



Personality correlates with contextual plasticity in a free-living, long-lived seabird

Jacquelyn K. Grace* and David J. Anderson

Department of Biology, Wake Forest University, Winston Hall, 1834 Wake Forest Rd.,
Winston-Salem, NC 27109, USA

*Corresponding author's e-mail address: gracjk7@wfu.edu

Accepted 6 February 2014; published online 7 March 2014

Abstract

Despite a rapid increase in animal personality research, critical gaps remain. We have little knowledge of the long-term (as opposed to short-term) consistency of personality and the relationships between personality and behavioural flexibility in wild, free-living animals. This study investigates personality (i.e., consistent inter-individual differences in behaviour) and contextual plasticity (i.e., plasticity in behaviour between contexts) of these traits in the Nazca booby (*Sula granti*), a long-lived, free-living seabird. We tested birds for personality in the field, during incubation, using a human intruder test, two novel object tests, and a social stimulus test (mirror), and determined repeatability of behavioural traits both within-season (short-term) and between several years (long-term). We found high short- and long-term repeatability of aggressive and anxiety-related behaviours when confronted with a nest intruder and novel objects, but lower repeatability during social stimulation. Contextual plasticity was highly repeatable across years for aggressive behaviours, and low for anxiety-related behaviours. Plasticity did not correlate across behaviours into a 'meta-personality' trait. Contextual plasticity in behavioural traits was highly correlated with the expression of those traits, suggesting that individuals are able to both react strongly and modify their behaviour depending on context, perhaps due to consistent differences in individual quality. Model selection exercises using structural equation models evaluated the relationships between personality factors, indicating a behavioural syndrome in which anxiety- or agitation-related behaviours covary strongly and positively, and both covary weakly and negatively with aggression.

Keywords

personality, behavioural syndrome, plasticity, Nazca booby (*Sula granti*), seabird.

1. Introduction

Variability in behaviour between individuals in a population has traditionally been considered biological noise (Wolf & Weissing, 2012). Yet highly

structured and consistent behavioural differences have been identified over the past decade in a variety of organisms, from amoebas to humans, that are maintained across time and situation/context (Réale et al., 2010). The recognition of such 'animal personality' and behavioural syndromes prompted a recent explosion of behavioural research focused on individuals instead of populations (Wolf & Weissing, 2012). Here, we take a broad approach, first evaluating behavioural consistency of multiple traits in a wild, long-lived vertebrate, the Nazca booby (*Sula granti*). We then integrate contextual plasticity and personality, investigating the relationship between the two and evaluating plasticity as a 'meta-personality' trait (Stamps & Groothuis, 2010).

1.1. A note on terminology

For this study we followed the definitions of Stamps & Groothuis (2010), where 'context' describes all external stimuli surrounding an individual when it performs a behaviour; this definition encompasses both 'situation' and 'context' as defined by Sih et al. (2004), and 'personality' refers to individual behavioural consistency across time and context compared to all other individuals in the study population (also satisfying requirements for a 'behavioural syndrome', which requires only temporal or contextual consistency).

1.2. Behavioural consistency in wild, long-lived animals

Most personality research on wild animals involves capture and testing in a laboratory (Bell et al., 2009; Garamszegi et al., 2012), which can result in questionable ecological validity (Herborn et al., 2010). With the rapid expansion of animal personality research, studies of wild animals tested in the field or in ecologically valid settings are becoming more common (e.g., Sapolsky & Ray, 1989; Gosling, 1998; Réale & Festa-Bianchet, 2003; Kralj-Fiser et al., 2007; Briffa et al., 2008; Garamszegi et al., 2008, 2012; Herborn et al., 2010; Dammhahn & Almeling, 2012; Montiglio et al., 2012; Jennings et al., 2013; Patrick et al., 2013; Petelle et al., 2013), and some species show high correlation between laboratory tests of personality and behaviour in the wild (e.g., blue tits, *Cyanistes caeruleus*: Herborn et al., 2010), although this assumption has rarely been tested.

Recent work has indicated that life history strategy is related to personality (Wolf et al., 2007; Biro & Stamps, 2008; Stamps & Groothuis, 2010; von

Merten & Siemers, 2012) and that long-lived organisms may be more risk-averse due to high future reproductive potential compared to shorter-lived organisms and to those with lower future reproductive potential (von Merten & Siemers, 2012). Seabirds are among the longest-lived animals, with greatly delayed senescence compared to other birds (Ricklefs, 1998), suggesting that life-history tradeoffs may be more pronounced in this group (Patrick et al., 2013). They also exhibit long-term social relationships in dense nesting colonies, which may impose additional selection constraints on personality. While these attributes make seabirds attractive models to expand our understanding of life history and personality, very little research on personality has been conducted in seabirds, although repeatability and heritability of boldness appear to be high in at least one species (Patrick et al., 2013). Apart from basic scientific interest, seabirds are also of conservation and management interest due to their interaction with fisheries and a generally high rate of population decline relative to other bird groups (Croxall et al., 2012). Anthropogenic selection pressures may be acting on variation in behavioural traits including response to novelty, exploration, and aggression, and understanding that variation can inform conservation decisions (McDougall et al., 2006). Our motivation for determining behavioural consistency, or ‘personality’, in Nazca boobies is two-fold: to add to the relatively small number of studies on personality in long-lived animals, and to determine the ecological validity of high repeatability of behavioural traits observed in laboratory or captive-tested animals, by using a wild animal tested in the field.

1.3. Plasticity and personality

Integration of consistent between-individual variation in behaviour (i.e., repeatability: proportion of total variance that is between individuals, Westneat et al., 2011) and within-individual variation in behaviour (i.e., plasticity or a reaction norm; Dingemanse et al., 2010; Westneat et al., 2011) is a burgeoning field of interest. Contextual plasticity (i.e., plasticity between contexts, Stamps & Groothuis, 2010) can be related to personality in many ways, including correlation of the mean expression of a behaviour with contextual plasticity in that same behaviour (i.e., elevation–slope covariance); two hypotheses exist to predict the direction of this specific relationship. The ‘individual quality hypothesis’ predicts a positive elevation–slope covariance due to underlying differences in ‘quality’ of an individual (*sensu* Wilson & Nussey, 2010; Bergeron et al., 2011): high-quality individuals perform more

of a behaviour and can adjust their level of performance of that behaviour better across contexts (for example, across an environmental gradient) under this hypothesis, showing higher contextual plasticity in that trait (Betini & Norris, 2012). The ‘individual quality hypothesis’ assumes that high magnitude of behavioural response and high contextual plasticity each enhance fitness, but are also expensive, so only high-quality individuals are able to accomplish both, while low-quality individuals are constrained to a low magnitude of behavioural response and low contextual plasticity (Betini & Norris, 2012). High contextual plasticity could be expensive due to information sampling, and/or to neural mechanisms necessary for decision making and information analyses (Stamps & Groothuis, 2010; Mathot et al., 2012). The ‘compensatory hypothesis’ also assumes that behavioural response and plasticity are expensive. However, under this hypothesis, high responsiveness is always desirable and low-quality individuals are able to compensate by trading off responsiveness and plasticity, while highly responsive individuals do not need high plasticity, resulting in a negative elevation–slope relationship (Betini & Norris, 2012). We point out that high responsiveness may not always be beneficial, but this trade-off could also apply to highly responsive individuals, who may be constrained to low plasticity (regardless of whether high responsiveness is beneficial) because they have invested in a high magnitude of response. This would result in the same prediction of a negative elevation–slope relationship. Finally, plasticity could be unrelated to personality if high-quality individuals can have both a high mean level of behavioural responsiveness and adjust that responsiveness across contexts (without assuming that high responsiveness is always desirable), while lower-quality individuals are able to compensate for low responsiveness with high plasticity. This would result in uniformly high plasticity in a population.

1.4. General predictions

In this study we investigate short- and long-term behavioural consistency across contexts, the relationship between behavioural traits, and integrate personality with contextual plasticity, in wild Nazca boobies. Specifically, we ask: (1) Do Nazca boobies display consistent differences between individuals in behaviour within contexts (tests) across (a) time (both short- and long-term), and (b) between contexts? (2) Do different behaviours covary in elevation of behavioural response (e.g., are highly aggressive individuals also

highly anxious)? (3) Do Nazca boobies differ consistently in contextual plasticity in a behaviour, across time? (4) Is contextual plasticity in one behaviour correlated with contextual plasticity in another behaviour, making some birds more plastic generally than others (i.e., is plasticity a ‘meta-personality’ trait, Stamps & Groothuis, 2010)? (5) Do Nazca boobies exhibit an elevation–slope relationship between contextual plasticity and personality? Because a variety of animal species, including a long-lived seabird, display consistent personality traits (Réale et al., 2010; Patrick et al., 2013), we expect: (1) consistent differences between adults in behaviour within contexts (tests) across (a) time (both short- and long-term), and (b) between contexts. Recognizing the ‘shy-bold’ or ‘reactive-proactive’ axis observed in many other species (see Cockrem, 2007), we expect (2) negative covariance between elevation of behavioural response in aggressive behaviours and elevation of response in anxiety-related behaviours. Following the suggestion that contextual plasticity in a behaviour can be a personality trait (reviewed by Dingemanse et al., 2010; Mathot et al., 2012), we expect (3) consistent differences between individuals in contextual plasticity within a behaviour. If plasticity is a personality trait, (4) contextual plasticities of different behaviours should be positively correlated (i.e., some birds are more plastic than others, regardless of behaviour), identifying plasticity as a ‘meta-personality’ trait (Stamps & Groothuis, 2010). Alternatively, selection may act more strongly on plasticity in some behaviours than others, preventing a ‘meta-personality trait’ from forming. Finally, we predict (5) a positive correlation between elevation of behavioural response and slope (plasticity) of that response between contexts, supporting the ‘individual quality hypothesis’ outlined, above. Because our tests involve varying degrees of threat to offspring, high responsiveness should be favourable. Thus, if responsiveness and plasticity are expensive, high quality birds should exhibit high responsiveness and high plasticity (in accordance with degree of threat), while low quality birds will be constrained to lower responsiveness and lack the reserves to alter that response.

2. Materials and methods

All research reported here was permitted under the regulations of the Wake Forest University Institutional Animal Care and Use Committee and the Galápagos National Park Service, and adheres to NIH standards for animal use in research.

2.1. Study animals

The Nazca booby is a long-lived, colonial, pelagic seabird (Anderson & Apanius, 2003). At our long-term study site on Isla Española, Galápagos (Apanius et al., 2008), they nest on the ground in open habitat, permitting easy observation at the nest site. Nazca boobies can live at least 26 years (unpublished data). They are socially and genetically monogamous within a breeding season (Anderson & Boag, 2006; Maness & Anderson, 2007), with biparental incubation and care of offspring (Anderson & Ricklefs, 1992; Apanius et al., 2008). They spend approximately half of the year at sea, but must return to land to breed and are highly natally philopatric (Huyvaert & Anderson, 2004). Individuals are identified by permanent metal leg bands bearing unique numbers.

For Nazca boobies, and most pelagic seabirds, most time spent on land is devoted to pair formation and offspring care (approx. 43 days of egg incubation, approx. 15 days of nestling brooding). During these first 15 days highly altricial Nazca booby nestlings are constantly brooded due to their inadequate thermoregulatory ability and extremely low mobility. We avoided testing birds with chicks unless a chick hatched after testing began. Thus, the majority of birds were incubating eggs throughout testing (for the first round of testing, nest intruder: 444 birds on eggs only, 35 on chicks; first novel object: 418 on eggs only, 61 on chicks; second novel object and social stimulus: 409 on eggs only, 70 on chicks). Because care of very young nestlings (<15 days old) is behaviourally very similar to incubation of eggs (and in many cases birds were simultaneously incubating an egg and brooding a nestling), we did not differentiate between incubating and brooding birds in analyses. Males and females alternate incubation, resulting in typical incubation stints of 3 days during this study. Incubating adults rarely leave the nest, but do interact with non-breeding and incubating conspecifics, and with other species present in the colony. We exploited this relatively high level of activity during incubation to evaluate Nazca booby personality (or lack thereof). Females were not tested for personality during their first incubation bout (immediately following egg-laying) because the risk of abandonment due to disturbance is higher at this time, and up-regulated hormones and neuropeptides associated with egg laying (Li et al., 1996) could influence behaviour.

At our long-term study site, Nazca boobies tolerate the presence of humans well (Apanius et al., 2008). During incubation, birds typically do not respond to humans at a distance of two meters or greater from the nest site.

Table 1.

Description and interpretation of the function of behaviours analysed in this study.

Behaviour	Description	Interpretation
Gardening	Picking up and moving nest material.	Mate advertisement (males), territorial display (both sexes), anxiety- or agitation-related displacement behaviour.
Shaking	Large body shakes, head shakes, and small shivers (extremely quick shakes of the head and neck).	Aggressive signal (social context), settling of feathers, anxiety- or agitation-related displacement behaviour.
Aggression	Biting and jabbing at nest intruder, novel object, or simulated conspecific (mirror).	Aggressiveness or boldness in both social and non-social contexts.

Appendix A provides additional detail for gardening and shaking.

At approximately two meters distance, the bird may begin to watch the human, although many birds show no visible signs of disturbance. Vigilance behaviours, territorial behaviours, and possible anxiety-related behaviours (gardening and shaking, see Table 1) exhibited by incubating adults increase as a human, or any other vertebrate, approaches the nest site. Once a human, or other vertebrate (conspecific or heterospecific), is within the nest site (approximately 0.5 m from the incubating bird), many birds will begin to exhibit anxiety-related behaviours and/or aggressive behaviours such as jabbing at the intruder.

2.2. Behavioural tests

We assayed behavioural traits in 479 adult Nazca boobies, using a nest intruder test, two novel object tests, and a social stimulus test. These four tests represented three different contexts in which we could analyse consistency of behavioural response between and within contexts (the two novel object tests). Behaviours analysed were gardening, shaking, and aggression (described in Table 1). Other behaviours included wing flaps, preening, inter- or intraspecific interactions, sleeping, repositioning of incubation posture, stretching, calling, yes-no headshakes (aggressive signal), and head turns. These behaviours were not analysed because they occurred at very low frequency, were highly sex-specific (e.g., yes-no headshakes are primarily performed by males), or were not performed during all tests, preventing

cross-context comparison. Birds were not tested if they were engaged in intra- or inter-specific interactions. Much of the terminology that we used to identify behaviours is derived from the Electronic Supplementary Material of Maness & Anderson, 2008, and Nelson's work on this same species (Nelson, 1978), which was referred to at that time as the 'masked' booby (*Sula dactylatra*). The eastern Pacific 'masked' boobies that Nelson studied were later recognized as a distinct species, the Nazca booby (American Ornithologists' Union, 2000; Friesen et al., 2002).

J.K.G. performed and recorded behaviours for all tests. Before each test, J.K.G. first approached the focal adult momentarily to within 2 m to induce vocalization to determine the sex of the incubating bird (Nazca boobies have sexually dimorphic voices; Nelson, 1978). Following the vocalization check, the observer retreated approx. 20 m from the nest site and observed the bird for five minutes; focal birds typically showed no recognizable reaction to the observer's presence. Tests were conducted between 07:15–12:00 h and 13:30–17:00 h to avoid the hot mid-day and the socially active dawn and dusk periods of the day.

2.2.1. Nest intruder test

J.K.G. approached the nest until standing at a distance of approximately 1 m and recorded behaviour for 1 min. After 1 min at 1 m, J.K.G. stepped forward and placed her booted foot approximately 13 cm from the focal bird, toe up, and recorded behaviour for 30 s.

2.2.2. First novel object test

A Red Bull™ can filled with pebbles (considered to be unfamiliar to the subjects), attached to a 3 m long pole, was placed on the ground at the nest site, approx. 13 cm from focal bird, and the focal bird's reaction was recorded for 5 min.

2.2.3. Second novel object and social stimulus tests

A plastic crate containing an upright mirror covered with a black cloth curtain was placed approximately 0.5 m away from the focal bird. Exact distance from the nest depended upon nest topography. J.K.G. then retreated approx. 6 m and recorded behavioural reaction to the crate (second novel object) for 5 min. Following this, J.K.G. pulled a string attached to the curtain covering the mirror (social stimulus), thereby revealing it, and recorded the subject's reaction to its reflection for 5 min. If a bird did not react, J.K.G. crouched behind the bird following the test, to ensure that it could see its reflection. In

cases of doubt that the reflection was visible (28 out of 722 cases), the bird was re-tested following a rest period of 2 min.

2.2.4. *Test sequence*

In November–January of 2008–2009 (behaviour session 1), 2009–2010 (behaviour session 2) and 2011–2012 (behaviour session 4, all during the egg incubation part of the annual breeding cycle), and March–April 2010 (behaviour session 3, during late egg incubation and early chick-rearing). Visits to nest sites were randomized, and each bird was presented with one test per day (first round) in the following order: nest intruder, first novel object and second novel object/social stimulus test (social stimulus directly followed the second novel object test). We repeated this series (second round) a minimum of three days after the last test of the first round to evaluate short-term temporal consistency of behaviour. The specific test–retest interval depended on each individual's nest attendance pattern (for all tests: minimum test–retest interval = 6 days, maximum = 24 days, $N = 193$; nest intruder test: test–retest interval mean \pm SD = 9.75 ± 2.95 days; first novel object test: 9.90 ± 3.19 days; second novel object and social stimulus tests: 9.92 ± 3.01 days). On average, birds completed both rounds of testing within 14 days (mean \pm SD = 13.82 ± 3.78 days, min = 8 days, max = 26 days).

Two testing groups were established: (1) those that received a nest intruder test in the morning, first novel object test in the afternoon, and second novel object and social stimulus tests in the morning, (2) those that received a nest intruder test in the afternoon, first novel object test in the morning, and second novel object and social stimulus tests in the afternoon. Morning and afternoon groups remained the same across the two testing rounds because of possible effects of time of day. Behaviour sessions 1–3 were used to examine short-term repeatability (within session, between rounds) and no birds were re-tested for long-term repeatability between these sessions. We repeated these tests on the same birds, in behaviour session 4, following the same protocol and keeping birds in the same time of day group, to evaluate long-term repeatability of behavioural traits between sessions (Figure 1).

2.3. *Statistical analyses*

Table 2 summarizes analyses, statistical methods and sample sizes.

2.3.1. *Behaviours examined*

Only birds that completed the first round of a behaviour session were used in analyses. Gardening, shaking, and aggressive behaviours were the only

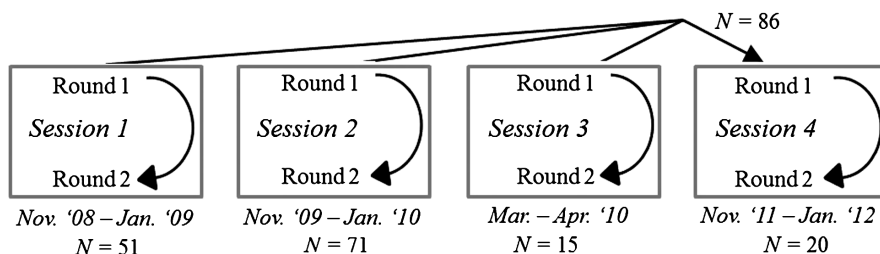


Figure 1. Test sequence for temporal repeatability of behaviour. Curved arrows indicate short-term consistency repeatability (within session, between rounds), straight arrows indicate long-term consistency repeatability (between sessions). Sample sizes are provided for each session (the sum of which is the sample size for short-term repeatability) and between sessions (long-term repeatability).

behaviours performed consistently across all tests by both sexes, although specific aggressive behaviour did vary by test (Table 1). For further description of behaviours see Appendix A in the online edition of this journal, which can be accessed via <http://booksandjournals.brillonline.com/content/journals/1568539x>.

2.3.2. Data transformations and effects of sex, time of day, and session

Initial examination of data suggested that behaviour counts were not normally distributed; thus, each behaviour count was square-root transformed after adding one (to avoid different transformation of counts between 0 and 1, and counts above 1, Osborne, 2002), which resulted in an approximately normal distribution, then *z*-scored within test (mean count subtracted from individual count, then divided by standard deviation) to prevent weighting results toward tests that resulted in high mean activity. Preliminary data analysis revealed significant effects of session, sex, morning or afternoon group, and their interactions on transformed behavioural data. Thus, for all analyses, except those calculating short-term repeatability within test, *z*-scores were corrected for these main effects and all their interactions by saving the residuals from a factorial analysis of variance (see Results: Effects of sex, time of day, and session). We also calculated relative importance of these effects using a saturated model set (each predictor alone, and with every other predictor, and all possible interactions) and summing the AICc weights (a.k.a. model probabilities) of each model that contained the predictor (Anderson, 2008). Because sex, session, and morning/afternoon group did not change from test to re-test, short-term repeatability within test (context) was

Table 2.

Summary of analyses, statistical methods, sample sizes and associated results tables/figures to address the questions in the text.

Question	Analysis	Statistical method	<i>N</i>	Results
1a	Short-term temporal repeatability of behaviours within context between rounds, within session	Intraclass correlation coefficient (Lessells & Boag, 1987)	157 (147)	Figure 3A
1a	Long-term temporal repeatability of behaviours within context and round, between sessions	Intraclass correlation coefficient (Lessells & Boag, 1987)	86	Figure 3B
1b	Repeatability of behavioural response between tests, within round and session	Intraclass correlation coefficient (Lessells & Boag, 1987)	479	Figure 3C
2	Covariance in elevations of behavioural response for different behaviours, within round and session	Structural Equation Modelling with AICc model comparison	479	Table 3, Figure 4
3	Long-term repeatability of contextual plasticity, between sessions	Coefficient of Relative Plasticity (Dingemanse et al., 2010)	86	Figure 3B, Table 4
4	Correlation between contextual plasticity of different behaviours, between sessions	Coefficient of Relative Plasticity (Dingemanse et al., 2010), Structural Equation Modelling, and Pearson's product–moment correlation	479	Text
5	Relationship between elevation of behavioural traits and contextual plasticity, within round and session	Pearson's product–moment correlation	65–292	Table 4, Appendix D

A value in parentheses indicates that a smaller subset was used for some analyses within the group.

the only analysis that did not correct the z -scores for these main effects or their interactions.

2.3.3. Consistent differences between individuals in behaviour within contexts across time (1a)

We calculated 'consistency repeatabilities' (Nakagawa & Schielzeth, 2010), following Lessells & Boag (1987), of individual behaviours within contexts (different tests), and not of latent variable scores (described below, these combined scores in different contexts), because we were interested in assessing whether short-term repeatability of behaviours were affected by context. Repeatability values (a form of intraclass correlation coefficient) are essentially the proportion of variance explained by between-individual effects divided by the total variance. We calculated ANOVA-based repeatability values and approximate corresponding 95% confidence intervals based on F ratios using the *rptR* package in R (Nakagawa & Schielzeth, 2010). To determine whether between individual effects explained level of behavioural response better than the null (no individual effect), ANOVA models including an individual effect and not including an individual effect (null models) were evaluated using Akaike's Information Criterion corrected for small sample size (AICc), where $AICc = N \times \log(RSS/N) + (2K(K+1))/(N-K-1) + 2K$; N is the sample size, RSS is the residual sums of squares for the model and K is the number of parameters, including error (Burnham et al., 2011). Thus, AICc analysis balances the variance explained against the complexity of the model (see Table 4 for details).

To calculate short-term temporal repeatability, we calculated consistency repeatability values and their corresponding 95% confidence intervals (Nakagawa & Schielzeth, 2010) for each behaviour within test and between rounds (Figure 1), after transformation to approximate normality, and z -scoring, as described above, for the 157 birds that completed both rounds of all personality tests ($N = 147$ for aggressive behaviours in the first novel object test). Z -scores were not corrected for confounding factors because they did not change between test and re-test.

For long-term temporal repeatability, we calculated consistency repeatability (Lessells & Boag, 1987; Nakagawa & Schielzeth, 2010) of behaviours (following transformation and z -scoring) performed only during the first round (within test), across sessions, for the 86 incubating birds that could be relocated during behaviour session 4 (Figure 1). Data from birds tested

in either the first or second behaviour sessions were grouped as the ‘early’ behaviour sessions, and were compared with their data from behaviour session 4. Within these groups, behaviour counts from round 1 were square-root-transformed after adding one, and then were z -transformed. Z -scores were corrected for confounding variables and all their interactions by saving the residuals from a factorial analysis of variance including ‘session’ (only for the early behaviour sessions because all birds were assayed for the second time in session 4), morning/afternoon group, sex of individual, and all of their interactions.

2.3.4. Between-individual differences in behaviour between contexts (tests) (1b)

Consistency of between-individual differences in behavioural response between contexts (tests) was measured by calculating consistency repeatability (Lessells & Boag, 1987; Nakagawa & Schielzeth, 2010) of transformed and z -scored behaviour counts across tests, within session, after controlling confounding factors as described above, for the 479 birds that completed at least round 1. The second novel object test (crate) occurred immediately before the social stimulus test (reflection), and this temporal association could inflate repeatability, so cross-context repeatabilities were calculated across the nest intruder and both novel object tests, and separately across the nest intruder, the first novel object, and social stimulus tests.

2.3.5. Covariance in elevations of behavioural response for different behaviours (2)

We used Structural Equation Models (SEM) to examine the relationships between personality traits. Compared to Factor or Principal Component Analysis, SEM has the advantage of allowing covariance between latent variables (similar to ‘factors’ in a Factor Analysis) to be examined within one model (see Appendix B in the online edition of this journal, which can be accessed via <http://booksandjournals.brillonline.com/content/journals/1568539x> for further discussion of SEM and Factor Analysis) SEM was conducted with the *lavaan* package in R (Rosseel, 2012) using covariance matrices. We generated SEM models to evaluate the relationships between shaking, gardening, and aggressive behaviours for the 479 birds that completed round 1 of all personality tests, in any session. Gardening, shaking, and aggressive behaviours for each test (context), and contextual Coefficient of Relative

Plasticity values ('CRP'; within-individual variance/total variance, calculated between tests, within round 1; Dingemanse et al., 2010) for each behaviour type were included in initial models. Contextual CRP values could not converge into a latent variable, suggesting low covariance between CRP values for different behaviour types, and prevented convergence when included independently and so were omitted from model sets. We tested two different forms of latent variables, grouping transformed behavioural data into latent variables first by 'test' and then by 'behaviour', which we used to evaluate whether context ('test') drove covariance between behaviours, or whether covariances were context-independent (grouped by 'behaviour' not 'test'). None of the 'test' latent variable models could converge, suggesting low covariance between different behaviour types within tests and providing no support for context driving covariances between behaviours; thus, only 'behaviour' latent variable models were included in the model set. Because the second novel object and social stimulus tests were not independent, a saturated set of eight models (including all possible combination of covariances between latent variables) were generated for the nest intruder and both novel object tests, and another saturated set of eight models for the nest intruder, first novel object, and social stimulus tests (see Appendix C in the online edition of this journal, which can be accessed via <http://booksandjournals.brillonline.com/content/journals/1568539x> for graphic depiction of models). Models for both sets of tests were compared by AICc alternative model selection. Top models were those within the 95% confidence set of Akaike weights (AICc weight), and covariance and regression coefficients for top models were averaged to produce the final model for both sets of tests (Anderson, 2008).

2.3.6. Consistent difference in contextual plasticity in a behaviour, across time (3)

Contextual plasticity was calculated using the Coefficient of Relative Plasticity (Dingemanse et al., 2010) of behaviour counts following transformation and after controlling for confounding factors, as described above, within the first round, between contexts (tests), for the 479 birds that completed the first round of all personality tests. Repeatability of these CRP values was calculated across sessions (long-term), as described above to determine temporal consistency ($N = 86$).

2.3.7. Correlation across behaviours in plasticity (i.e., are some birds more plastic than others, regardless of behaviour) (4)

CRP values were calculated as described above, and analysed using SEM and Pearson's product-moment correlations for all birds that completed the first round of all personality tests ($N = 479$).

2.3.8. Relationship between elevation-slope and contextual plasticity and personality (5)

We were unable to explicitly examine the slope (contextual plasticity) and elevation (personality) of each bird's behaviour, because our data did not lend themselves to analysis with a behavioural reaction norm (our tests were not arrayed along a gradient, but instead were designed to test birds in different contexts). Instead, we inferred the relationship between slope and elevation for each behaviour within session with Pearson's product-moment correlations between individual contextual CRP values, as calculated above, and individual factor scores obtained during round 1 of any session for shaking, gardening, and aggression across the nest intruder, and both novel object tests, and the nest intruder, first novel object, and social stimulus tests. However, a positive correlation between behaviour level (elevation) and plasticity (slope) of these birds could be a statistical artefact of higher behavioural responsiveness allowing increased variation compared to birds with very low behavioural responsiveness, who are bounded by being unable to perform less than zero of a behaviour. To address this concern, we compared these results to analyses that include only those who performed a target behaviour (shaking, gardening, or aggression) at least once in each test and thus were not constrained by a lower zero limit. Alternatively, a positive correlation between CRP values and behavioural factor scores could be due to a high level of behavioural response from some birds during the first test, then habituation of all birds to tests, in general, during the first round of testing. If so, we expect higher responsiveness in all behaviours in the first test, compared to all subsequent tests. To address this, we converted behavioural counts to a rate per minute of testing, then square-root-transformed these rates after adding one to approximate a normal distribution, and compared the mean transformed rates between tests.

3. Results

3.1. *Effects of sex, time of day, and session on mean behaviour*

3.1.1. *Shaking*

During the nest intruder test, and first novel object test, the null model with no predictors (just the intercept) was included in the 95% confidence set, suggesting that little variation is explained by time of day, behaviour session, or sex. Of the information explained, predictor relative importances suggest that time of day affects shaking more than sex does during the nest intruder test (Figure 2A), with birds shaking slightly more in the afternoon than the morning, and males displayed slightly more shaking behaviour than females.

During the second novel object test and the social stimulus test, relative importances suggest that behaviour session was the most important predictor of shaking behaviour (Figure 2A). Birds shook less during behaviour session 3 than any other session.

3.1.2. *Gardening*

Following AICc model selection of factorial ANOVA models, sex was among all the top models (those within the 95% confidence set, see Appendix D in the online edition of this journal, which can be accessed via <http://booksandjournals.brillonline.com/content/journals/1568539x>) predicting gardening, and had a relative importance of >0.9 for all models predicting gardening, well above that of behaviour session (<0.31) or time of day (<0.37 , Figure 2B), with males gardening more than females.

3.1.3. *Aggression*

Aggression was best predicted by sex and behaviour session (no interaction; Figure 2C, Appendix D). Males performed more aggressive behaviours than females. In the nest intruder and both novel object tests, all birds displayed fewer aggressive behaviours during behaviour session 3 (the only March session), while this trend was reversed for the social stimulus test (i.e., highest aggression during sessions 3 and 4).

3.2. *Do Nazca boobies display consistent between-individual differences in behaviour within contexts across time (both short- and long-term) (1a)?*

Short-term consistency repeatabilities of behaviours (within-context and across rounds within a session) were generally high (Figure 3A), with two exceptions: aggressive behaviour during the social stimulus test, and gardening during the second novel object test. Birds displayed essentially no

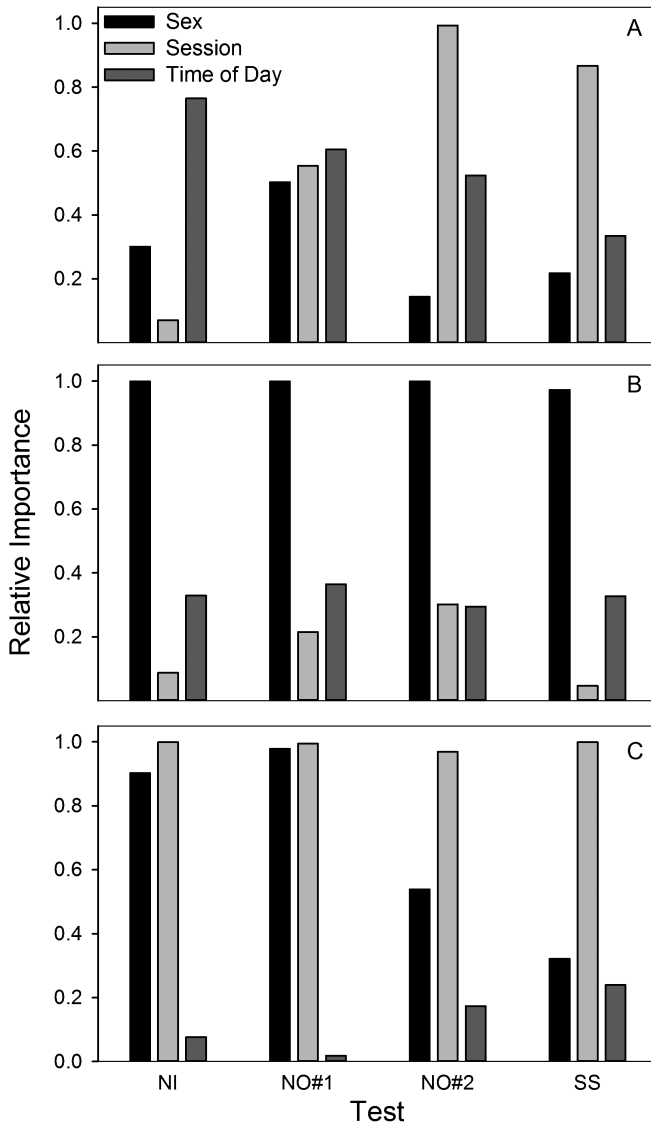


Figure 2. Relative importance of the predictor variables: sex, session, and time of day on (A) shaking, (B) gardening and (C) aggression, separated by context ($N = 479$). Models were factorial analyses of variance (ANOVA), and the model set included each predictor by itself, and in all possible combinations, including interactions, with the other predictors. This resulted in 12 models that were compared using AICc (see Materials and methods). Relative importance for each predictor was calculated by summing the Akaike weights (AICc weights) of each model that contained that predictor (Anderson, 2008).

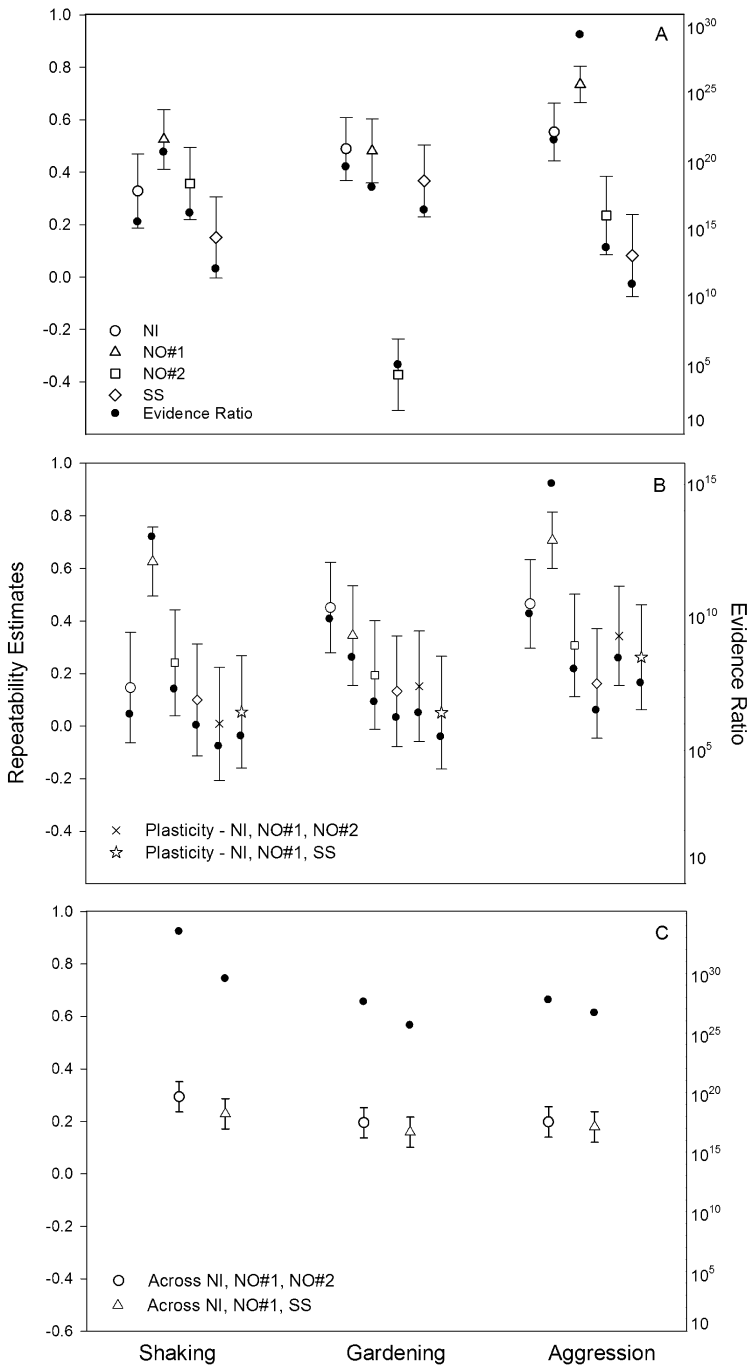
repeatability in these two cases ($R = -0.372$ and 0.082 , respectively). Aggressive behaviours during the nest intruder and first novel object test had the highest consistency repeatability ($R = 0.553$ and 0.735 , respectively), and behaviours performed during the first novel object test were generally most repeatable.

Long-term consistency repeatabilities of behaviours (within test and across session) were also generally high, with the exception of the social stimulus test, gardening during the second novel object test, and shaking during the nest intruder test (Figure 3B). The first novel object test generally had the highest consistency repeatability across sessions, suggesting that the initial reaction to a novel object is a very stable trait.

3.3. Do Nazca boobies display consistent between-individual differences in behaviour between contexts (1b)?

Within round and session, consistency repeatabilities between contexts (tests) were lower than within test (between time points), although models including an individual effect were at least 10^{25} times better than null models with no individual effect (Figure 3C). Shaking was the most repeatable behaviour across contexts, and behaviours across the nest intruder and both novel object tests had slightly higher consistency repeatabilities than across the nest intruder, first novel object, and social stimulus tests (Figure 3C).

Figure 3. Short-term (A), long-term (B) and cross-context (C) consistency repeatabilities (R , intraclass correlation coefficient) of shaking, gardening, and aggressive behaviours, corresponding 95% confidence intervals, and AICc evidence ratios. NI indicates the Nest Intruder test, NO#1 the first novel object test, NO#2 the second novel object test and SS the social stimulus test. Short-term repeatabilities (A) were calculated within test, between rounds ($N = 157$, except $N = 147$ birds for aggressive behaviour in the first novel object test). Long-term repeatabilities (B) of behaviours and contextual plasticity values were calculated within round 1, between sessions ($N = 86$). Repeatabilities of behaviours between contexts (C) were calculated within session, across contexts: nest intruder and both novel object tests, or the nest intruder, first novel object, and social stimulus tests ($N = 479$). ANOVA models including and not including an individual effect were evaluated by AICc model selection, and in all cases the model including an individual effect was the top model. Evidence ratios (ratio between the AICc weight of the top model and the null model) of these models are presented. Thus, the model for shaking during the human intruder test with individual as a predictor is 4.3×10^{15} times better than the null model with no individual effect. Symbols common to all three panels are as shown in the legend of Panel A.



3.4. Do different behaviours covary in elevation of behavioural response (e.g., are highly aggressive individuals also highly anxious) (2)?

Shaking, gardening and aggressive behaviours converged across contexts (tests) into latent variables for each behaviour, suggesting that personality traits are not specific to context. SEM modelling of behaviours across the nest intruder and both novel object tests produced five models with AICc weights within the 95% confidence set, all of which include gardening correlated with shaking (Table 3). Modelling of behaviours across the nest intruder, first novel object, and social stimulus tests produced six top models with AICc weights within the 95% confidence set, comprising the same models as those for the nest intruder and both novel object tests, but also including a model with shaking uncorrelated with aggression or gardening (Table 3). Covariances between gardening, shaking and aggression for these models were averaged (Anderson, 2008) for the nest intruder and both novel object tests, and the nest intruder, first novel object and social stimulus tests, indicating positive covariance between gardening and shaking, and negative covariance of both with aggression (Figure 4).

3.5. Do Nazca boobies differ consistently in contextual plasticity in a behaviour, across time (3)?

CRP values were calculated for behaviours between tests, and only aggression was highly repeatable in the long-term, between session ($R = 0.008\text{--}0.343$). However, all ANOVA models for contextual CRP values between behavioural sessions that included an individual effect were at least 10^5 times better than the null that contained no individual effect. This suggests that although repeatability values may be generally low for plasticity, individual identity affects these values (Figure 3B).

3.6. Is plasticity in one behaviour correlated with contextual plasticity in another behaviour (i.e., are some birds more contextually plastic than others, regardless of behaviour) (4)?

CRPs for behaviours calculated between contexts (tests), within session, were unable to converge into a latent variable with SEM, suggesting that no underlying 'plasticity' latent variable exists, and plasticity is behaviour-specific (e.g., birds that were highly plastic in aggression across tests were

Table 3.

AICc comparison of SEM models of personality syndromes across the nest intruder and both novel object tests, or the nest intruder, first novel object and social stimulus tests for 479 Nazca booby adults.

Model	AICc	Δ	Model likelihood	AICc weight	Evidence ratio
Nest intruder and both novel object tests					
Gardening~~Shaking, Aggression*	11 403.55	0.00	1	0.37	1
Gardening~~Shaking, Shaking~~Aggression*	11 404.48	0.94	0.63	0.23	1.60
Gardening~~Shaking, Gardening~~Aggression*	11 405.03	1.49	0.48	0.18	2.11
Full*	11 405.67	2.13	0.35	0.13	2.90
Orthogonal*	11 408.83	5.28	0.07	0.03	14.02
Aggression~~Shaking, Gardening	11 409.61	6.06	0.05	0.02	20.70
Aggression~~Gardening, Shaking	11 410.10	6.55	0.04	0.01	26.44
Aggression~~Gardening, Aggression~~Shaking	11 411.10	7.56	0.02	0.01	43.78
Nest intruder, first novel object, and social stimulus tests					
Gardening~~Shaking, Aggression*	11 410.73	0.00	1	0.35	1
Gardening~~Shaking, Gardening~~Aggression*	11 411.69	0.97	0.62	0.21	1.62
Gardening~~Shaking, Shaking~~Aggression*	11 411.97	1.25	0.54	0.19	1.87
Full*	11 412.80	2.08	0.35	0.12	2.82
Orthogonal*	11 414.61	3.88	0.14	0.05	6.96
Aggression~~Gardening, Shaking*	11 415.44	4.71	0.10	0.03	10.54
Aggression~~Shaking, Gardening	11 415.56	4.83	0.09	0.03	11.19
Aggression~~Gardening, Aggression~~Shaking	11 416.70	5.98	0.05	0.02	19.87

Full models are those in which all latent variables (i.e., gardening, shaking, aggression) are allowed to covary. Orthogonal models are those in which no latent variables covary. Waves (~~) indicate covariation between two variables, Δ indicates the difference between the AICc values of the top model and the model in question, model likelihood is calculated by $\exp(-(1/2)\Delta)$, AICc weight is the model likelihood divided by the sum of all model likelihoods, and the evidence ratio is the ratio between the weight of the top model to the model in question. Thus, the model with gardening and shaking covarying, and aggression independent of both, is 14-times more likely than the orthogonal model.

* Models with AICc weights within the 95% confidence set.

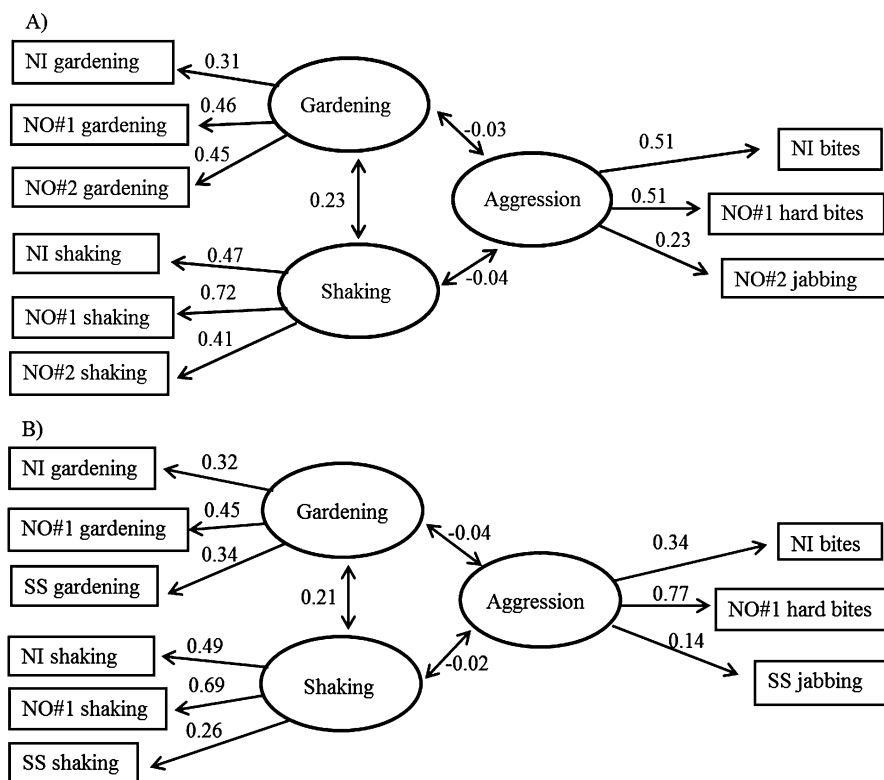


Figure 4. Covariances between latent variables produced by SEM analysis of behaviours during (A) the nest intruder (NI) and both novel object tests (NO#1 and NO#2), or (B) NI, NO#1 and SS tests. Covariances (double headed arrows), and regression estimates (single headed arrows) presented are averages of the three top models for each set of tests ($N = 479$). Circles represent latent variables, squares represent observed variables. Any covariances, estimates, or variables not depicted were not present in the top models.

not likewise highly plastic in gardening, or shaking). Pearson's product-moment correlations between CRP values for shaking, gardening, and aggression all produced correlation coefficients less than 0.05 ($p > 0.1$), supporting the interpretation that contextual plasticity in one behaviour was unrelated to contextual plasticity in another. Calculations of repeatabilities of cross-context CRP values for aggression, shaking, and gardening confirmed this suggestion (for nest intruder and both novel object tests: $R = 0.016$, 95% CIs = $-0.037, 0.069$; nest intruder, first novel object test, and social stimulus: $R = -0.011$, 95% CIs = $-0.062, 0.04$).

3.7. Do Nazca boobies exhibit an elevation–slope relationship between contextual plasticity and personality (5)?

Initial analyses including all individuals revealed a strong positive correlation between CRP values for behaviours between contexts and their associated behavioural factor scores (all r values ≥ 0.30 , $p < 0.0001$ for all Pearson product–moment correlations). For the subset of birds that performed a target behaviour at least once, CRP values for behaviours between contexts were also all significantly positively correlated with their associated behavioural factor scores (all r values ≥ 0.23 , $p < 0.002$ for all Pearson product–moment correlations; Table 4, see Appendix E in the online edition of this journal, which can be accessed via <http://booksandjournals.brillonline.com/content/journals/1568539x>), thus the relationship is likely not a statistical artefact due to a lower zero limit on behavioural response. The positive correlation between CRP values and behavioural factor scores also does not appear to be due to a high level of behavioural response from some birds during the first test, then habituation of all birds to tests, in general. The first test (nest intruder) elicited a higher behaviour rate only in aggression, while the second test (first novel object) elicited the highest mean shaking and gardening responses. Furthermore, standard deviations overlapped for most behaviours between most tests (nest intruder: shaking $\bar{x} \pm \text{SD} = 1.49 \pm 0.53$, gardening 1.19 ± 0.47 , aggression 3.49 ± 1.37 ; first novel object: shaking 1.86 ± 0.49 , gardening 1.47 ± 0.49 , aggression 1.67 ± 0.86 ; second novel object: shaking 1.50 ± 0.30 , gardening 1.28 ± 0.36 , aggression 1.56 ± 0.67 ; social stimulus: shaking 1.44 ± 0.31 , gardening 1.3 ± 0.43 , aggression 1.33 ± 0.62).

Table 4.

Pearson's product–moment correlation coefficients (r) and associated p -values for correlations between contextual CRP values of behaviours and individual factor scores generated by Structural Equation Modelling for those same behaviours.

Tests	Behaviour	r	p	N
Nest intruder and both novel object tests	Gardening	0.44	<0.0001	85
	Shaking	0.41	<0.0001	276
	Aggression	0.35	<0.0001	292
Nest intruder, first novel object, and social stimulus tests	Gardening	0.39	<0.0001	65
	Shaking	0.34	0.0013	261
	Aggression	0.23	0.0013	188

4. Discussion

Our results indicate that incubating Nazca booby adults display personality traits that are contextually general, and also repeatable across time both in the short-term and the long-term. Gardening and shaking are suspected anxiety-related traits (see Appendix A) that covary positively with each other, and show a slight negative covariance with aggression. Temporal consistency in contextual plasticity (the ratio of within-individual variance in performance of behaviours within test to total population variance) was high for aggressive behaviours, and lower for gardening and shaking. Low covariance between individual contextual plasticity scores for different behaviours suggests that plasticity in this system is not a 'meta-personality' trait (Stamps & Groothuis, 2010). Contextual plasticity in a behaviour was positively correlated with the elevation of response in that behaviour, providing tentative support for the 'individual quality hypothesis' (Betini & Norris, 2012), although further work must be done to determine if birds with higher responsiveness are of higher 'quality' (an axis correlated with fitness, see Bergeron et al., 2011; Wilson & Nussey, 2010).

4.1. Do Nazca boobies display consistent between-individual differences in behaviour within contexts across time (both short- and long-term), and between contexts (tests) (1)?

Our results indicate that aggression and suspected anxiety-related behaviours (gardening and shaking) are generally repeatable within a breeding season (across weeks), and across years, in these long-lived birds, and that these individual differences in elevation of behaviour are maintained across context (test). The high short- and long-term repeatability of three behavioural traits across different contexts complements findings in another long-lived seabird, the wandering albatross (*Diomedea exulans*) in which boldness in response to human approach (similar to our nest intruder test) is also highly repeatable (Patrick et al., 2013).

Until recently, human approach (Patrick et al., 2013) and novel objects were uncommon on seabird colonies, and our results suggest that these contexts induce a strong individually-fixed reaction. However, social interaction with conspecifics is frequent and our findings indicate that individual reactions may be more flexible than reaction to heterospecifics or novel objects, because repeatabilities of behaviours were lower during the social stimulus test, the only test that assayed social interaction. Personality in social

interactions may depend on current energy status and/or habituation to neighbouring nesting birds, which varies within and between years with proximity of neighbours. Analysed population variance during this test did not differ from other tests because we analysed z -scores of square-root-transformed behaviours (i.e., population variance was close to one for all behaviours in all tests, variance was not exactly one because repeatability estimates were calculated from the subset of birds that were tested twice). Therefore, differences in population variance are not driving the difference in repeatability values.

Although our tests presented Nazca boobies with different experimental contexts (nest intruder, novel object, social stimulus), further work is needed to determine if these personality traits are maintained outside of the incubation state.

4.2. *Do elevations of response for different behaviours covary (2)?*

Gardening and shaking covaried positively during incubation ($0.21 < \text{cov} < 0.23$), and both have a slight negative covariance with aggression ($-0.02 < \text{cov} < -0.04$). The negative covariance between aggression (a typically ‘bold’ behaviour) and gardening/shaking suggests that gardening and shaking may be ‘shy’ behaviours. However, the negative covariance is so small that it is more likely that these behaviours represent different personality axes in Nazca boobies, rather than a single axis. The covariance estimate of 0.23 between shaking and gardening implies that there is a tendency to perform these behaviours together, likely because they are both probable anxiety-related behaviours (see Appendix A).

Alternatively, our behavioural tests actually measured only ‘activity’ in response to a stressor (i.e., some birds perform more of all behaviours because they are more ‘active’) and not aggression or anxiety *per se* (see Sinn & Moltschaniwskyj, 2005 for a discussion of this issue). If we had only measured activity, we would expect to see high positive covariance between all behaviours, because they are all measures of ‘activity’. Because we did not observe this, a single underlying ‘activity’ behavioural axis does not explain our results.

4.3. *Do Nazca boobies differ consistently in contextual plasticity in a behaviour, across time (3)?*

While contextual plasticity is suggested to be a stable personality trait (Stamps & Groothuis, 2010), the repeatability of plasticity has rarely been

tested in wild populations, and most studies examine plasticity in only one trait (reviewed by Dingemanse et al., 2010; Mathot et al., 2012). However, if between-individual differences in plasticity are highly inconsistent across time, the relationship between personality traits and contextual plasticity in those traits may be ephemeral and direct selective pressures on such plasticity may be weak, or limited to short periods of time. We found relatively low long-term repeatabilities of contextual plasticity, with the notable exception of aggressive behaviours. This suggests that, for incubating adult Nazca boobies, contextual plasticity in aggression may be a stable, long-term personality trait, but plasticity in anxiety/agitation-related behaviours appears to be less stable.

4.4. Is contextual plasticity in one behaviour correlated with plasticity in another behaviour (4)?

We found no evidence that some birds were consistently more contextually plastic than others, regardless of behaviour. Thus, contextual plasticity is behaviour-specific for Nazca boobies, and we found no support for contextual plasticity as a ‘meta-personality trait’ (Stamps & Groothuis, 2010) that is independent of behaviour type. This may be due to the slight negative covariance between aggression and anxiety-related traits, and the positive correlation between contextual plasticity of a behaviour and expression of that behaviour, discussed below (e.g., highly aggressive birds are also highly plastic in aggression, but tend to perform less anxiety-related traits and are also less plastic in these traits).

4.5. Do Nazca boobies exhibit an elevation–slope relationship between contextual plasticity and personality (5)?

Contextual plasticity in behavioural traits was positively correlated with the expression of those traits. For example, very aggressive individuals also exhibited high contextual plasticity in aggression. The same is true for gardening and shaking. Habituation to testing does not appear to explain this relationship because the first test did not elicit higher behavioural response, except for aggression. Because the nest intruder test is the most invasive of our tests and a strong aggressive response to heterospecific intrusion into the nest site is typical for this species, habituation probably did not cause the observed decreased aggressive behavioural rate from the first test to the ones following it. Instead, this test apparently genuinely elicited a strong aggressive response.

Two explanations are plausible for the observed positive correlation between elevation of behavioural response and contextual plasticity: exogenous variables may have induced high reactivity in birds, arbitrarily, in one of the test (contexts), or more-reactive individuals were more able to alter their behaviour between contexts (i.e., the individual quality hypothesis). The first explanation is unlikely for two reasons. First, aggression, gardening, and shaking were factor scores resulting from structural equation models of three tests. Hence, birds that reacted strongly in only one test, perhaps reflecting recent social interactions or their nutritional status, could have a high contextual CRP, but would not have a high factor score for that trait because the factor score integrates all three tests. Second, behavioural traits were generally repeatable across weeks and years, indicating elevation of behavioural response is a long-term characteristic of an individual.

Instead, our high positive correlation between plasticity (slope) and magnitude of behavioural response (elevation) is support for the ‘individual quality’ hypothesis, similar to results from tree swallows (Betini & Norris, 2012), great tits (Dingemanse et al., 2012) and Ural owls (*Strix uralensis*, Kontiainen et al., 2009). Our results provide no support for the ‘compensatory hypothesis’, which predicts a negative elevation–slope covariance, because low-responding individuals need to compensate by being able to respond to variation in environmental conditions (Betini & Norris, 2012), or for generally high plasticity with low-responding individuals compensating by increasing contextual plasticity in these behaviours. For Nazca boobies, both aggressive behaviour and plasticity in aggression were generally repeatable across years, suggesting that these traits are relatively stable and may reflect individual quality, or pleiotropic (Sinervo & Svensson, 2002) or developmental effects (Duckworth, 2010), but not immediate condition. Further research is needed to determine whether highly reactive Nazca boobies are also higher quality, which may result in higher fitness, or if this positive correlation is unrelated to fitness (and perhaps individual quality), as was found in tree swallows (Betini & Norris, 2012).

Acknowledgements

We thank the Galápagos National Park Service for permission to work in the Park; the Charles Darwin Research Station, and TAME Airlines for logistic support; Wake Forest University for research funding; and the members of

the Anderson lab group, A. Carter and P.-O. Montiglio for comments on the manuscript. This material is based upon work supported under a National Science Foundation Graduate Research Fellowship to J.K.G. and National Science Foundation Grant DEB 0842199 to D.J.A.

References

- American Ornithologists' Union (2000). Forty-second supplement to the American Ornithologists' Union Check-list of North American Birds. — *Auk* 117: 847-858.
- Anderson, D.J. & Apanius, V. (2003). Actuarial and reproductive senescence in a long-lived seabird: preliminary evidence. — *Exp. Gerontol.* 38: 757-760.
- Anderson, D.J. & Boag, P.T. (2006). No extra-pair fertilization observed in Nazca booby (*Sula granti*) broods. — *Wilson J. Ornithol.* 118: 244-247.
- Anderson, D.J. & Ricklefs, R.E. (1992). Brood size and food provisioning in masked and blue-footed boobies (*Sula* spp.). — *Ecology* 73: 1363-1374.
- Anderson, D.R. (2008). Model based inference in the life sciences: a primer on evidence. — Springer, New York, NY.
- Apanius, V., Westbrock, M.A. & Anderson, D.J. (2008). Reproduction and immune homeostasis in a long-lived seabird, the Nazca booby (*Sula granti*). — *Ornithol. Monogr.* 65: 1-46.
- Bell, A.M., Hankison, S.J. & Laskowski, K.L. (2009). The repeatability of behaviour: a meta-analysis. — *Anim. Behav.* 77: 771-783.
- Bergeron, P., Baeta, R., Pelletier, F., Réale, D. & Garant, D. (2011). Individual quality: tautology or biological reality? — *J. Anim. Ecol.* 80: 361-364.
- Betini, G.S. & Norris, D.R. (2012). The relationship between personality and plasticity in tree swallow aggression and the consequences for reproductive success. — *Anim. Behav.* 83: 137-143.
- Biro, P.A. & Stamps, J.A. (2008). Are animal personality traits linked to life-history productivity? — *Trends Ecol. Evol.* 23: 361-368.
- Briffa, M., Rundle, S.D. & Fryer, A. (2008). Comparing the strength of behavioural plasticity and consistency across situations: animal personalities in the hermit crab *Pagurus bernhardus*. — *Proc. Roy. Soc. Lond. B: Biol. Sci.* 275: 1305-1311.
- Burnham, K.P., Anderson, D.R. & Huyvaert, K.P. (2011). AIC model selection and multi-model inference in behavioral ecology: some background, observations, and comparisons. — *Behav. Ecol. Sociobiol.* 65: 23-35.
- Cockrem, J.F. (2007). Stress, corticosterone responses and avian personalities. — *J. Ornithol.* 148: 169-178.
- Croxall, J.P., Butchart, S.H.M., Lascelles, B., Stattersfield, A.J., Sullivan, B., Symes, A. & Taylor, P. (2012). Seabird conservation status, threats and priority actions: a global assessment. — *Bird Conserv. Int.* 22: 1-34.
- Dammhahn, M. & Almeling, L. (2012). Is risk taking during foraging a personality trait? A field test for cross-context consistency in boldness. — *Anim. Behav.* 84: 1131-1139.

- Dingemanse, N.J., Bouwman, K.M., van de Pol, M., van Overveld, T., Patrick, S.C., Matthysen, E. & Quinn, J.L. (2012). Variation in personality and behavioural plasticity across four populations of the great tit *Parus major*. — J. Anim. Ecol. 81: 116-126.
- Dingemanse, N.J., Kazem, A.J.N., Réale, D. & Wright, J. (2010). Behavioural reaction norms: animal personality meets individual plasticity. — Trends Ecol. Evol. 25: 81-89.
- Duckworth, R.A. (2010). Evolution of personality: developmental constraints on behavioral flexibility. — Auk 127: 752-758.
- Friesen, V.L., Anderson, D.J., Steeves, T.E., Jones, H. & Schreiber, E.A. (2002). Molecular support for species status of the Nazca booby (*Sula granti*). — Auk 119: 820-826.
- Garamszegi, L.Z., Eens, M. & Török, J. (2008). Birds reveal their personality when singing. — PLoS One 3: e2647.
- Garamszegi, L.Z., Rosivall, B., Rettenbacher, S., Markó, G., Zsebők, S., Szöllősi, E., Eens, M., Potti, J. & Török, J. (2012). Corticosterone, avoidance of novelty, risk-taking and aggression in a wild bird: no evidence for pleiotropic effects. — Ethology 118: 621-635.
- Gosling, S.D. (1998). Personality dimensions in spotted hyenas (*Crocuta crocuta*). — J. Comp. Psychol. 112: 107-118.
- Herborn, K.A., Macleod, R., Miles, W.T.S., Schofield, A.N.B., Alexander, L. & Arnold, K.E. (2010). Personality in captivity reflects personality in the wild. — Anim. Behav. 79: 835-843.
- Huyvaert, K.P. & Anderson, D.J. (2004). Limited dispersal by Nazca boobies, *Sula granti*. — J. Avian Biol. 35: 46-53.
- Jennings, D.J., Hayden, T.J. & Gammell, M.P. (2013). Personality and predictability in fallow deer fighting behaviour: the relationship with mating success. — Anim. Behav. 86: 1041-1047.
- Kontiainen, P., Pietiainen, H., Huttunen, K., Karell, P., Kolunen, H. & Brommer, J.E. (2009). Aggressive Ural owl mothers recruit more offspring. — Behav. Ecol. 20: 789-796.
- Kralj-Fiser, S., Scheiber, I.B.R., Blejec, A., Moestl, E. & Kotrschal, K. (2007). Individualities in a flock of free-roaming greylag geese: behavioral and physiological consistency over time and across situations. — Horm. Behav. 51: 239-248.
- Lessells, C.M. & Boag, P.T. (1987). Unrepeatable repeatabilities: a common mistake. — Auk 104: 116-121.
- Li, D., Tsutsui, K., Muneoka, Y., Minakata, H. & Nomoto, K. (1996). An oviposition-inducing peptide: isolation, localization, and function of avian galanin the quail oviduct. — Endocrinology 137: 1618-1626.
- Maness, T.J. & Anderson, D.J. (2007). Serial monogamy and sex ratio bias in Nazca boobies. — Proc. Roy. Soc. Lond. B: Biol. Sci. 274: 2047-2054.
- Maness, T.J. & Anderson, D.J. (2008). Mate rotation by female choice and coercive divorce in Nazca boobies, *Sula granti*. — Anim. Behav. 76: 1267-1277.
- Mathot, K.J., Wright, J., Kempnaers, B. & Dingemanse, N.J. (2012). Adaptive strategies for managing uncertainty may explain personality-related differences in behavioural plasticity. — Oikos 121: 1009-1020.

- McDougall, P.T., Réale, D., Sol, D. & Reader, S.M. (2006). Wildlife conservation and animal temperament: causes and consequences of evolutionary change for captive, reintroduced, and wild populations. — *Anim. Conserv.* 9: 39-48.
- Montiglio, P.-O., Garant, D., Pelletier, F. & Réale, D. (2012). Personality differences are related to long-term stress reactivity in a population of wild eastern chipmunks, *Tamias striatus*. — *Anim. Behav.* 84: 1071-1079.
- Nakagawa, S. & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. — *Biol. Rev. Camb. Philos. Soc.* 85: 935-956.
- Nelson, J.B. (1978). The Sulidae: gannets and boobies. — Oxford University Press, Oxford.
- Osborne, J.W. (2002). Notes on the use of data transformations. — *Pract. Assessment, Res. Eval.* 8: 1-11.
- Patrick, S.C., Charmanier, A. & Weimerskirch, H. (2013). Differences in boldness are repeatable and heritable in a long-lived marine predator. — *Ecol. Evol.* 3: 4291-4299.
- Petelle, M.B., McCoy, D.E., Alejandro, V., Martin, J.G. & Blumstein, D.T. (2013). Development of boldness and docility in yellow-bellied marmots. — *Anim. Behav.* 86: 1147-1154.
- Réale, D. & Festa-Bianchet, M. (2003). Predator-induced natural selection on temperament in bighorn ewes. — *Anim. Behav.* 65: 463-470.
- Réale, D., Dingemanse, N.J., Kazem, A.J.N. & Wright, J. (2010). Evolutionary and ecological approaches to the study of personality. — *Philos. Trans. Roy. Soc. Lond. B: Biol. Sci.* 365: 3937-3946.
- Ricklefs, R.E. (1998). Evolutionary theories of aging: confirmation of a fundamental prediction, with implications for the genetic basis and evolution of life span. — *Am. Nat.* 152: 24-44.
- Sapolsky, R. & Ray, J. (1989). Styles of dominance and their endocrine correlates among wild olive baboons. — *Am. J. Primatol.* 18: 1-13.
- Sih, A., Bell, A.M., Johnson, J.C. & Ziemba, R.E. (2004). Behavioral syndromes: an integrative overview. — *Q. Rev. Biol.* 79: 241-277.
- Sinervo, B. & Svensson, E. (2002). Correlational selection and the evolution of genomic architecture. — *Heredity (Edinb)*. 89: 329-338.
- Sinn, D.L. & Moltschaniwskyj, N.A. (2005). Personality traits in dumpling squid (*Euprymna tasmanica*): context-specific traits and their correlation with biological characteristics. — *J. Comp. Psychol.* 119: 99-110.
- Stamps, J. & Groothuis, T.G.G. (2010). The development of animal personality: relevance, concepts and perspectives. — *Biol. Rev. Camb. Philos. Soc.* 85: 301-325.
- von Merten, S. & Siemers, B.M. (2012). Exploratory behaviour in shrews: fast-lived *Sorex* versus slow-lived *Crocidura*. — *Anim. Behav.* 84: 29-38.
- Westneat, D.F., Hatch, M.I., Wetzel, D.P. & Ensminger, A.L. (2011). Individual variation in parental care reaction norms: integration of personality and plasticity. — *Am. Nat.* 178: 652-667.
- Wilson, A.J. & Nussey, D.H. (2010). What is individual quality? An evolutionary perspective. — *Trends Ecol. Evol.* 25: 207-214.

- Wolf, M., van Doorn, G.S., Leimar, O. & Weissing, F.J. (2007). Life-history trade-offs favour the evolution of animal personalities. — *Nature* 447: 581–584.
- Wolf, M. & Weissing, F.J. (2012). Animal personalities: consequences for ecology and evolution. — *Trends Ecol. Evol.* 27: 1–10.

Appendix A: Shaking and gardening in Nazca boobies

Earlier work identified gardening as a dual-function behaviour for males. Males may garden as an advertisement to attract females, or as a territorial behaviour during conflicts with neighbours (Nelson, 1978; Maness & Anderson, 2008). We observed that gardening increased in response to a mild stressor, such as the first novel object test, compared to behaviour immediately prior to introduction of the novel object (Wilcoxon Matched Pairs test: $T = 1088.5$, $p < 0.00$, $N = 126$). Both sexes gardened during and following social and non-social disturbance (e.g., novel object tests), suggesting that social interaction is not necessary to produce gardening behaviour. Instead, gardening appears to have three functions: mate advertisement (for males), territorial display (for both sexes), and displacement activity following/during disturbance (for both sexes). Agitation-related behaviours include increased inappropriate activity that is not explained by needs (Fugate et al., 1997; Twelftree & Qazi, 2006), and one view of agitation is that it expresses underlying anxiety (Fugate et al., 1997). For the purposes of this paper we will refer to behaviours as anxiety-, not agitation-related, because this is the more common term in animal models. Thus, gardening, which is an inappropriate activity when confronted with a heterospecific nest intruder or novel object, is probably an anxiety-related behaviour in some contexts, similar to displacement behaviours in other species (Maestriperi et al., 1992).

Shaking behaviour in our study included large body and head shakes, as well as small ‘shivers’ that are extremely quick and ephemeral shakes of the head and neck. Shaking was described by Nelson (1978) in several different forms: the ‘oblique headshake’, ‘rotary headshake’, and ‘dogshake’. The dogshake is a tail, body, and head shake, and was scored as three different simultaneous shakes in our study. Nelson attributed no signal value to the dogshake or rotary headshake and suggested that oblique headshakes were mate appeasement or aggressive social displays. We did not differentiate between oblique and rotary headshakes because the slight tilt of the bill was difficult to discern in the field. Like gardening, shaking increased following

introduction of the first novel object (Wilcoxon Matched Pairs test: $T = 294.5$, $p < 0.00$, $N = 127$) and was performed during all tests, not only the social stimulus test. Thus, the performance of this behaviour outside of a social context suggests three functions, the first two of which were suggested by Nelson (1978): an aggressive signal, settling of the feathers (maintenance), and also an anxiety- or agitation-related displacement activity following disturbance, similar to body shakes in primates (Maestripieri et al., 1992).

References

- Costello, A.B. & Osborne, J.W. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. — *Pract. Assessment Res. Eval.* 10: 1-9.
- Fugate, L.P., Spacek, L.A., Kresty, L.A., Levy, C.E., Johnson, J.C. & Mysiw, W.J. (1997). Definition of agitation following traumatic brain injury: I. A survey of the brain injury special interest group of the american academy of physical medicine and rehabilitation. — *Arch. Phys. Med. Rehabil.* 78: 917-923.
- Maestripieri, D., Schino, G., Aureli, F. & Troisi, A. (1992). A modest proposal: displacement activities as an indicator of emotions in primates. — *Anim. Behav.* 44: 967-979.
- Maness, T.J. & Anderson, D.J. (2008). Mate rotation by female choice and coercive divorce in Nazca boobies, *Sula granti*. — *Anim. Behav.* 76: 1267-1277.
- Nelson, J.B. (1978). *The Sulidae: gannets and boobies*. — Oxford University Press, Oxford.
- Twelftree, H. & Qazi, A. (2006). Relationship between anxiety and agitation in dementia. — *Aging Ment. Health* 10: 362-367.

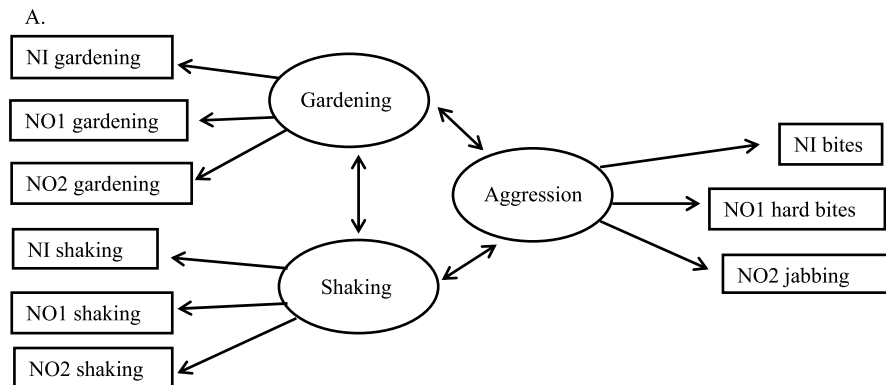
Appendix B: Structural equation modelling and factor analysis

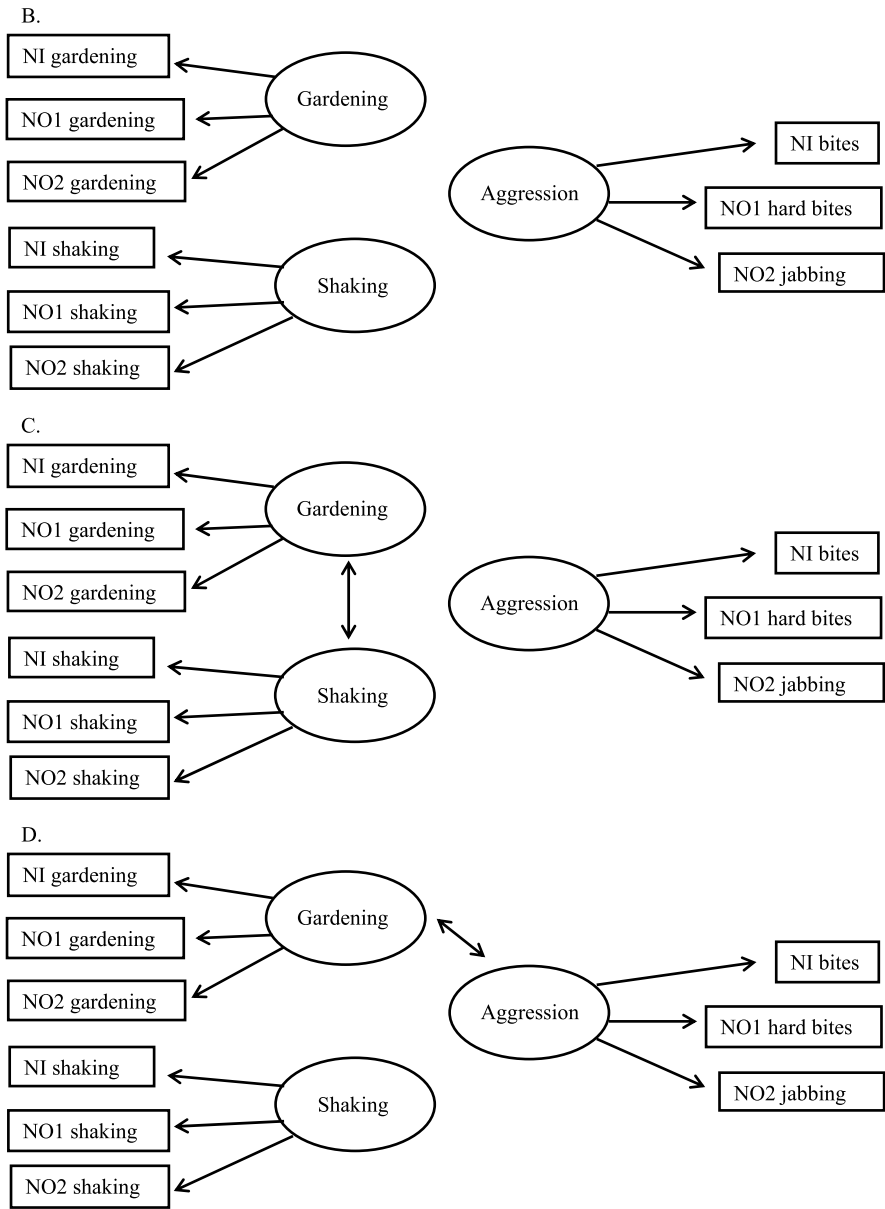
In this study, we used Structural Equation Modelling (SEM) to analyse covariance between personality traits. Structural Equation Modelling (SEM) has the advantage of allowing covariances between latent variables (similar to ‘factors’ in a Factor Analysis) to be examined within one model, instead of analysing a series of covariances between factors in separate tests. Additionally, SEM allows direct and indirect effects of observed variables going to and coming from latent variables, an aspect which was not used in this study but makes SEM a useful analytical technique for personality studies where effects may be added in the future. In any case, Factor Analysis was inappropriate for our data. For comparison with the SEM approach, we attempted Factor Analysis in SPSS (version 15.0; SPSS, Chicago, IL, U.S.A.)

using only the first round of the first session for each bird (following transformations to normalize and correct confounding factors as described in the Methods section of the text), using Principal Axis extraction followed by Oblimin rotation with Kaiser normalization. Communalities were low (0.19–0.51) and the Kaiser–Meyer–Olkin Measure of Sampling Adequacy was barely met at 0.519. A scree test (examination of eigenvalues) suggested no natural break point in factor eigenvalues (Costello & Osborne, 2005). When the number of factors retained was forced to only those above an eigenvalue of 0.7, factors were not readily interpretable, with most being driven primarily by a single behaviour or test. For these reasons, Factor Analysis was unsuitable for our data and was not conducted for comparison with SEM results.

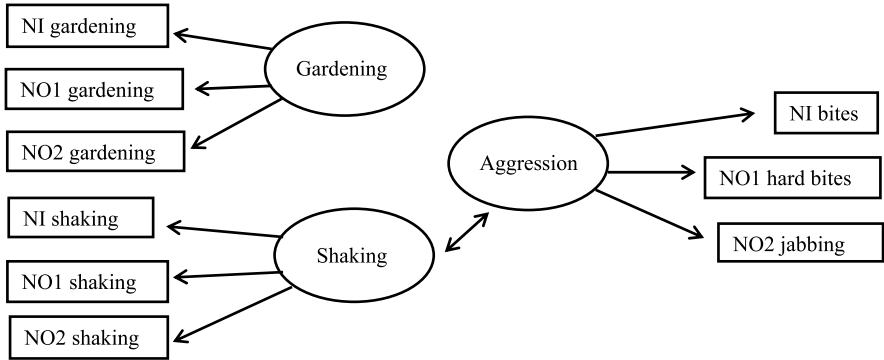
Appendix C: SEM models of the relationships between behaviours in the nest intruder (NI), novel object 1 (NO1) and novel object 2 tests (NO2)

This same set of models was tested for the NI, NO1 and social stimulus SS tests, with SS replacing NO2. Covariances are represented by double-headed arrows, and regressions by single-headed arrows. Top models were determined using 95% confidence sets of AICc weights. Gardening, shaking, and aggressive behaviours were the only behaviours performed consistently across all tests by both sexes, and so were the only behaviour types analyzed. Models A–H converged, while models I–K were unable to converge, and so not considered in the model set for AICc comparison.

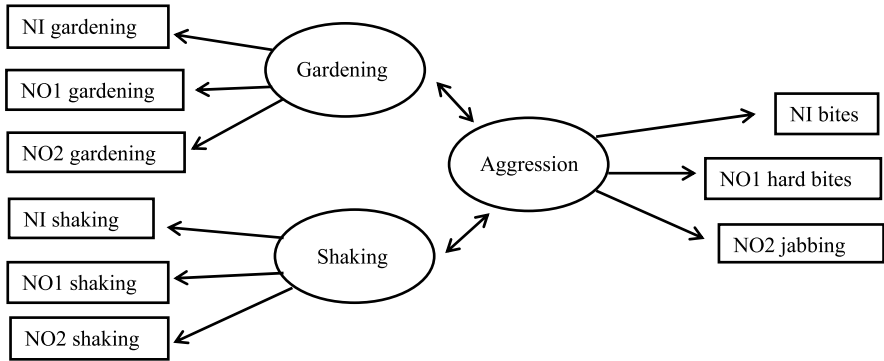




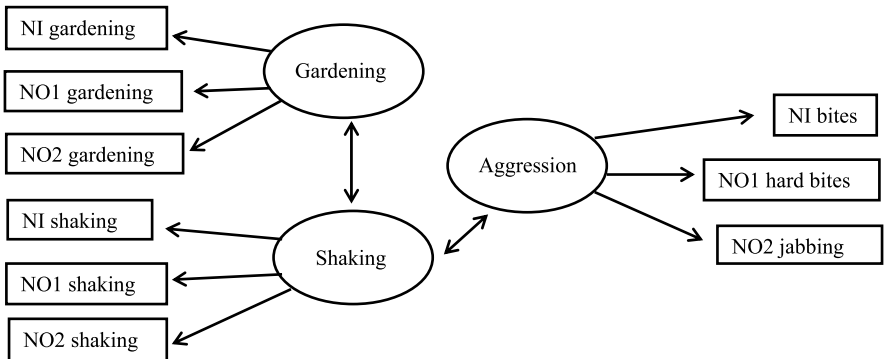
E.



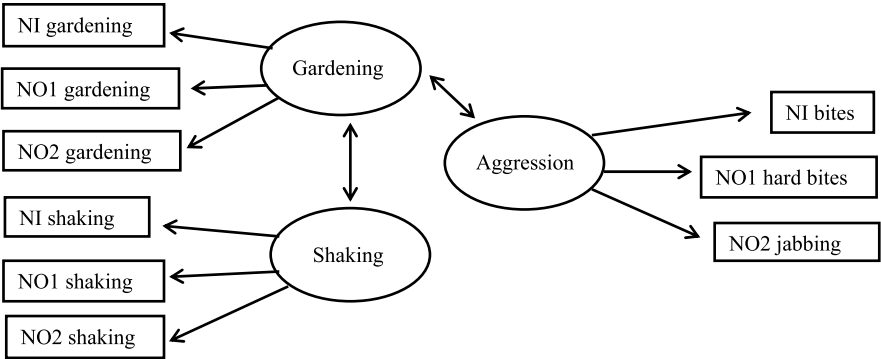
F.



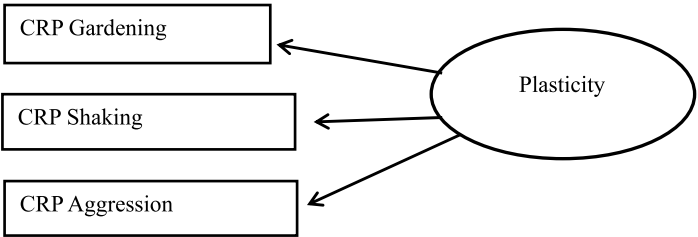
G.



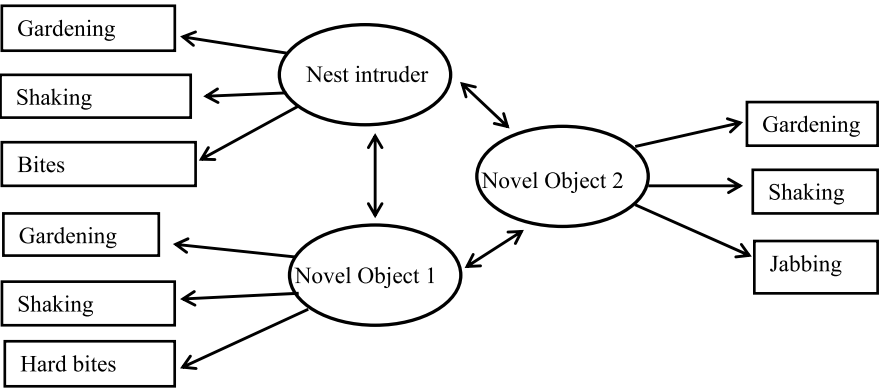
H.

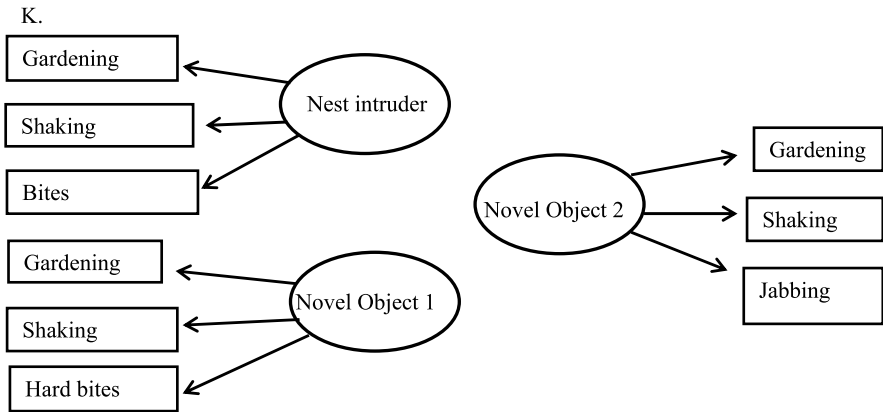


I. This same model was used for Coefficient of Relative Plasticity (CRP) values calculated between the nest intruder and both novel object tests, and the nest intruder, first novel object test, and social stimulus tests.



J.





Appendix D.

Models predicting gardening, shaking, and aggression during the nest intruder (NI), first novel object (NO1), second novel object (NO2) and social stimulus tests (SS).

Model	Gardening				Shaking				Aggression			
	NI	NO1	NO2	SS	NI	NO1	NO2	SS	NI	NO1	NO2	SS
Null					x	x			x			
Sex + Session + Time, all interactions												
Sex + Session + Time, Sex × Session Sex × Time, Session × Time												
Sex + Session, Sex × Session									x	x	x	x
Sex + Time, Sex × Time	x	x	x	x	x	x						
Session + Time, Session × Time							x	x	x			
Sex	x	x	x	x	x							
Session						x	x	x			x	x
Time					x	x		x				
Session + Sex	x	x	x	x		x	x	x	x	x	x	x
Time + Sex	x	x	x	x	x							
Time + Session					x	x	x	x			x	x

An 'x' indicates that the model was included in the 95% confidence set calculated using AICc weights. 'Time' is time of day tested (morning or afternoon), 'Session' is behaviour session (see Figure 1), and 'Sex' is sex of the focal bird. × indicates an interaction, + indicates an additive effect.

Appendix E: Correlations between contextual plasticity for behaviours and their associated latent variables scores.

Correlations between contextual plasticity for behaviours ('Plasticity of Behaviour') and their associated latent variables scores ('Behaviour') generated by Structural Equation Modelling of z -scored behavioural counts during the nest intruder (NI), novel object 1 (NO#1) and novel object 2 (NO#2) tests, or during the NI, NO#1 and SS tests.

