ORIGINAL ARTICLE

Personality traits are consistent when measured in the field and in the laboratory in African striped mice (*Rhabdomys pumilio*)

C. H. Yuen[1] · N. Pillay[1] · M. Heinrichs[2,3] · I. Schoepf[1] · C. Schradin[1,4,5,6]

Abstract

Personality in free-living individuals has predominantly been measured under standardized laboratory conditions. Such measurements have been then linked to life-history traits, fitness and survival. Yet, it remains unclear how such personality measurements reflect the variation shown by free-living individuals, if the same measurements were taken directly in their natural environment. Here, we used free-living African striped mice to test whether the personality traits of activity, boldness, exploration and aggression are consistent when measured in the laboratory and in the field contexts. First, we established whether personality traits were repeatable and consistent within one context. Next, we compared measurements across the two different contexts. Additionally, we established whether personality traits were correlated with one another in behavioural syndromes and assessed whether the resulting syndromes were consistent across the two contexts. All personality traits in the laboratory were measured using classical personality tests. The same tests were then modified and used to measure personality of the same individuals in the field. All personality traits were highly repeatable and consistent within the same context. In addition, individuals behaved consistently for all the behaviours measured both in the laboratory and in the field. Further, we found that the presence of two correlated context-specific separate latent variables (one for the field and one for the laboratory) underpinned all the behaviours measured, indicating that there is a context-specific syndrome in this species. Overall, our results confirm that measurements of personality traits of wild striped mouse individuals recorded in the laboratory environment are consistent with the traits that the same individuals show under natural conditions.

Significance statement

Animal personality research has been mostly focused on measuring traits of wild individuals under standardized laboratory conditions. Most recently, however, such approaches have come under increased scrutiny as it is unclear whether personality measured in the laboratory reflects the variation that the same individuals would show if the same measurements were taken directly in their natural habitat. Here, we used wild African striped mice to test whether the personality traits of activity, boldness, exploration and aggression are consistent when measured in the laboratory and in the field contexts. Using similar protocols, we established whether personality traits were repeatable and consistent within one context and then compared measurements across the two different contexts. We showed that personality measures from standardized laboratory conditions are consistent with field measurements for all personality traits considered.

✉ C. H. Yuen
bt8205@yahoo.co.uk

[1] School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, Johannesburg, South Africa

[2] Department of Psychology, Laboratory for Biological and Personality Psychology, University of Freiburg, D-79104 Freiburg, Germany

[3] Freiburg Brain Imaging Center, University Medical Center, University of Freiburg, D-79106 Freiburg, Germany

[4] Department of Animal Behaviour, Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland

[5] IPHC-DEPE, Université de Strasbourg, 23 rue Becquerel, 67087 Strasbourg, France

[6] CNRS, UMR7178, 67087 Strasbourg, France

Springer

Introduction

Understanding how and why individual variation in behaviour is maintained in a population is an important area of research in animal behaviour. Variation among individuals that is consistent over time and across contexts is defined as personality (Sih et al. 2004; Bell 2007). Personality studies have described the predictable manner in which individuals maintain consistent differences in the face of environmental challenges (van Overveld and Matthysen 2013). Such research has often involved capturing free-living individuals and measuring their behaviour under standardized conditions in the laboratory, and then relating personality measurements to varying life-history and fitness parameters (Dingemanse et al. 2004; Bell 2005). These studies have been extremely valuable for demonstrating that individual variation in one behaviour is often linked with variation in other behaviours, creating behavioural syndromes (Sih et al. 2004), which are ecologically relevant, such as dispersal (Cote et al. 2010), foraging (Quinn et al. 2012), predator-avoidance (Jones et al. 2009) and space-use (Kurvers et al. 2010). While testing individuals under standardized conditions eliminates several extrinsic influences on the data (Campbell et al. 2009), it is unclear whether the personality of free-living individuals measured in a laboratory environment is truly representative of the behaviour that the same individuals would show if the measurements were taken directly in their natural environment (Herborn et al. 2010; Niemela and Dingemanse 2014).

Behaviour measured under laboratory conditions can be adversely affected by stress brought about by the artificial environment, which may result in a modification to their gene expression and behaviour (individual × environment and genes × environment; Hodgins-Davis and Townsend 2009; Niemela and Dingemanse 2014). Biro (2012) demonstrated that initial tests in a novel laboratory setting did not relate to later tests within the same settings, and that individuals tested in a familiar environment display different behaviour when tested in a novel but also artificial environment, leading him to question whether personality measured from a single assay under artificial conditions can be reliably used to infer personality in nature. Similarly, Carter et al. (2012) showed that multiple assays meant to measure a single trait may not always relate to each other. Consequently, laboratory tests may produce behavioural differences between behavioural types that are not present in nature and vice versa (Herborn et al. 2010), especially in studies where wild individuals are particularly sensitive to being handled and housed in captivity. Studies of great tits (*Parus major*) performed in the laboratory have, for example, found an overall negative correlation between dominance rank and exploratory tendency (Verbeek et al. 1999), but when the relationship between dominance and exploration was investigated in the wild, it was only negative in non-territorial juvenile males

(Dingemanse and De Goede 2004). Psychologists have shown that humans show low consistency in their behavioural traits when these are measured in different contexts. For example, an early study found that honesty of school children was not consistent across different situations (e.g. at home or at school; Hartshorne and May 1928). In another study of 300 college students, there was no consistency for the personality trait punctuality across different situations (Dudycha 1936). These studies show that measuring personality in the laboratory only may be misleading and limit the ability to predict the ecological significance of personality traits in captivity (Herborn et al. 2010). It would therefore be timely and necessary to test whether wild behavioural types can extend to the laboratory, particularly in the light of how environmental sensitivity can affect gene expression and behaviour (Niemela and Dingemanse 2014).

While the importance of measuring personality of individuals directly in their natural environment is widely recognized (Bell 2012; Niemela and Dingemanse 2014), this often remains difficult to achieve practically because obtaining reliable measures necessitates individuals being captured and handled multiple times. To date, a few studies have successfully compared results obtained from individuals tested in their natural environment with results obtained from the same individuals under captive conditions (Coleman and Wilson 1998; Brown et al. 2005; Wilson and McLaughlin 2007; Briffa et al. 2008; Hollander et al. 2008; Herborn et al. 2010; Cole and Quinn 2014). While these studies underline the importance of comparing laboratory with field tests, their strength is often constrained by a lack of validation or biased by the use of different types of tests in the field and in captivity to measure the same trait. For example, Herborn et al. (2010) investigated personality in free-living blue tits (*Cyanistes caeruleus*) by cleverly adapting the exploration test of Verbeek et al. (1994) and the novel object test of Greenberg (1984) developed in the laboratory to measure individual variation in exploratory tendency and neophobia in nature. Their results showed that personality measures obtained in captivity can uncover differences among individuals in their natural behaviour and demonstrated that the personality of individuals can be consistent over different contexts even in nature (Herborn et al. 2010). While their results provide important validation for captive versus free-living personality measures, each bird was tested in a non-random way (first in captivity and then in nature) which could have affected the findings. Further, there seems to be a general discrepancy and a lack of consistency between findings from the laboratory and the field by different authors. For example, Herborn et al. (2010) found positive relationships for two behaviours they measured, and Boon et al. (2008) have also confirmed that assays in the laboratory relate to similar behaviours in the field. However, recent work by Fisher et al. (2015) found relationships for activity and exploration, but not for boldness in field crickets. Similarly, Boyer et al. (2010) and van Overveld and Matthysen (2010) found relationships between different behaviours between the laboratory and the field. Thus, it

still remains unclear whether individual-level correlations measured in captivity remain consistent when measured in nature across a range of taxa.

In the present study, we investigated whether personality and behavioural syndromes observed in African striped mice (*Rhabdomys pumilio*) in the field also occurred in the laboratory. Striped mice are socially flexible, with individuals of both sexes following alternative reproductive tactics (Schradin et al. 2012). In a previous study, we showed that wild-caught African striped mice show consistency in personality traits when measured under standardized conditions in a field laboratory (Yuen et al. 2015), using a battery of classical personality tests (i.e. open field, novel object and novel conspecific tests; Verbeek et al. 1996; van Oers et al. 2004; Réale et al. 2007). However, we do not know whether striped mice show consistency in personality traits under natural conditions, and whether this is correlated to personality traits measured under standardized laboratory conditions. Here, we examined whether the personality traits of activity, boldness, exploration and aggression were consistent across the laboratory-field context. To ensure that we measured the same behaviour in both the laboratory and the field, we used classical personality tests previously employed to study personality in striped mice in the laboratory (Yuen et al. 2015) and adapted them to the field. First, we tested whether personality was present within contexts, i.e. in the laboratory and in the field (within context comparisons) by repeatedly measuring individuals within the same context. Second, we correlated personality measures from the laboratory with the same measures from the same individuals tested in the field (across context comparisons). Finally, we tested whether the different personality traits were correlated with each other in behavioural syndromes and whether the laboratory and the field setting resulted in similar behavioural syndromes. To do so, we follow procedures outlined in Araya-Ajoy and Dingemanse (2014) and tested four a priori hypotheses: (a) was each behavioural type underpinned by a separate factor (the null model; Fig. 1a); (b) was a single latent variable affecting all behaviours in both the field and the laboratory environment (Fig. 1b); (c) were two context-specific separate latent variables underpinning all the behaviours (Fig. 1c); and (d) were two correlated context-specific separate latent variables underpinning all the behaviours (Fig. 1d)?

Materials and methods

### Study area and field techniques

Data were collected in the non-breeding seasons (December–April) between 2008 and 2012 on a field site located in the Goegap Nature Reserve, in the Succulent Karoo biome, South Africa (S41.56, E1.60). In the semi-arid Succulent Karoo, striped mice are typically group-living, with each group consisting of one breeding male, two to four breeding females and their philopatric offspring (Schradin and Pillay 2004). However, if population density is low during the breeding season, philopatrics leave their natal group and start to breed solitarily (Schradin et al. 2010; Schoepf and Schradin 2012a). Trapping, behavioural observations and radio-tracking were used to identify striped mice within the study site and to determine social tactics and group composition (Schradin and Pillay 2004, 2005; Schradin et al. 2010). Striped mice were trapped with Sherman-like metal traps ($26 \times 9 \times 9$ cm) baited with a mixture of bran flakes, currants, sea salt, and salad oil (Schradin 2005). Traps were set directly at striped mouse nests in the early morning and were checked 45 min later (Schradin 2005). Each trapped mouse was weighed, sexed and received a permanent ear-tag (National Band and Tag Co., Newport, KY, U.S.A.). Additionally, individuals were marked with a non-toxic hair dye (Inecto Rapido, Pinetown, South Africa), which aided with individual recognition during behavioural observations and field personality tests. Striped mice at our field site are habituated to our presence and readily enter traps once they are set. This allowed us to easily capture individuals that were observed during field tests for testing in the laboratory. Trapping and behavioural tests did not have any adverse effects on individuals' behaviour (Yuen et al. 2015). Behavioural observations were made at each group nest in the morning and evening to determine individual affiliation to specific groups. In addition, at least one breeding female from each group was fitted with a radio-collar (Holohil, Carp, Ontario, Canada; 2.5–4.4 g) and was radio-tracked to determine the nesting site location of the group (Schradin and Pillay 2005). Radio-tracking was carried out using an AOR 8000 wide range receiver (Tokyo, Japan), an H-antenna (Africa Wildlife Tracking, Pretoria, South Africa) and a global positioning system (GPS) navigation device (eTrex Venture, GARMIN International, USA) with an accuracy of ±5 m. All striped mice fitted with a transmitter were radio-tracked twice a day to determine ranging areas and sleeping sites.

Personality tests under standardized laboratory conditions

Test subjects were all adult. Captive and field tests were randomized so that half of the individuals were first tested in the laboratory while the other half were first tested in the field. All laboratory tests were performed within 2 weeks of the tests in the field and vice versa so that all individuals were measured under the same conditions (e.g. age, reproductive status, season). Multiple-samples per individuals were obtained by selecting specific individuals so that each individual was tested within a week of one another. Striped mice are diurnal, with peak activity in the early morning and evening (Schradin and Pillay 2004). Therefore, all individuals were tested in the early morning. Mice were trapped directly at their nests as they emerged to bask and were taken to the research station, where they were transferred to a type III Perspex

Fig. 1 Four models (hypotheses) explaining syndrome structure among the different behavioural types (activity, boldness, exploration and aggression) assayed in free-living striped mice in the laboratory and in the field. Model (*a*) predicted that each behavioural type was underpinned by a separate factor (the null model). Model (*b*) predicted that a single latent variable (referred to as "*R. pumilio* syndrome") affected all behavioural types. Model (*c*) predicted that two context-specific separate latent variables underpinned all the behavioural types. Model (*d*) predicted that two correlated context-specific separate latent variables underpinned all the behaviours types. For each model, we provide the ΔAIC as well as its associated Akaike weight (w_i)
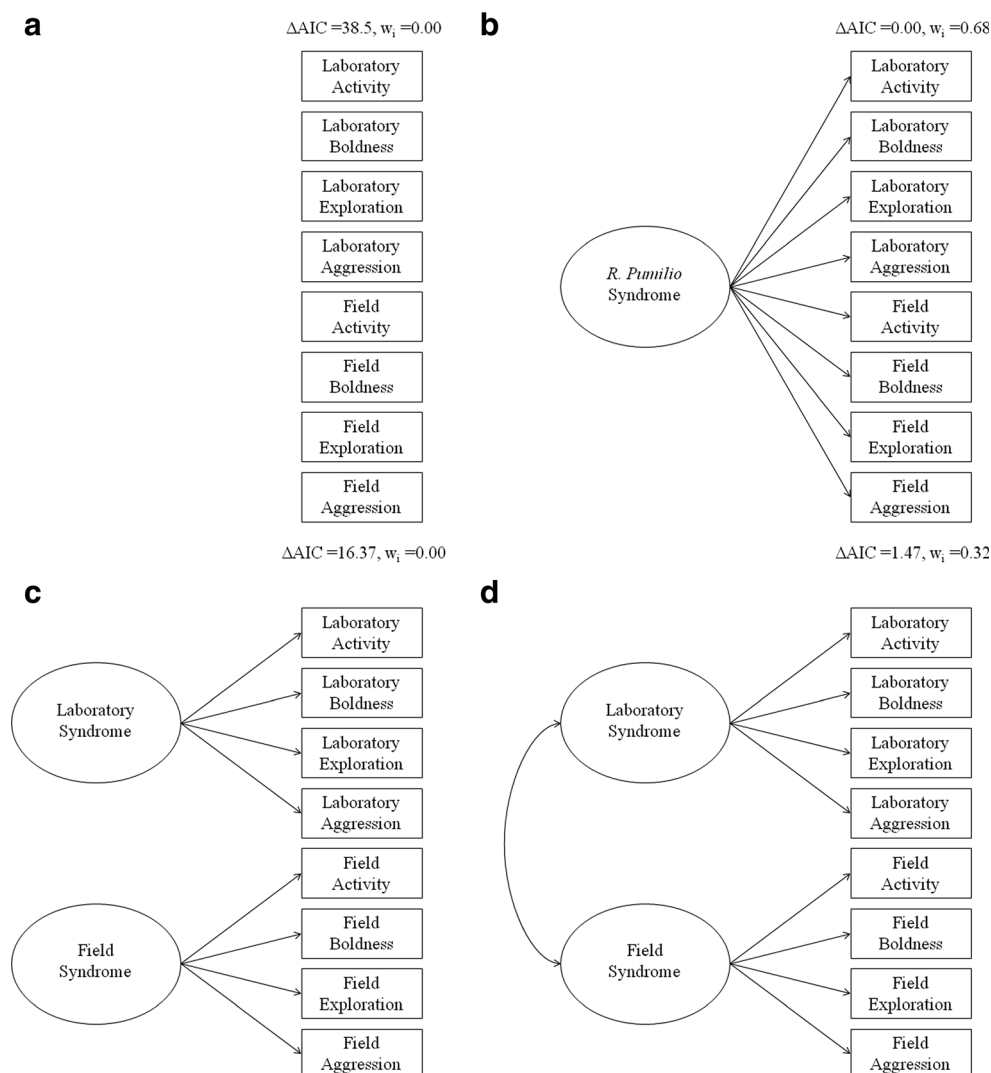
cage ($38 \times 22 \times 15$ cm). Each cage was provided with bedding (sand) and food (10 sun flower seeds) to account for hunger during tests. Mice were left to settle for a period of 10 min in the test room before being transferred individually to a neutral presentation arena made of wood chip (80×65 cm and 94 cm high, with a partition in the middle), similar to the one used in previous personality studies in striped mice (Schoepf and Schradin 2012b; Yuen et al. 2015). The presentation arena was cleaned with a mixture of odourless disinfectant (Dis-Chem Pharmacies, Northriding, South Africa) and water after each mouse had been tested. For all tests in the laboratory, we followed the same procedure that we validated previously (Yuen et al. 2015). Specifically, each focal mouse was sequentially tested for (a) activity and boldness, (b) exploration, and (c) aggression.

Activity was measured using an open field test (Wilson et al. 1976; Réale et al. 2007). During this test, a focal individual was placed in a corner of the arena for a period of 5 min. Activity was recorded every 15 s using instantaneous focal sampling (Martin and Bateson 1993) as the number of times

an individual spent being active. In the same open field test, boldness was recorded using continuous focal sampling (Martin and Bateson 1993) as the total time (in seconds) an individual spent at least half-a-mouse length away from the wall of the arena (estimated by sight).

Exploration was tested using a novel object test (Birke and Archer 1983; Greenberg 1984; Verbeek et al. 1994) which lasted for 5 min. A fixed object, consisting of a small plastic animal toy ($115 \times 20 \times 44$ mm), which was secured to the floor of the arena and could not be moved by the test subject, was set at the far side of the arena, in the opposite corner to where the focal individual was located. Exploration was measured as the latency (in seconds) it took the focal mouse to physically come into contact with the fixed object.

Aggression was tested in dyadic encounters with a novel conspecific test (Verbeek et al. 1994; Benus and Rondigs 1996), during which we tested the focal mouse against a stimulus individual of the same sex (the stimulus) from our captive colony, which was permanently maintained at the research

station. Stimulus individuals were always at least 3 g (but never more than 7 g) lighter than the focal animal. Because body mass has a positive influence on the outcome of aggressive encounters (Schradin 2004), we expected the focal mouse to initiate interactions. Aggression tests were performed using standard procedures previously used for striped mice (Schoepf and Schradin 2012b). Aggression was measured as the total number of aggressive encounters initiated by the focal individual during a period of 5 min. In addition to aggression, we also recorded sniffing the stimulus mouse, body contact between the dyad, allo-grooming and activity, but these behaviours occurred too infrequently for statistical analysis and were not considered any further.

To minimize the effect that the captive environment could have on personality, we kept mice in the laboratory for a maximum of 2 h before release (Yuen et al. 2015). A maximum of three individuals were tested in a day. Once tests ended, all mice were returned in good condition to the field and released in the same place where they were captured. To minimize observer bias, a blind protocol was adopted when all behavioural data were recorded and/or analyzed. A total of 41 individuals were measured for activity and boldness, 48 for exploration and 20 for aggression in the laboratory.

Personality tests in the field

To assess activity in the field, we used a modified version of the "whole-day follow" (Schradin 2006). All focal mice were fitted with radio-collars and followed for a period of 3 h during their peak activity times in the early morning and for another 3 h in the early evening (6 h total observation time per mouse). Activity was recorded as the frequency of all the "active" behaviours displayed by the focal individual (e.g. travelling, foraging, self-grooming). We recorded whether the mouse had been active or inactive in the past minute, and then calculated the percentage of the 180 recordings from the 6-h observation that the individual had been active. From the same observations, boldness was recorded as the time an individual spent in the open at least one mouse length away from the nearest shrub.

To assess exploration, we presented a novel object in front of individual nests (Fig. 2). The novel object was a plastic animal toy (115 × 20 × 44 mm), which was fixed to the ground and was the same as the one used in the neutral presentation arena tests in captivity. To be consistent with data collected under captive conditions, the novel object was placed at a distance of 70 cm away from the entrance of the nest. Exploration was recorded as the latency (in seconds) it took focal individuals to approach the novel object. Recording started as soon as an individual was seen outside of its nest. The novel object was cleaned between tests. As exploration in both the laboratory and the field was measured during a 5-min trial, the maximum value for exploration was always 300 s. Measures of exploration obtained in this way indicated that

Fig. 2 African striped mouse during an exploration test in the field, showing a striped mouse mounting a novel plastic toy (photograph by CHY)

individuals with high values were the least explorative. To facilitate interpretation of the results, we subtracted all exploration data from a value of 300 so that individuals with the highest score were the most explorative.

To assess aggression, we placed a food-scented box (the same as the one used for boldness tests) at the boundaries between two different group territories. Individual striped mice from two different groups were attracted by the scent from the box at the territory boundary. Aggression was measured as the total number of aggressive encounters between individuals belonging to different groups and the same aggressive behaviour patterns as in the laboratory were recorded.

To correct for difference in test length between the field and the laboratory and thus enable comparisons of data between the two contexts, all data were converted into behaviour/ minute prior to analysis. The same individuals were scored in both captive and field studies (i.e. a total of 41 individuals were measured for activity and boldness, 48 for exploration and 20 for aggression in the field). As such, each individual was assayed four times: twice in the laboratory and twice in the field. Among all individuals sampled, 18 individuals were measured in all tests and were used to determine the existence of potential behavioural syndromes among the different personality traits. Two individuals that were measured for aggression were not sampled for measurements of activity and boldness in the field and were thus excluded from the behavioural syndrome analysis.

Data analysis

Data analysis was performed using R version 3.0.2 (The R Foundation for Statistical Computing, Vienna, Austria). We checked for the normal distribution of the data using the Shapiro-Wilk test. To reach normality, activity, boldness and exploration were log-transformed whereas aggression was square-rooted. We used random intercept models to evaluate the degree of among-individual variation. Random intercept models were fitted using linear mixed effects models (LMMs, lmer; Package *lme4*; Bates et al. 2014). Each LMM was a

univariate model consisting of one of the behavioural traits (activity, boldness, exploration or aggression) as the response variable, while testing sequence (first, second) was the fixed factor. Individual ID was entered as a random factor in each model. Univariate models were calculated separately for each behaviour within each context. To check whether the degree of among-individual variance was significant at the 95 % level, we compared models that included the random effect of individual ID with simpler models without it, while maintaining the same fixed factors structure using likelihood ratio tests (Crawley 2007; Zuur et al. 2009). We used the package *RLRsim* (Scheipl 2010) using the exactLRT function to calculate accurate *P* values when comparing models with a single random effect to models with no random effect (*P* values were based on 10 000 simulated values; Crainiceanu and Ruppert 2004). For all our models, we report the R^2 (adjusted), as calculated following Nakagawa and Schielzeth (2013). We verified our model selection by (1) plotting the model residuals versus the fitted values, (2) checking the normal distribution of the model residuals using normal probability plots, (3) checking for heteroscedasticity, and (4) leverage (Crawley 2007). To assess the proportion of phenotypic variation attributable to between-individual variation, we calculated the coefficient of repeatability *R* and estimated the 95 % confidence intervals (CI) around the repeatability estimates for each behaviour in each context (laboratory, field) separately (Nakagawa and Schielzeth 2010). Repeatabilities (adjusted) were calculated for each model as the between-individual variance divided by the sum of the between-individual and the residual variance (Nakagawa and Schielzeth 2010).

Additional linear models were used to assess whether personality measured under standardized conditions in the laboratory were good predictors of personality measured in the field. Each of these models were constructed following Herborn et al. (2010) and Fisher et al. (2015) and had one of the behavioural scores measured in the wild (e.g. activity in the field) as the response variable, and the corresponding measured score for that individual's behaviour in the arena (e.g. activity in the laboratory) as the fixed factor. Individual ID was included in each model as the random factor to control for possible bias arising when repeated measures were taken from the same individual.

We used structural equation models (SEM; Package *lavaan*; Rosseel 2012) to investigate whether the different personality traits resulted in behavioural syndromes (Dochtermann and Jenkins 2007). To do so, we followed procedures outlined in Araya-Ajoy and Dingemanse (2014) and tested the four above-mentioned a priori hypotheses. Support for each model was determined by calculating Akaike information criteria (AIC). We selected the model that best fitted our data by selecting the model that yielded the lowest AIC (Dochtermann and Jenkins 2007). Repeated measures taken from individuals within each context were averaged prior to all SEM analysis. Pair-wise Spearman rank correlations (rs) were additionally calculated between the different behavioural characteristics to further elucidate syndrome structure. Because we conducted multiple comparisons, all the *P* values were adjusted using the Benjamini-Hochberg method (Benjamini and Hochberg 1995). All tests were two-tailed. For all tests, a significance level (α) of 0.05 was selected. Data are presented as mean and confidence intervals. Data were *z* transformed prior to analysis.

Results

Consistency and repeatability of personality traits in the laboratory

Including the random effect of mouse ID in our LMMs improved model fit for activity (ExactLRT: L.Ratio = 60.13, *P* < 0.0001; R^2adj. = 0.88; Tables 1 and 2), boldness (ExactLRT: L.Ratio = 30.81, *P* < 0.0001; R^2adj. = 0.73; Tables 1 and 2), exploration (ExactLRT: L.Ratio = 17.15, *P* < 0.0001; R^2adj. = 0.56; Tables 1 and 2) and aggression (ExactLRT: L.Ratio = 7.01, *P* = 0.006; R^2adj. = 0.54; Tables 1 and 2), suggesting that there was an inter-individual difference in the level of activity, boldness, exploration and aggression within the laboratory environment. Individuals displayed significant repeatability in all four behaviours when they were measured in the arena (activity: *P* < 0.0001; boldness: *P* < 0.0001; exploration: *P* = 0.0005; aggression: *P* = 0.03; Fig. 3).

Table 1 Mean and confidence interval for activity, boldness, exploration and aggression observed in each of the two contexts (laboratory, field) in free-living striped mice in the Succulent Karoo (South Africa)

Behavioural Trait	Laboratory		Field	
	First	Second	First	Second
Activity	2.29 (1.88, 2.71)	2.52 (2.16, 2.89)	0.40 (0.32, 0.47)	0.48 (0.40, 0.59)
Boldness	4.49 (3.13, 5.85)	5.60 (3.96, 7.23)	0.18 (0.12, 0.23)	0.15 (0.11, 0.19)
Exploration	14.69 (8.22, 21.15)	21.60 (14.81, 28.39)	17.99 (10.77, 25.22)	19.04 (11.96, 26.11)
Aggression	0.54 (0.25, 0.83)	0.64 (0.20, 1.08)	0.05 (0.02, 0.07)	0.03 (0.01, 0.05)

Each behaviour was measured in behaviour/min

Table 2 Summary of the results obtained from univariate mixed-effect models on each behavioural type within each context

		Activity	Boldness	Exploration	Aggression
Laboratory	Fixed effects				
	Intercept	1.61 (1.55, 1.67)	0.54 (0.40, 0.68)	−0.76 (−1.07, −0.46)	0.91 (0.71, 1.10)
	Test	−0.03 (−0.06, 0.00)	0.10 (0.00, 0.20)	0.37 (0.08, 0.67)	−0.01 (−0.20, 0.18)
	Random effects				
	Mouse ID	0.03 (−0.15, 0.23)	0.14 (−0.29, 0.49)	0.65 (−0.57, 1.06)	0.11 (−0.17, 0.50)
	Residual	0.01 (−0.06, 0.09)	0.05 (−0.19, 0.29)	0.54 (−0.60, 0.90)	0.09 (−0.22, 0.41)
Field	Fixed effects				
	Intercept	0.49 (0.31, 0.67)	0.24 (0.06, 0.43)	0.47 (0.44, 0.52)	0.93 (0.72, 1.13)
	Test	0.11 (−0.04, 0.26)	−0.03 (−0.23, 0.16)	−0.01 (−0.05, 0.11)	−0.07 (−0.23, 0.08)
	Random effects				
	Mouse ID	0.22 (−0.35, 0.62)	0.16 (−0.24, 0.56)	0.46 (0.33, 0.52)	0.15 (−0.26, 0.57)
	Residual	0.12 (−0.28, 0.43)	0.19 (−0.36, 0.55)	0.14 (−0.16, 0.25)	0.06 (−0.18, 0.34)

Estimates were derived separately for each behaviour. Each model included mouse ID as the random factor, while test (first or second) was the fixed effect. For each model, we report point estimates for the fixed (mean) and the random parameters (variance) along with their 95 % confidence intervals

Consistency and repeatability of personality traits in the field

We found that the random effect of mouse ID in our LMMs was significant for activity (ExactLRT: L.Ratio = 22.76, $P < 0.0001$; R^2adj. = 0.66; Tables 1 and 2), boldness (ExactLRT: L.Ratio = 9.35, $P = 0.002$; R^2adj. = 0.45; Tables 1 and 2), exploration (ExactLRT: L.Ratio = 49.20, $P < 0.0001$; R^2adj. = 0.77; Tables 1 and 2) and aggression (ExactLRT: L.Ratio = 14.60, $P = 0.0001$; R^2adj. = 0.72; Tables 1 and 2), suggesting that there was an inter-individual difference in the level of activity, boldness, exploration and aggression within the field environment. Individuals displayed significant
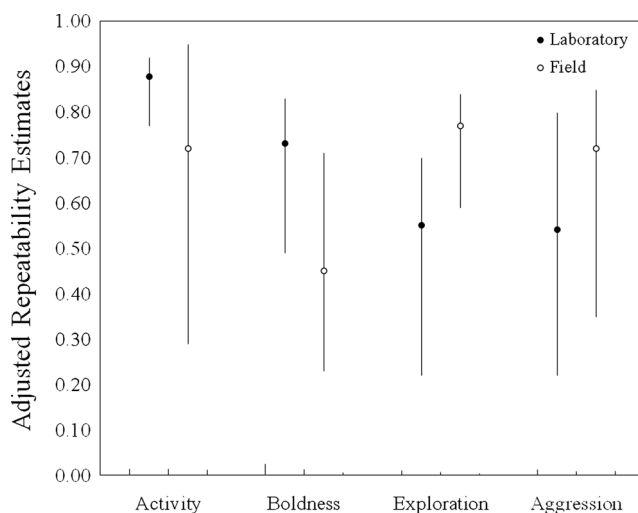

Fig. 3 Estimates of the repeatability measures obtained when running repeated tests within the laboratory and the field context. *Black circles* represent adjusted repeatability measures from the arena. *White circles* represent adjusted repeatability measures from the field. Adjusted repeatability measures are reported together with their 95 % confidence intervals

repeatability in all four behaviours when they were measured in the field (activity: $P = 0.004$; boldness: $P = 0.01$; exploration: $P < 0.0001$; aggression $P = 0.04$; Fig. 3).

Comparisons of personality traits between the laboratory and the field

Personality measured in captivity was a good predictor for personality measured in the field for all the behavioural characteristics measured. Specifically, models, which included the fixed effect of captivity better explained our data than models without them (activity: $\chi^2 = 9.64$, $P = 0.002$; R^2adj. = 0.65; boldness: $\chi^2 = 6.21$, $P = 0.01$; R^2adj. = 0.45; exploration: $\chi^2 = 25.99$, $P < 0.0001$; R^2adj. = 0.81; aggression: $\chi^2 = 24.17$, $P < 0.0001$; R^2adj. = 0.72; Table 3).

Behavioural syndromes in the laboratory and in the field

The comparison of our four a priori hypotheses using structural equation modelling (SEM) resulted in two models with similar AICs, which could have potentially explained our data: model 2 (AIC = 393.93) and model 4 (AIC = 395.39). However, the cross-context correlation between the latent variables of field and laboratory was rather strong ($z = 5.16$, $P < 0.0001$), suggesting that model 4 was a better fit for our data (Fig. 2). This model predicted that two correlated context-specific separate latent variables underpinned all the behaviours (Fig. 4). Pair-wise correlations of the different personality traits showed the existence of a negative behavioural syndrome between boldness and exploration in both the laboratory and in the field (Table 4), indicating that the boldest individuals took the least amount of time to approach the novel object in both environments. Two further behavioural syndromes were found in the laboratory: a behavioural

Table 3 Summary of the results obtained from univariate mixed-effect models to test whether personality measured under standardized conditions in the laboratory were good predictors of personality measured in the field

	Activity	Boldness	Exploration	Aggression
Fixed effects				
Intercept	2.34 (1.18, 3.53)	0.03 (−0.21, 0.27)	0.69 (0.57, 0.81)	1.10 (0.97, 1.24)
Test	0.08 (−0.07, 0.22)	−0.07 (−0.27, 0.12)	−0.03 (−0.11, 0.04)	−0.09 (−0.27, 0.08)
Laboratory score	−1.15 (−1.88, −0.44)	0.40 (0.09, 0.70)	0.17 (0.10, 0.23)	0.36 (0.29, 0.49)
Random effects				
Mouse ID	0.21 (−0.33, 0.59)	0.13 (−0.17, 0.50)	0.12 (−0.27, 0.44)	0.003 (0.00, 0.24)
Residual	0.10 (−0.26, 0.40)	0.20 (−0.36, 0.55)	0.03 (−0.15, 0.22)	0.08 (−0.20, 0.35)

Each of these models had one of the behavioural scores measured in the wild (e.g. activity) as the response variable and the corresponding measured score for that individual's behaviour in the arena (e.g. activity in the laboratory) as the fixed factor. Individual ID was entered as the random factor in each model. For each model, we report point estimates for the fixed (mean) and the random parameters (variance) along with their 95 % confidence intervals

syndrome between (1) activity and boldness, indicating that the most active individuals were also the boldest, and (2) activity and exploration, indicating that the most active individuals took the least amount of time to approach the novel object (Table 4). No evidence of behavioural syndromes was found when any of the other personality traits were correlated using either field or laboratory data (Table 4).

Discussion

We showed that personality traits of individual African striped mice tested under standardized conditions in the laboratory were consistent with measurements of personality traits from the same individuals in their natural habitat. We showed that all personality traits were consistent and repeatable both within and between the laboratory and the field, thereby demonstrating

that personality measures collected under artificial laboratory conditions did reflect natural behavioural tendencies, regardless of the sequence of the testing. Moreover, we found that the presence of two correlated context-specific separate latent variables (one for the field and one for the laboratory) explained all the behaviours measured, indicating that there is a context-specific syndrome in this species.

Several studies have described personality variation in wild animals tested under standardized laboratory conditions (Bell and Sih 2007; Cote and Clobert 2007; Johnson and Sih 2007). Most recently, however, the urgency of establishing whether the behaviour observed in the laboratory are reliably representing the behaviour of individuals under natural conditions have been highlighted (Bell 2012), especially in the light of the fact that the environment in which an individual is tested may end up modifying its behaviour (Hodgins-Davis and Townsend 2009; Niemela and Dingemanse 2014). As a consequence, several

Fig. 4 Parameter estimates of the structural equation model that best fitted our data and thus considered to be representative of the behavioural syndrome structure for *R. pumilio*. Factors loadings together with their 95 % confidence intervals as well as variance estimates explained by the latent variables are reported
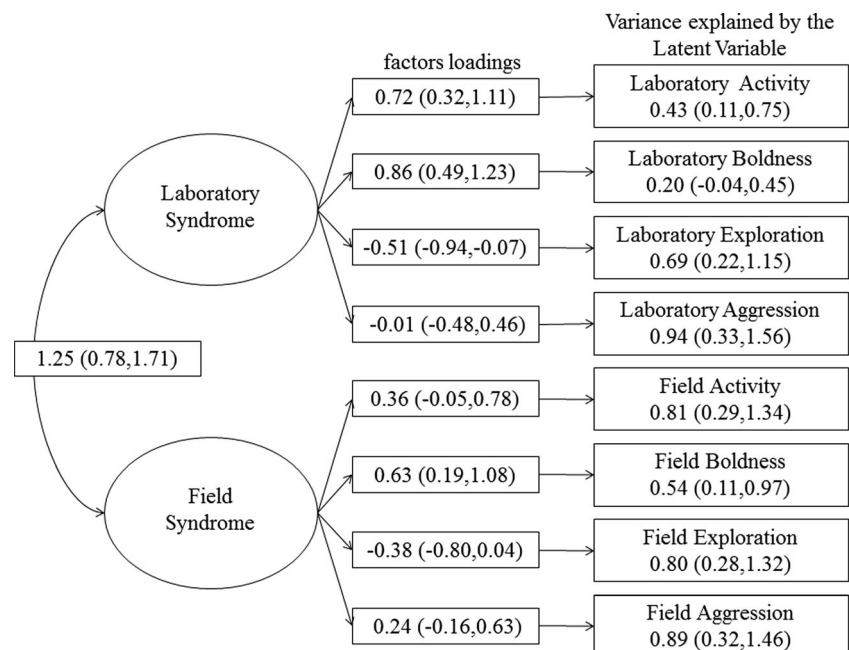
Table 4 Correlations between different personality traits of striped mice tested in a neutral arena in captivity and in the field, indicating behavioural syndromes

Behavioural syndrome	Laboratory			Field		
	n	rho	P	n	rho	P
Activity-boldness	18	0.62	*0.01*	18	0.23	0.36
Activity-exploration	18	−0.78	*0.0001*	18	−0.07	0.76
Activity-aggression	18	0.14	0.58	18	−0.19	0.44
Boldness-exploration	18	−0.65	*0.004*	18	−0.55	*0.02*
Boldness-aggression	18	−0.04	0.87	18	−0.18	0.46
Exploration-aggression	18	0.21	0.41	18	−0.09	0.73

Only individuals that were tested twice in the arena and twice in the field were included in the analysis. Numbers in *italics* indicate significant differences

authors have started to investigate whether wild behavioural types can also spill over to the laboratory, resulting in conflicting reports regarding the degree of consistency between the two contexts. For example, Herborn et al. (2010) and Boon et al. (2008) found consistency in all the behavioural traits measured between the laboratory and the wild, whereas Fisher et al. (2015), Boyer et al. (2010) and van Overveld and Matthysen (2010) only found consistency for some measures but not others. Our results support the former, as we found that in striped mice all the behaviours we measured (activity, boldness, exploration and aggression) were consistent across the field-laboratory context. From a methodological point of view, our results are important because they show that (1) the tests we employed to measure the different behaviour were representative of the target behaviour we measured in both contexts, and (2) classical personality tests, such as open field, novel object and dyadic encounters with a novel conspecific, typically used to measure individuals in a neutral presentation arena can be successfully transposed to the field, at least for striped mice. Our results are also significant as they show that the laboratory environment in which striped mice were tested did not adversely affect them, as was expected in the literature (Hodgins-Davis and Townsend 2009; Niemela and Dingemanse 2014). This might have been the direct result of either our sampling protocol, which restricted the time each individual spent in the laboratory, or could have been due to the fact that striped mice at our field site are habituated to our presence because each individual is trapped and handled several times per month (Yuen et al. 2015). This is further supported by the fact that corticosterone levels, which are typically elevated in individuals experiencing a stressful event, remained similar before and after individuals were tested for personality in the laboratory (CHY, unpublished data).

We found repeatability to be higher in certain behaviours more than others. Activity and boldness were found to be highly repeatable in the laboratory, but less so in the field and the relationship of both behaviours between the field

and the arena was the weakest. The lower repeatability observed for activity and boldness in the field compared to the laboratory suggests that these behaviours might be more easily affected by external stimuli, such as weather conditions or temperature or by the type of environment in which they are assayed (novel versus unfamiliar). Though significant, aggression was found to be the least repeatable trait within both the field and the laboratory, but the relationship between field and laboratory measures was among the strongest. Aggression, in contrast to activity and boldness, was always performed in a neutral setting whether it be in the wild (at the border between territories) or in the laboratory (in the neutral presentation arena), which could have accounted for the high strength in the relationship between aggression measured between the two contexts. However, aggression was also most likely affected by the type of stimulus presented, with individuals always being presented with different stimulus mice whether in the field or in the laboratory. Aggressive encounters may be affected by within contest decision-making and information gathering, and are also highly energetically demanding, resulting in post-contest changes in behaviour (Briffa et al. 2015). Further, as individuals will engage in more than one contest over their life, their behaviour will be affected both by the opponent's identity and behaviour (Briffa et al. 2015) as well as by their own previous experience and familiarity with that opponent. As all of these factors will bear on aggression by varying degrees, different individuals will elicit different aggressive responses, which might explain why repeatability for this behaviour was not as high as for the other behaviours. In this respect, our results are consistent with other studies that have shown aggression to have a low repeatability overall (Briffa et al. 2015). On the other hand, exploration was highly repeatable in both contexts and had the strongest relationship between the field and the laboratory; exploring a novel environment is particularly important for dispersing individuals (Schoepf and Schradin 2012b). Further, of all the behaviours measured, exploration was tested in the most similar way in both the field and the laboratory, further highlighting the need of carefully designing tests that are as similar as possible when doing across context comparisons.

Different personality traits are often correlated with each other, creating behavioural syndromes (Sih et al. 2004), which can be present both in the captive environment and in nature (Dochtermann and Jenkins 2007; Adriaenssens and Johnsson 2013). While some studies have shown that syndromes can be stable over time (Chapman et al. 2013) and across different ecological conditions (Mowles et al. 2012), some authors have found syndromes to differ among conditions, populations or over time (Bell and Stamps 2004; Dingemanse et al. 2007; Clobert et al. 2009). In our study, we found that two separate, but correlated; latent variables affected all the behaviours, pointing to the presence of a context-specific syndrome structure in this species, although the support for this model was

rather weak. Closer inspection of the estimates obtained from the SEM model revealed that the two latent variables loaded most heavily on activity, boldness and exploration in the laboratory context and on boldness in the field context respectively. Further analysis using pair-wise correlations revealed that boldness-exploration behavioural syndrome was consistent in both contexts. Specifically, we found that the boldest individuals, which approached the novel object fastest in the arena, were also the boldest and approached the novel object fastest in the field, indicating that the boldest individuals were the most exploratory in both the laboratory and in the field. Surprisingly, however, we found no consistency between the other behavioural syndromes in the two different contexts. This is intriguing because it would be expected that if all the personality traits measured in isolation are present between the captive and the natural environment, the correlations between such personality traits should also be present. Herborn et al. (2010) suggested that boldness and exploration might be perceived as two measures of a single approach–avoidance trait, with risk-prone, fast-exploring individuals at the one extreme and risk-averse, slow-exploring individuals at the other, thus indicating that the open field test developed by Wilson et al. (1976) and the novel object test developed by Greenberg (1984) to be regarded as approach–avoidance in a novel and a familiar environment, respectively (Clark and Ehlinger 1987; Wilson et al. 1993; Johnson and Sih 2007). Another equally plausible explanation for why we did not find a relationship between activity and boldness in the field is that in the laboratory activity assays were conducted within a very short time of each other. In contrast, activity and exploration in the field were tested separately with a greater time interval between them. This temporal separation might have weakened the correlation between activity and exploration in the field. Similarly, the lack of a boldness-activity correlation in the field might have been the result of a time discrepancy between the measures of boldness, because in the laboratory, boldness was measured during a period of 5 min whereas in the field it was measured during a period of 6 h. Another possibility could be that our results reflect the small sample number of individuals at our disposal. In a previous study, we showed that male and female striped mice differ in their personality traits (Yuen et al. 2015), with females being consistent for activity, boldness and exploration and males being consistent for exploration and aggression, even after adopting a new alternative reproductive tactic. In the present study, 18 individuals were available to test for behavioural syndromes, which included both males (nine) and females (nine). If the two sexes display different behavioural syndromes, the low number of samples might have constrained detecting sex differences. This could also explain why some of the behavioural syndromes did not match between the field and the arena.

Several studies have measured personality of wild-caught individuals in captivity and used these measures to explain individual differences in fitness observed in nature (e.g. Dingemanse et al. 2004). Most recently, however, several authors have started to investigate whether wild behavioural types can also spill over to the laboratory because concerns have been raised regarding the effect of the environment on behaviour. This has resulted in a surge of studies testing for consistency among the two context, with different authors often reporting different levels of consistency between the field and the captive environment (Boon et al. 2008; Boyer et al. 2010; Herborn et al. 2010; van Overveld and Matthysen 2010; Cole and Quinn 2014; Fisher et al. 2015). Ours is the first study that has measured four of the most common behaviours (activity, boldness, exploration and aggression) typically researched in personality by using similar protocols both in the laboratory and in the field. We showed that personality measures from standardized laboratory conditions can reflect field measurements, at least in striped mice. Furthermore, to our knowledge, ours is one of the few studies that have investigated whether behavioural syndromes measured in captivity can be related to behavioural syndromes measured in nature. Our methodological approach validates previous field studies and confirms that personality traits of free-living individuals measured under standardized laboratory conditions reflect the natural variation related to important life-history parameters, such as reproductive fitness and survival.

Compliance with ethical standards

Ethical statement   All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. Fieldwork for this study was carried out under the necessary licences and was in accordance with the relevant animal welfare regulations. Research permits were obtained from the Department of Tourism, Environment and Conservation of the Northern Cape. We received ethical clearance from the Animal Ethics Committee and the University of the Witwatersrand (AESC: 2007/38/04).

Conflict of interest   The authors declare that they have no conflict of interest.

Ethical approval   All applicable international, national, and/or institutional guidelines for the care and use of animals were followed.

Informed consent   This article does not contain any studies with human participants.

References

Adriaenssens B, Johnsson JI (2013) Natural selection, plasticity and the emergence of a behavioural syndrome in the wild. Ecol Lett 17:47–55

Araya-Ajoy YG, Dingemanse NJ (2014) Characterizing behavioural 'characters': an evolutionary framework. Proc R Soc B 281: 20132645

Bates D, Mächler M, Bolker BM, Walker SC (2014) Fitting linear mixed-effects models using lme4. http://CRAN.R-project.org/package=lme4

Bell AM (2005) Behavioural differences between individuals and two populations of stickleback (Gasterosteus aculeatus). J Evol Biol 18:464–473

Bell AM (2007) Future directions in behavioural syndromes research. Proc R Soc Lond B 274:755–761

Bell AM (2012) Personality in the wild. Nature 491:341–342

Bell AM, Sih A (2007) Exposure to predation generates personality in three-spined sticklebacks (Gasterosteus aculeatus). Ecol Lett 10: 828–834

Bell AM, Stamps JA (2004) Development of behavioural differences between individuals and populations of sticklebacks, Gasterosteus aculeatus. Anim Behav 68:1339–1348

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 57:289–300

Benus RF, Rondigs M (1996) Patterns of maternal effort in mouse lines bidirectionally selected for aggression. Anim Behav 51:67–75

Birke LIA, Archer J (1983) Some issues and problems in the study of animal exploration. In: Archer J, Birke LIA (eds) Exploration in animals and humans. Nostrand, New York, pp 1–21

Biro PA (2012) Do rapid assays predict repeatability in labile (behavioural) traits? Anim Behav 83:1295–1300

Boon AK, Réale D, Boutin S (2008) Personality, habitat use, and their consequences for survival in North American red squirrels Tamiasciurus hudsonicus. Oikos 117:1321–1328

Boyer N, Réale D, Marmet J, Pisanu B, Chapuis JL (2010) Personality, space use and tick load in an introduced population of Siberian chipmunks Tamias sibiricus. J Anim Ecol 79:538–547

Briffa M, Rundle SD, Fryer A (2008) Comparing the strength of behavioural plasticity and consistency across situations: animal personalities in the hermit crab. Pagurus bernhardus. Proc R Soc B 275: 1305–1311

Briffa M, Sneddon LU, Wilson AJ (2015) Animal personality as a cause and consequence of contest behaviour. Biol Lett 11:20141007

Brown C, Jones F, Braithwaite V (2005) In situ examination of boldness-shyness traits in the tropical poeciliid, Brachyraphis episcopi. Anim Behav 70:1003–1009

Campbell DLM, Weiner SA, Starks PT, Hauber ME (2009) Context and control: behavioural ecology experiments in the laboratory. Ann Zool Fenn 46:112–123

Carter AJ, Marshall HH, Heinsohn R, Cowlishaw G (2012) How not to measure boldness: novel object and antipredator responses are not the same in wild baboons. Anim Behav 84:603–609

Chapman BB, Hegg A, Ljungberg P (2013) Sex and the syndrome: individual and population consistency in behaviour in rock pool prawn Palaemon elegans. PLoS ONE 8:e59437

Clark A, Ehlinger TJ (1987) Pattern and adaptation in individual behavioral differences. Persp Ethol 7:1–47

Clobert J, Le Galliard JF, Cote J, Meylan S, Massot M (2009) Informed dispersal, heterogeneity in animal dispersal syndromes and the dynamics of spatially structured populations. Ecol Lett 12:197–209

Cole EF, Quinn JL (2014) Shy birds play it safe: personality in captivity predicts risk responsiveness during reproduction in the wild. Biol Lett 10:20140178

Coleman K, Wilson DS (1998) Shyness and boldness in pumpkinseed sunfish: individual differences are context-specific. Anim Behav 56: 927–936

Cote J, Clobert J (2007) Social personalities influence natal dispersal in a lizard. Proc R Soc Lond B 274:383–390

Cote J, Fogarty S, Weinersmith K, Brodin T, Sih A (2010) Personality traits and dispersal tendency in the invasive mosquitofish (Gambusia affinis). Proc R Soc Lond B 277:1571–1579

Crainiceanu C, Ruppert D (2004) Likelihood ratio tests in linear mixed models with one variance component. J R Stat Soc Ser B 66:165–185

Crawley MJ (2007) The R Book. Wiley, Chichester

Dingemanse NJ, de Goede P (2004) The relation between dominance and exploratory behavior is context-dependent in wild great tits. Behav Ecol 15:1023–1030

Dingemanse NJ, Both C, Drent PJ, Tinbergen JM (2004) Fitness consequences of avian personalities in a fluctuating environment. Proc R Soc Lond B 271:847–852

Dingemanse NJ, Wright J, Kazem AJN, Thomas DK, Hicking R, Dawnay N (2007) Behavioural syndromes differ predictably between 12 populations of three-spined stickleback. J Anim Ecol 76: 1128–1138

Dochtermann NA, Jenkins SH (2007) Behavioural syndromes in Merriam's kangaroo rats (Dipodomys merriami): a test of competing hypotheses. Proc R Soc Lond B 274:2343–2349

Dudycha GJ (1936) An objective study of punctuality in relation to personality and achievement. Arch Psychol 204:1–53

Fisher DN, James A, Rodrıguez-Munoz R, Tregenza T (2015) Behaviour in captivity predicts some aspects of natural behaviour, but not others, in a wild cricket population. Proc R Soc B 282:20150708

Greenberg R (1984) Neophobia in the foraging-site selection of a neotropical migrant bird: an experimental study. Proc Natl Acad Sci U S A 81:3778–3780

Hartshorne H, May MA (1928) Studies in the nature of character, vol 1. Studies in deceit. Macmillan, New York

Herborn KA, Macleod R, Miles WTS, Schofield ANB, Alexander L, Arnold KE (2010) Personality in captivity reflects personality in the wild. Anim Behav 79:835–843

Hodgins-Davis A, Townsend JP (2009) Evolving gene expression: from G to E to G x E. Trends Ecol Evol 24:649–658

Hollander FA, Overveld TV, Tokka I, Matthysen E (2008) Personality and nest defence in the great tit (Parus major). Ethology 114:405–412

Johnson JC, Sih A (2007) Fear, food, sex and parental care: a syndrome of boldness in the fishing spider, Dolomedes triton. Anim Behav 74: 1131–1138

Jones KA, Krebs JR, Whittingham MJ (2009) Heavier birds react faster to predators: individual differences in the detection of stalking and ambush predators. Behav Ecol Sociobiol 63:1319–1329

Kurvers RHJM, Prins HHT, Wieren SE, van Oers K, van Nolet BA, Ydenberg RC (2010) The effect of personality on social foraging: shy barnacle geese scrounge more. Proc R Soc Lond B 277:601–608

Martin P, Bateson P (1993) Measuring behaviour: an introductory guide, 2nd edn. Cambridge University Press, Cambridge

Mowles SL, Cotton PA, Briffa M (2012) Consistent crustaceans: the identification of stable behavioural syndromes in hermit crabs. Behav Ecol Sociobiol 66:1087–1094

Nakagawa S, Schielzeth H (2010) Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. Biol Rev 85:935–956

Nakagawa S, Schielzeth H (2013) A general and simple method for obtaining R^2 from generalized linear mixed-effects models. Methods Ecol Evol 4:133–142

Niemela P, Dingemanse N (2014) Artificial environments and the study of 'adaptive' personalities. Trends Ecol Evol 29:245–247

Quinn JL, Cole EF, Bates J, Payne RW, Cresswell W (2012) Personality predicts individual responsiveness to the risks of starvation and predation. Proc R Soc Lond B 279:1919–1926

Réale D, Reader SM, Sol D, McDougall PT, Dingemanse NJ (2007) Integrating animal temperament within ecology and evolution. Biol Rev 82:291–318

Rosseel Y (2012) lavaan: an R package for structural equation modeling. J Stat Softw 48:1–36

Scheipl F (2010) Rlrsim: exact (restricted) likelihood ratio tests for mixed and additive models. http://CRAN.R-project.org/package=Rlrsim

Schoepf I, Schradin C (2012a) Better off alone! Reproductive competition and ecological constraints determine sociality in the African striped mouse (Rhabdomys pumilio). J Anim Ecol 81:649–656

Schoepf I, Schradin C (2012b) Differences in social behaviour between group-living and solitary African striped mice, Rhabdomys pumilio. Anim Behav 84:1159–1167

Schradin C (2004) Territorial defense in a group living solitary forager: who, where, against whom? Behav Ecol Sociobiol 55:439–446

Schradin C (2005) When to live alone and when to live in groups: ecological determinants of sociality in the African striped mouse (Rhabdomys pumilio, Sparrman, 1784). Belg J Zool 135:77–82

Schradin C (2006) Whole-day follows of striped mice (Rhabdomys pumilio), a diurnal murid rodent. J Ethol 24:37–43

Schradin C, Pillay N (2004) The striped mouse (Rhabdomys pumilio) from the Succulent Karoo, South Africa: a territorial group-living solitary forager with communal breeding and helpers at the nest. J Comp Psychol 118:37–47

Schradin C, Pillay N (2005) Intraspecific variation in spatial and social organization of the African striped mouse. J Mammal 86:99–107

Schradin C, König B, Pillay N (2010) Reproductive competition favours solitary living while ecological constraints impose group-living in African striped mice. J Anim Ecol 79:515–522

Schradin C, Lindholm AK, Johannesen J, Schoepf I, Yuen CH, König B, Pillay N (2012) Social flexibility and social evolution in mammals: a

case study of the African striped mouse (Rhabdomys pumilio). Mol Ecol 21:541–553

Sih A, Bell AM, Johnson JC (2004) Behavioral syndromes: an ecological and evolutionary overview. Trends Ecol Evol 19:372–378

van Oers K, Drent PJ, de Goede P, van Noordwijk AJ (2004) Realized heritability and repeatability of risk-taking behaviour in relation to avian personalities. Proc R Soc Lond B 271:65–73

van Overveld T, Matthysen E (2010) Personality predicts spatial responses to food manipulations in free-ranging great tits (Parus major). Biol Lett 6:187–190

van Overveld T, Matthysen E (2013) Personality and information gathering in free-ranging great tits. PLoS ONE 8:e54199

Verbeek MEM, Drent PJ, Wiepkema PR (1994) Consistent individual differences in early exploratory behaviour of male great tits. Anim Behav 48:1113–1121

Verbeek MEM, Boon A, Drent PJ (1996) Exploration, aggressive behaviour and dominance in pair-wise confrontations of juvenile male great tits. Behaviour 133:945–963

Verbeek MEM, de Goede P, Drent PJ, Wiepkema PR (1999) Individual behavioural characteristics and dominance in aviary groups of great tits. Behaviour 136:23–48

Wilson DS, McLaughlin RL (2007) Behavioural syndromes in brook charr, Salvelinus fontinalis: prey-search in the field corresponds with space use in novel laboratory situations. Anim Behav 74:689–698

Wilson RC, Vacek T, Lanier DL, Dewsbury DA (1976) Open-field behaviour in muroid rodents. Behav Biol 17:495–506

Wilson DS, Coleman K, Clark AB, Biederman L (1993) Shy-bold continuum in pumpkinseed sunfish (Lepomis gibbosus): an ecological study of a psychological trait. J Comp Psychol 107:250–260

Yuen CH, Pillay N, Heinrichs M, Schoepf I, Schradin C (2015) Personality does not constrain social and behavioural flexibility in African striped mice. Behav Ecol Sociobiol 69:1237–1249

Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM (2009) Mixed effects models and extensions in ecology with R. In: Gail M, Krickeberg K, Samet JM, Tsiatis A, Wong W (eds) Springer, New York, pp 116–120