

**Fig. 1** Bats' foraging behavior was organized in clear bouts. The time intervals between two visits at a flower had a bimodal distribution when logarithmically transformed. Each curve with gray transparent area underneath is based on the cumulative visits of an individual tagged bat. We selected the time interval of 120 seconds (vertical dashed line) as a criterion to distinguish visits within a bout from visits between bouts.

### ***Behavioral consistency***

Based on prior experiments in the laboratory, we expected the individual differences between bats to remain stable over time. In order to test this, we performed repeatability analyses (Bell et al. 2009; Nakagawa and Schielzeth 2010) separately for males ( $N = 30$ ) and females ( $N = 21$ ). We used a multivariate random intercept and random slope model with *relative intensity* (eq. 1) as a fixed effect and with visits per bout, visit duration, and flowers visited as independent variables (Hadfield 2010; Dingemanse and Dochtermann 2013; *MCMCglmm* package in R 3.5.1, R Development Core Team 2018). We used the Gaussian family for the parameters visits per bout and visit duration, and the Poisson family for flowers visited. Fitting was based on Markov chain Monte Carlo methods. For this, and each further analysis unless otherwise specified, we used 130,000 iterations with a thinning interval of 100 and a burn-in phase of 30,000, obtaining 1,000 samples for each estimate. Different priors (inverse Wishart, flat covariances, and parameter expanded; Hadfield 2010; Mutzel et al. 2013) did not result in qualitative changes in the output. From the output of these models we calculated the adjusted repeatabilities, that is the repeatabilities after controlling for

confounding effects (Nakagawa and Schielzeth 2010). From the same models we also obtained estimates for the between-individual and within-individual correlations of the three behavioral measures, along with their 95% credible intervals.

In order to test which of the three behavioral parameters (visits per bout, visit duration, and flowers visited) were influenced by relative intensity, sex, and body size, we also ran three univariate random intercept and random slope models with relative intensity, sex, and forearm length as fixed factors, and bat individual as random factor on the pooled data set for males and females (*MCMCglmm* package in R 3.5.1, R Development Core Team 2018). For these tests we corrected the *P* values using the false discovery rate correction (Benjamini and Hochberg 1995).

### ***Behavioral syndrome structure***

We compared the posterior estimates for the between-individual correlation pattern with *a priori* hypotheses about behavioral syndrome structures (Figure 2; Dochtermann and Jenkins 2007; Dingemanse et al. 2010). The models are described below.

*Model 1.* Behavioral independence. This was the null model with assumed lack of relationship between the three behavioral parameters.

*Model 2.* A link between visits per bout and visit duration (Figure 2, only path ‘a’ activated), both interpreted as indicators of the *activity* dimension (Réale et al. 2007). Flowers visited is independent from the other two measures.

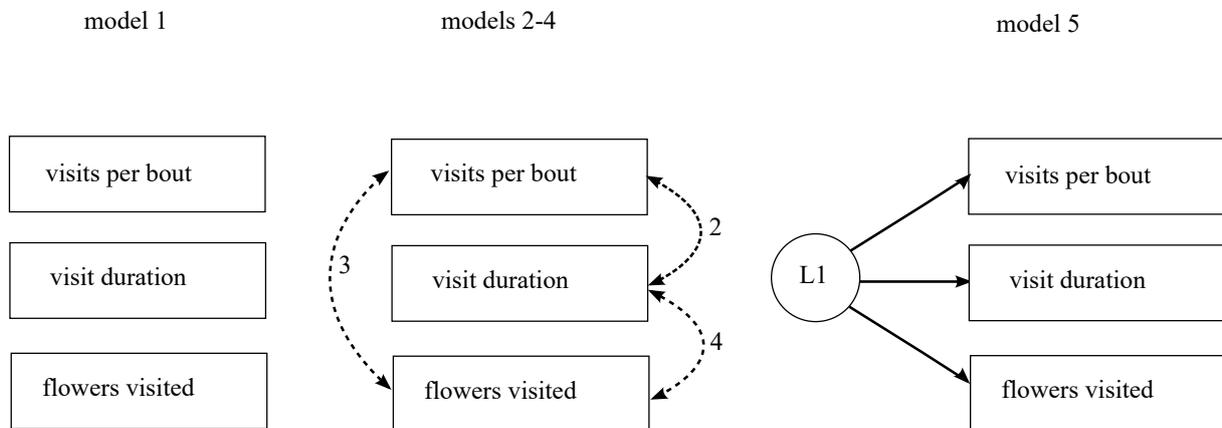
*Model 3.* A link between visits per bout and flowers visited (Figure 2, only path ‘b’ activated). In this model visit duration is independent from the other variables and not an indicator of any particular personality dimension, whereas visits per bout and flowers visited both reflect the *activity* dimension.

*Model 4.* A link between flowers visited and visit duration and (Figure 2, only path ‘c’ activated), both interpreted as probable indicators of the *proactivity-reactivity* dimension (Koolhaas et al. 2007). The rationale behind this hypothesis is that an animal might make visits to a larger number of different flowers to sample their nectar concentration and abort the visits prematurely when the concentration is of the lower type. In this model visits per bout is the sole indicator of the *activity* dimension.

*Model 5.* Full domain-general syndrome with all measures indicators of the same dimension, (e.g. *activity*).

For each sex we extracted the full posterior distributions of covariance estimates generated from the MCMC procedure and fitted each SEM model to these covariance matrices ( $N = 1000$  for each sex) with the appropriate sample size ( $N = 30$  males and  $N = 21$

females) using the lavaan package (Rosseel 2012). The goodness of fit of the models was compared using AIC scores. A lower AIC score indicates higher explanatory power and models with  $\Delta AIC > 2$  (i.e. with AIC scores differing by more than 2 from the score of the model with the lowest AIC score) are generally considered to be statistically unsupported (Dingemanse et al. 2010; Dochtermann and Jenkins 2007; Dochtermann and Jenkins 2011). For each model and each sex, we obtained the distribution of AIC distances from the best model and then ranked the models based on the proportion of samples (out of 1000) that were within  $\Delta AIC$  of 2 from the best model. We considered the top-ranked model to represent the behavioral syndrome structure of *G. commissarisi*. In order to prevent negative variance estimates in model 5, we constrained the estimates to values  $> 0.001$ .



**Fig. 2** Models (1-5) of different syndrome structures. See *Methods* for model descriptions. Continuous unidirectional lines represent a causal relationship from a latent variable (L1) to behavioral measures. Dashed bidirectional arrows represent correlations between behavioral measures expressed in particular syndrome structures. In model 2 only path ‘2’ is active, in model 3 only path ‘3’ is active and in model 4 only path ‘4’ is active.

### *Choice performance*

Foraging performance is often modeled using temporal difference (TD) learning. Here, the expectations about finding a positive reward in an environment with resource locations of yet unknown quality are continually updated with ongoing experience. Choice performance then depends primarily on the learning rate (Glimcher 2011) and also the balance between exploration and exploitation, i.e. the sampling of locations with uncertain quality vs. the visitation of locations with a recently experienced positive outcome (Daw et al. 2006; Mathot et al. 2012). According to the speed-accuracy trade-off hypothesis, animals are expected to either make fast but superficial choices or slow but accurate choices (Sih and Del Giudice 2012). Consequently, we used two different measures of choice performance, bouts to

criterion and error rate (see below), both calculated from the seven nights on which the most extreme differences in sugar concentrations were presented to the bats (5% vs. 20%, 10% vs. 25%, and 15% vs. 30%). These pairs of sugar concentrations were chosen because they were associated with the highest, nearly perfect discrimination performance by bats (Nachev and Winter 2012). Under these conditions the effect of individual differences in perception of sugar concentrations is minimized, so that differences in learning rates and non-perceptual error rates (e.g. due to exploration) can be better estimated. For these measures the sample size was  $N = 38$  bats (23 males and 15 females), because not all bats were detected on at least three of these seven nights. Again, if a bat made fewer than 24 visits on a given night, this data point was not included.

**Visits to criterion** – the number of visits a bat made on a particular night, until the average proportion of visits to the higher sugar concentration flowers (discrimination performance) reached 0.8 or higher. In order to calculate this parameter, we used a change point algorithm (Gallistel et al. 2004; Nachev 2018) that separates a sequence of visits into chunks with significantly different choice preferences. We considered only chunks of at least ten visits. The visits to criterion measure was taken as the number of the first visit of the chunk in which the bat reached the criterion. Most bats reached this criterion on every night on which they were detected, except for three males, each on only a single night. Thus, the average number ( $\pm$  SD) of repeated measures per bat for this parameter in our data set was 6 nights  $\pm$  1.6. All else being equal, a high value for this measure indicates that a bat was faster in avoiding options with lower sugar concentrations and therefore gained a higher energy intake per visit. The probability of making a mistake and the overall number of visits needed to obtain a reliable estimate of the flower quality for each flower both scale with the total number of flowers visited. Therefore, we expected that bats visiting overall a smaller number of different flowers to reach criterion within a lower number of visits. On the other hand, we expected, all else being equal, for bats to reach the criterion just as fast, regardless of how they distributed their visits between bouts. In other words, we expected a bat making 50 visits in a single bout and a bat making ten bouts of five visits each (but visiting the exact same flowers in the exact same sequence) both to reach the criterion in the same number of visits. Therefore, we expected a lack of correlation between visits per bout and visits to criterion. Finally, we did not have a prior expectation for a relationship between visits to criterion and visit duration.

**Error rate** – a measure of the relative frequency of errors (in this case, visits to low concentration flowers) due to factors of a non-perceptual nature, e.g. information-gathering,

exploratory behavior or sampling (Sih and Del Giudice 2012). Error rate was calculated from the data after a bat had reached its nightly asymptotic choice behavior phase. The error rate is the proportion of visits to low concentration flowers after the last change point (Gallistel et al. 2004; Nachev 2018) in preference. We expected our stable experimental conditions (fixed sugar concentrations and volumes during a night) to favor bats with more routine behavior. Once several advantageous flowers had been found a bat did not gain from further exploration. Therefore, bats visiting only a few different flowers were expected to have lower error rates. Since we measured error rate from asymptotic behavior, we expected error rate to be unaffected by visits per bout. We had no prior expectation for the relationship with visit duration. However, according to the speed-accuracy trade-off hypothesis (Sih and Del Giudice 2012), the error rate should positively correlate with the visits to criterion measure of choice performance. As mentioned above, error rate was only determined from those experimental conditions where concentration differences were of high salience.

We ran another multivariate random intercept and random slope model as the one described in *Behavioral consistency*, but added visits to criterion and error rate as dependent variables (*MCMCglmm* package in R 3.5.1, R Development Core Team 2018). We used the Poisson family for visits to criterion and the multinomial family for the error rate. This allowed us to estimate the repeatabilities for the two measures of choice performance, as well as the between and within-individual correlations for all behavioral measures. Because of the larger number of estimated parameters, we used 1,300,000 iterations with a thinning interval of 1000 and a burn-in phase of 30,000, obtaining 1,000 samples for each estimate. We ran separate analyses for males and females, but because of the smaller sample sizes, we also performed a pooled analysis with sex as a fixed factor.

### ***Data availability***

The datasets generated during and/or analysed during the current study, including all statistical tests are available online: <https://doi.org/10.5281/zenodo.2537944>.

## **RESULTS**

Some 50–80 *G. commissarisi* bats foraged simultaneously during the course of the two-alternative choice experiment, making on average one visit per minute per individual. The average number ( $\pm$  SD; range) of flower visits by individual tagged bats per night was  $2900 \pm 1600$  (400-9700). The univariate glmm model revealed no significant differences between males and females in visit duration (estimate = -11.8, credible interval = -41.0, 20.9, hereafter reported as  $-41.0 \leq -11.8 \leq 20.9$ ,  $P = 0.55$ , all  $P$  values corrected for false discovery rate; Table 1). However, males made on average fewer visits per bout ( $-1.66 \leq -0.95 \leq -0.16$ ,

$P = 0.024$ ; Table 1) and visited fewer different flowers than females did ( $-1.67 \leq -1.55 \leq -1.42$ ,  $P = 0.005$ ; Table 1). Forearm length did not significantly affect visits per bout ( $-0.53 \leq -0.07 \leq 0.32$ ,  $P = 0.86$ ), or flowers visited ( $-0.05 \leq 0.00 \leq 0.05$ ,  $P = 0.94$ ), but bats with longer forearms on average made visits with shorter durations ( $-35 \leq -18 \leq -1.3$ ,  $P = 0.045$ ). However, the effect size was small with visits being shorter only by 18 ms per mm difference in forearm length. On average, visits per bout increased with the relative intensity of the difference in available concentrations ( $4.59 \leq 9.74 \leq 14.79$ ,  $P = 0.005$ ), but visit duration ( $-561 \leq -288 \leq -28.9$ ,  $P = 0.072$ ) or the number of flowers visited did not significantly depend on relative intensity of the sugar concentration difference ( $-1.20 \leq 1.34 \leq 2.46$ ,  $P = 0.435$ ). A test of the effect of reproductive status in female bats revealed that pregnant individuals made on average visits with longer durations ( $108 \leq 148 \leq 192$ ,  $P < 0.001$ ,  $N = 8$  pregnant and 13 non-pregnant females) and fewer visits per bout ( $-5.13 \leq -3.91 \leq -2.71$ ,  $P < 0.001$ ) compared to non-pregnant females. There was no significant difference in the number of flowers visited ( $-0.09 \leq 0.02 \leq 0.13$ ,  $P = 0.718$ ) between pregnant and non-pregnant females.

The individual behavioral consistency of all three measures in the 51 tagged individuals was high, as indicated by the individual repeatability estimates for visits per bout (males:  $\text{adj.}R = 0.40$ , 95% credible interval = 0.26, 0.54,  $N = 30$ ; hereafter reported as  $0.26 \leq 0.40 \leq 0.54$ ; females:  $0.26 \leq 0.39 \leq 0.58$ ,  $N = 21$ ), visit duration (males:  $0.59 \leq 0.69 \leq 0.81$ ,  $N = 30$ ; females:  $0.37 \leq 0.57 \leq 0.70$ ,  $N = 21$ ), and flowers visited (males:  $0.29 \leq 0.46 \leq 0.61$ ,  $N = 30$ ; females:  $0.34 \leq 0.49 \leq 0.71$ ,  $N = 21$ ).

**Table 1** Descriptive statistics and phenotypic correlation (Pearson's  $r$ ) coefficients for behavioral parameters of the 51 bats

	Visits per bout	Visit duration (ms)	Flowers visited
<b>Males (<math>N = 30</math>)</b>			
Visits per bout	7.73 (3.13) <sup>a</sup>		
Visit duration (ms)	0.05 (0.80) <sup>b</sup>	659 (214)	
Flowers visited	0.33 (0.08)	<b>0.37 (0.04)</b>	8.28 (3.09)
<b>Females (<math>N = 21</math>)</b>			
Visits per bout	8.86 (3.83)		
Visit duration (ms)	<b>-0.64 (0.002)</b>	662 (150)	
Flowers visited	0.18 (0.42)	0.24 (0.29)	13.10 (3.70)

<sup>a</sup> Values on main diagonals give mean measures with standard deviations in parentheses. These mean measures are calculated from the individual means

<sup>b</sup> Correlations are based on individual means calculated over the whole data set, excluding nights in which bats made fewer than 24 visits, and bats that were detected on fewer than three nights. The corresponding two-tailed  $p$  values are given in parentheses. Significant correlations are shown in bold

***Relationship between behavioral measures and measures of choice performance***

The two measures of choice performance were both significantly repeatable within individuals (visits to criterion:  $0.18 \leq 0.30 \leq 0.54$ ,  $N = 23$  males and  $0.21 \leq 0.40 \leq 0.65$ ,  $N = 15$  females; error rate:  $0.32 \leq 0.49 \leq 0.71$ ,  $N = 23$  males and  $0.28 \leq 0.45 \leq 0.71$ ,  $N = 15$  females). The glmm revealed that bats that made more visits per bout and visited more

**Table 2** Between-individual and within-individual correlation coefficients for behavioral parameters and measures of choice performance

	Between individuals		Within individuals	
	Visits to criterion	Error rate	Visits to criterion	Error rate
<b>Males</b> ( $N = 23$ )				
Error rate	$-0.45 \leq 0.05 \leq 0.49$		$-0.22 \leq 0.13 \leq 0.39$	
Visits per bout	<b><math>0.03 \leq 0.58 \leq 0.83</math></b>	$-0.47 \leq -0.09 \leq 0.47$	$-0.21 \leq 0.11 \leq 0.36$	$-0.51 \leq -0.11 \leq 0.20$
Visit duration	$-0.53 \leq 0.08 \leq 0.45$	$-0.61 \leq -0.13 \leq 0.35$	$-0.31 \leq -0.10 \leq 0.19$	<b><math>-0.58 \leq -0.35 \leq -0.01</math></b>
Flowers visited	$-0.20 \leq 0.21 \leq 0.60$	$-0.41 \leq 0.05 \leq 0.46$	$-0.20 \leq 0.32 \leq 0.60$	$-0.32 \leq -0.02 \leq 0.34$
<b>Females</b> ( $N = 15$ )				
Error rate	$-0.48 \leq 0.01 \leq 0.55$		$-0.10 \leq 0.25 \leq 0.53$	
Visits per bout	$-0.30 \leq 0.27 \leq 0.71$	$-0.69 \leq -0.24 \leq 0.29$	$-0.09 \leq 0.17 \leq 0.53$	<b><math>-0.71 \leq -0.50 \leq -0.06</math></b>
Visit duration	$-0.35 \leq 0.18 \leq 0.65$	$-0.33 \leq 0.16 \leq 0.63$	$-0.50 \leq -0.24 \leq 0.12$	<b><math>-0.70 \leq -0.50 \leq -0.07</math></b>
Flowers visited	$-0.15 \leq 0.28 \leq 0.71$	$-0.44 \leq -0.08 \leq 0.49$	$-0.15 \leq 0.34 \leq 0.71$	$-0.47 \leq -0.18 \leq 0.07$
<b>Pooled<sup>a</sup></b> ( $N = 38$ )				
Error rate	$-0.35 \leq -0.13 \leq 0.44$		$-0.11 \leq 0.18 \leq 0.36$	
Visits per bout	<b><math>0.01 \leq 0.36 \leq 0.66</math></b>	$-0.51 \leq -0.10 \leq 0.26$	$-0.01 \leq 0.17 \leq 0.41$	$-0.56 \leq -0.26 \leq 0.01$
Visit duration	$-0.27 \leq 0.17 \leq 0.50$	$-0.41 \leq -0.01 \leq 0.36$	$-0.28 \leq -0.04 \leq 0.12$	<b><math>-0.59 \leq -0.42 \leq -0.16</math></b>
Flowers visited	<b><math>0.03 \leq 0.34 \leq 0.63</math></b>	$-0.31 \leq 0.03 \leq 0.37$	<b><math>0.03 \leq 0.47 \leq 0.63</math></b>	$-0.28 \leq -0.10 \leq 0.21$

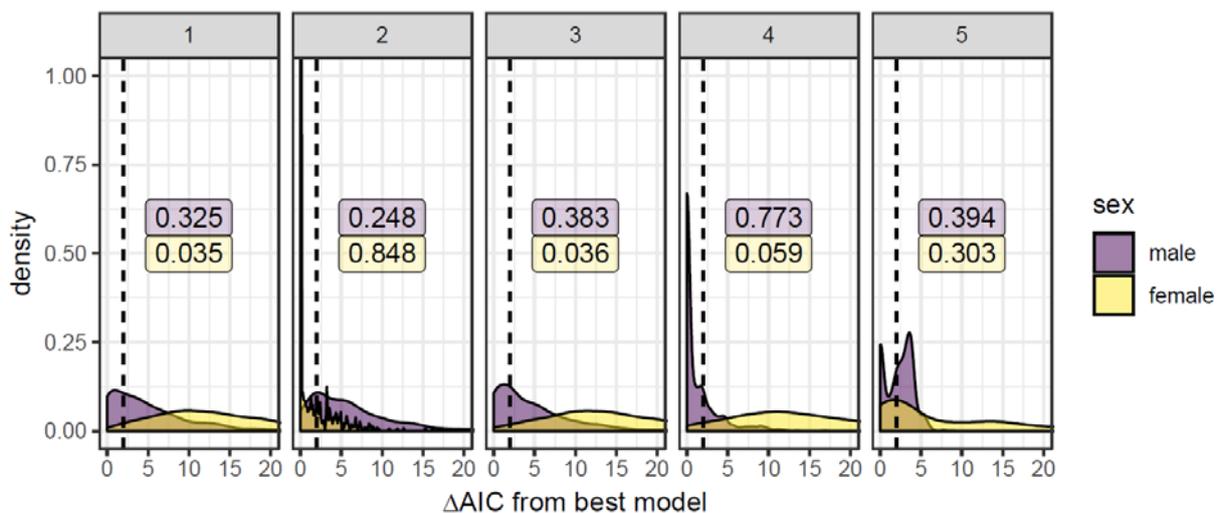
Correlations are based on 1000 posterior samples per individual. The values are given as lower 95% credible interval  $\leq$  posterior mode  $\leq$  upper 95% credible interval. Between-individual correlations are given to the left and within-individual correlations are given to the right. Values in bold do not overlap with 0

<sup>a</sup> In the pooled model sex was used as a fixed factor

flowers on average took more visits to reach the criterion of 80% discrimination performance (Table 2). However, contrary to the expectation based on a speed-accuracy trade-off, there was no significant between-individual correlation between visits to criterion and error rate (Table 2). Within individuals, there was a significant negative correlation between error rate and visit duration in both males and females, and a negative correlation between error rate and visits per bout in females only (Table 2). An analysis on the pooled data from both sexes was comparable to the results in male bats, except for a stronger support for positive between- and within-individual correlations between visits to criterion and flowers visited (Table 2).

### *Analysis of behavioral syndrome structure*

From the five different models tested (Fig. 2), model 4 was best supported in males, while model 2 was best supported in females (Fig. 3, Table 3). The ranking of best supported models was 4, 5, 3, 1, 2 in males and 2, 5, 4, 3, 1 in females. Although model 5 ranked second-best, it was also the only model in which failure of convergence was observed and restriction of the variance estimates to positive numbers was necessary, indicating a poor fit.



**Fig. 3** AIC model comparisons for the five *a priori* models of syndrome structure using structural equation modeling. See Fig. 2 and *Methods* for further model descriptions. The distributions of the distances ( $\Delta$ AIC) from the model with the lowest AIC (best model) for the five models (panels 1-5) are based on the between-individual covariance matrices from the posterior sample ( $n = 1000$ ) for males (purple) and females (yellow). Models with  $\Delta$ AIC  $> 2$  (to the right of the vertical dashed line) were considered statistically unsupported. If a model was the best model in a particular sample, its distance to the best model was 0. The proportion of samples to the left of the dashed line ( $\Delta$ AIC  $\leq 2$ ) is given in each panel for males (purple labels) and females (yellow labels). Improper AIC estimates due to failure of convergence only occurred for model 5, at a proportion of 0.08 in males and 0.24 in females, and were excluded from the analysis.

In males there was a significant positive within-individual correlation between visits per bout and flowers visited, whereas in females all the within-individual correlations were significant and positive (Table 3).

**Table 3** Correlation coefficients for three behavioral parameters between-individuals (below diagonal) and within-individuals (above diagonal) for 51 bats

	Visits per bout	Visit duration	Flowers visited
<b>Males (<math>N = 30</math>)</b>			
Visits per bout		$-0.15 \leq -0.09 \leq 0.00$	<b><math>0.52 \leq 0.61 \leq 0.67</math></b>
Visit duration	$-0.35 \leq 0.07 \leq 0.38$		$-0.04 \leq 0.09 \leq 0.17$
Flowers visited	$-0.18 \leq 0.23 \leq 0.56$	$-0.01 \leq 0.42 \leq 0.67$	
<b>Females (<math>N = 21</math>)</b>			
Visits per bout		<b><math>0.03 \leq 0.13 \leq 0.21</math></b>	<b><math>0.49 \leq 0.59 \leq 0.66</math></b>
Visit duration	<b><math>-0.88 \leq -0.73 \leq -0.39</math></b>		<b><math>0.18 \leq 0.29 \leq 0.42</math></b>
Flowers visited	$-0.42 \leq 0.11 \leq 0.48$	$-0.24 \leq 0.07 \leq 0.59$	

Correlations are based on 1000 posterior samples per individual. The values are given as lower 95% credible interval  $\leq$  posterior mode  $\leq$  upper 95% credible interval. Between-individual correlations are given to the left and below the main diagonal and within-individual correlations are given to the right and above. Values in bold do not overlap with 0

## DISCUSSION

Consistent with our observations from the laboratory, the wild *G. commissarisi* in this study exhibited individual behavioral consistency in the number of visits per bout they made to a patch of artificial flowers, their mean visit duration, and the number of different flowers they visited.

The individual behavioral consistency of the three parameters could not be explained by their interdependence within a full domain-general syndrome. Instead males and females exhibited two different patterns of between-individual correlations (Table 3, Fig. 3). In females the data supported the existence of an activity dimension that describes a continuum between bats that make a high number of visits per bout but only hover briefly in front of each flower and bats that make fewer visits with longer durations (model 2, Fig. 2, Fig. 3). However, due to the smaller sample size and the heterogeneous group structure (including pregnant and non-pregnant individuals) this pattern needs to be replicated before it can be generalized. Surprisingly, the within-individual correlation between visits per bout and visit duration was significant, but in the opposite direction compared to the between-individual correlation. This means that although in general the females that tend to make longer visits also tend to make fewer visits per bout, a typical female on any given night tends to make either more longer visits per bout, fewer shorter visits per bout or something in between.

Since both visits per bout and visit duration correlated negatively with error rate within individuals, it appears that females tend to cut both their visits and bouts short, when encountering flowers of inferior quality.

In males we found support for a correlation between the number of flowers visited and the visit duration measures, but the correlation was positive, rather than the negative correlation we predicted (model 4, Fig. 2, Fig. 3). However, this pattern is still consistent with a proactivity-reactivity dimension, if we take visit duration as an indication of exploration thoroughness, rather than shorter visit durations indicating rejections of inferior quality flowers, as we had originally suggested.

The only consistent positive finding in both male and female bats was a within-individual correlation between visits per bout and flowers visited (Table 3). This means that when individuals made a higher number of visits per bout during a particular night, they were also more likely to visit a higher number of flowers. It was not the case, however, that bats that on average made more visits per bout also visited a higher number of different flowers on average, as indicated by the lack of strong between-individual correlations in either sex. In other words, even when bats made a high number of visits within a bout, some bats spread these visits over many different flowers, while others kept returning to only a few. All in all, our findings demonstrate that accounting for within- and between-individual variance separately (Table 3) reveals details about behavior that cannot be captured from phenotypic correlations alone (Table 1; Dingemanse et al. 2012; Dingemanse and Dochtermann 2013).

Without a consistent pattern of correlations for males and females and without knowing what other behaviors the behavioral measures from this study correlate with, it is not possible to determine the mechanisms leading to the observed behavioral consistency. In addition to latent personality traits, behavioral repeatability may arise from differences in state (age, roost location, etc.) that do interact on a feedback loop with personality traits. For example, the distance bats travelled from their (night-time) roosts to the flower array is a potential uncontrolled confounding factor that could account for the repeatability of the number of visits per bout and potentially other behavioral measures. However, personal observations (VN) of known day-time roost locations revealed a fluctuating number of individuals, suggesting roost lability rather than roost fidelity. Furthermore, we first became aware of the repeatability of visits per bout in our laboratory studies, where all bats were kept in the same room with the flowers and considerations of travel distances were not applicable. To provide a stronger support for personality traits, the roosting behavior would also need to be analyzed and age effects would need to be controlled for.

In the following, we provide some tentative interpretations and ideas for future studies. Concerning the pattern of resource exploitation, some bats consistently visited only a few of the available flowers, whereas others spread their activity over more than half of the flower array (range of mean number of flowers visited: 4-20; Table 1). We suggest that this difference may be a difference in the degree of behavior routinization (Koolhaas et al. 2007; see also Wolf et al. 2008). Our flowers always delivered rewards and the sugar concentrations of their nectars were stable within nights and only varied from night to night, thus favoring the development of stable choice patterns and penalizing unnecessary information-gathering. It remains to be shown that number of flowers visited is linked to known behaviors from the *proactivity-reactivity* (Koolhaas et al. 2007, e.g. aggressive interactions with conspecifics) or *shyness-boldness* continuum (responses to non-novel risk situations, Réale et al. 2007, e.g. delay to resume foraging after a perceived predator attack).

Our results provide some support to the hypothesis that different behavioral types may be better adapted to different environmental conditions (Guillette et al. 2011), more specifically, to different resource qualities and distributions. Though we did not assess fitness directly, differences in foraging efficiency can be positively correlated with fitness (Ritchie 1990; Lemon 1991; Jeanniard-du-Dot et al. 2017). The differences in both measures of choice performance were significantly repeatable. The measure visits to criterion was correlated with both flowers visited and visits per bout (Table 2). Although the correlation between visits per bout and visits to criterion was unexpected, we had based our prediction on the assumption that bats differing in visits per bout do not also differ in the spread of visits over the flower array *per bout*. However, if we define a new measure, flowers per bout, as the average number of different flowers a bat visits per bout during a single night, we can see that there is a significant correlation between flowers per bout and visits per bout (between-individual  $r = 0.36 < 0.60 < 0.77$ , within-individual  $r = 0.74 < 0.77 < 0.79$ ,  $N = 51$ , pooled over males and females). Thus, since bats that make more visits within a bout are also more likely to visit more different flowers, they are slowed down in their learning and have to make a larger number of visits to reach criterion. This interpretation is consistent with a speed-accuracy trade-off. Although the lack of correlation between visits to criterion and error rate is not consistent with a speed-accuracy trade-off, the learning rate and the exploration-exploitation balance are two independent parameters in theoretical treatments of reinforcement learning (e.g. Daw et al. 2006). Under conditions of higher reward uncertainty, we would expect bats that invest more in information-gathering to have better chances of detecting the locations of more profitable sources of nectar. On the other hand, especially at flowers with high nectar

secretion rates, a more routine behavior may confer fitness benefits through ‘defense by exploitation’ (Paton and Carpenter 1984; Ohashi and Thomson 2005). Foragers employing this strategy maintain high activity rates and exploit a limited number of replenishable food sources therefore keeping the mean resource standing crops low. This can reduce resource competition, because competitors using different strategies may perceive the shared food sources as unprofitable and leave to forage elsewhere. In our experiments the behavior of males seems consistent with defense by exploitation (a positive correlation between visit duration and number of flowers visited, Fig. 3). Alternatively, the smaller numbers of visits per bout and flowers visited compared to females might also be an indication of agonistic interactions at the flowers.

Realistic models of bat foraging need to take into account the observed repeatable differences in frequency and distribution of flower visits. However, these differences also have important implications for the fitness of the plants that bats pollinate. For example, bats with different propensities to make revisits to the same plant or flower can exert very different selection pressures on plants, especially if the plants are self-incompatible. On the other hand, bats that make more visits per bout probably remove and deposit more pollen from flowers, since pollen can be ingested during grooming in the pauses between bouts. Finally, although hovering flight duration has been shown to be uncorrelated with pollen transfer (Muchhala and Thomson 2009), it may be that clinging visits result in higher pollen transfer than hovering visits. Thus, both the quality and quantity of pollination service provided by a pollinator may depend on its foraging strategy.

### **Acknowledgments**

We thank Arne Jungwirth for fieldwork assistance and Alexej Schatz for software programming, as well as Jerry Wilkinson and the anonymous reviewers, whose comments greatly helped improve this manuscript. This is a post-peer-review, pre-copyedit version of an article published *Behavioral Ecology and Sociobiology*. The final authenticated version is available online at: <https://doi.org/10.1007/s00265-019-2637-4>

### **ETHICAL STATEMENT**

#### **Funding**

This work was supported by the National Geographic Society (8579-08), the Deutsche Forschungsgemeinschaft (Exc257, Exc277), and the Volkswagen Foundation (84915 to VN).

#### **Conflict of Interest**

The authors declare that they have no conflict of interest.

## Ethical Approval

This is a reanalysis of a previously published study. Treatment of the experimental animals in that study complied with the national laws on animal care and experimentation.

## REFERENCES

- Beekman M, Jordan LA (2017) Does the field of animal personality provide any new insights for behavioral ecology? *Behav Ecol* 28:617–623. doi: [10.1093/beheco/arx022](https://doi.org/10.1093/beheco/arx022)
- Bell AM (2007) Future directions in behavioural syndromes research. *Proc R Soc Lond B* 274:755–761. doi: 10.1098/rspb.2006.0199
- Bell A, Hankison S, Laskowski K (2009) The repeatability of behaviour: a meta-analysis. *Anim Behav* 77:771–783. doi: 10.1016/j.anbehav.2008.12.022
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 57:289–300
- Benus RF, Bohus B, Koolhaas JM, van Oortmerssen GA (1991) Heritable variation for aggression as a reflection of individual coping strategies. *Experientia* 47:1008–1019
- Coppens CM, de Boer SF, Koolhaas JM (2010) Coping styles and behavioural flexibility: towards underlying mechanisms. *Phil Trans R Soc B* 365:4021–4028. doi: 10.1098/rstb.2010.0217
- Daw ND, O’Doherty JP, Dayan P, Seymour B, Dolan RJ (2006) Cortical substrates for exploratory decisions in humans. *Nature* 441:876–879. doi: 10.1038/nature04766
- Dingemanse NJ, Dochtermann N, Wright J (2010) A method for exploring the structure of behavioural syndromes to allow formal comparison within and between data sets. *Anim Behav* 79:439–450. doi: 10.1016/j.anbehav.2009.11.024
- Dingemanse NJ, Dochtermann NA (2013) Quantifying individual variation in behaviour: mixed-effect modelling approaches. *J Anim Ecol* 82:39–54. doi: 10.1111/1365-2656.12013
- Dingemanse NJ, Dochtermann NA, Nakagawa S (2012) Defining behavioural syndromes and the role of “syndrome deviation” in understanding their evolution. *Behav Ecol Sociobiol* 66:1543–1548. doi: 10.1007/s00265-012-1416-2
- Dingemanse NJ, Réale D (2005) Natural selection and animal personality. *Behaviour* 142:1159–1184. doi: 10.1163/156853905774539445
- 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. doi: 10.1098/rspb.2007.0622
- Dochtermann NA, Jenkins SH (2011) Multivariate methods and small sample sizes. *Ethology* 117:95–101. doi: 10.1111/j.1439-0310.2010.01846.x

Dochtermann NA, Schwab T, Sih A (2015) The contribution of additive genetic variation to personality variation: heritability of personality. *Proc R Soc B* 282:20142201. doi: 10.1098/rspb.2014.2201

Fisher DN, Brachmann M, Burant JB (2018) Complex dynamics and the development of behavioural individuality. *Anim Behav* 138:e1–e6. doi: [10.1016/j.anbehav.2018.02.015](https://doi.org/10.1016/j.anbehav.2018.02.015)

Gallistel CR, Fairhurst S, Balsam P (2004) The learning curve: Implications of a quantitative analysis. *P Natl Acad Sci USA* 101:13124–13131. doi: 10.1073/pnas.0404965101

Glimcher PW (2011) Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *P Natl Acad Sci USA* 108:15647–15654. doi: 10.1073/pnas.1014269108

Gosling SD (2001) From mice to men: What can we learn about personality from animal research? *Psychol Bull* 127:45–86. doi: 10.1037/0033-2909.127.1.45

Gould TD, Gottesman II (2006) Psychiatric endophenotypes and the development of valid animal models. *Genes Brain Behav* 5:113–119. doi: 10.1111/j.1601-183X.2005.00186.x

Guillette LM, Reddon AR, Hoeschele M, Sturdy CB (2011) Sometimes slower is better: slow-exploring birds are more sensitive to changes in a vocal discrimination task. *Proc R Soc Lond B* 278:767–773. doi: 10.1098/rspb.2010.1669

Hadfield JD (2010) MCMC Methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J Stat Softw* 33:1–22. doi: [10.18637/jss.v033.i02](https://doi.org/10.18637/jss.v033.i02)

Jeanniard-du-Dot T, Trites AW, Arnould JPY, Guinet C (2017) Reproductive success is energetically linked to foraging efficiency in Antarctic fur seals. *PLoS ONE* 12:e0174001. doi: [10.1371/journal.pone.0174001](https://doi.org/10.1371/journal.pone.0174001)

Koolhaas JM, de Boer SF, Buwalda B, van Reenen K (2007) Individual variation in coping with stress: a multidimensional approach of ultimate and proximate mechanisms. *Brain Behav Evol* 70:218–226. doi: 10.1159/000105485

Lemon WC (1991) Fitness consequences of foraging behaviour in the zebra finch. *Nature* 352:153–155. doi: 10.1038/352153a0

Mathot KJ, Wright J, Kempnaers B, Dingemanse NJ (2012) Adaptive strategies for managing uncertainty may explain personality-related differences in behavioural plasticity. *Oikos* 121:1009–1020. doi: 10.1111/j.1600-0706.2012.20339.x

Muchhala N, Thomson JD (2009) Going to great lengths: selection for long corolla tubes in an extremely specialized bat-flower mutualism. *Proc R Soc Lond B* 276:2147–2152. doi: 10.1098/rspb.2009.0102

Mutzel A, Dingemanse NJ, Araya-Ajoy YG, Kempenaers B (2013) Parental provisioning behaviour plays a key role in linking personality with reproductive success. *Proc R Soc B* 280:20131019. doi: 10.1098/rspb.2013.1019

Nachev V (2018) ontogenerator/cpdetectorrr: Blue Opal. Zenodo, doi: 10.5281/zenodo.1316925

Nachev V, Stich KP, Winter C, Bond A, Kamil A, Winter Y (2017) Cognition-mediated evolution of low-quality floral nectars. *Science* 355:75–78. doi: [10.1126/science.aah4219](https://doi.org/10.1126/science.aah4219)

Nachev V, Winter Y (2012) The psychophysics of uneconomical choice: non-linear reward evaluation by a nectar flower. *Anim Cogn* 15:393–400. doi: 10.1007/s10071-011-0465-7

Nachev V, Winter Y (2018) Repeatability and foraging performance in wild free-flying nectarivorous bats (*Glossophaga commissarisi*). Zenodo, doi: 10.5281/zenodo.2537944

Nachev V, Stich KP, Winter Y (2013) Weber’s law, the magnitude effect and discrimination of sugar concentrations in nectar-feeding animals. *PLoS ONE* 8:e74144. doi: 10.1371/journal.pone.0074144

Nakagawa S, Schielzeth H (2010) Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biol Rev* 85:935–956. doi: 10.1111/j.1469-185X.2010.00141.x

Ohashi K, Thomson JD (2005) Efficient harvesting of renewing resources. *Behav Ecol* 16:592–605. doi: 10.1093/beheco/ari031

Paton DC, Carpenter FL (1984) Peripheral foraging by territorial rufous hummingbirds: defense by exploitation. *Ecology* 65:1808–1819. doi: 10.2307/1937777

Penke L, Denissen JJA, Miller GF (2007) The evolutionary genetics of personality. *Eur J Personality* 21:549–587. doi: 10.1002/per.629

R Development Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, <http://www.R-project.org>

Réale D, Reader SM, Sol D, McDougall PT, Dingemanse NJ (2007) Integrating animal temperament within ecology and evolution. *Biol Rev* 82:291–318. doi: 10.1111/j.1469-185X.2007.00010.x

Ritchie ME (1990) Optimal foraging and fitness in Columbian ground squirrels. *Oecologia* 82:56–67. doi: 10.1007/BF00318534

Roche DG, Careau V, Binning SA (2016) Demystifying animal ‘personality’ (or not): why individual variation matters to experimental biologists. *J Exp Biol* 219:3832–3843. doi: 10.1242/jeb.146712

Rosseel Y (2012) lavaan: An R Package for Structural Equation Modeling. *J Stat Softw* 48:1–36. doi: 10.18637/jss.v048.i02

- Sih A, Bell AM, Johnson JC (2004) Behavioral syndromes: an ecological and evolutionary overview. *Trends Ecol Evol* 19:372–378. doi: 10.1016/j.tree.2004.04.009
- Sih A, Del Giudice M (2012) Linking behavioural syndromes and cognition: a behavioural ecology perspective. *Phil Trans R Soc B* 367:2762–2772. doi: 10.1098/rstb.2012.0216
- Sih A, Mathot KJ, Moirón M, Montiglio P-O, Wolf M, Dingemanse NJ (2015) Animal personality and state–behaviour feedbacks: a review and guide for empiricists. *Trends Ecol Evol* 30:50–60. doi: 10.1016/j.tree.2014.11.004
- Thiele J (2006) Nahrungssuchstrategien der nektarivoren Fledermaus *Glossophaga commissarisi* (Phyllostomidae) im Freiland - eine individuenbasierte Verhaltensstudie unter Verwendung von Transpondertechnik. Dissertation, Ludwig Maximilian University, <http://edoc.ub.uni-muenchen.de/5566/>
- Tschapka M (2004) Energy density patterns of nectar resources permit coexistence within a guild of Neotropical flower-visiting bats. *J Zool* 263:7–21. doi: 10.1017/S0952836903004734
- Winter Y, Stich KP (2005) Foraging in a complex naturalistic environment: capacity of spatial working memory in flower bats. *J Exp Biol* 208:539–548. doi: 10.1242/jeb.01416
- Wolf M, van Doorn GS, Leimar O, Weissing FJ (2007) Life-history trade-offs favour the evolution of animal personalities. *Nature* 447:581–584. doi: 10.1038/nature05835
- Wolf M, van Doorn GS, Weissing FJ (2008) Evolutionary emergence of responsive and unresponsive personalities. *P Natl Acad Sci USA* 105:15825–15830. doi: 10.1073/pnas.0805473105