2016

Evaluation of boldness assays and associated behavioral measures in a social parrot, monk parakeet (*Myiopsitta monachus*)

Kaan Kerman
Virginia Polytechnic Institute and State University, kkerman@vt.edu

Kathryn E. Sieving
University of Florida

Colette S. Mary
University of Florida

Michael L. Avery
USDA Wildlife Services, National Wildlife Research Center, michael.l.avery@aphis.usda.gov

Evaluation of boldness assays and associated behavioral measures in a social parrot, monk parakeet (*Myiopsitta monachus*)

Kaan Kerman [a,*], **Kathryn E. Sieving** [b], **Colette S. Mary** [c] **and Michael L. Avery** [d]

[a] Department of Biological Sciences, Derring Hall, 1405 Perry Street, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA
[b] Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, FL 32611, USA
[c] Department of Biology, University of Florida, Gainesville, FL 32611, USA
[d] USDA Wildlife Services, National Wildlife Research Center, Florida Field Station, Gainesville, FL 32641, USA
[*]Corresponding author's e-mail address: kkerman@vt.edu

Abstract

Boldness reflects consistent individual differences in risk-taking behavior across various contexts. However, evaluating this basic assumption has largely been neglected in birds. In a captive monk parakeet population (*Myiopsitta monachus*; $N = 33$), we undertook an analysis of 7 measures across 3 commonly used boldness assays (i.e., novel object, emergence, and predator-exposure tests). Using principal component analysis, we derived 3 components (PCs). PC-2 loaded strongly with measures from emergence and predator-exposure tests; we interpreted it as the closest approximation of boldness. PC-1 and PC-3 described different aspects of feeding such as foraging activity and rate, respectively. Finally, we assessed the predictive power of each measure that loaded significantly on the boldness axis. We found that no single metric explained even %55 of the variation in PC-2, nor could more than %50 individuals at the extremes of the spectrum be predicted. Our results demonstrate the utility of an inclusive approach in personality research.

Keywords

boldness, behavioral assays, monk parakeet.

1. Introduction

Animal personality describes behavioural phenotypes that are predictable over time and across different environmental contexts (Dingemanse et al., 2010). In many species, individuals exhibit consistent intra-specific differences in their reaction to potentially risky situations (Boissy, 1995; Greenberg & Mettke-Hofmann, 2001; Carere & Maestripieri, 2013). Such variation in risk-taking behaviour is sometimes described as an individuals' relative 'boldness'; a well-studied animal personality trait (Sloan Wilson et al., 1994; Arnold et al., 2007; Jones & Godin, 2010; Dammhahn & Almeling, 2012). Although the term boldness has sometimes been used synonymously with behavioural phenomena such as 'exploratory activity' (Verbeek et al., 1994; Minderman et al., 2009) and 'neophobia' (Boogert et al., 2006), here we follow the conceptual definition discussed by Carter et al. (2012a): boldness captures the willingness to utilize resources such as food or habitat in the presence of threatening stimuli (Bell, 2005; Campo et al., 2015; Finkler & Terkel, 2015).

A diverse array of experimental set-ups for manipulating risk-taking behaviour can be found in the literature (Reale et al., 2007), which are generally grouped under three broad categories (Vazire et al., 2007; Carter et al., 2012a). (1) Novel object tests simulate an encounter with an unfamiliar item, generally accompanying a food resource in the test arena (Wilson et al., 2010; Garamszegi et al., 2012). (2) Emergence tests focus on the propensity of individuals to leave a safe location and enter an unknown, potentially dangerous environment (Brown & Braithwaite, 2005; Miller et al., 2006). (3) Simulated predator exposure tests consist of either a real or artificial predator presented in a controlled environment to elicit predator inspection or alarm calls under predation risk (Wilson & Godin, 2009). Generally, individuals are subjected to controlled testing environments in each assay with the assumption that observed responses reflect a uniform boldness phenotype, which is expressed consistently across risk-related settings (Jolles et al., 2013; Ingley et al., 2014). However, animal risk-taking behaviour can often be expressed in a context-specific manner. Implications of this for describing bold–shy phenotype are that behavioural expression may not be consistent across assays and, therefore, a general boldness phenotype may be lacking or extend to only some settings (Coleman & Wilson, 1998; Beckmann & Biro, 2013). In addition to context-specific boldness expression, other distinct per-

sonality traits may be expressed in assays traditionally used for detecting boldness (Carter et al., 2012d; Andersson et al., 2014).

The choice of the best assay for the behavioural trait of interest is further complicated by the fact that there are multiple measures associated with each experimental setting. As boldness is one of the most extensively studied personality traits in animals, research on risk-taking behaviour has engendered a small set of distinct metrics that are assumed to represent functionally the same underlying behavioural trait (Sloan Wilson et al., 1994; Reale et al., 2007; Carter et al., 2012a, c). Moreover, selecting a single measure from a given behavioural assay appears to be a common methodological approach, particularly if there is baseline knowledge concerning the personality axes in the study organism (Highcock & Carter, 2014; Moscicki & Hurd, 2015). In other cases, investigators select several measures implemented in closely related species and adapt these metrics to their biological model (e.g., Seaman & Briffa, 2015). However, given the increasingly frequent detection of context-specificity in boldness assays (Noer et al., 2015), single measures will likely not be equally reliable within or across animal species, and their uncritical adoption based on precedence alone may limit precision in depicting bold–shy variation (Carter et al., 2012a; Beckmann & Biro, 2013). A clear lesson from recent studies is that the use of a single 'standard' measure to characterize boldness types within a population precludes the determination of its functional accuracy (Carter et al., 2012c). This situation merits a continuous effort to describe boldness and other latent behavioural traits in a robust and standardized manner.

Here, we compare the reliability of the simplest methods in common use today with a more comprehensive approach for better understanding and diagnosis of boldness in a species where the presence of a boldness phenotype has not previously been assessed (Carter et al., 2012a; Dall & Griffith, 2014). Specific objectives of this study were to (1) evaluate the congruence of three standardized behavioural assays in describing boldness, where each of the assays includes multiple behavioural measures, and (2) assess the predictability of each of seven behavioural measures in ranking individuals according to their boldness levels. We selected a captive population of the monk parakeet as our model organism (MOPA; *Myiopsitta monachus*). According to our knowledge, no empirical work on the personality traits of MOPA has been published.

The study population was subjected to three assays: novel object, emergence, and predation-exposure risk. Using a principal component analysis (PCA), we examined the consistency of behaviour across the assays. We predicted that, if risk-taking behaviour is represented equally well by all three standardized behavioural tests, we should observe a single principal component (PC) defined by strong correlations among all seven measures. In other words, this would be interpreted as evidence of an overall boldness personality in MOPA that is well captured by all the measures. However, if several principal components emerge, separating metrics or especially assays into different clusters, it can be interpreted as a challenge to the uniformity of these three behavioural assays in characterizing boldness for this species, and that certain assays could instead be linked to other personality traits. We would conclude that MOPA exhibit a boldness personality trait if we identified a single axis of risk taking behaviour that was loaded by measures from at least two assays, reflecting that risk-taking behaviour was consistent across at least two contexts. Finally, we sought to characterize the variation in the value of single behavioural measures in estimating the boldness values identified in the full principal component analysis. Hence, we compared the predictability of each behavioural measure with (a) boldness scores across all individuals and (b) scores that represented the boldest and shyest (extreme) ends of the axis identified by the PCA (Reale et al., 2007; Scheid & Noë, 2010; Kurvers et al., 2012b). Measures that performed well can then be used confidently as accurate metrics by themselves to capture boldness in MOPA.

2. Material and methods

2.1. Study organism

A member of the parrot family (*Psittacidae*), MOPA is native to southeastern South America, and has established successful populations outside of its native range (Bucher et al., 1990; Collar, 1997; South & Pruett-Jones, 2000; Russello et al., 2008; Avery et al., 2012). Its primary habitat is dry vegetation with subtropical woodland cover, but MOPA readily utilizes anthropogenic landscapes such as agricultural fields, ranches, and orchards (Burger & Gochfeld, 2005).

2.2. Trapping and maintenance of monk parakeet

Birds in this study ($N = 33$; 21 females and 12 males) were trapped in Miami-Dade and Broward Counties in South Florida in 2008. Trappers

placed large nets over entrances to communal nest structures where all individuals were roosting at night (Tillman et al., 2004). This method avoided biases that could arise from active baiting, where bold individuals are more likely than shy ones to explore and eventually enter the trap (Biro & Dingemanse, 2009; Wilson et al., 2011; Carter et al., 2012b). Captured individuals were then transported to the USDA Wildlife Research Center in Gainesville, FL, USA. Upon arrival at the facility, birds were sexed, banded, weighed, and housed in communal cages ($1.8 \times 1.2 \times 1.2$ m) within a roofed outdoor aviary. Each cage had a nesting platform (30×30 cm), and short sticks and branches were supplied ad libitum as nesting material. Birds were maintained on a standard diet of fresh fruits, mixed seeds and lettuce. At the time of the study, all individuals were in the captive environment for 5 years; and were healthy, with no apparent signs of disease.

2.3. Experimental setting and cage design

We applied three behavioural assays from the animal personality literature, with a specific focus on experimental settings developed for testing birds (see Table 1 for references). A total of 7 boldness measures were obtained from the assays. Biological relevance of these metrics were based on their commonality in avian personality research. Behavioural assays were conducted from May to August 2013; starting with the novel object test. It was followed by emergence and simulated predation risk tests, which were implemented consecutively in the same experimental set-up. Novel object tests were conducted 2 weeks prior to the emergence-simulated predation risk tests to minimize any carry-over effects of movement and human disturbance (Bell, 2012). Behavioural output was recorded by using two Panasonic HC-V100M cameras. Measures were extracted and quantified by KK.

Novel object assays were performed in a roofed outdoor aviary with a cement floor, supported by wooden pole frames. The aviary contained 10 test cages ($1.8 \times 1.2 \times 1.2$ m) to house birds individually. In each cage, we provided 4 branches with approximately equal length and thickness, and two brown plant saucers were placed to supply food and water. Visual contact between neighbouring cages was prevented by hanging opaque sheets between them. Although we could not obstruct vocal interactions between the neighbouring cages, we did not observe any noticeable changes in call patterns when the food cup with the novel object was presented. Since the space was limited to hold only 10 cages simultaneously, we randomly assigned individuals into 4 test blocks. Prior to each behavioural assay, we

Table 1.

List of behavioural measures used, the experimental setting and the method of transformation.

Behavioural measure	Experimental setting	Transformation
Latency to enter the novel environment and perch on a branch after the opening of the release cage door. Quantified as total number of seconds.	Emergence test	Logarithmic
Latency to consume once the food bowl, with a novel object situated in the middle, is placed in the test cage. Quantified as total number of seconds until the bird pecks at seeds for the first time.	Novel object test	Logarithmic
Proportion of time spent feeding during the test period. Feeding activity involved handling and consuming while in close proximity to the food bowl. Quantified as total number of seconds spent feeding, divided by the total test period in seconds.	Novel object test	Arcsine square root
Pecking frequency was quantified as the total number of pecks at the food source over the time spent foraging during the trials. Individuals bobbed their heads in a distinct pattern when pecking consecutively, facilitating the quantification process.	Novel object test	Logarithmic
Total number of feeding approaches over the entire test period. Counting of feeding approaches started when the bird made its first successful feeding approach. Each feeding approach ended when birds flied to a branch or moving approximately 30 cm away from the food bowl.	Novel object test	Square root
Total number of vocalizations while exposed to a predator during the test. Quantified number of distress calls when the predator is in the visual range.	Predator exposure test	Square root
Number of flights when exposed to a predator; quantifying each distinct flights and hops started when the bird left its perch and ended when bird landed on another perch.	Predator exposure test	Square root

Behavioural assays and relevant metrics were adapted from the following studies: Greenberg (1989); Martella & Bucher (1990); Webster & Lefebvre (2001); Martin & Fitzgerald (2005); Stowe et al. (2006a); Hollander et al. (2008); Campler et al. (2009); Nilsson et al. (2010); David et al. (2011); Feenders et al. (2011); Rockwell et al. (2012); Schuett et al. (2012).

selected one of the test blocks, and allocated individual birds in this block to their test cages in a randomized fashion. An acclimatization period lasting 4 days followed the transfer of test birds. We routinely placed food bowls inside the cages at 0830, and removed them at 1630, and provided the same maintenance diet as the birds received in their home cages to control for any changes in the nutritional state of individuals. On test day, we put a novel object, a colourful plush toy 5 cm in diameter inside each food bowl with the food; and recorded the relevant behavioural reactions during a 90-min sampling period (Table 1). We ran a generalized linear mixed effect model with Poisson distribution to detect any significant decrease in latency from day 1 to day 5. Latency to start feeding exhibited a significant downward trend from day 1 to day 4 (mean ± SE: Day 1, 85.1 ± 18.3; Day 2, 65.2 ± 20.4; Day 3, 56.2 ± 10.3; Day 4, 72.1 ± 18.9). However, when compared to the first four days, the latency increased considerably on the test day when familiar food bowls with novel objects were introduced into the cages (Day 5: 738.8 ± 244.3). We concluded that individuals recognized plush toys as novel objects, as depicted in other taxa (Stowe et al., 2006b). After completing the behavioural assay, birds from the same block were transferred back to their home cages, and the procedure was repeated until all the test blocks were completed. Boxplots for each behavioural measure (before transformation) are shown in Figure A1 in the Appendix to this article.

For the remaining behavioural assays, we stationed test birds in a cage (1.2 × 1.8 × 2.4 m) within a small aviary. Eight branches of equal length and width were provided inside the test cage, increasing the appeal of the interior as a novel environment. One side of the cage consisted of a wooden door and a small opening with Plexiglas cover (30.5 × 30.5 cm). A small release cage, covered with opaque material, was mounted in the opening. For each trial, we randomly selected a test bird and moved it to the test arena in carriers with no visual contact with the outside environment. We then placed the bird inside the release cage for a 5-min habituation period. Tests began by remotely releasing the door with a small rope to allow the bird to enter the novel environment. We recorded the latency to leave the release cage as the metric for the emergence test. Once the bird was inside the novel environment cage, we implemented a 20-min habituation period before initiating the simulated predation risk tests by revealing an owl model that was hidden behind a cover. We measured the behavioural reaction to the

dummy predator for 1 min by recording two metrics (Table 1). Emergence-simulated predation risk assays were conducted between 0830 and 1130, testing 2–3 individuals per day for 3 consecutive weeks.

2.4. Statistical analysis

All statistical analyses were done in R (R Core Team, 2014). For PCA, we used the packages 'FactoMiner' and 'Psych' (Lê et al., 2008; Revelle, 2013).

2.4.1. Identifying axes

We implemented PCA in order to reduce behavioural measures into fewer sets of orthogonal components, each potentially representing a personality trait (Budaev, 2010). Since PCA is sensitive to non-normality and relative scaling of variables, all measures were subjected to transformation and scaling (Maindonald & Braun, 2010). This method resulted in normal distribution for each variable. After the analysis, we excluded principal components (PCs) with Eigenvalues < 1 (Kaiser, 1960). We rotated the axes of the remaining PCs (Varimax method) to obtain a better visualization of variable loadings (Kaiser, 1959; Cote et al., 2010). Following the rotation, variables with loadings < 0.4 were treated as having no substantial influence on relevant PCs; variables with larger loadings were used to interpret the behavioural axis each PC represented (Stevens, 1992; van den Brink et al., 2012).

2.4.2. Behavioural metrics as predictors of boldness axes

We evaluated the reliability of each single measure to accurately describe the relative boldness rankings of individuals by calculating the proportion of variance (R^2) explained by each metric. This allowed us to obtain an estimate of how single behavioural measures differ in their boldness ranking utility.

Lastly, we were interested to see whether the same predictive power would apply if we restricted our observation to the extreme ends of the distribution (i.e., bolder and shyer individuals). We started our analysis by rank ordering individuals from bold to shy according to their scores obtained from each behavioural measure as well as PC-2 (our axis describing boldness). Once relative rankings of individuals were ordered that way, we selected 5 individuals from the boldest and shyest ends of the spectrum for each of the 8 behavioural measures and PC-2. We calculated percentages of individuals that were found in the same group between each measure and the PC. Higher

percentage scores indicated that a particular behavioural measure was suc-
cessful in matching the relative boldness rankings from the representative PC
axis.

2.4.3. Gender effects on boldness

Expression of risk-taking behaviour can be influenced by sex-specific factors
(Schuett & Dall, 2009; Small & Schoech, 2015). We did not expect this,
given the monomorphic nature of MOPA. However, since the study included
individuals from both sexes, we investigated sex differences in the expression
of both individual behavioural measures and the PC scores using the Welsh
two-sample t-test.

3. Results

3.1. Context specificity of boldness

We observed no significant effect of sex on individual behavioural measures
or any of the composite PC scores (the lowest p value = 0.08, $t = 1.82$).
Therefore, we continued using the pooled data from both sexes (Table 2).

The PCA produced 3 PCs with Eigenvalues > 1. All 3 PCs explained
69% of the cumulative variation (Table 3). In PC-1, individuals who were
quicker in approaching and consuming from the food bowl also made more
frequent visits, and pecked at the food item more frequently throughout the
behavioural trial. In other words, this axis described individual differences
in willingness to consume food from the food cup. Therefore, we called this

Table 2.
Summary statistics of behavioural measures.

Behavioural measure	Mean ± SE	Range (min–max)
Latency to enter the novel environment (seconds)	866.7 ± 211.7	5079 (5–5084)
Latency to consume food when novel object is present (seconds)	738.8 ± 244.3	5664 (21–5685)
Proportion of time spent feeding	0.13 ± 0.03	0.8 (0.02–0.82)
Pecking frequency when novel object is present	0.57 ± 0.17	4.03 (0.02–4.05)
Total number of feeding approaches	5.85 ± 0.96	32 (1–33)
Total number of vocalizations when exposed to a predator	7.4 ± 1.2	24 (1–24)
Total number of flights when exposed to a predator	4.4 ± 0.7	16 (1–16)

Table 3.

Component loadings with Varimax rotation from the principal component analysis (PCA) on 7 behavioural measures from 3 boldness assays.

Behavioural measure	PC-1 (Foraging activity)	PC-2 (Risk avoidance)	PC-3 (Foraging intensity)
Latency to enter the novel environment	−0.13	0.58*	0.27
Latency to consume food when novel object is present	−0.85*	0.04	−0.06
Proportion of time spent feeding	0.04	0.13	0.93*
Pecking frequency when novel object is present	0.81*	−0.01	0.47*
Total number of feeding approaches	0.70*	0.12	−0.30
Total number of vocalizations when exposed to a predator	0.16	0.74*	0.08
Total number of flights when exposed to a predator	0.01	0.88*	0.13
Eigenvalues	1.90	1.68	1.27
Percentage of variance explained	27%	24%	18%
Cumulative percentage of variance explained	27%	51%	69%

* Behavioural measures with significant loadings ($\geqslant 0.4$).

axis foraging activity. In PC-2, birds with greater latency to enter an unfamiliar environment escaped from the predator sooner and gave more distress calls when the predator was visible. It appeared to capture an underlying risk-avoidance behaviour, correlated across contexts. In PC-3, birds that spent a higher proportion of their time foraging, also pecked more frequently at the food source. We surmised that PC-3 may have demonstrated individual variation in foraging rate during a feeding bout.

We found that all novel object measures were clustered in two components (i.e., PC-1 and PC-3); while all measures from the predator exposure and emergence tests were merged in a single component (i.e., PC-2). We interpret PC-2 as the axis that best represents boldness in MOPA because it reflects consistent variation in risk-avoidance across two of our three contexts.

3.2. Predictability of behavioural measures

Three of the behavioural measures, when used alone, demonstrated a meaningful predictive ability in ranking individuals according to their boldness

scores, as characterized by PC-2 (Figure 1). However, the best predictor explained only about 50% of variation in PC-2 (Table 4). Total number of flights when exposed to a predator, which had the highest loading on the boldness axis, demonstrated the best fit with the boldness axis ($R^2 = 0.53$; $p < 0.0001$). It was followed by the remaining measures that also had strong loadings on PC-2; total number of vocalizations when exposed to a predator ($R^2 = 0.40$; $p < 0.0001$), and latency to enter the novel environment ($R^2 = 0.31$; $p < 0.001$).

Focusing on the extreme ends of the spectrum revealed a similar pattern (Figure 1). None of the measures succeeded in predicting more than 60% of the individuals at either end of the distribution (Table 4). Indeed, taking both extremes together, no single measure predicted more than 50% of the extreme individuals. Behavioural measures that loaded significantly on PC-2 had greater accuracy in reflecting the rankings of individuals, but did only slightly better than the best of variables that did not load on PC-2.

4. Discussion

4.1. Comparison of behavioural assays

Our results showed that not all assays traditionally used for characterizing boldness in birds captured the same underlying behavioural pattern. Although risk-sensitivity has been considered a major constraint on foraging behaviour for most species (Real & Caraco, 1986; Kie, 1999; Dammhahn & Almeling, 2012), and latency to approach the novel object has been one of the most common measures used to reflect boldness (Short & Petren, 2008; Dammhahn & Almeling, 2012; Kurvers et al., 2012a), we did not observe any association between the behavioural output in the novel object test and the remaining boldness tests. In novel object assays, the unusual item is placed in a familiar feeding tray, and this is hypothesized to invoke an elevated perception of risk (Greenberg, 1984). However, this context could also reflect underlying variation in resting metabolic rate (RMR). Resting metabolic rates (RMR; not measured here) has been linked with activity levels in several species (Biro & Stamps, 2010). Since the feeding schedule was the same for all birds in our test conditions (see Methods), PC-1 and PC-3, which characterize differences in foraging activity and feeding rates, may instead have captured underlying variation in RMR. Using a measure of RMR
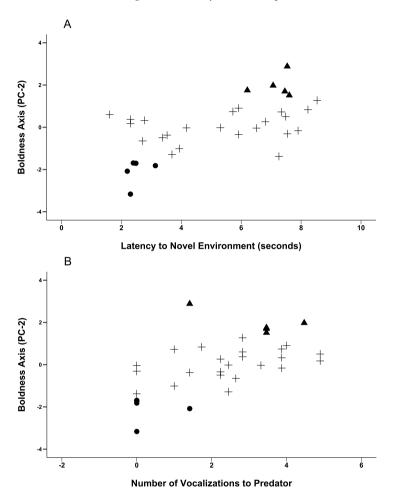
Figure 1. Linear relationship between individual scores from the boldness axis and each of the 3 behavioral measures (transformed values) that loaded significantly on that axis: (A) latency to enter novel environment, (B) total number of vocalizations when exposed to a predator, and (C) total number of flights when exposed to a predator. Individuals that are boldest, shyest, and in-between are represented with different shapes to show how they are clustered (▲, 5 boldest individuals according to PC-2; ●, 5 shyest individuals according to PC-2; +, individuals that are in-between according to PC-2).

as a covariate in behavioural typing may be something to consider in future studies of MOPA (Killen et al., 2011), and perhaps other species.

The observed incongruity between novel object and the remaining assays could be due to the unnatural conditions in the test setting. In personality
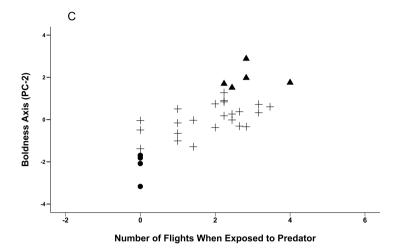
C

Figure 1. (Continued.)

research, animal subjects are usually tested in isolation to obtain individual boldness scores, therefore eliminating any potentially confounding effects of their social environment (Webster & Ward, 2011). As a species that lives in

Table 4.
Results of the assessment of the predictability of 7 behavioural measures.

Behavioural measure	Coefficient of determination (R^2)	Matching success (%)		
		Boldest	Shyest	Both
Latency to enter the novel environment	0.31***	40	20	30
Latency to consume food when novel object is present	0.11	20	40	30
Proportion of time spent feeding	0.003	0	20	10
Pecking frequency when novel object is present	0.10	20	0	10
Total number of feeding approaches	0.03	0	60	30
Total number of vocalizations when exposed to a predator	0.40****	60	20	40
Total number of flights when exposed to a predator	0.53****	60	40	50

Coefficient of determination (R^2) and p-values for comparisons between each behavioural measure and PC-2. Matching success represent the percent overlap in membership of each behavioural measure against the PC scores for boldest, shyest, and both ends of the distribution. Significant values are indicated with asterisks (***$0.0001 < p < 0.001$; ****$p < 0.0001$).

complex social groups that are centered on their communal nests (Hobson et al., 2014), absence of any flock members during the novel object assay may very well have influenced our results (Webster & Ward, 2011). Similar to the presence or absence of behavioural consistency across functional contexts (i.e., assays), correlation among behavioural traits might vary depending on the social environment within which the animals are tested (Dzieweczynski & Crovo, 2011; Carter et al., 2012d). We therefore assume that replicating the study design in a social setting, and measuring the repeatability over a temporal scale should clarify whether novel object assay accurately represent another type of behavioural strategy by MOPA to manage risky situations.

4.2. Using single measures in personality research

Emergence and predator exposure tests yielded measures that captured risk-avoidance behaviour in MOPA, whereas measures associated with novel object assay did not exhibit any association with the rest of the measures. However, despite high loadings of three measures from these assays on a definitive risk-avoidance axis (PC-2), no single variable explained even 60% of the variation in that axis nor were they successful enough in identifying individuals at the extremes of the boldness axis. For example, the highest success in predicting the individuals at both extremes of the distribution was 50%; and the variation explained by single measures were no higher than approx. 0.5. Since personality research relies on relative ranks of individuals within a group, such discrepancies in ranking bold and shy individuals between single measures and composite scores could lead to erroneous conclusions. Therefore, grouping many measures by implementing PCA is potentially the most appropriate approach.

Our conclusion has been further supported by several studies that adopted a similar comparative approach. For instance, Carter et al. (2012d) observed in baboons (*Papio ursinus*) latency to handle the food item was not correlated with other novel object measures such as time spent inspecting and handling the food item. Moreover, these measures did not show any association with anti-predator response to a snake model. Also, Beckmann & Biro (2013) found no correlation between the latency to emerge into the novel environment and reaction to a predation threat in damselfish (*Pomacentrus* spp.). Echoing other researchers in the field of animal personality, we believe that studies should facilitate a thorough assessment of all major assumptions inherent in personality typing until we establish appropriate guidelines for the design and suitability of testing environments for a wide range of species.

4.3. Conclusions

Boldness is assumed to be a universal behavioural trait with important ecological significance, and investigators generally face a key trade-off between comprehensiveness and efficiency in behavioural assays to achieve a useful ranking of individuals on this trait axis. Relying heavily on previous work to use a 'tried and true' single measure and single assay to effectively achieve the goal is a common solution. Our results showed that focusing on one measure from a single assay may not actually represent the trait of interest, leading to degradation of understanding rather than enhanced insights, and highlight the potential limitations on the general applicability of standard assays in diagnosing boldness types within populations that have yet to be assessed (Coleman & Wilson, 1998; Wilson & Stevens, 2005; Fox et al., 2009; Carter et al., 2012d).

Furthermore, initial characterization of boldness in a previously unstudied species certainly requires a more robust and confident assessment. As we encountered, for species never before assayed for boldness, characterizing their risk-taking behaviour or seeking simply to rank individuals in a study group with respect to their relative boldness scores, raises problematic empirical issues. If investigators automatically deploy standards of testing that are reliable for a few well-characterized model species, risks of erroneous labelling of personality traits may be significant (Biro & Dingemanse, 2009b; Carter et al., 2012b). For example, an ornithologist solely implementing a novel object assay to describe boldness in another parrot species may face the risk of mislabelling the trait as boldness rather than calling it a form of foraging activity. We therefore think that a comprehensive approach (multi-measure and multi-assay) can reveal context-dependencies that standard application of tests from model organisms might fail to detect. A preliminary study incorporating all the relevant aspects of boldness, depending on the question that is of interest to the investigator, should also be important to provide groundwork for future, more robust analyses.

Similarly, the growing interest of studying personality traits in the wild will benefit from investigations done in a more controlled environment. For example, latency to approach a regular feeder in the presence of a novel object has been utilized as a measure of neophobia in the field (Greenberg, 1989; Herborn et al., 2010). Number of flights and alarm calls were quantified as metrics for risk-taking in the presence of an intruder in birds (Barnett et al., 2012; Hyman et al., 2013). Similarly, the likelihood to enter into a

trap was applied as a metric for boldness in lizards (Carter et al., 2012b). In field settings such as these, measures that can be accurately assessed may be limited in number, constrained by particular settings (trapping, feeders, etc.) in which the animals can be observed while free living. We suggest that research conducted in a captive setting can facilitate decisions concerning the choice of boldness metrics that are both reliable and practical for field settings. Adapting the best experimental procedures for field-oriented research programs is of high importance in generating understanding of the functions of boldness in ecological contexts (Dammhahn & Almeling, 2012). Moreover, captive testing involving multiple metrics — including those used in the wild — is necessary to assess whether the same behavioural patterns are being assessed in both captive and free-ranging population (Carter et al., 2012d; Forss et al., 2015).

Acknowledgements

References

Andersson, A., Laikre, L. & Bergvall, U.A. (2014). Two shades of boldness: novel object and anti-predator behavior reflect different personality dimensions in domestic rabbits. — J. Ethol. 32: 1-14.

Arnold, K.E., Ramsay, S.L., Donaldson, C. & Adam, A. (2007). Parental prey selection affects risk-taking behaviour and spatial learning in avian offspring. — Proc. Roy. Soc. Lond. B: Biol. Sci. 274: 2563-2569.

Avery, M., Tillman, E., Keacher, K., Arnett, J. & Lundy, K. (2012). Biology of invasive monk parakeets in South Florida. — Wilson J. Ornithol. 124: 581-588.

Barnett, C.A., Thompson, C.F. & Sakaluk, S.K. (2012). Aggressiveness, boldness and parental food provisioning in male house wrens (*Troglodytes aedon*). — Ethology 118: 984-993.

Beckmann, C. & Biro, P.A. (2013). On the validity of a single (boldness) assay in personality research. — Ethology 119: 937-947.

Bell, A. (2012). Randomized or fixed order for studies of behavioral syndromes? — Behav. Ecol. 24: 16-20.

Bell, A.M. (2005). Behavioural differences between individuals and two populations of stickleback (*Gasterosteus aculeatus*). — J. Evol. Biol. 18: 464-473.

Biro, P. & Dingemanse, N. (2009). Sampling bias resulting from animal personality. — Trends Ecol. Evol. 24: 66-67.

Biro, P.A. & Stamps, J.A. (2010). Do consistent individual differences in metabolic rate promote consistent individual differences in behavior? — Trends Ecol. Evol. 25: 653-659.

Boissy, A. (1995). Fear and fearfulness in animals. — Q. Rev. Biol. 70: 165-191.

Boogert, N., Reader, S.M. & Laland, K.N. (2006). The relation between social rank, neophobia and individual learning in starlings. — Anim. Behav. 72: 1229-1239.

Brown, C. & Braithwaite, V.A. (2005). Effects of predation pressure on the cognitive ability of the poeciliid *Brachyraphis episcopi*. — Behav. Ecol. 16: 482-487.

Bucher, E.H., Martin, L.F., Martella, M.B. & Navarro, J.L. (1990). Social behaviour and population dynamics of the Monk Parakeet. — Proc. Int. Ornithol. Congr. 20: 681-689.

Budaev, S.V. (2010). Using principal components and factor analysis in animal behaviour research: caveats and guidelines. — Ethology 116: 472-480.

Burger, J. & Gochfeld, M. (2005). Nesting behavior and nest site selection in monk parakeets (*Myiopsitta monachus*) in the Pantanal of Brazil. — Acta Ethologica 8: 23-34.

Campler, M., Jöngren, M. & Jensen, P. (2009). Fearfulness in red junglefowl and domesticated white Leghorn chickens. — Behav. Process. 81: 39-43.

Campo, J.L., Dávila, S.G., Gil, M.G., Torres, O. & Moreno, J.S. (2015). Fear and stress measurements in pure and hybrid red-legged partridges. — Appl. Anim. Behav. Sci. 166: 123-130.

Carere, C. & Maestripieri, D. (eds) (2013). Animal personalities: behavior, physiology, and evolution. — University of Chicago Press, Chicago, IL.

Carter, A.J., Feeney, W.E., Marshall, H.H., Cowlishaw, G. & Heinsohn, R. (2012a). Animal personality: what are behavioural ecologists measuring? — Biol. Rev. 88: 465-475.

Carter, A.J., Heinsohn, R., Goldizen, A.W. & Biro, P.A. (2012b). Boldness, trappability and sampling bias in wild lizards. — Anim. Behav. 83: 1051-1058.

Carter, A.J., Marshall, H.H., Heinsohn, R. & Cowlishaw, G. (2012c). Evaluating animal personalities: do observer assessments and experimental tests measure the same thing? — Behav. Ecol. Sociobiol. 66: 153-160.

Carter, A.J., Marshall, H.H., Heinsohn, R. & Cowlishaw, G. (2012d). How not to measure boldness: novel object and antipredator responses are not the same in wild baboons. — Anim. Behav. 84: 603-609.

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

Collar, N.J. (1997). Family Psittacidae (parrots). — In: Handbook of the birds of the world, Vol. 4 (del Hoyo, J., Elliott, A., Sargatal, J. & Cabot, J., eds). Lynx Edicions, Barcelona, p. 280-477.

Cote, J., Fogarty, S., Weinersmith, K., Brodin, T. & Sih, A. (2010). Personality trait and dispersal tendency in the invasive mosquitofish (*Gambusia affinis*). — Proc. Roy. Soc. Lond. B: Biol. Sci. 277: 1571-1579.

Dall, S.R.X. & Griffith, S.C. (2014). An empiricist guide to animal personality variation in ecology and evolution. — Behav. Evol. Ecol. 2: 3.

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.

David, M., Auclair, Y. & Cézilly, F. (2011). Personality predicts social dominance in female zebra finches, *Taeniopygia guttata*, in a feeding context. — Anim. Behav. 81: 219-224.

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.

Dzieweczynski, T.L. & Crovo, J.A. (2011). Shyness and boldness differences across contexts in juvenile three-spined stickleback *Gasterosteus aculeatus* from an anadromous population. — J. Fish Biol. 79: 776-788.

Feenders, G., Klaus, K. & Bateson, M. (2011). Fear and exploration in European starlings (*Sturnus vulgaris*): a comparison of hand-reared and wild-caught birds. — PLoS ONE 6: e19074.

Finkler, H. & Terkel, J. (2015). The relationship between individual behavioural styles, dominance rank and cortisol levels of cats living in urban social groups. — Appl. Anim. Behav. Sci. 173: 22-28.

Forss, S.I.F., Schuppli, C., Haiden, D., Zweifel, N. & Van Schaik, C.P. (2015). Contrasting responses to novelty by wild and captive orangutans. — Am. J. Primatol. 77: 1109-1121.

Fox, R.A., Ladage, L.D., Roth, T.C. & Pravosudov, V.V. (2009). Behavioural profile predicts dominance status in mountain chickadees, *Poecile gambeli*. — Anim. Behav. 77: 1441-1448.

Garamszegi, L.Z., Rosivall, B., Rettenbacher, S., Marko, G., Zsebok, S., Szollosi, E., Eens, M., Potti, J. & Torok, J. (2012). Corticosterone, avoidance of novelty, risk-taking and aggression in a wild bird: no evidence for pleiotropic effects. — Ethology 118: 621-635.

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

Greenberg, R. (1989). Neophobia, aversion to open space, and ecological plasticity in song and swamp sparrows. — Can. J. Zool. 67: 1194-1199.

Greenberg, R. & Mettke-Hofmann, C. (2001). Ecological aspects of neophobia and neophilia in birds. — Curr. Ornithol. 16: 119-178.

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.

Highcock, L. & Carter, A.J. (2014). Intraindividual variability of boldness is repeatable across contexts in a wild lizard. — PLoS ONE 9: e95179.

Hobson, E.A., Avery, M.L. & Wright, T.F. (2014). The socioecology of Monk Parakeets: insights into parrot social complexity. — Auk. 131: 756-775.

Hollander, F.A., van Overveld, T., Tokka, I. & Matthysen, E. (2008). Personality and nest defence in the great tit (*Parus major*). — Ethology 114: 405-412.

Hyman, J., Myers, R. & Krippel, J. (2013). Personality influences alarm calling behaviour in song sparrows. — Behaviour 150: 1147-1164.

Ingley, S.J., Rehm, J. & Johnson, J.B. (2014). Size doesn't matter, sex does: a test for boldness in sister species of *Brachyrhaphis* fishes. — Ecol. Evol. 4: 4361-4369.

Jolles, J.W., Ostojić, L. & Clayton, N.S. (2013). Dominance, pair bonds and boldness determine social-foraging tactics in rooks, *Corvus frugilegus*. — Anim. Behav. 85: 1261-1269.

Jones, K. & Godin, J. (2010). Are fast explorers slow reactors? Linking personality type and anti-predator behaviour. — Proc. Roy. Soc. Lond. B: Biol. Sci. 277: 625-632.

Kaiser, H.F. (1959). Computer program for varimax rotation in factor analysis. — Educ. Psychol. Meas. 19: 413-420.

Kaiser, H.F. (1960). The application of electronic computers to factor analysis. — Educ. Psychol. Meas. 20: 141-151.

Kie, J.G. (1999). Optimal foraging and risk of predation: effects on behavior and social structure in ungulates. — J. Mammal. 80: 1114-1129.

Killen, S.S., Marras, S., Ryan, M.R., Domenici, P. & McKenzie, D.J. (2011). A relationship between metabolic rate and risk-taking behaviour is revealed during hypoxia in juvenile European sea bass. — Funct. Ecol. 26: 134-143.

Kurvers, R.H.J.M., Nolet, B.A., Prins, H.H.T., Ydenberg, R.C. & van Oers, K. (2012a). Boldness affects foraging decisions in barnacle geese: an experimental approach. — Behav. Ecol. 23: 1155-1161.

Kurvers, R.H.J.M., van Santen de Hoog, S., van Wieren, S.E., Ydenberg, R.C. & Prins, H.H.T. (2012b). No evidence for negative frequency-dependent feeding performance in relation to personality. — Behav. Ecol. 23: 51-57.

Lê, S., Josse, J. & Husson, F. (2008). FactoMineR: an R package for multivariate analysis. — J. Stat. Softw. 25: 1-18.

Maindonald, J. & Braun, W.J. (eds) (2010). Data analysis and graphics using R: an example-based approach, 3rd edn. — Cambridge University Press, New York, NY.

Martella, M.B. & Bucher, E.H. (1990). Vocalizations of the monk parakeet. — Bird Behav. 8: 101-110.

Martin, L.B. & Fitzgerald, L. (2005). A taste for novelty in invading house sparrows, *Passer domesticus*. — Behav. Ecol. 16: 702-707.

Miller, K.A., Garner, J.P. & Mench, J.A. (2006). Is fearfulness a trait that can be measured with behavioural tests? A validation of four fear tests for Japanese quail. — Anim. Behav. 71: 1323-1334.

Minderman, J., Reid, J.M., Evans, P.G.H. & Whittingham, M.J. (2009). Personality traits in wild starlings: exploration behavior and environmental sensitivity. — Behav. Ecol. 20: 830-837.

Moscicki, M.K. & Hurd, P.L. (2015). Sex, boldness and stress experience affect convict cichlid, *Amatitlania nigrofasciata*, open field behaviour. — Anim. Behav. 107: 105-114.

Noer, C.L., Needham, E.K., Wiese, A.-S., Balsby, T.J.S. & Dabelsteen, T. (2015). Context matters: multiple novelty tests reveal different aspects of shyness–boldness in farmed American mink (*Neovison vison*). — PLoS ONE 10: e0130474.

R Core Team (2014). R: a language and environment for statistical computing. — R Foundation for Statistical Computing, Vienna. Available online at http://cran.r-project.org.

Real, L. & Caraco, T. (1986). Risk and foraging in stochastic environments. — Annu. Rev. Ecol. Evol. Syst. 17: 371-390.

Reale, D., Reader, S., Sol, D., McDougall, P. & Dingemanse, N. (2007). Integrating animal temperament within ecology and evolution. — Biol. Rev. 82: 291-318.

Revelle, W. (2015). psych: procedures for psychological, psychometric, and personality research. — R package version 1.5.4. Available online at http://cran.r-project.org/web/packages/psych/index.html.

Rockwell, C., Gabriel, P.O. & Black, J.M. (2012). Bolder, older, and selective: factors of individual-specific foraging behaviors in Steller's jays. — Behav. Ecol. 23: 676-683.

Russello, M., Avery, M. & Wright, T. (2008). Genetic evidence links invasive monk parakeet populations in the United States to the international pet trade. — BMC Evol. Biol. 8: 217.

Scheid, C. & Noë, R. (2010). The performance of rooks in a cooperative task depends on their temperament. — Anim. Cogn. 13: 545-553.

Schuett, W. & Dall, S.R.X. (2009). Sex differences, social context and personality in zebra finches, *Taeniopygia guttata*. — Anim. Behav. 77: 1041-1050.

Schuett, W., Laaksonen, J. & Laaksonen, T. (2012). Prospecting at conspecific nests and exploration in a novel environment are associated with reproductive success in the jackdaw. — Behav. Ecol. Sociobiol. 66: 1341-1350.

Seaman, B. & Briffa, M. (2015). Parasites and personality in periwinkles (*Littorina littorea*): infection status is associated with mean-level boldness but not repeatability. — Behav. Process. 115: 132-134.

Short, K. & Petren, K. (2008). Boldness underlies foraging success of invasive *Lepidodactylus lugubris* geckos in the human landscape. — Anim. Behav. 76: 429-437.

Sloan Wilson, D., Clark, A.B., Coleman, K. & Dearstyne, T. (1994). Shyness and boldness in humans and other animals. — Trends Ecol. Evol. 9: 442-446.

Small, T.W. & Schoech, S.J. (2015). Sex differences in the long-term repeatability of the acute stress response in long-lived, free-living Florida scrub-jays (*Aphelocoma coerulescens*). — J. Comp. Physiol. B 185: 119-133.

South, J. & Pruett-Jones, S. (2000). Patterns of flock size, diet, and vigilance of naturalized Monk Parakeets in Hyde Park, Chicago. — Condor 102: 848-854.

Stevens, J. (1992). Applied multivariate statistics for the social sciences. — Lawrence Erlbaum Associates, Hillsdale, NJ.

Stowe, M., Bugnyar, T., Heinrich, B. & Kotrschal, K. (2006a). Effects of group size on approach to novel objects in raven (*Corvus corax*). — Ethology 112: 1079-1088.

Stowe, M., Bugnyar, T., Loretto, M.C., Schloegl, C., Range, F. & Kotrschal, K. (2006b). Novel object exploration in ravens (*Corvus corax*): effects of social relationships. — Behav. Process. 73: 68-75.

Tillman, E.A., Genchi, A.C., Lindsay, J.R., Newman, J.R. & Avery, M.L. (2004). Evaluation of trapping to reduce monk parakeet populations at electric utility facilities. — Vertebr. Pest Conf. 21: 126-129.

van den Brink, V., Henry, I., Wakamatsu, K. & Roulin, A. (2012). Melanin-based coloration in juvenile kestrels (*Falco tinnunculus*) covaries with anti-predatory personality traits. — Ethology 118: 673-682.

Vazire, S., Gosling, S.D., Dickey, A.S. & Schapiro, S.J. (2007). Measuring personality in nonhuman animals. — In: Handbook of research methods in personality psychology. p. 190-206.

Verbeek, M.E.M., Drent, P.J. & Wiepkema, P.R. (1994). Consistent individual differences in early exploratory behaviour of male great tits. — Anim. Behav. 48: 1113-1121.

Webster, S.J. & Lefebvre, L. (2001). Problem solving and neophobia in a columbiform–passeriform assemblage in Barbados. — Anim. Behav. 62: 23-32.

Wilson, A.D.M. & Godin, J.G.J. (2009). Boldness and behavioral syndromes in the bluegill sunfish, *Lepomis macrochirus*. — Behav. Ecol. 20: 231-237.

Wilson, A.D.M. & Stevens, E.D. (2005). Consistency in context-specific measures of shyness and boldness in rainbow trout, *Oncorhynchus mykiss*. — Ethology 111: 849-862.

Wilson, A.D.M., Whattam, E.M., Bennett, R., Visanuvimol, L., Lauzon, C. & Bertram, S.M. (2010). Behavioral correlations across activity, mating, exploration, aggression and antipredator contexts in the European house cricket, *Acheta domesticus*. — Behav. Ecol. Sociobiol. 64: 703-715.

Wilson, A.D.M., Binder, T.R., McGrath, K.P., Cooke, S.J. & Godin, J.G.J. (2011). Capture technique and fish personality: angling targets timid bluegill sunfish, *Lepomis macrochirus*. — Can. J. Fish. Aquat. Sci. 68: 749-757.
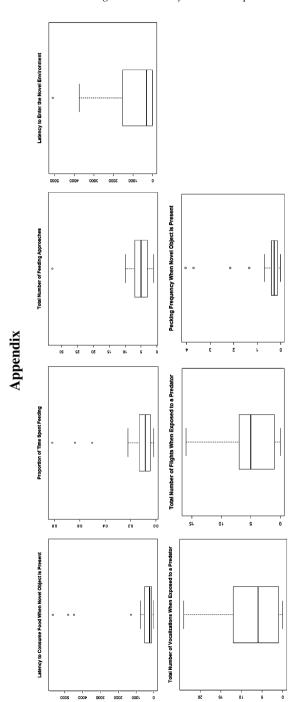
Appendix