., & Strube, M. J. (2005). Personality traits in captive lion-tailed macaques (*Macaca silenus*). *American Journal of Primatology, 67,* 177–198. http://dx.doi.org/10.1002/ajp.20176

Schülke, O., Bhagavatula, J., Vigilant, L., & Ostner, J. (2010). Social bonds enhance reproductive success in male macaques. *Current Biology, 20,* 2207–2210. http://dx.doi.org/10.1016/j.cub.2010.10.058

Seyfarth, R. M., & Cheney, D. L. (2013). Affiliation, empathy, and the origins of theory of mind. *Proceedings of the National Academy of Sciences of the United States of America, 110,* 10349–10356. http://dx.doi.org/10.1073/pnas.1301223110

Seyfarth, R. M., Silk, J. B., & Cheney, D. L. (2012). Variation in personality and fitness in wild female baboons. *Proceedings of the National Academy of Sciences of the United States of America, 109,* 16980–16985. http://dx.doi.org/10.1073/pnas.1210780109

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428. http://dx.doi.org/10.1037/0033-2909.86.2.420

Silk, J. B., Beehner, J. C., Bergman, T. J., Crockford, C., Engh, A. L., Moscovice, L. R., . . . Cheney, D. L. (2010). Female chacma baboons form strong, equitable, and enduring social bonds. *Behavioral Ecology and Sociobiology, 64,* 1733–1747. http://dx.doi.org/10.1007/s00265-010-0986-0

Sussman, A. F., Ha, J. C., Bentson, K. L., & Crockett, C. M. (2013). Temperament in rhesus, long-tailed, and pigtailed macaques varies by species and sex. *American Journal of Primatology, 75,* 303–313. http://dx.doi.org/10.1002/ajp.22104

Thierry, B., Singh, M., & Kaumanns, W. (2004). *Macaques societies - a model for the study of social organization*. Cambridge, United Kingdom: Cambridge University Press.

Tkaczynski, P. J., Ross, C., MacLarnon, A., Mouna, M., Majolo, B., & Lehmann, J. (2018). Measuring personality in the field: An In Situ comparison of personality quantification methods in wild Barbary macaques (*Macaca sylvanus*). *Journal of Comparative Psychology*. Advance online publication. http://dx.doi.org/10.1037/com0000163

Uher, J. (2008a). Comparative personality research: Methodological approaches. *European Journal of Personality, 22,* 427–455. http://dx.doi.org/10.1002/per.680

Uher, J. (2008b). Three methodological core issues of comparative personality research. *European Journal of Personality, 22,* 475–496. http://dx.doi.org/10.1002/per.688

Uher, J. (2011). Personality in nonhuman primates: What can we learn

from human personality psychology? In A. Weiss, J. E. King, & L. Murray (Eds.), *Personality and temperament in nonhuman primates* (pp. 41–76). New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4614-0176-6_3

Uher, J. (2015). Conceiving "personality": Psychologist's challenges and basic fundamentals of the Transdisciplinary Philosophy-of-Science Paradigm for Research on Individuals Integrative. *Psychological and Behavioral Science, 49,* 398–458. http://dx.doi.org/10.1007/s12124-014-9283-1

Uher, J. (2018). Taxonomic models of individual differences: A guide to transdisciplinary approaches. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*. http://dx.doi.org/10.1098/rstb.2017.0171

Uher, J., Addessi, E., & Visalberghi, E. (2013). Contextualised behavioural measurements of personality differences obtained in behavioural tests and social observations in adult capuchin monkeys (*Cebus apella*). *Journal of Research in Personality, 47,* 427–444. http://dx.doi.org/10.1016/j.jrp.2013.01.013

Uher, J., & Asendorpf, J. B. (2008). Personality assessment in the Great Apes: Comparing ecologically valid behavior measures, behavior ratings, and adjective ratings. *Journal of Research in Personality, 42,* 821–838. http://dx.doi.org/10.1016/j.jrp.2007.10.004

Uher, J., Asendorpf, J. B., & Call, J. (2008). Personality in the behaviour of great apes: Temporal stability, cross-situational consistency and coherence in response. *Animal Behaviour, 75,* 99–112. http://dx.doi.org/10.1016/j.anbehav.2007.04.018

Uher, J., & Visalberghi, E. (2016). Observations versus assessments of personality: A five-method multi-species study reveals numerous biases in ratings and methodological limitations of standardised assessments. *Journal of Research in Personality, 61,* 61–79. http://dx.doi.org/10.1016/j.jrp.2016.02.003

Uher, J., Werner, C. S., & Gosselt, K. (2013). From observations of individual behaviour to social representations of personality: Developmental pathways, attribution biases, and limitations of questionnaire methods. *Journal of Research in Personality, 47,* 647–667. http://dx.doi.org/10.1016/j.jrp.2013.03.006

van Aken, M. A. G., & Asendorpf, J. B. (2018). Personality and peer relationships. In K. H. Rubin, W. M. Bukowski, & B. Laursen (Eds.), *Handbook of peer interactions, relationships, and groups* (pp. 159–176). New York, NY: Guilford Press.

Vazire, S., Gosling, S. D., Dickey, A. S., & Schapiro, S. J. (2007). *Measuring personality in nonhuman animals*. New York, NY: Guilford Press.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika, 41,* 321–327. http://dx.doi.org/10.1007/BF02293557

von Borell, C., Kulik, L., & Widdig, A. (2016). Growing into the self: The development of personality in rhesus macaques. *Animal Behaviour, 122,* 183–195. http://dx.doi.org/10.1016/j.anbehav.2016.10.013

Weiss, A. (2017). Exploring factor space (and other adventures) with the Hominoid Personality Questionnaire. In J. Vonk, A. Weiss, & S. A. Kuczaj (Eds.), *Personality in nonhuman animals* (pp. 19–38). New York, NY: Springer. http://dx.doi.org/10.1007/978-3-319-59300-5_2

Weiss, A., Adams, M. J., Widdig, A., & Gerald, M. S. (2011). Rhesus macaques (*Macaca mulatta*) as living fossils of hominoid personality and subjective well-being. *Journal of Comparative Psychology, 125,* 72–83. http://dx.doi.org/10.1037/a0021187

Weiss, A., Inoue-Murayama, M., Hong, K.-W., Inoue, E., Udono, T., Ochiai, T., . . . King, J. E. (2009). Assessing chimpanzee personality and subjective well-being in Japan. *American Journal of Primatology, 71,* 283–292. http://dx.doi.org/10.1002/ajp.20649

Streich, W., Widdig, A., Bercovitch, F., Nürnberg, P., & Krawczak, M. (2002). Affiliation and aggression among adult female rhesus macaques: A genetic analysis of paternal cohorts. *Behaviour, 139,* 371–391. http://dx.doi.org/10.1163/156853902760102717

Wilcox, R. R. (1994). The percentage bend correlation coefficient. *Psychometrika, 59,* 601–616. http://dx.doi.org/10.1007/BF02294395

Wolf, M., & Weissing, F. J. (2012). Animal personalities: Consequences for ecology and evolution. *Trends in Ecology and Evolution, 27,* 452–461. http://dx.doi.org/10.1016/j.tree.2012.05.001

Received September 22, 2018
Revision received June 6, 2019
Accepted June 9, 2019 ■