Sex differences in longitudinal personality stability in chimpanzees

Bruce Rawlings*[1,2,3], Emma Flynn[4], Hani Freeman[3], Lisa Reamer[3], Steven J Schapiro[3,5], Susan Lambeth[3], & Rachel L Kendal[1].

1. Durham Cultural Evolution Research Centre, Department of Anthropology, Durham University, UK.
2. Department of Psychology, University of Texas at Austin, Texas, USA.
3. National Center for Chimpanzee Care, Michale E. Keeling Center for Comparative Medicine and Research, The University of Texas MD Anderson Cancer Center, Bastrop, TX, USA
4. School of Psychology, Queen's University, Belfast, UK
5. Department of Experimental Medicine, University of Copenhagen, Copenhagen, Denmark

*Corresponding author: Bruce Rawlings: bruce.rawlings@utexas.edu

Abstract

Personality factors analogous to the Big Five observed in humans are present in the great apes. However, few studies have examined the long-term stability of great ape personality, particularly using factor-based personality instruments. Here, we assessed overall group, and individual-level, stability of chimpanzee personality by collecting ratings for chimpanzees (*N* = 50) and comparing them to ratings collected approximately 10 years previously, using the same personality scale. The overall mean scores of three of the six factors differed across the two time points. Sex differences in personality were also observed, with overall sex differences found for three traits, and males and females showing different trajectories for two further traits over the 10-year period. Regardless of sex, rank-order stability analysis revealed strong stability for dominance; individuals who were dominant at the first time point were also dominant 10 years later. The other personality factors exhibited poor to moderate rank-order stability indicating that individuals were variable in their rank-position consistency over time. As many studies assessing chimpanzee cognition rely on personality data collected several years prior to testing, these data highlight the importance of collecting current personality data when correlating them with cognitive performance.

Key words: chimpanzees; personality, longitudinal; sex-differences

Introduction

The turn of this century saw an unprecedented interest in nonhuman animal (hereafter animal) personality. Numerous animal species are now known to display consistent individual variation in behaviour across time and contexts. This individual variation is known to have a wide-ranging impact on nonhuman animals, including on measures of fitness and welfare (Dall, Houston, & McNamara, 2004; Dingemanse & Wolf, 2010; Gosling, 2001; McCowan, Rollins, & Griffith, 2014)1) and cognition (Lermite, Peneaux, & Griffin, 2016).

Understanding animal personality augments our knowledge of the origins of human personality, and comparative studies of personality help us understand development in human personality by providing non-humancentric perspectives (Weiss, Inoue-Murayama, King, Adams, & Matsuzawa, 2012). Empirical studies examining the comparability of animal and human personality afford insights about the evolutionary trajectory of specific personality traits, as cross-species similarities likely indicate evolutionarily preserved dispositions (Gosling, 2001). Chimpanzees' phylogenetic proximity to humans make them a particularly valuable study species in this context, and factor-based instruments similar to those applied to humans have convincingly been applied to chimpanzees. Such studies established that chimpanzees (and bonobos) display personality differences in traits analogous to the 'Big Five', which incorporates agreeableness, conscientiousness, extraversion, neuroticism and openness to experience (Freeman et al., 2013; King & Figueredo, 1997; Weiss et al., 2012, 2015). Moreover, ratings on these factor-based instruments predict individual differences in great ape cognition (Altschul, Wallace, Sonnweber, Tomonaga, & Weiss, 2017; Hopper et al., 2014), long-term survival (Altschul et al., 2018) and even brain structure (Latzman, Hecht, Freeman, Schapiro, & Hopkins, 2015), providing further validation of their use.

Despite the recent interest in animal personality, one topic that remains understudied - particularly in great apes - is that of personality stability over substantial time periods. Understanding whether personality remains consistent across the lifespan of great apes allows researchers to document species-specific personality maturation, and to make comparisons with the development and stability of human personality. Cross-sectional studies of great apes reveal that in chimpanzees, bonobos and gorillas, older individuals are rated as less extraverted than younger individuals (King, Weiss, & Sisco, 2008; Kuhar, Stoinski, Lukas, & Maple, 2006; Staes, Eens, Weiss, & Stevens, 2016; Weiss & King, 2015) – patterns broadly comparable with studies of human personality changes over time (Roberts, Walton, & Viechtbauer, 2006; Srivastava, John, Gosling, & Potter, 2003). Likewise, as with humans, older chimpanzees and bonobos show increased agreeableness (Dutton, 2008; King et al., 2008; Staes et al., 2016; Weiss & King, 2015) and conscientiousness, and decreased neuroticism (King et al., 2008) compared to younger individuals.

Humans and chimpanzees also show some overlap regarding sex differences in age-related variations in personality factors. For instance, in humans (Srivastava et al., 2003; Weisberg, Deyoung, & Hirsh, 2011) and chimpanzees (King et al., 2008; Weiss & King, 2015), females score higher than males on ratings of agreeableness, and show stronger age-related increases in agreeableness than males. Sex differences in personality are thought to reflect differences in sexual selection (Schmitt, Realo, Voracek, & Allik, 2008) and social factors or life events, such as status competition and cooperation (de Waal, 2000; King et al., 2008; Srivastava et al., 2003), as well as sex differences in human cultural norms and social inequality (Brandt & Henry, 2012; Wood & Eagly, 2002). Hence, while further research is needed, the above data suggest some personality factors reflect evolutionary continuity between humans and chimpanzees (Weiss & King, 2015).

Few studies have taken a longitudinal approach to measure great ape personality, particularly those using factor-based instruments analogous to the human Big Five. In a recent study, 24 chimpanzees from Gombe were rated on the Hominoid Personality Questionnaire (HPQ); a nonhuman primate-adapted version of the Big Five, plus dominance. These ratings were compared to ratings taken almost 40 years earlier with the same chimpanzees on the Emotions Profile Index (EPI) (Weiss et al., 2017). Several dimensions were significantly correlated across the two instruments and time periods. For instance, EPI ratings of trustful, aggressive and gregarious were significantly positively correlated with HPQ ratings of agreeableness, neuroticism and extraversion, respectively, while timid and depressed (EPI) were negatively correlated with openness and agreeableness (HPQ), respectively. These correlations suggest convergent validity between different measures and may indicate that some traits, such as aggressiveness and gregariousness, remained stable over time. However, it is difficult to directly assess the stability of personality traits using instruments based on different ratings systems, and this may explain why some expected correlations were not manifest (e.g., a negative correlation between distrustful and agreeableness), and some unexpected correlations appeared (e.g. between gregariousness and agreeableness).

Among captive chimpanzees, Dutton (2008) found that correlations were strong for individual traits over a three-year period for 23 chimpanzees, but for some traits (persistent, adaptable, avoids aggression, moody, socially withdrawn and fearful) stability was comparatively weak. Similarly, King et al. (2008) rated 51 chimpanzees over a mean interval of 6.8 years on an instrument containing the Big Five plus dominance, finding relative stability over the intervals, with some evidence that conscientiousness and extraversion decreased over time. As with Dutton (2008), males exhibited a stronger increase in

dominance over the study period, though females showed a stronger increase in agreeableness than males. The mixed findings and methods outlined above from longitudinal research means drawing firm conclusions, for comparison with cross-sectional data, remains difficult.

When considering behavioural measures of personality (rather than ratings), chimpanzees appear to show stability over short, intermediate and longer time points. For instance, chimpanzees displayed temporal consistencies over two-week (Uher, Asendorpf, & Call, 2008) and three-year (Massen, Antonides, Arnold, Bionda, & Koski, 2013) periods, for various experimentally induced situations (e.g., approaching novel stimuli or foods, reactions to humans, problem solving, tool use behaviours). Similarly, over a six- to eight-year period, individual differences in post-conflict consolation behaviours of captive chimpanzees remained moderately consistent (Webb, Romero, Franks, & de Waal, 2017). Further work is required, however, assessing behavioural stability over longer time points to verify these findings.

Another important reason for establishing personality consistency in animals is to assess the reliability of using previously collected personality data when testing for relationships between personality and other variables. Personality data across a range of animal species has been applied to study topics including disease immunity (Capitanio, 2011; Koolhaas, 2008; Wallis, Szabó, Erdélyi-Belle, & Kubinyi, 2018), welfare and conservation (Boissy & Erhard, 2014; Gartner & Weiss, 2018) and sociality (Koski, 2011; Massen & Koski, 2014; Planas-Sitjà, Nicolis, Sempo, & Deneubourg, 2018; von Merten, Zwolak, & Rychlik, 2017). Recently, there has been particular focus on examining whether animal personality predicts cognitive performance (for a review, see Dougherty & Guillette, 2018). Great ape studies,

using personality data collected (often several) years prior to measurement of the cognitive performance variable, have reported a relationship between personality and participation on cognitive touchscreen tasks (Altschul et al., 2017; Herrelko, Vick, & Buchanan-Smith, 2012), response to inequity (Brosnan et al., 2015), puzzle-box interaction success (Hopper et al., 2014) and interaction/success with tools and tool-use tasks (Massen et al., 2013). Although these studies highlight the importance of considering personality when drawing conclusions from cognitive experiments in general (Altschul et al., 2017; Morton, Lee, & Buchanan-Smith, 2013), it is apparent that the original personality data may not be representative of the individuals at the time of cognitive investigation.

The present study is a longitudinal assessment of stability of personality in a population of captive chimpanzees. The personality instrument used in the current study measured six personality factors based on the Big Five: agreeableness (being considerate, consoling and protective), dominance (being bold, agonistic and dominant), extraversion (being active, playful, affiliative and sociable), methodical (being goal-orientated and self-caring), openness (being curious, inventive, exploratory and intelligent) and reactivity/undependability (being manipulative, jealous, temperamental and impulsive). These are the same chimpanzees and the same personality instrument that have been examined in previous studies of the relationship between personality and cognitive behaviours (Brosnan et al., 2015; Hopper et al., 2014). Further, the chimpanzees in question are known to exhibit consistent individual differences in social learning behaviours over an overlapping 12-year period (Watson et al., 2018).

The four broad aims of this study were to: 1) provide further longitudinal data to increase knowledge, regarding great ape personality stability over time, particularly assessing factors

analogous to the Big Five. 2) Grant insights into how factors change over time among males and females, and how this compares to humans. 3) Produce richer insights into chimpanzee personality using a variety of methodical approaches to assess long-term stability. 4) Assess the suitability of drawing conclusions informed by personality data collected several years prior to cognitive testing. Based on previous studies of great apes' personality stability, we considered two main hypotheses. First, we hypothesised that personality traits would show changes over time, predicting that chimpanzees would be rated as more dominant, and less extraverted, on the later assessment than on the first (King et al., 2008; Weiss & King, 2015; Weiss, King, & Murray, 2011). Second, there would be sex differences in overall ratings and the trajectory of personality traits, predicting that a) males would be rated as more dominant and more extraverted than females (King et al., 2008; Weiss & King, 2015), and b) females would be rated as higher in openness and agreeableness than males (Weisberg et al., 2011) and c) would show an increase in agreeableness over the time period, while males would not (King et al., 2008; Weiss & King, 2015).

Methods

Subjects

We studied 50 chimpanzees (25 males) housed in multiple social groups at the National Center for Chimpanzee Care (NCCC), Bastrop, Texas. Most chimpanzees were captive-born and mother-reared and housed at the facility for the entire 10-year study period. The chimpanzees' personality was rated at two separate time points. First (T1), between April 2006-December 2008 (Freeman et al., 2013) when all participants had been housed at the facility for several years, and second (T2) between September 2015-December 2016. At the start of T1 (April 2006), chimpanzees ranged from 5.09 to 39.27 years old (M = 18.45 years, SD = 7.50), and at the start of T2 (September 2015), the chimpanzees ranged from 14.51 to

50.70 years old (*M* = 28.12 years, *SD* = 8.04). The breakdown of mean age by sexes is as follows: T1: males *M* = 18.00 (*SD* = 7.39), females *M* = 18.89 (*SD* = 7.72), T2: males *M* = 27.42 (*SD* = 7.39), females *M* = 28.82 (*SD* = 8.73).

During the approximately 10-year period between T1 and T2, some subjects traversed age categories (see Supplementary Information, SI 1.1). Specifically, at T1, four individuals were classed as juveniles, 20 as adolescents and 26 as adults (in all categories the number of males and females were exactly evenly split). At T2, all subjects were classified as adults (i.e. 16 years or older). Further, all subjects experienced changes in group dynamics (either new members added, existing members moved to other groups or deceased, and/or a combination of these). At T1, the sizes of the study groups ranged from 3 to 14 subjects (*M* = 6.33, SD = 3.00), while at T2, group sizes ranged from 8 to 10 subjects (*M* = 8.33, *SD* = 0.82). At T2, subjects were housed with a mean of 4.48 group members that differed from T1 (*SD* = 2.06, range = 1-8 different members) and with a mean of 4.55 same group members as T1 (*SD* = 3.08, range = 0-9 same members). At T1, chimpanzees came from nine groups, and made up an average of 48% of each group (range = 13-90%). AT T2, chimpanzees came from six groups and all members of all groups are included (i.e., the study sample was all members of each of the six groups).

Materials and Procedure

Personality Instrument

Chimpanzees were rated by human carestaff on a 40-item, seven-point Likert scale questionnaire developed by Freeman et al. (2013). The questionnaire measured six overall traits; agreeableness, dominance, extraversion, methodical, openness and reactivity/undependability. The scale was generated from data collected on the NCCC

chimpanzees across a two-stage process between April 2006 and December 2008 (T1). First, a broad corpus of descriptors was produced, based on chimpanzee ethograms, previous research and expert knowledge. Next, to minimise redundancy, three experts selected 41 of the items to comprise the final scale (Table 1). The trait 'predictable' was initially included in the instrument but was subsequently removed due to low reliability, leaving 40 items (Freeman et al., 2013). The six factors obtained though principle component analysis were then validated (at T1) with independently collected behavioural measurements (Freeman et al., 2013). For instance, extraversion was positively correlated with contact aggression, sexual behavior, begging, and play, while dominance was positively correlated with aggressive and displaying behaviours and negatively correlated with submissive behaviors. Agreeableness positively correlated with affiliation and negatively correlated with displace and solicit. Methodical negatively correlated with intervene, reactivity/undependability was positively associated with aggressive behaviors such as display, intervene, and sexual behavior, and was negatively associated with post-conflict affiliation. Finally, openness positively correlated with submissive and playful, and negatively correlated with proximity and social groom (for full details of the behavioural validation process, see Freeman et al. 2013). AT T2, ratings were collated and compared to the ratings collected on the same 40 item instrument approximately 10 years previously. The six factors based on Table 1 were obtained by using a process in which only the items that loaded most heavily on a particular factor were counted towards that factor (Brosnan et al., 2015; Hopper, Cronin, & Ross, 2018; Hopper et al., 2014; Reamer et al., 2014). For instance, inventive loaded most heavily on to openness, and active loaded most heavily on to extraversion and so on (for all trait-factor loadings from T1, see SI 2.2).

Insert Table 1 about here

Agreeableness	Dominance	Extraversion	Methodical	Openness	Reactivity/ Undependability
Considerate	Anxious (-)	Active	Methodical	Affectionate/Friendly	Aggressive
Protective	Bold	Affiliative	Self-Caring	Human Orientated	Autistic
	Cautious (-)	Depressed (-)		Inquisitive/Curious	Bullying
	Dependent (-)	Playful		Intelligent	Calm (-)
	Dominant	Sexual		Inventive	Deceptive
	Fearful (-)	Solitary (-)		Persistent	Defiant
	Relaxed				Eccentric
	Timid (-)				Excitable
					Impulsive
					Irritable
					Jealous
					Manipulative
					Mischievous
					Socially-inept
					Stingy
					Temperamental/Moody

Table 1: *The six personality factors with their corresponding traits, based on highest trait loadings from Freeman et al. (2013). (-) denotes negative loadings such that these traits negatively correlated with their factors, e.g., the trait 'anxious' negatively correlated with the factor Dominance. The trait 'predictable' was initially included in the instrument but was subsequently removed from the due to low reliability* (Freeman et al., 2013).

These six factors (agreeableness, dominance, extraversion, methodical, openness and reactivity/undependability) are largely comparable to the Human Big Five (agreeableness, conscientiousness, extraversion, openness to experience and neuroticism). In human research, agreeableness captures being kind, considerate and prosocial, extraversion captures being active, social and assertive, openness to experience captures being creative, curious and exploratory while neuroticism captures being emotionally unstable, temperamental and irritable. These human based factors show strong overlap with the factors agreeableness, extraversion, openness and reactivity/undependability used in this study. In human research, conscientiousness denotes being goal-orientated, organized and plan, which shows some overlap with methodical. Dominance is not typically found on measures of human

255 personality, but captures a combination of extraversion (low caution, bold, assertive), and
256 low neuroticism (low fear and anxiety).
257
258 *Personality Ratings*
259 Ratings for T1 and T2 were collected during weekly staff meetings. Raters were either care-
260 staff or supervisory staff, all of whom had worked daily with the chimpanzees for at least six
261 months. At T1, the 17 raters had worked with the chimpanzees for 6 months-21 years, and
262 rated 8 to 10 chimpanzees each week as part of a study investigating personality in a larger
263 number of the NCCC chimpanzees (Freeman et al., 2013). At T2, the 8 raters had worked
264 with the chimpanzees for 6 months-19 years and rated 3-5 chimpanzees each week. Four
265 raters were present at both T1 and T2, providing some consistency in raters across time
266 points. All raters at T1 and T2 rated all chimpanzees in this study. Raters were instructed to
267 rate chimpanzees based on their overall experience of a chimpanzees' typical behaviours and
268 interactions, rather than specific and/or recent experiences, and were explicitly instructed not
269 to discuss ratings with each other (see SI 1.2 for the questionnaire used).
270
271 There are two main approaches to measure personality consistency over time. Group-level
272 stability measures the extent to which populations of individuals change over time on
273 personality dimensions. In contrast, rank-order stability reflects the extent to which groups
274 (in this case the entire study population) of individuals maintain similar rank ordering (i.e.
275 ordinal positions) on personality dimensions over time. To assess personality stability at the
276 global and individual-levels, we examined both the mean and individual-level stability.
277

Statistical analysis

We first report the reliability of ratings for T1 and T2 separately, before reporting the mean rank, rank-order stability and individual stability data as measures of consistency over time. For reliability measures, consistent with other studies on nonhuman primate personality (Freeman et al., 2013), intra-class correlation coefficients (ICCs) are provided to give a measure of inter-rater reliability between chimpanzee care-staff on all factors, where values closer to 1 suggest stronger reliability between raters. To allow comparison with the T1 data, we use two methods, ICC (3,1), which estimates reliability ratings of one individual, and ICC (3,k), where reliability is calculated using the average of the k raters' ratings (see S2.2 for information on how ICC (3,1) and (3,k) are each calculated). Following Koo and Li (2016), we interpret ICCs as follows: less than .05 as poor reliability, 0.5-.75 as moderate, 0.75-0.9 as good and greater than 0.9 as excellent reliability.

To compare the stability of the six personality factors across the two time points, overall mean rater scores for each of the six factors (based on the highest trait loadings) were calculated for all chimpanzees (Freeman et al., 2013; Latzman et al., 2015). Specifically, each chimpanzee was given a mean score (ranging from 1-7) for each of the six factors, which was the mean score of the respective traits loading on to each of the six factors, as defined by Freeman et al. (2013). To prevent alpha inflation arising from multiple comparisons, we used a false discovery rate control (Storey, 2002), set at 10% (as recommended by McDonald, 2009), which calculates the expected proportion of false positives (rejections of the null hypotheses) from all discoveries. False discovery rate 'families' were selected to match their lines of analyses, such that overall mean rank stability reflected a family, as did both assessment of sex differences and rank-order stability analysis.

Group-level stability was assessed by comparing overall mean scores for each of the six traits at T1 and T2 such that if a mean rating of a trait changed from (for example) 4.1 to 4.6, this would represent an increase of 0.5 on the scale. Mixed effects ANOVAs were conducted; the two time points were the within-subjects independent variable, sex was the between-subjects independent variable, and personality rating was the dependent variable. We first report the main effects of whether each of the six personality factors remained stable and then, for each factor, sex differences are examined by analysing both overall main effects of sex and sex by time interactions. We finish by reporting stability of personality for males and females separately.

To assess rank-order stability, we examined intra-class correlations between individuals across the two rating periods (Dingemanse & Dochtermann, 2013; Dutton, 2008; King et al., 2008; Koski, 2011; Uher, 2013). To account for variance in ratings due to different raters rating subjects at T1 and T2, we calculated ICCs (3,k) for all raters combined (N = 50 chimpanzees), for those chimpanzees who were rated by the same raters at both time points (N = 14), and for chimpanzees (N = 36) whose raters differed at T1 and T2 (King et al., 2008).

For further analysis of individual-level stability, we also calculated the reliable change index (RCI) (Jacobson & Truax, 1991). The RCI is used to distinguish individual change that is statistically significant from change that may have occurred due to measurement error. For each individual subject, the difference in ratings from T2-T1 were compared to the distribution of change scores expected solely by measurement error (RCI = (T2 score -T1 score)/standard error of the measurement of the difference; see SI 2.1 for further information on the RCI calculation). Using a 95% confidence interval, for each factor individuals were

classified as having either ‘increased’, ‘decreased’, or stayed the ‘same’ on each factor (Pullmann, Raudsepp, & Allik, 2006).

Results

Reliability of ratings

For T1 (Freeman et al., 2013), the ICC (3,1) and (3,k) were as follows: agreeableness (0.37, 0.51), dominance (0.48, 0.64), extraversion (0.48, 0.65), methodical (0.28, 0.36), Openness (0.49, 0.63) and reactivity/undependability (0.48, 0.61), For T2, the ICC (3,1) and (3,k) were as follows: agreeableness (0.57, 0.72), dominance (0.43, 0.84), extraversion (0.24, 0.61), methodical (0.25, 0.41), openness (0.43, 0.79) and reactivity/undependability (0.37, 0.90). See SI 2.2 for the intra-class correlation coefficients values (3,1) and (3,k) for all individual traits at T1 and T2.

Mean-rating consistency

Main effects over time

Table 2 provides a breakdown of the overall mean scores for the six factors at T1 and T2. Mean scores of agreeableness ($F_{1,48} = 6.33$ $p = .015$) and reactivity/undependability ($F_{1, 48} = 54.08$, $p < .001$) decreased significantly overall from T1 to T2. There was also a significant increase in mean scores of dominance ($F_{1,48} = 43.83$, $p < .001$) from T1 to T2, whereas extraversion, methodical and openness did not differ between T1 and T2.

Insert Table 2 about here

	Overall mean scores			Males			Females		
Factor	T1	T2	T1-T2	T1	T2	T1-T2	T1	T2	T1-T2
Agreeableness	4.32 (.46)	4.11 (.62)	-.21*	4.25 (.44)	3.78 (.53)	-.48***	4.39 (4.8)	4.44 (.65)	+.05
Dominance	4.19 (.59)	4.64 (.76)	+.45***	4.46 (.53)	4.99 (.51)	+.53***	3.91 (.53)	4.29 (.80)	+.38**
Extraversion	4.78 (.47)	4.69 (.44)	-.08	4.93 (.42)	4.85 (.42)	-.08	4.61 (.47)	4.53 (.40)	-.08
Methodical	4.65 (.38)	4.51 (.67)	+.03	4.63 (.44)	4.54 (.43)	-.09	4.66 (.33)	4.70 (.42)	+.04
Openness	4.73 (.53)	4.75 (.62)	+.02	4.81 (.52)	4.65 (.65)	-.16	4.64 (.53)	4.84 (.60)	+.20
Reactivity/ Undependability	3.90 (.47)	3.33 (.59)	-.57***	4.01 (.43)	3.48 (.57)	-.53***	3.80 (.51)	3.18 (.60)	-.62***

Table 2: *Mean scores (SD) of each of the six factors at T1 (April 2006-December 2008) and T2 (September 2015-December 2016), overall and for males and females. Mean-order stability demonstrates the group-level T1 and T2 scores (on a scale of 1-7) and change over the ten-year time point for each factor. Significant differences between T1 and T2 indicated as * $p < .05$, ** $p < .01$ and *** $p < .001$.*

Sex differences

Table 2 provides a breakdown of the overall mean scores for the six factors at T1 and T2 for males and females. To examine sex differences in personality, we looked at main effects of sex, time by sex interactions and where appropriate, within-sex effects for each factor.

Agreeableness: Males were rated as significantly less agreeable than females across T1 and T2 combined ($F_{1, 48}$, = 10.63, p = .002). There was also a significant interaction between time and sex (Figure 1), such that males exhibited a decrease of 0.48 and females displayed a slight increase of 0.05 ($F_{1, 48}$ = 9.77, p = .003). The decrease in male agreeableness from T1 to T2 was significant ($F_{1, 24}$ = 20.41, p < .001) but the increase in females was not.

Dominance: Males were rated as more dominant than females across T1 and T2 combined ($F_{1, 48}$ = 9.74, p < .001). There was no significant interaction between time and sex but both male and female ratings of dominance increased significantly (males: $F_{1, 24}$ = 57.23, p < .001;

370 females: $F_{1, 24} = 10.12$, $p = .004$). *Extraversion*: Males were rated as more extraverted than
371 females across T1 and T2 combined ($F_{1, 48} = 9.53$, $p = .003$). There was no sex by time
372 interaction, nor did male or female ratings differ between T1 and T2. *Openness*: There was
373 no main effect of sex but there was a sex by time interaction ($F_{1, 48} = 4.67$, $p = .036$) such that
374 males exhibited a decrease of 0.16 and females an increase of 0.20 from T1 to T2. The
375 decrease in male openness only approached significance ($F_{1, 24} = 4.02$, $p = .056$), while the
376 increase in females was not significant. *Reactivity/undependability*: There was no main effect
377 of sex or a sex by time interaction. However, ratings decreased significantly from T1 to T2
378 for both sexes (males: $F_{1, 24} = 24.46$, $p < .001$; females: $F_{1, 24} = 32.14$, $p < .001$). *Methodical*:
379 There were no significant effects.
380
381 To assess whether individuals changed more within or between age category, we conducted
382 additional analysis looking at time by age category interactions. Although small sample sizes
383 preclude making firm conclusions, no time by age category interactions were significant (all
384 *p*s >.05; see SI Table 4 for means for T1 and T2, and T2-T1 by age category for each factor).
385
386 *Insert Table 3 about here*

	Rank order stability (Raters)			Reliable change index (RCI) (%)		
Factor	All (n = 50)	Same (n = 14)	Different (n = 36)	Increased	Same	Decreased
Agreeableness	.535	.551	.309	2	88	10
Dominance	.854	.529	.551	18	82	0
Extraversion	.712	.479	.589	2	98	0
Methodical	.493	.515	.025	4	94	2
Openness	.596	.493	.824	10	82	8
Reactivity/ Undependability	.661	.631	.559	2	42	56

Table 3: *Overview of results from individual analyses. For rank-order stability, 'All' represents ICC correlations for all raters combined (N = 50 chimpanzees), 'same' represents chimpanzees who were rated by the same raters at both time points (N = 14), and 'different' represents chimpanzees whose raters differed at T1 and T2. The reliable change index (RCI) provides the percentage of individuals that significantly increased, stayed the same or decreased over the study period according to the RCI calculation.*

Insert Figure 1 about here

Figure 1*: Results revealed significant sex by time interactions for agreeableness (A) and openness (B).*

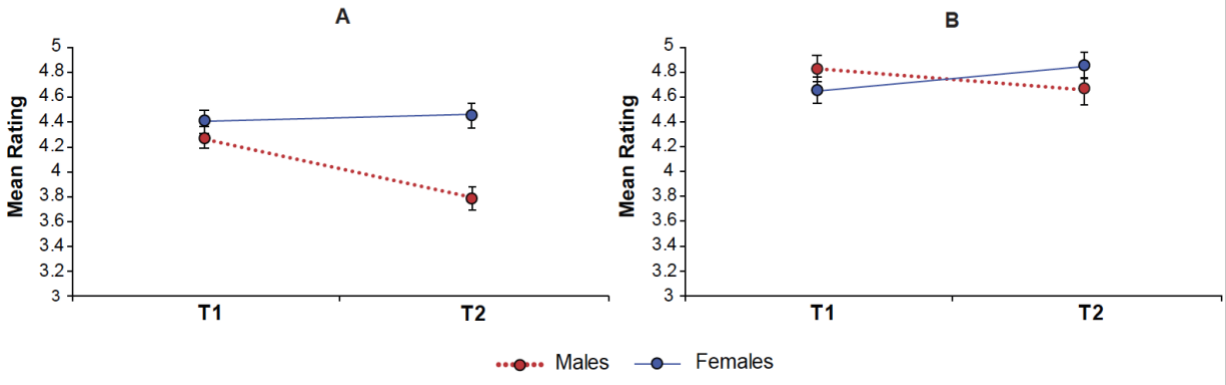


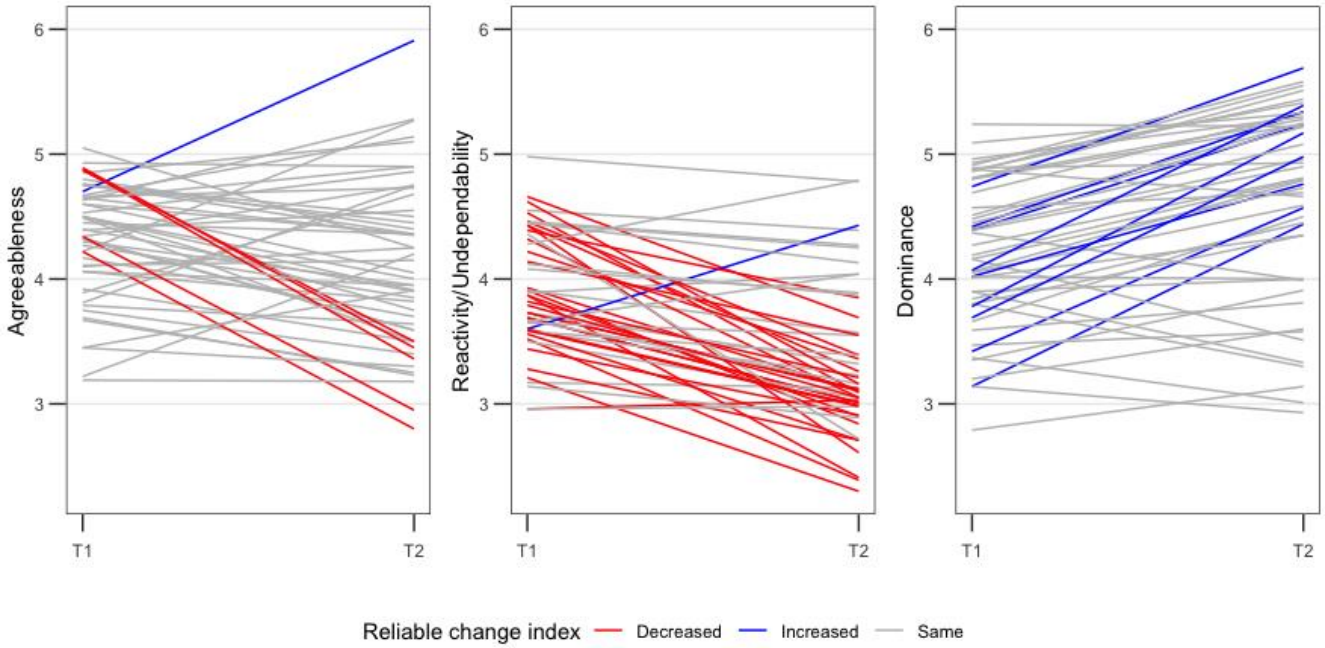
Figure 1: Results revealed significant sex by time interactions for agreeableness (A) and openness (B).

Rank-order stability

Table 3 presents the rank-order stability results. When all raters were combined, dominance (ICC 3,k = .854) showed the highest (good) intra-class correlation coefficient between T1 and T2, and methodical (ICC 3,k = .493), showed the lowest (poor) rank-order stability. The other four factors all showed moderate rank-order stability (ICC 3,k range = .535-.712), suggesting individuals were variable in their rank-order position over time. For four of six factors, ICCs were stronger when analysis was restricted to raters who were present at both time points (ICC 3,k range: = .479-.631) compared to the case where raters differed (ICC 3,k range = .025-.824).

Reliable Change Index (RCI)

The extent to which individuals' change (T1 to T2) was over the RCI threshold varied by factor. Reactivity/undependability was the factor for which most individuals changed, with 58% passing the RCI threshold in either direction. Dominance (18%), openness (18%) and agreeableness (12%) showed lower individual-level change, and methodical (6%) and extraversion (2%) showed the lowest rates of individual change over time. Thus, while reactivity/undependability, agreeableness and dominance all showed overall (group) mean level change from T1 to T2, only for reactivity/undependability did the majority (and by a small margin) of individuals show significant change according to the RCI. Table 3 presents the group RCI scores and Table 4 presents RCI group by the sexes, and for a full breakdown of RCI scores by age category see Table SI 5.



Insert Figure 2 about here

Figure 2: *Individual RCI values for agreeableness, dominance and reactivity/undependability, which all showed significant mean level change over time. Red lines show individuals whose RCI value significantly decreased, blue lines indicate individuals whose RCI value significantly increased and grey lines indicate individual's whose RCI value did not change significantly over the time points.*

Insert Table 4 about here

	Reliable change index (RCI) (%)					
	Males			**Females**		
Factor	Increased	Same	Decreased	Increased	Same	Decreased
Agreeableness	2	90	8	0	98	2
Dominance	10	90	0	8	92	0
Extraversion	2	98	0	0	100	0
Methodical	0	98	2	4	96	0
Openness	2	94	4	8	88	4
Reactivity/ Undependability	2	70	28	0	72	28

Table 4: *Breakdown of Reliable Change Index scores by sexes.*

Discussion

Stability of chimpanzees' personality over time; overall and sex differences

We examined the stability of multiple chimpanzee personality traits by measuring changes in factors across an approximately 10-year period using the same instrument, revealing consistencies and differences with previous work. Analysis of mean rank stability revealed that consistent with previous findings and with our prediction, overall, chimpanzees showed increased dominance with age. Approximately half of the study subjects traversed age categories during the study period (predominantly moving from adolescence into adulthood). Our findings largely fit cross-sectional data on personality development showing that adult chimpanzees are more dominant than juvenile and younger chimpanzees (King et al., 2008) - although we note that no age category by time interactions were found for any of the six factors. Contrary to our prediction, chimpanzees did not show an overall significant decrease in extraversion over time. The chimpanzees were also rated as significantly less reactive/undependable over time - a finding that was also not predicted.

Analyses of sex differences in personality traits also indicated that males and females differed for agreeableness, openness, dominance and extraversion. In line with our predictions, males were rated as more dominant and extraverted than females, and females showed an increase in agreeableness and openness while males did not. The finding that males actually decreased in these factors was not, however, predicted. Further, in contrast to other cross-sectional findings, the chimpanzees were rated as less agreeable over time - though this decrease was driven by males.

Comparisons with sex differences in chimpanzee and human personality stability

The results revealed some sex differences in personality traits that contrast with previous chimpanzee studies but correspond with findings in humans. For instance, King et al. (2008) found that chimpanzees decline in openness with age, whereas in the present study, while males significantly declined in openness, females increased by a similar margin. Although these findings contrast with those of King and colleagues, they are consistent with findings that human females score higher on openness to experience than males (Weisberg et al., 2011) and that this pattern continues throughout development (Gjerde & Cardilla, 2009). In humans, females score particularly high on the facets of warmth, openness to feelings and aesthetics (Chapman, Duberstein, Sörensen, & Lyness, 2007; Costa, Terracciano, & McCrae, 2001). It is important to note, however, that vast majority of longitudinal studies of personality in humans are based on Western populations, for whom our human comparisons are based on (and thus limited to), and further research is needed to measure cultural differences in personality stability over time (see Costa, McCrae, & Löckenhoff, 2019 and SI3 for additional discussion on variance and invariance of personality across age groups and cultures).

One potential explanation for the contrasting findings between this study and others is that, the present study, unlike King et al.'s included affectionate/friendly as a facet of openness, and thus could contribute to the sex differences found here. Similarly, intelligent and persistent loaded on to openness for our instrument, while both loaded on to dominance in King et al.'s (for a breakdown of the traits used in both studies, see SI Tables 6 and 7). However, despite these there is also large overlap between the two instruments. For instance, there was large similarity in the traits loading on to the factors agreeableness (protective, kind), dominance (non-fearful and non-submissive, dominant), extraversion (sociable, affiliative, playful, non-depressed) and openness (inventive, inquisitive).

Reactivity/Undependability also showed overlap with King et al.'s consciousnesses (irritable, jealous, impulsive).

Similarly, it is important to consider differences in age categories, group composition and environmental factors when comparing these data with those of other studies, particularly breeding populations. Here, all study subjects experienced changes in group members and group sizes across the study period, and many experienced relocations to new enclosures (on-site). Personality has been shown to correlate with individual differences in stress response in young chimpanzees (Anestis, Bribiescas, & Hasselschwert, 2006) and it has been found that nonhuman primate social dynamics including individual and group level affiliative and aggressive behaviours are disrupted by enclosure relocation and changes to group demographics (Dufour, Sueur, Whiten, & Buchanan-Smith, 2011; Schel et al., 2013), but that such behaviours and group dynamics begin to return to pre-disruption levels within a year (Schel et al., 2013; Yamanashi et al., 2016). Given there were no major alterations to group demographics or relocations for the study subjects for several years prior to the second data collection period, it is not clear whether the effects of relocation had a major bearing on ratings.

In turn, these findings can contribute to the development of a theoretical framework in which to empirically examine specific hypotheses about chimpanzee personality over time, particularly with regards to ecological and life history changes. For example, future research could examine how individuals high or low in social based traits such as dominance, agreeableness and extraversion are shaped by adjustments to group dynamics. Tools such as social network analysis have proven useful for helping facilitate and monitor the integration of different groups or relocation of nonhuman primates (Dufour et al., 2011; Schel et al.,

2013) and chimpanzees display ‘friendships’ based on personality homophily (Massen & Koski, 2014). Thus personality instruments may be an important tool for group formations or relocations (Schapiro, 2017). Further, given that these data indicated that reactivity/undependability showed high levels of mean and rank-order decreases, it may be that individuals high in this trait exhibit lower stability over time than those scoring low in it. These questions would be well suited to longitudinal personality data over multiple time points. Such data, coupled with documentation of major events, including changes to social environments, would allow these types of assessments, and in turn comparisons with analogous human data (Ying & Han, 2006).

Likewise, evolutionary theory suggests that if changes in personality over time are an evolutionary preserved feature of chimpanzees there should be corresponding fitness benefits (Blaszczyk, 2020). While extraversion itself has been linked with longer survival in wild gorillas (Weiss, Gartner, Gold, & Stoinski, 2013), there has been a striking lack of empirical research assessing fitness benefits of nonhuman animal personality instability (Blaszczyk, 2020; Trillmich, Müller, & Müller, 2018). It is possible, for examples that females - who are the socially dispersing sex in chimpanzees - become more agreeable over adulthood to maximise social bonds. It is important for evolutionary models of personality, for researchers to document the association between changes in nonhuman animal personality over time with fitness benefits so such hypotheses can be tested.

Our findings also afford comparisons with other ape species and humans. Assessments of bonobo personality has shown both overlap and differences with human and chimpanzee personality data. For example, while there are similarities in the factors found in bonobos, there are contrasting patterns of sex differences to chimpanzees and humans. Female bonobos

score higher on traits such as assertive and extraversion than male bonobos and receive less aggression (Staes et al., 2016). Higher female assertiveness and extraversion reflects the fact that, unlike chimpanzees, they are more socially dominant and maintain close relationships with other group members compared to male bonobos (Staes et al., 2016; Vervaecke, De Vries, & Van Elsacker, 2000). Similarly, as with humans, orangutans - for whom factor based personality traits have also been validated - show age related declines in extraversion and neuroticism. Male orangutans, like chimpanzees, also score higher in dominance than females (Weiss & King, 2015). Comparisons across different ape species are crucial for understanding evolutionary continuity of personality (Weiss & King, 2015).

Individual-level change over time: multiple approaches to assessing long-term stability

Investigation of rank-order stability revealed comparatively strong stability for ratings of dominance; individuals who were rated as scoring highly in the factor dominance at T1 were also rated as scoring highly in the factor dominance 10 years later. This finding is perhaps expected: dominance exhibited the strongest rank order stability in other studies (e.g., King et al., 2008b). Extraversion also exhibited relatively high rank-order stability compared to the other traits, also suggesting that individuals high (or low) remained high (or low) in this factor. The other four traits overall exhibited lower rank-order stability, indicating that individuals were variable in their ordinal rank-position consistency when compared at T1 and T2. That methodical displayed the least rank-order consistency (regardless of whether the raters were the same, different or combined) is not surprising. In the initial study by Freeman and colleagues (Freeman et al., 2013), showed methodical to have the lowest reliability and failed to correlate with factors from other instruments measuring chimpanzee personality (and caution should be exercised when interpreting from this factor, as noted by Freeman and colleagues in the initial study).

When assessing individual-level change using the reliable change index (RCI), despite overall mean changes in dominance, agreeableness and reactivity/undependability, only in the latter trait did most individuals exhibit a change that was considered 'reliable'. For dominance and agreeableness under 20% of individuals exhibited a statistically significant change over time. This may be because reactivity/undependability included traits such as being excitable, impulsive, aggressive, mischievous, eccentric and calm (negatively loaded) – all traits that perhaps change to a greater extent as subjects traverse age categories to those within dominance and agreeableness.

An understanding of individual-level changes occurring over time compliments our understanding of population changes. Population-level changes of personality may either be driven by a subset of individuals or represent a general group-level trend in change over time (or a combination of both). Discrepancies between population-level and individual-level changes over time have important implications for future research and the conclusions that can be drawn from longitudinal assessments of personality. First, researchers should be cautious when drawing conclusions about population-level changes in personality over time. Although data may indicate that personality may significantly change overtime at the population level, this may be driven by certain individuals. Second, presenting individual and population data on all subjects is important to provide a complete picture of the data and how personality changes over time – an approach taken in very few studies. Third, in line with studies with other nonhuman animals, these findings may indicate that key individuals, in terms of personality scores, may have significant impact on group behaviours (Aplin et al., 2013; Brown & Irving, 2014; Farine, Montiglio, & Spiegel, 2015)

In addition to providing insights regarding how group and individual-level changes in personality interact, our findings build on the existing, yet limited, longitudinal data using factor-based instruments to assess chimpanzee personality. For instance, despite increasing the time scale compared to King et al. (2008) (6.8 years versus 10 years here), when all raters were combined, most of the correlation coefficients are similar to those obtained in their study: 0.85 vs. 0.74 for dominance; 0.66 vs. 0.51 for reactivity/undependability vs. dependability/conscientiousness; 0.60 vs. 0.70 for openness; 0.54 vs. 0.39 for agreeableness; and 0.71 vs. 0.48 for extraversion. Further, as with King et al. (2008) at least half of the traits studied exhibited higher correlation coefficients for data from raters who were present at both time points compared to data from raters who differed. Such closely matched coefficients and findings are indicative of robust validity in findings across measures and chimpanzee populations.

Drawing conclusions based on personality data collected years prior to cognitive testing. These findings also have implications for the use of personality ratings obtained prior to other types of empirical tests (e.g. cognitive assessments). For example, much recent work has highlighted the importance of openness in chimpanzee problem solving, study participation and success (Altschul et al., 2017; Herrelko et al., 2012; Hopper et al., 2014), and performance on inequity tasks (Brosnan et al., 2015). These studies relied on the personality ratings collected several years prior to the cognitive testing sessions, and indeed, two of these studies used the same subjects and same personality instrument as this study (Brosnan et al., 2015; Hopper et al., 2014). Here, we found that males significantly decreased in openness over several years, while female ratings increased by a similar (although non-significant) margin. This may suggest, depending on the timeframe between rating collection and experimental testing, that the personality ratings may not always accurately reflect the

individuals at the time of study participation. Although rating data requires much effort and valuable time from care-staff, we encourage, where possible, 1) authors use or collect recent personality data when conducting personality-based assessments of cognitive performance or other empirical measurements, or 2) researchers consider temporal instability in personality measures when drawing conclusions regarding the predictive power of personality for cognitive measures.

Our data revealed important insights regarding stability in chimpanzee personality over an approximately 10-year period. We found group-level changes in three of six personality factors measured (an increase in dominance and decreases in agreeableness and reactivity/undependability), overall sex differences found for three traits (males rated higher than females in dominance and extraversion but lower in agreeableness), and males and females showing different trajectories for two further traits (males decreasing and females increasing in agreeableness and openness) over the 10-year period. Given that several personality factors showed group level changes *and* variable individual stability over time we suggest, researchers measuring the relationship between personality and cognitive performance in nonhuman primates obtain the most current personality data possible. The reported sex differences converge with studies of Western humans, providing new longitudinal evidence for an evolutionary basis for the human pattern of age-related fluctuations in male and female personality traits. In turn, these findings lay the foundation of an exciting suite of questions about how environmental and social changes influences chimpanzees with specific personality profiles, and how this compares to data on human personality and environmental and social changes.

Ethics statement

This study was approved by the UTMDACC Institutional Animal Care and Use Committee (IACUC approval number 0894-RN01).

References

Altschul, D. M., Hopkins, W. D., Herrelko, E. S., Inoue-Murayama, M., Matsuzawa, T., King, J. E., … Weiss, A. (2018). Personality links with lifespan in chimpanzees. *ELife*, *7*. https://doi.org/10.7554/eLife.33781

Altschul, D. M., Wallace, E. K., Sonnweber, R., Tomonaga, M., & Weiss, A. (2017). Chimpanzee intellect: personality, performance and motivation with touchscreen tasks. *Royal Society Open Science*, *4*(5), 170169. https://doi.org/10.1098/rsos.170169

Anestis, S. F., Bribiescas, R. G., & Hasselschwert, D. L. (2006). Age, rank, and personality effects on the cortisol sedation stress response in young chimpanzees. *Physiology & Behavior*, *89*(2), 287–294. https://doi.org/10.1016/j.physbeh.2006.06.010

Aplin, L. M., Farine, D. R., Morand-Ferron, J., Cole, E. F., Cockburn, a, & Sheldon, B. C. (2013). Individual personalities predict social behaviour in wild networks of great tits (Parus major). *Ecology Letters*, *16*(11), 1365–1372. https://doi.org/10.1111/ele.12181

Blaszczyk, M. B. (2020). Primates got personality, too: Toward an integrative primatology of consistent individual differences in behavior. *Evolutionary Anthropology: Issues, News, and Reviews*, *29*(2), 56–67. https://doi.org/10.1002/evan.21808

Boissy, A., & Erhard, H. W. (2014). How Studying Interactions Between Animal Emotions, Cognition, and Personality Can Contribute to Improve Farm Animal Welfare. In *Genetics and the Behavior of Domestic Animals* (pp. 81–113). Elsevier Inc. https://doi.org/10.1016/B978-0-12-394586-0.00003-2

Brandt, M. J., & Henry, P. J. (2012). Gender Inequality and Gender Differences in Authoritarianism. *Personality and Social Psychology Bulletin*, *38*(10), 1301–1315. https://doi.org/10.1177/0146167212449871

Brosnan, S. F., Hopper, L. M., Richey, S., Freeman, H., Talbot, C. F., Gosling, S. D., … Schapiro, S. J. (2015). Personality influences responses to inequity and contrast in chimpanzees. *Animal Behaviour*, *101*, 75–87. https://doi.org/10.1016/J.ANBEHAV.2014.12.019

Brown, C., & Irving, E. (2014). Individual personality traits influence group exploration in a feral guppy population. *Behavioral Ecology*, *25*(1), 95–101. https://doi.org/10.1093/beheco/art090

Capitanio, J. P. (2011). Nonhuman Primate Personality and Immunity: Mechanisms of Health and Disease. In *Personality and Temperament in Nonhuman Primates* (pp. 233–255). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4614-0176-6_9

Chapman, B. P., Duberstein, P. R., Sörensen, S., & Lyness, J. M. (2007). Gender differences in Five Factor Model personality traits in an elderly cohort. *Personality and Individual Differences*, *43*(6), 1594–1603. https://doi.org/10.1016/j.paid.2007.04.028

Costa, P. T., McCrae, R. R., & Löckenhoff, C. E. (2019). Personality Across the Life Span. *Annual Review of Psychology*, *70*(1), 423–448. https://doi.org/10.1146/annurev-psych-010418-103244

Costa, P. T., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, *81*(2), 322–331. https://doi.org/10.1037/0022-3514.81.2.322

Dall, S. R. X., Houston, A. I., & McNamara, J. M. (2004). The behavioural ecology of personality: Consistent individual differences from an adaptive perspective. *Ecology Letters*, *7*(8), 734–739. https://doi.org/10.1111/j.1461-0248.2004.00618.x

de Waal, F. B. M. (2000). *Chimpanzee politics: Power and sex among apes*. Johns Hopkins University Press. https://doi.org/10.1037/022265

Dingemanse, N. J., & Dochtermann, N. A. (2013). Quantifying individual variation in

687 behaviour: Mixed-effect modelling approaches. *Journal of Animal Ecology*, *82*(1), 39–
688 54. https://doi.org/10.1111/1365-2656.12013
689 Dingemanse, N. J., & Wolf, M. (2010). Recent models for adaptive personality differences: a
690 review. *Philosophical Transactions of the Royal Society of London B: Biological*
691 *Sciences*, *365*(1560).
692 Dougherty, L. R., & Guillette, L. M. (2018, September 26). Linking personality and cognition:
693 A meta-analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*.
694 Royal Society Publishing. https://doi.org/10.1098/rstb.2017.0282
695 Dufour, V., Sueur, C., Whiten, A., & Buchanan-Smith, H. M. (2011). The impact of moving
696 to a novel environment on social networks, activity and wellbeing in two new world
697 primates. *American Journal of Primatology*, *73*(8), 802–811.
698 https://doi.org/10.1002/ajp.20943
699 Dutton, D. M. (2008). Subjective assessment of chimpanzee (Pan troglodytes) personality:
700 reliability and stability of trait ratings. *Primates*, *49*(4), 253–259.
701 https://doi.org/10.1007/s10329-008-0094-1
702 Farine, D. R., Montiglio, P. O., & Spiegel, O. (2015). From Individuals to Groups and Back:
703 The Evolutionary Implications of Group Phenotypic Composition. *Trends in Ecology and*
704 *Evolution*. https://doi.org/10.1016/j.tree.2015.07.005
705 Freeman, H. D., Brosnan, S. F., Hopper, L. M., Lambeth, S. P., Schapiro, S. J., & Gosling, S.
706 D. (2013). Developing a comprehensive and comparative questionnaire for measuring
707 personality in chimpanzees using a simultaneous top-down/bottom-up design. *American*
708 *Journal of Primatology*, *75*(10), 1042–1053. https://doi.org/10.1002/ajp.22168
709 Gartner, M. C., & Weiss, A. (2018). Studying primate personality in zoos: implications for the
710 management, welfare and conservation of great apes. *International Zoo Yearbook*, *52*(1),
711 79–91. https://doi.org/10.1111/izy.12187

712 Gjerde, P. F., & Cardilla, K. (2009). Developmental implications of openness to experience in
713 preschool children: gender differences in young adulthood. *Developmental Psychology*,
714 *45*(5), 1455–1464. https://doi.org/Doi 10.1037/A0016714
715 Gosling, S. D. (2001). From mice to men: What can we learn about personality from animal
716 research? *Psychological Bulletin*, *127*(1), 45–86. https://doi.org/10.1037/0033-
717 2909.127.1.45
718 Herrelko, E. S., Vick, S.-J. J., & Buchanan-Smith, H. M. (2012). Cognitive research in zoo-
719 housed chimpanzees: Influence of personality and impact on welfare. *American Journal*
720 *of Primatology*, *74*(9), 828–840. https://doi.org/10.1002/ajp.22036
721 Hopper, L. M., Cronin, K. A., & Ross, S. R. (2018). A multi-institutional assessment of a short-
722 form personality questionnaire for use with macaques. *Zoo Biology*, *37*(5), 281–289.
723 https://doi.org/10.1002/zoo.21439
724 Hopper, L. M., Price, S. a, Freeman, H., Lambeth, S. P., Schapiro, S. J., & Kendal, R. (2014).
725 Influence of personality, age, sex, and estrous state on chimpanzee problem-solving
726 success. *Animal Cognition*, *17*(4), 835–847. https://doi.org/10.1007/s10071-013-0715-y
727 Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining
728 meaningful change in psychotherapy research. *Journal of Consulting and Clinical*
729 *Psychology*, *59*(1), 12–19. https://doi.org/10.1037/0022-006X.59.1.12
730 King, J. E., & Figueredo, A. J. (1997). The five-factor model plus dominance in chimpanzee
731 personality. *Journal of Research in Personality*, *31*(2), 257–271.
732 https://doi.org/10.1006/jrpe.1997.2179
733 King, J. E., Weiss, A., & Sisco, M. M. (2008). Aping Humans: Age and Sex Effects in
734 Chimpanzee (Pan troglodytes) and Human (Homo sapiens) Personality. *Journal of*
735 *Comparative Psychology*, *122*(4), 418–427. https://doi.org/10.1037/a0013125
736 Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation

737 Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163.
738 https://doi.org/10.1016/j.jcm.2016.02.012
739 Koolhaas, J. M. (2008, July 1). Coping style and immunity in animals: Making sense of
740 individual variation. *Brain, Behavior, and Immunity*. Academic Press.
741 https://doi.org/10.1016/j.bbi.2007.11.006
742 Koski, S. E. (2011). Social personality traits in chimpanzees: Temporal stability and structure
743 of behaviourally assessed personality traits in three captive populations. *Behavioral*
744 *Ecology and Sociobiology*, *65*(11), 2161–2174. https://doi.org/10.1007/s00265-011-
745 1224-0
746 Kuhar, C. W., Stoinski, T. S., Lukas, K. E., & Maple, T. L. (2006). Gorilla Behavior Index
747 revisited: Age, housing and behavior. *Applied Animal Behaviour Science*, *96*(3–4), 315–
748 326. https://doi.org/10.1016/j.applanim.2005.06.004
749 Latzman, R. D., Hecht, L. K., Freeman, H. D., Schapiro, S. J., & Hopkins, W. D. (2015).
750 Neuroanatomical correlates of personality in chimpanzees (Pan troglodytes): Associations
751 between personality and frontal cortex. *NeuroImage*, *123*, 63–71.
752 https://doi.org/10.1016/j.neuroimage.2015.08.041
753 Massen, J. J. M., Antonides, A., Arnold, A.-M. K., Bionda, T., & Koski, S. E. (2013). A
754 behavioral view on chimpanzee personality: exploration tendency, persistence, boldness,
755 and tool-orientation measured with group experiments. *American Journal of Primatology*,
756 *75*(9), 947–958. https://doi.org/10.1002/ajp.22159
757 Massen, J. J. M., & Koski, S. E. (2014). Chimps of a feather sit together: chimpanzee
758 friendships are based on homophily in personality. *Evolution and Human Behavior*, *35*(1),
759 1–8. https://doi.org/10.1016/j.evolhumbehav.2013.08.008
760 McCowan, L. S. C., Rollins, L. A., & Griffith, S. C. (2014). Personality in captivity: more
761 exploratory males reproduce better in an aviary population. *Behavioural Processes*, *107*,

150–157. https://doi.org/10.1016/j.beproc.2014.08.020

McDonald, J. (2009). *Handbook of Biological Statistics*. Sparky House Publishing.

Morton, F. B., Lee, P. C., & Buchanan-Smith, H. M. (2013). Taking personality selection bias seriously in animal cognition research: A case study in capuchin monkeys (Sapajus apella). *Animal Cognition*, *16*(4), 677–684. https://doi.org/10.1007/s10071-013-0603-5

Planas-Sitjà, I., Nicolis, S. C., Sempo, G., & Deneubourg, J.-L. (2018). The interplay between personalities and social interactions affects the cohesion of the group and the speed of aggregation. *PLOS ONE*, *13*(8), e0201053. https://doi.org/10.1371/journal.pone.0201053

Pullmann, H., Raudsepp, L., & Allik, J. (2006). Stability and change in adolescents' personality: A longitudinal study. *European Journal of Personality*, *20*(6), 447–459. https://doi.org/10.1002/per.611

Reamer, L. A., Haller, R. L., Thiele, E. J., Freeman, H. D., Lambeth, S. P., & Schapiro, S. J. (2014). Factors affecting initial training success of blood glucose testing in captive chimpanzees (Pan troglodytes). *Zoo Biology*, *33*(3), 212–220. https://doi.org/10.1002/zoo.21123

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *132*(1), 1–25. https://doi.org/10.1037/0033-2909.132.1.1

Schapiro, S. J. (2017). *Handbook of primate behavioral management* (1st ed.). Boca Raton, Florida, United States: CRC Press.

Schel, A. M., Rawlings, B., Claidière, N., Wilke, C., Wathan, J., Richardson, J., … Slocombe, K. (2013). Network analysis of social changes in a captive chimpanzee community following the successful integration of two adult groups. *American Journal of Primatology*, *75*(3), 254–266. https://doi.org/10.1002/ajp.22101

Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a

woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology*, *94*(1), 168–182. https://doi.org/10.1037/0022-3514.94.1.168

Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in early and middle adulthood: set like plaster or persistent change? *Journal of Personality and Social Psychology*, *84*(5), 1041–1053. https://doi.org/10.1037/0022-3514.84.5.1041

Staes, N., Eens, M., Weiss, A., & Stevens, J. M. G. (2016). Bonobo personality: age and sex effects and links with behavior and dominance. *Bonobos: Unique in Mind, Brain and Behavior*.

Storey, J. D. (2002). A direct approach to false discovery rates. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*. Wiley Royal Statistical Society. https://doi.org/10.2307/3088784

Trillmich, F., Müller, T., & Müller, C. (2018). Understanding the evolution of personality requires the study of mechanisms behind the development and life history of personality traits. *Biology Letters*, *14*(2), 20170740. https://doi.org/10.1098/rsbl.2017.0740

Uher, J. (2013). Personality Psychology: Lexical Approaches, Assessment Methods, and Trait Concepts Reveal Only Half of the Story—Why it is Time for a Paradigm Shift. *Integrative Psychological and Behavioral Science*, *47*(1), 1–55. https://doi.org/10.1007/s12124-013-9230-6

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. https://doi.org/10.1016/j.anbehav.2007.04.018

Vervaecke, H., De Vries, H., & Van Elsacker, L. (2000). Dominance and its behavioral measures in a captive group of bonobos (Pan paniscus). *International Journal of Primatology*, *21*(1), 47–68. https://doi.org/10.1023/A:1005471512788

812 von Merten, S., Zwolak, R., & Rychlik, L. (2017). Social personality: a more social shrew
813 species exhibits stronger differences in personality types. *Animal Behaviour*, *127*, 125–
814 134. https://doi.org/10.1016/j.anbehav.2017.02.021
815 Wallis, L. J., Szabó, D., Erdélyi-Belle, B., & Kubinyi, E. (2018). Demographic change across
816 the lifespan of pet dogs and their impact on health status. *Frontiers in Veterinary Science*,
817 *5*(AUG), 200. https://doi.org/10.3389/fvets.2018.00200
818 Watson, S. K., Vale, G. L., Hopper, L. M., Dean, L. G., Kendal, R., Price, E. E., … Whiten, A.
819 (2018). Chimpanzees demonstrate individual differences in social information use.
820 *Animal Cognition*, *21*(5), 639–650. https://doi.org/10.1007/s10071-018-1198-7
821 Webb, C. E., Romero, T., Franks, B., & de Waal, F. B. M. (2017). Long-term consistency in
822 chimpanzee consolation behaviour reflects empathetic personalities. *Nature*
823 *Communications*, *8*(1), 292. https://doi.org/10.1038/s41467-017-00360-7
824 Weisberg, Y. J., Deyoung, C. G., & Hirsh, J. B. (2011). Gender differences in personality
825 across the ten aspects of the Big Five. *Frontiers in Psychology*, *2*, 178.
826 https://doi.org/10.3389/fpsyg.2011.00178
827 Weiss, A., Gartner, M. C., Gold, K. C., & Stoinski, T. S. (2013). Extraversion predicts longer
828 survival in gorillas: an 18-year longitudinal study. *Proceedings. Biological Sciences / The*
829 *Royal Society*, *280*(1752), 20122231. https://doi.org/10.1098/rspb.2012.2231
830 Weiss, A., Inoue-Murayama, M., King, J. E., Adams, M. J., & Matsuzawa, T. (2012). All too
831 human? Chimpanzee and orang-utan personalities are not anthropomorphic projections.
832 *Animal Behaviour*, *83*(6), 1355–1365. https://doi.org/10.1016/j.anbehav.2012.02.024
833 Weiss, A., & King, J. E. (2015). Great ape origins of personality maturation and sex
834 differences: A study of orangutans and chimpanzees. *Journal of Personality and Social*
835 *Psychology*, *108*(4), 648–664. https://doi.org/10.1037/pspp0000022
836 Weiss, A., King, J. E., & Murray, L. (2011). *Personality and temperament in nonhuman*

primates. (A. Weiss, J. E. King, & L. Murray, Eds.). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4614-0176-6

Weiss, A., Staes, N., Pereboom, J. J. M., Inoue-Murayama, M., Stevens, J. M. G., & Eens, M. (2015). Personality in Bonobos. *Psychological Science*, *26*(9), 1430–1439. https://doi.org/10.1177/0956797615589933

Weiss, A., Wilson, M. L., Collins, D. A., Mjungu, D., Kamenya, S., Foerster, S., & Pusey, A. E. (2017). Personality in the chimpanzees of Gombe National Park. *Scientific Data*, *4*, 1–18. https://doi.org/10.1038/sdata.2017.146

Wood, W., & Eagly, A. H. (2002). A cross-cultural analysis of the behavior of women and men: Implications for the origins of sex differences. *Psychological Bulletin*, *128*(5), 699–727. https://doi.org/10.1037/0033-2909.128.5.699

Yamanashi, Y., Teramoto, M., Morimura, N., Hirata, S., Inoue-Murayama, M., & Idani, G. (2016). Effects of relocation and individual and environmental factors on the long-term stress levels in captive chimpanzees (Pan troglodytes): monitoring hair cortisol and behaviors. *PLOS ONE*, *11*(7), e0160029. https://doi.org/10.1371/journal.pone.0160029

Ying, Y.-W., & Han, M. (2006). The contribution of personality, acculturative stressors, and social affiliation to adjustment: A longitudinal study of Taiwanese students in the United States. *International Journal of Intercultural Relations*, *30*(5), 623–635. https://doi.org/10.1016/J.IJINTREL.2006.02.001

Acknowledgements

We are grateful to Hannah Roome, Lara Wood, Alex Weiss, Sally Street and Josep Call for helpful comments on an earlier draft and to Samuel Gosling, Sarah Brosnan and Lydia Hopper for advice and for assistance with initial data collection. We thank the NCCC chimpanzee carestaff for their contributions.

Financial support:

Bruce Rawlings was funded by an Economic and Social Research Council (ESRC) studentship from the North East Doctoral Training Centre, number 1449189.

Author contributions

BR, HF, EGF and RK conceived of the study. BR, HF and LR collected data and BR conducted statistical analyses. All authors contributed to manuscript writing.

Conflicts of interest

The authors declare no conflicts of interest.

Data availability

Data is uploaded to Dryad Repository (doi:10.5061/dryad.xksn02vc0)