Animal Cognition
https://doi.org/10.1007/s10071-018-1174-2

ORIGINAL PAPER

Proactive behavior, but not inhibitory control, predicts repeated innovation by spotted hyenas tested with a multi-access box

Lily Johnson-Ulrich[1,2,3] · **Zoe Johnson-Ulrich**[4] · **Kay Holekamp**[1,2,3]

Received: 10 October 2017 / Revised: 26 February 2018 / Accepted: 2 March 2018

Abstract

Innovation is widely linked to cognitive ability, brain size, and adaptation to novel conditions. However, successful innovation appears to be influenced by both cognitive factors, such as inhibitory control, and non-cognitive behavioral traits. We used a multi-access box (MAB) paradigm to measure repeated innovation, the number of unique innovations learned across trials, by 10 captive spotted hyenas (*Crocuta crocuta*). Spotted hyenas are highly innovative in captivity and also display striking variation in behavioral traits, making them good model organisms for examining the relationship between innovation and other behavioral traits. We measured persistence, motor diversity, motivation, activity, efficiency, inhibitory control, and neophobia demonstrated by hyenas while interacting with the MAB. We also independently assessed inhibitory control with a detour cylinder task. Most hyenas were able to solve the MAB at least once, but only four hyenas satisfied learning criteria for all four possible solutions. Interestingly, neither measure of inhibitory control predicted repeated innovation. Instead, repeated innovation was predicted by a proactive syndrome of behavioral traits that included high persistence, high motor diversity, high activity and low neophobia. Our results suggest that this proactive behavioral syndrome may be more important than inhibitory control for successful innovation with the MAB by members of this species.

Keywords Innovation · Inhibitory control · Cognition · Problem-solving · Behavioral syndromes

Introduction

Innovation is the ability to invent novel solutions to existing problems and solve novel problems (Reader et al. 2016). A growing body of research has shown that innovation is adaptive in a wide array of species for invading novel environments (Lefebvre et al. 2004). Innovation in the wild is related to brain size, behavioral flexibility, general intelligence, culture, and even anatomical evolution and speciation (for review see Reader et al. 2016). Spontaneous innovation in the wild can be very difficult to observe in some species. For this reason, many researchers have begun to use experimental problem-solving paradigms to measure innovative abilities in both captive and wild animals (Auersperg et al. 2011; Benson-Amram et al. 2013; Borrego and Dowling 2016). Innovative problem-solving paradigms typically require a subject to perform a specific motor pattern to obtain food from an apparatus and are sometimes referred to as "extractive foraging tasks." However, experimental assessments of innovation via problem-solving paradigms can be problematic because many traits, some of which are "non-cognitive," affect problem-solving, and variation in any of these traits might explain variation in success. Although the traits that underlie problem-solving are also those predicted to underlie innovation (Griffin and Guez 2014), the influence of non-cognitive traits calls into question the validity of problem-solving as a measure of cognition. For example, a recent study suggests that problem-solving ability in dogs may not actually be a separable construct from temperament (Bray

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10071-018-1174-2) contains supplementary material, which is available to authorized users.

✉ Lily Johnson-Ulrich
john3923@msu.edu

[1] Department of Integrative Biology, Michigan State University, 288 Farm Lane Rm 203, East Lansing, MI 48824, USA

[2] Ecology, Evolutionary Biology, and Behavior Program, Michigan State University, East Lansing, MI 48824, USA

[3] BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, MI 48824, USA

[4] Department of Psychology, Oakland University, Rochester, MI 48309, USA

Published online: 06 March 2018

et al. 2017). To further investigate the relationship between problem-solving and cognition, we aimed to concurrently examine several traits, including inhibitory control, that may underlie problem-solving using a multi-access box paradigm that requires repeated innovation and learning.

Inhibitory control is one important cognitive trait thought to underlie innovation and problem-solving; it is the ability to resist performing a prepotent or previously learned behavior when it is inappropriate, such that the behavior will yield no reward (MacLean et al. 2014). Inhibitory control is widely believed to be an important component of innovative problem-solving because individuals must inhibit previously learned responses and innate biases in order to develop a novel behavioral solution to a familiar problem (Manrique et al. 2013). Inhibitory control is well established as an important executive function in human cognition (e.g. Mischel et al. 1989), and it is also a well-studied aspect of animal cognition (Kabadayi et al. 2018). Overall, problem-solving and inhibitory control are often discussed together, with failures in problem-solving often attributed to inadequate inhibitory control (Taylor et al. 2009; Thornton and Samson 2012; Manrique et al. 2013). Although inhibitory control is strongly correlated with success in various cognitive tests among primates (Herrmann et al. 2007; Burkart et al. 2017), two studies that have independently measured both inhibitory control and problem-solving in other animals failed to find straightforward relationships between them. First, Shaw (2017) found no correlation between performance on an inhibitory control task and a problem-solving task in wild North Island robins. Second, problem-solving performance by domestic dogs was related to their performance in two inhibitory control tasks, but the relationship was negative for one task and positive for another (Müller et al. 2016). Thus, the relationship between an individual's level of inhibitory control and its problem-solving ability remains unclear.

Other non-cognitive behavioral traits that have been measured in relation to problem-solving success include persistence, motivation, motor diversity, neophobia, efficiency and activity (Sih and Del Giudice 2012; Griffin and Guez 2014; Chow et al. 2016). Persistence typically correlates positively with problem-solving success; individuals (and perhaps also species) that engage more with tasks are more likely to solve them (Benson-Amram et al. 2016; Griffin and Guez 2014). Interestingly, motivation does not typically seem to correlate with problem-solving success separately from its influence on persistence (Griffin and Guez 2014). Motor diversity also has a strong positive relationship with problem-solving success; the number of different motor actions that an animal uses predicts success (Griffin and Guez 2014). Low neophobia sometimes predicts success between individuals, though evidence on this is mixed (Griffin and Guez 2014). Efficiency, or the amount of time taken to solve a task, is often used as the dependent measure in problem-solving studies (Chow et al. 2016). However, individuals that are faster at solving novel problems are sometimes less accurate across trials or tasks (Sih and Del Giudice 2012). Activity, the general physical activity level of an individual, is thought to be related to high efficiency in problem-solving, but also less flexibility in relation to tasks like reversal learning (Sih and Del Giudice 2012; Brust et al. 2013; Schuster et al. 2017).

The multi-access box (MAB) paradigm is ideal for addressing questions about the relationship between cognitive and non-cognitive behavioral traits in problem-solving studies. A MAB is a problem-solving apparatus that offers a novel way of assessing innovation. A MAB has multiple, unique entry points or solutions on a box to a common interior that is baited with food rewards (Auersperg et al. 2011; Manrique et al. 2013; Huebner and Fichtel 2015). Different MAB solutions may require sequential learning of skills needed to open the box (Huebner and Fichtel 2015), or they may require different cognitive skills (Auersperg et al. 2011) or motor actions (Manrique et al. 2013). A MAB can be used to measure both innovation and learning because subjects can be scored on several variables across successive trials. For example, innovation can be measured as finding a novel solution and learning can be measured as the number of solutions a subject learns to open reliably across multiple trials. Once a subject has learned one solution, it can be blocked so that the subject is required to learn a new solution to access further food rewards. Because subjects must inhibit using blocked solutions in order to learn new ones, the MAB paradigm allows researchers to directly measure inhibitory control by recording the amount of time spent on blocked solutions before discovering a new solution. As in traditional problem-solving tasks, other behavioral traits can also be assessed for their relative influence on performance with the MAB.

Here, we used a MAB (Auersperg et al. 2011) that we designed for use with mammalian carnivores. Our MAB had four solutions, each requiring a different motor action, all of which occur within the repertoires of most carnivore species. Repeated innovation was assessed as the number of these solutions learned. We use the term "repeated innovation" because, unlike binary measures of innovation, achieving a high score requires subjects to innovate multiple times and demonstrate learning through repeated use of each solution. We also used the MAB to measure persistence, motor diversity, motivation, activity, efficiency, neophobia, and inhibitory control.

In addition to measuring inhibitory control exhibited by subjects while interacting with the MAB, we also used a standard "cylinder task" to independently measure inhibitory control in the same subjects. The cylinder task is a detour-reaching task that requires an individual to inhibit

the impulse to go straight toward a food reward easily visible inside a clear cylinder, and instead detour to the opening on either end of the cylinder to retrieve the food (Kabadayi et al. 2018). Inhibitory control on the cylinder task requires subjects to use a previously learned detour response while inhibiting an impulse to reach straight for food. By contrast, inhibitory control with the MAB requires inhibiting a previously learned response to a blocked solution and trying a novel behavior. Thus, these may represent different types of inhibitory control, each bearing a unique relationship to innovation. While some previous research has shown that both types of inhibitory control positively correlate with one another (MacLean et al. 2014), other studies suggest that there may be multiple types of inhibitory control that do not necessarily correlate (Brucks et al. 2017).

We chose to assess repeated innovation with the MAB in captive spotted hyenas (*Crocuta crocuta*) because these animals have been established as good model organisms for testing hypotheses suggested to explain the evolution of intelligence (Holekamp et al. 2007); innovation has been previously assessed both in wild and captive subjects using a single-access puzzle box (Benson-Amram et al. 2013). Captive hyenas readily participate in cognitive tests, and they are fairly innovative relative to other carnivores (Benson-Amram et al. 2016). In addition, personality traits have been assessed in both captive (Gosling et al. 1998) and wild hyenas (Yoshida et al. 2016). The striking variation among hyenas in regard to personality traits makes them a good model for understanding the role that personality and other non-cognitive behavioral traits might play in relation to repeated innovation. Finally, problem-solving success is correlated with brain size in captive carnivores, including spotted hyenas, suggesting that problem-solving may be an adequate paradigm for assessing innovative problem-solving in these animals (Benson-Amram et al. 2016).

Our first goal was to investigate the relationship between repeated innovation and inhibitory control. Based on previous research (MacLean et al. 2014), we expected that our measures of inhibitory control from the MAB and the cylinder task would correlate with one another. In addition, based on the suggested link between inhibitory control and failures in problem-solving (Manrique et al. 2013), we predicted that both our measures of inhibitory control would be positively related to repeated innovation. Our second goal was to investigate the other behavioral traits associated with innovation and problem-solving, including persistence, motivation, motor diversity, activity, efficiency, and neophobia. In order to address the issue regarding whether successful problem-solving is affected by these "non-cognitive" traits, we chose to use a MAB. The MAB paradigm was designed to feature a battery of tasks based on the idea that performance on any single task is unlikely to be fully representative of an individual's cognitive abilities (Auersperg et al. 2011). While performance on each task individually is likely to be affected by various traits, performance across different tasks should be a strong indication of cognitive ability (Burkart et al. 2017). Here, subjects were required to innovate multiple times using a diverse array of motor actions and demonstrate learning of each innovation. Because repeated innovation scores on our MAB may be more indicative of cognitive ability than success on a single task, we predicted that our non-cognitive traits, persistence, motivation, motor diversity, and neophobia, would have no relationship to repeated innovation. Our only exception from this prediction was with respect to efficiency and activity, both of which are associated with reduced accuracy or flexibility. We predicted that both of these would have a negative relationship with repeated innovation, which requires both accuracy and flexibility to achieve a high score.

Methods

Subjects

We tested 10 captive spotted hyenas housed at two different institutions, 5 at the Oak Creek Zoological Conservatory (OCZC) (Madisonville, TX) and 5 at the Denver Zoo (Denver, CO). Subjects included 4 adult females, 3 adult males, 2 subadult females, and 1 subadult male. All subjects had been located at their institutions for at least a year before testing began with the exception of "Wibari," who had only been at OCZC for 2 months.

Apparatus

Our MAB was a square metal box 40.64 × 40.64 × 40.64 cm (length × width × height), weighing approximately 18 kg (Fig. 1). The interior of the MAB was baited with a food reward. Each vertical side of the MAB provided access to the interior of the box via a solution that required a different motor behavior to open. The solutions were as follows: (1) The push flap: this was a door 30.5 × 28 cm with a hinge on the top that could be pushed inwards to open. (2) The sliding door: this was a door 30.5 × 28 cm with protruding flanges that could be pushed or pulled sideways to slide open. (3) The pull flap: this was a door 30.5 × 28 cm with a hinge on the bottom that could be pulled outwards and downwards to open by grasping a doorknob near the top of the door. (4) The drawer: this was a drawer 10 cm in height that was flush against the bottom of the MAB and took up the entirety of the floor of the MAB; it could be pulled outwards to open. Magnets were used to create mild resistance on all solutions to ensure that they would not accidentally fall open; subjects were required to actively interact with the MAB to retrieve the food. All solutions could be accessed using either mouth

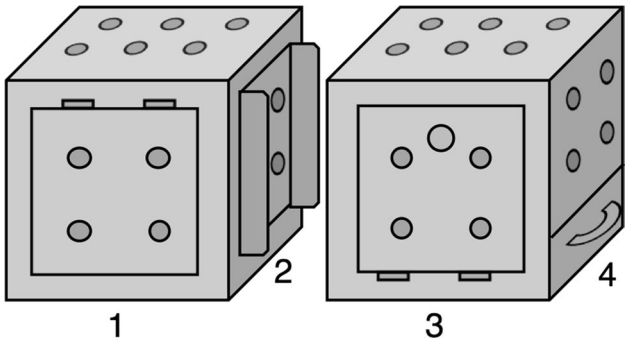

Fig. 1 MAB used in the current study. (1) The push flap solution; (2) the sliding door solution; (3) the pull flap solution; and (4) the drawer solution. Small filled gray circles indicate the approximate number and location of holes drilled through the wall of the MAB. Large gray circle on side 3 represents the location of the door knob. Small rectangles represent the location of hinges

or paws. The top of the MAB was removable for familiarization trials. The MAB had multiple 2.5-cm circular holes cut on every side except the bottom so that subjects could smell the food inside during trials. All four solutions could be blocked by bolting them shut such that blocked solutions could still be manipulated by subjects but would not open to allow food retrieval.

In the cylinder task, we used two hollow cylinders. In familiarization trials, we used an opaque cylinder made from PVC. In test trials, we used a transparent cylinder made from cast acrylic plastic. Both cylinders measured 46 cm in length and 30.5 cm in diameter with a wall thickness of 1 cm. During testing, we placed a cylinder horizontally on the ground with both ends open for the retrieval of food through either end. Unlike previous studies that have used an inhibitory control cylinder, these cylinders were drilled with holes 4 cm in diameter evenly spaced across the wall of the cylinder. Spotted hyenas are highly sensitive to olfactory cues and this was done to ensure that the scent of the food would not spontaneously lead subjects to the correct behavior before they had seen the food through the wall of the transparent cylinder.

Procedure

Each subject was tested alone in its home enclosure. Subjects were not fasted prior to testing; instead testing took place during either their morning (AM) or evening (PM) feeding times, and food rewards were part of their normal daily diet, which included, eggs, ground meat, and pieces of meat and bone. The size of the food reward was kept to roughly 200 g. Each subject was temporarily moved to an adjacent enclosure and the MAB or the cylinder was then placed between 1–2 m from the entrance of the test enclosure, in a location clearly visible to both human observers and the subject when it re-entered the test enclosure. All trials were videotaped. Trials began when the subject entered the test enclosure and ended upon successful retrieval of the food or after 15 min had elapsed, whichever came first.

Subjects were first given 2–5 familiarization trials with the MAB, during which all four solutions were accessible and the top of the box was removed. All solutions were left accessible, rather than blocked, during familiarization trials to prevent subjects from wrongly learning that a solution could not be opened if they attempted to do so, but no hyenas retrieved food via any solution during familiarization trials. Familiarization trials ended when the individual retrieved the food or 15 min elapsed, whichever came first. Subjects were required to successfully retrieve food from the MAB in under 3 min on at least 2 consecutive trials to progress to the first test phase.

There were four phases of MAB testing. During the initial phase, the top of the box was in place and all four solutions were accessible. Once a subject used the same solution in 3 out of 4 consecutive trials, that solution was considered learned and the subject would progress to the next phase of testing. In phase 2, the learned solution from phase 1 was made inaccessible by screws that held it closed. In phase 3, the two previously learned solutions were inaccessible. In the final phase of testing all solutions but one were inaccessible. Subjects could fail to progress to subsequent phases in two ways: either they timed out (15 min elapsed without successful food retrieval) in 4 consecutive trials across more than 1 testing session or they required more than 7 trials to reach learning criterion, excluding trials in which the subject timed out. Seven trials were used as the cut-off for learning because the chance of reaching learning criterion by chance alone exceeded 50% at 8 trials during phase 1. We continued to use this criterion in subsequent phases largely because reducing it further for each phase would have interfered with our learning procedure that allowed for at least 4 trials per subject per phase. Ultimately, only 1 subject reached the 7 trial limit on phase 2, 1 subject required 5 trials for phase 2, and all other subjects reached criterion in 3 or 4 trials on phase 2, rendering an even stricter criterion unnecessary (Fig. 2). Subjects were given a score of 0 through 4 for the number of solutions learned; this represented our measure of repeated innovation.

Before using the clear cylinder to test inhibitory control (hereafter inhibition-C), we gave subjects a minimum of 5 familiarization trials with the opaque cylinder. Subjects were required to pass at least 4 of 5 trials in order to proceed to test trials with the transparent cylinder. A "pass" was defined as successful retrieval of the food without touching the outside of the cylinder. Subjects were allowed to retrieve the food regardless of whether they passed or failed. After familiarization trials, each subject was given 10 test trials with the transparent cylinder. Inhibition-C was scored as

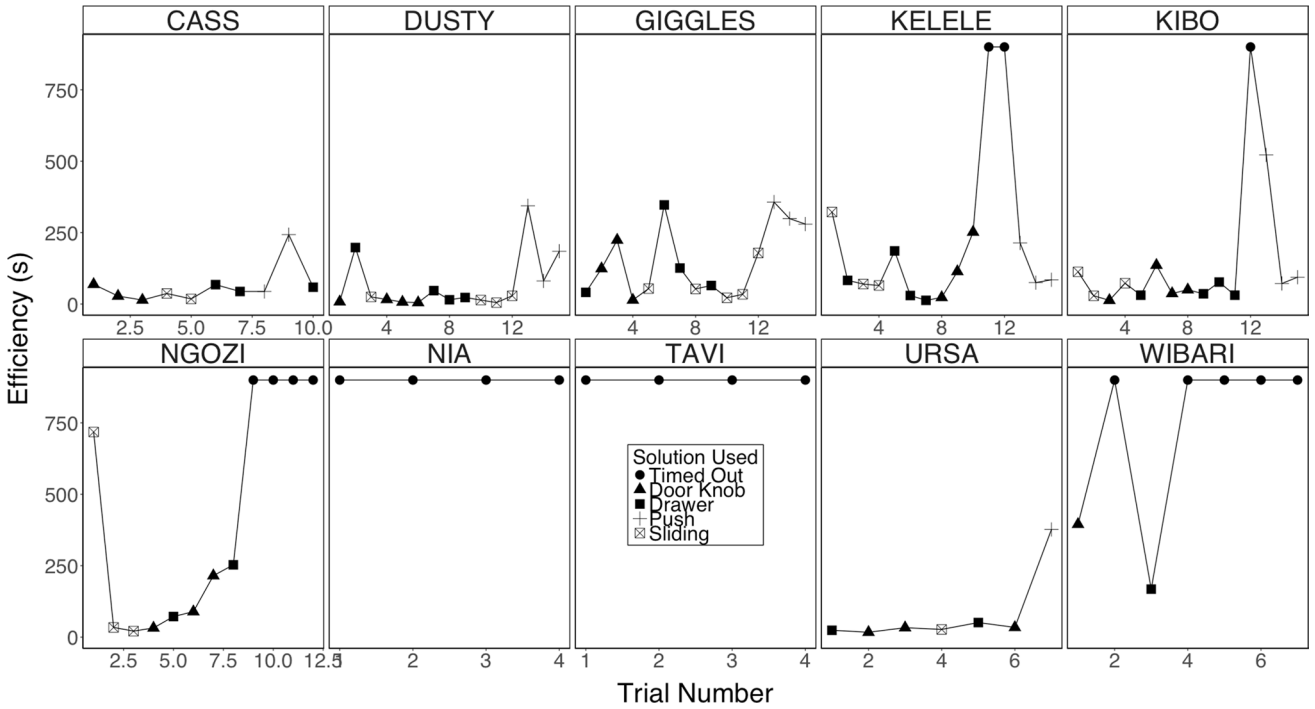

Fig. 2 Problem-solving efficiency across trials for each subject. Note that the x-axis scales vary among subjects. Efficiency was measured as the number of seconds from first contact to food retrieval. In trials where subjects failed to retrieve food a score of 900 s was applied

the number of passes a subject achieved out of 10 trials. A high inhibition-C score indicated good inhibitory control. The cylinder was always positioned perpendicular to the exhibit entrance to encourage an approach toward the long side of the cylinder such that subjects were actively required to detour to either side of the cylinder to obtain the food reward.

All subjects at the Denver Zoo were presented with the cylinder task first, and all subjects at the OCZC were presented with the MAB first. On the first day of testing, subjects were given three trials with each apparatus. On subsequent days, subjects received two to five trials with each apparatus, with a maximum of six trials total involving both apparatuses per day. Subjects were tested approximately 6 days each week until completion of testing. Testing for all 10 subjects was completed over the course of 2-week periods at both institutions.

Behavioral assays

All trial videos were coded using BORIS v. 2.97 (Friard and Gamba 2016). Repeated innovation was measured as the number of the four possible MAB solutions learned by each subject (Table 1, #9). We defined learning as solving the MAB with the same solution in three of four consecutive trials. Successfully solving the MAB was defined as a trial in which the subject opened the box and retrieved the bait using one of the four possible solutions. The learning criterion was based on preliminary observations of carnivores interacting with the MAB. We verified this criterion by recording the amount of time subjects spent in contact with the solution that they ultimately used to retrieve food from the MAB (work time, Table 1, #7), and inquiring how work time changed across successful trials.

Six specific measures were calculated once for each trial (Table 1, #1–6). Persistence is typically described as the extent to which subjects engage with an apparatus, usually measured as contact time, work time, or number of attempts (Griffin and Guez 2014; van Horik et al. 2017). Here, persistence was extracted as the ratio of time spent in contact with the MAB to the total duration of the trial. Contact was defined as touching the MAB with the snout or paws. Scores close to 1 indicated high persistence and scores close to 0 indicated low persistence. We specifically used a ratio to measure persistence, instead of total time in contact with the MAB, to avoid confounding persistence with efficiency.

Motivation, the degree to which an animal wants to engage with a task (typically because of desire for food rewards), is closely related to persistence in that persistence can be considered a measure of motivation to work on a task. Motivation is typically measured as either body condition, some other measure of hunger, or latency to approach an apparatus (Griffin and Guez 2014; van Horik et al. 2017). Here, motivation was extracted as the latency from trial start

Table 1 Summary of behavioral measures used in analysis

Term	Definition
Measures extracted once per trial	
1. Persistence	Ratio of time spent contacting the MAB to total trial time
2. Motivation	Latency from trial start to first contact with the MAB
3. Motor diversity	# of 5 behavior patterns used while contacting the box
4. Activity	Ratio of time spent contacting the box through biting or pawing to total time spent in contact
5. Efficiency	Latency from first contact to food retrieval; unsuccessful trials did not receive a work time score
6. Inhibition-M	Amount of time spent in contact with blocked solutions relative to total contact time
7. Work time[a]	Duration of time subjects spent in contact with the solution that they ultimately used to retrieve food from the MAB
Measures extracted once per subject	
8. Neophobia	Latency from trial start to food retrieval on first familiarization trial
9. Repeated Innovation	Score of 0–4 indicating the # of MAB solutions learned
10. Inhibition-C	Score of 0–10 with the inhibitory control cylinder representing the number of successful trials

Measures #1–6 were extracted from each individual trial resulting in multiple measures per subject. Measures #7–9 were taken once per subject

[a]Work time was only used to assess learning and was not compared to repeated innovation

to first contact. A high score here indicated a slow approach and low motivation, whereas a low score indicated a fast approach and high motivation. Although persistence can sometimes be evidence of motivation, measures of motivation do not always correlate with measures of persistence. Furthermore, unlike persistence, motivation is not always linked with problem-solving success (Griffin and Guez 2014). Because persistence and motivation appear to be distinct (though related) constructs, we chose to use measures of both persistence and motivation here. Because we measured persistence as a ratio of time spent in contact, and motivation as latency to first contact, our measures of motivation and persistence did not overlap, so any correlation might be due to persistence and motivation deriving from a shared latent factor (e.g., hunger) rather than confounding of measurement.

Motor diversity is the total number of unique motor patterns that a subject may exhibit in a problem-solving context (Griffin and Guez 2014; Diquelou et al. 2015). This has also been referred to as behavioral variety (Chow et al. 2016), exploration diversity (Benson-Amram et al. 2013), and behavioral diversity (Benson-Amram et al. 2016). Here, motor diversity was extracted by calculating a score of 0–5 for the number of different motor patterns used to contact the MAB. The 5 motor patterns were (a) sniffing or contacting the MAB with the snout (but not biting), (b) biting the MAB, (c) flipping the MAB, (d) using one or both paws to dig at the MAB, and (e) moving the MAB by pushing, dragging or carrying. A score of 5 indicated high motor diversity and a score of 0 indicated low motor diversity.

Activity, or the physical activity level of a subject (Sih and Del Giudice 2012; Brust et al. 2013), was measured in the context of exploration with the MAB. We extracted activity as the ratio of time spent using contact behaviors that involved the use of mouth or paw to the total time spent in contact (i.e. the ratio of motor patterns b–e to patterns a–e). Activity therefore referred to using behaviors, during contact with the MAB, that had the potential to lead to solutions. A high score indicated highly active exploration and a low score indicated more passive exploration of the MAB.

We used the term efficiency to mean the latency to problem-solving success in each trial (after Chow et al. 2016). Although this measure is often sometimes used as the dependent variable to measure success in many problem-solving studies (e.g. Chow et al. 2016), here, we analyzed it as an independent variable. Efficiency was extracted as the latency from first contact with the MAB to food retrieval (i.e. the amount of time it took to solve the MAB). Subjects who were not successful at retrieving food did not receive an efficiency score in that trial. A low score here indicated quick retrieval of the food and high efficiency, whereas a high score indicated slow retrieval and low efficiency.

Inhibitory control as measured in MAB trials (hereafter inhibition-M) was extracted as the amount of contact time spent on a blocked solutions relative to total contact time. A low inhibition-M score indicated good inhibitory control.

To further assess that these trial measures were independent, we examined the one-to-one correlation between all trial-level measures (Table 2). Only persistence and motivation were correlated at $p < 0.05$ (Spearman's $r = -0.51$), which was not unexpected and likely due to a shared latent factor (see above).

In addition to repeated innovation and inhibition-C, we extracted one more measure once per subject. Neophobia is typically described as aversion to novelty, normally measured as latency to approach novel objects or food (Greggor

Table 2 Correlation matrix of six raw per trial behavioral measures from Table 1

	1	2	3	4	5	6
1. Persistence	1.00	− 0.48*	0.12	− 0.16	− 0.11	− 0.13
2. Motivation		1.00	− 0.02	0.21	0.07	0.27
3. Motor diversity			1.00	− 0.01	0.47	0.24
4. Activity				1.00	− 0.14	0.11
5. Efficiency					1.00	0.18
6. Inhibition-M						1.00

*Indicates $p < 0.05$, spearman's rank correlation

et al. 2015). Here, neophobia was extracted on each subject's first familiarization trial as the latency from the start of the trial to food retrieval. This is distinct from motivation, which was measured on each test trial (but not on familiarization trials) as latency from the start of the trial to first contact, rather than food retrieval. Neophobia has previously been measured in spotted hyenas as latency to contact a novel object with or without bait (Benson-Amram et al. 2013; Greenberg and Holekamp 2017). Elsewhere, neophobia is often measured as latency to feed near a novel object (Greggor et al. 2015), and in the current study we measured neophobia as latency to feed from inside a novel object by measuring each subject's latency to feed from the MAB the first time it encountered the MAB during its initial familiarization trial (Table 1, #8). We thought this measure best assessed neophobia because most hesitancy our subjects displayed toward the MAB was in regard to putting their heads inside the MAB to feed. Here a high score indicated a slow approach toward the MAB and high neophobia, whereas a low score indicated a fast approach and low neophobia.

Several of these behavioral traits are often considered "personality traits" when they are repeatable across both time and context. Here, most of our measures were taken across multiple trials, which allowed us to test repeatability, but we only measured them in the context of the MAB. However, two traits, neophobia and inhibition-C, were only scored once. Ultimately, there were two categories of variables used in the final analysis: variables that were measured per trial while subjects interacted with the MAB (Table 1, #1–7), and variables measured once per subject (Table 1, #8–10).

Reliability

All videos were coded by LJU. A random subsample of 20% of videos were coded by an independent coder for reliabilities on durations and frequencies of raw behaviors ($r = 0.99$, Pearson's correlation coefficient). A separate random subsample of 20% of videos were coded by another independent coder for latency to first contact ($r = 0.98$, Pearson's correlation coefficient) and latency to food retrieval ($r = 0.99$, Pearson's correlation coefficient). Inter-observer reliability assessments were conducted on the raw durations, latencies, and frequencies of behaviors.

Statistical analysis

To validate our learning criterion, we used a linear mixed model (LMM) to examine how work time changed across trials. Work time was used as the dependent variable, and we included trial number for the solution used as a fixed effect and subject ID as a random effect in the model.

To examine the traits (#1–6, 8, and 10 in Table 1) influencing the number of MAB solutions learned, we used a generalized linear model (GLM) with a Poisson error and log link. To avoid over-parameterization of our model, we systematically pared down the number of predictor variables in four steps. First, we used *t* tests for unequal variance to check for any differences in the number of solutions learned based on zoo, sex, and age class. Where no differences existed, we pooled the data in our model for predicting the number of MAB solutions learned. Otherwise, we included them in our model to control for their effect.

Second, because number of MAB solutions learned was a subject-level variable, we converted the repeated trial-by-trial measures (#1–6, Table 1) to mean scores for each subject. Prior to obtaining means, to ensure that means were valid representations of each subject's behavior, we looked for individual differences and repeatability within individuals in our trial measures. To do so, we used a likelihood ratio test to compare linear mixed models (LMMs) with and without subject ID as a random effect. Each of these six (#1–6, Table 1) variables was used as a response variable in its own LMM, and as predictor variables we included age class, sex, AM or PM, zoo, phase, phase trial number, total trial number, and an interaction between phase and phase trial number to control for their effects. From the LMMs, we also calculated intraclass-correlation coefficients (ICCs) to examine repeatability. For those measures in which ID was both significant and repeatable, we took a mean score for each subject for eventual comparison to the number of solutions learned. Any measures in which ID was not significant in the LMM were excluded from the final model because this indicated to us that variation across trials within subjects

was larger than variation between subjects, and therefore that a mean score would not be a valid representation of an individual subject's performance.

Third, we checked for any correlations between independent measures to avoid multicollinearity when predicting the number of MAB solutions learned. As it is not recommended to include variables that correlate at higher than 0.70 (Tabachnick and Fidell 1996), where correlation coefficients exceeded 0.70, we used principal component analysis to extract a composite score, representing a syndrome of non-cognitive behaviors. Ultimately only two variables were included in our final model and we conducted no more paring down after this step. The "rule of 10" suggests regression models should not include more than one variable per ten subjects (Peduzzi et al. 1996). However, review of this rule with logistic models suggested that, although type II error increases with fewer than ten subjects per factor, type I error does not increase substantially. Therefore, null results must be interpreted cautiously, but significant results can be interpreted normally, albeit with a degree of caution regarding generality (Vittinghoff and McCulloch 2007). All statistics were run using R 3.2.2 GUI 1.66 Mavericks build (6996) (Bates et al. 2015; Kuznetsova 2017; Lüdecke 2018; R Core Team 2016; Wickham et al. 2017; Wickham 2009).

Results

Of the ten hyenas tested, eight opened the box at least once, six learned at least 1 solution and four learned all 4 solutions on the MAB (Table 3). In sum, hyenas participated in a total of 104 trials (mean trials total per subject ± SD = 10.40 ± 4.62, range = 4–15 trials per subject). The push solution consistently appeared to be the last solution learned, suggesting that it was harder than the other solutions. However, there was no consistent pattern to the order in which other solutions were used (Fig. 2). Inhibition-C scores varied from 3 to 9 successful trials of 10 total trials (Table 3). See Supplementary Materials Video 1–4 and Video 5 for videos of hyenas interacting with the MAB and the cylinder, respectively.

Learning

To assess learning we used a linear mixed model to examine how work time changed across trials. Work time significantly decreased across successful trials (LMM, $N = 84$, $p = 0.04$, Fig. 3). This showed that subjects became faster at solving the MAB with experience and suggested that our criterion of using the same solution on 3 out of 4 consecutive trials was sufficient for requiring that learning occurred.

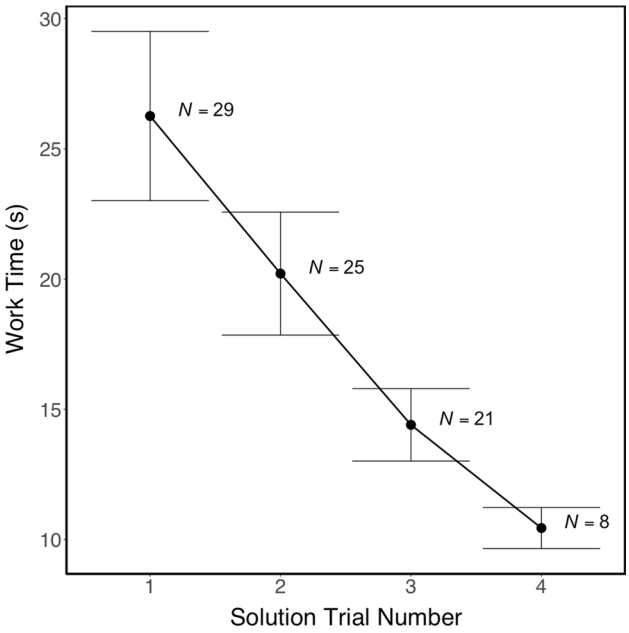

Fig. 3 Learning curve across trials. Work time (s) was calculated as the amount of time subjects spent in contact with the solution that was ultimately used to retrieve food. Solution trial number corresponds to the number of trials with that particular solution. Error bars indicate standard error. *N* varies across trials because subjects opened solutions variable numbers of times

Table 3 Scores for each subject on neophobia, inhibition-C, and # of solutions learned on the MAB

ID	MAB	Inhibition-C	Neophobia (s)	Age class	Sex	Institution
CASS (C)	1	5	13	Adult	Female	OCZC
DUSTY (D)	4	9	9	Adult	Male	OCZC
GIGGLES (G)	4	8	300	Adult	Female	OCZC
KELELE (KL)	4	8	6	Subadult	Male	Denver
KIBO (KB)	4	4	33	Adult	Male	Denver
NGOZI (NG)	2	8	14	Adult	Female	Denver
NIA (NI)	0	6	900	Subadult	Female	Denver
TAVI (T)	0	9	669	Subadult	Female	Denver
URSA (U)	0	5	17	Adult	Female	OCZC
WIBARI (W)	0	3	372	Adult	Male	OCZC

Zoo, age class, and sex differences

The number of solutions learned did not differ between institutions (two-sample t test assuming unequal variances $t(8) = -0.16$, $p = 0.88$), age classes (two-sample t test assuming unequal variances $t(3) = 0.54$, $p = 0.63$) or the sexes (two-sample t test assuming unequal variances $t(6) = -1.53$, $p = 0.18$). Therefore, data from all ten hyenas were pooled into a single analysis.

Repeatability of measures

Because we were interested in subject-level variation on each of our six trial measures for comparison to repeated innovation scores, prior to running any analyses we first checked for significant variation between subjects and repeatability within subjects on each measure using likelihood ratio tests ($N = 58$–104 trials). Of the 6 trial measures, only 3 showed significant variation among subjects across trials: persistence (Likelihood ratio test $\chi^2_{(1,N=104)} = 31.49$, $p < 0.001$, ICC = 0.60), motor diversity (Likelihood ratio test $\chi^2_{(1,N=104)} = 8.22$, $p < 0.01$, ICC = 0.21), and activity (Likelihood ratio test $\chi^2_{(1,N=104)} = 22.46$, $p < 0.001$, ICC = 0.40). Motivation (Likelihood ratio test $\chi^2_{(1,N=101)} = 0$, $p = 0.96$, ICC = 0.002), efficiency (Likelihood ratio test $\chi^2_{(1,N=84)} = 0$, $p = 1$, ICC = 0.00), and inhibition-M (Likelihood ratio test $\chi^2_{(1,N=58)} = 0.03$, $p = 0.87$, ICC = 0.00) did not show significant variation among subjects across trials (Fig. 4). This lack of significance, in addition to extremely low ICCs, indicates

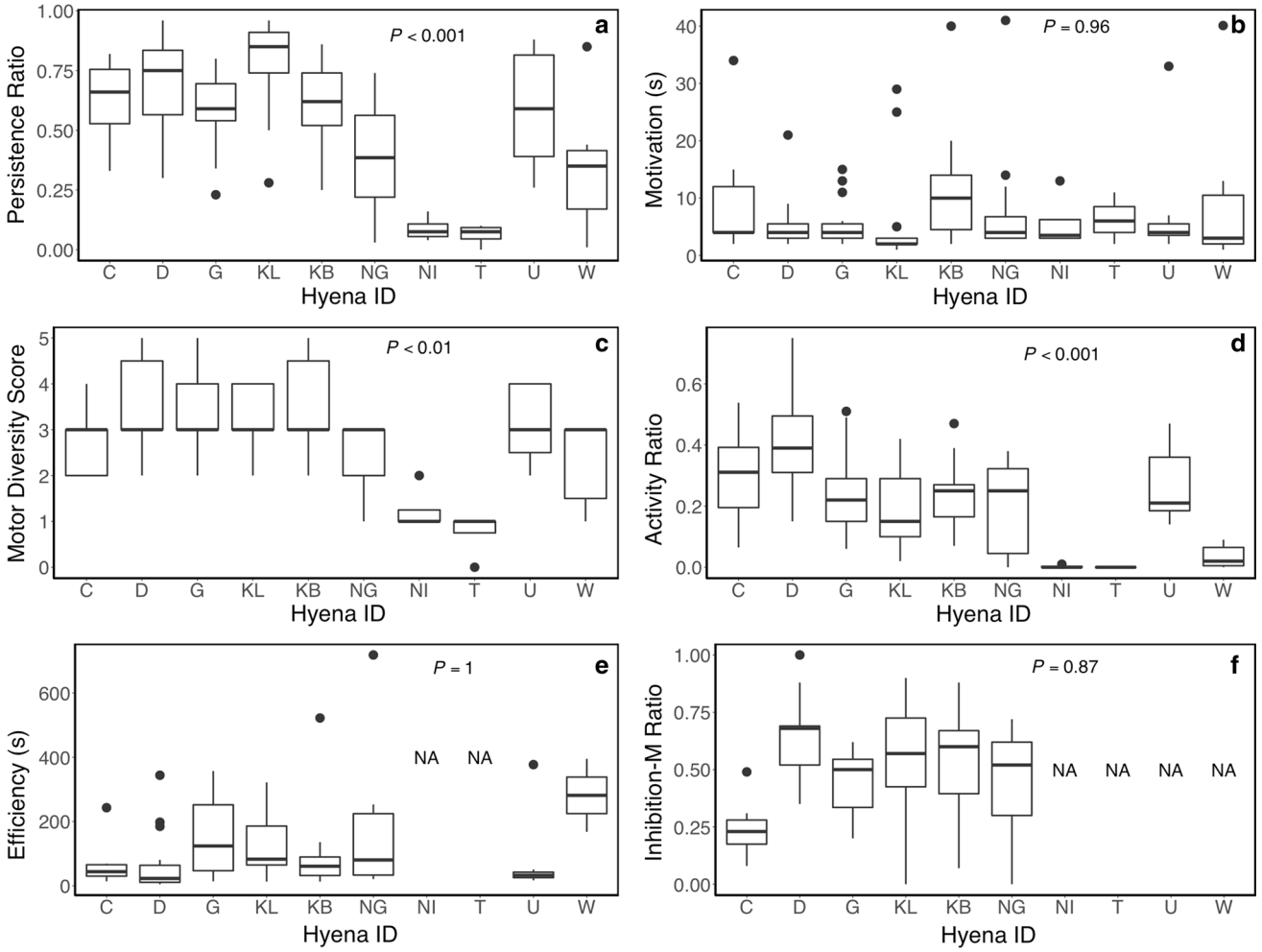

Fig. 4 Variation among subjects on the six trial measures (#1–6, Table 1). Box and whisker plots show mean (±SD) scores for measures #1–6 in Table 1. Black points represent outliers. p values are from likelihood ratio tests comparing models with and without subject ID while including control variables. Two subjects, NI and T, were not assigned scores (NAs) for work time because hyenas had to have at least one successful trial in order to achieve a work time score. Four subjects, NI, T, W, and U, were not assigned scores for inhibition-M because inhibition-M scores could only be assigned for trials in phases 2–4 where at least one solution was blocked

high variation and low repeatability within subjects on these variables, so they were excluded from further analysis. Mean scores for each subject on persistence, motor diversity, and activity were obtained for the next steps of analysis (Supplementary Table 1).

Independence of measures

Next, we checked for correlations between mean persistence, mean motor diversity, mean activity, neophobia and inhibition-C prior to including them in the model as predictors of the repeated innovation score. Mean persistence, mean motor diversity, mean activity, and neophobia, but not inhibition-C, were correlated at higher than 0.70 (Table 4). Therefore, we applied principal component analysis to mean persistence, mean motor diversity, mean activity, and neophobia scores. Both the Kaiser criterion (Kaiser 1960) and scree test (Cattell 1966) indicated that only the first component should be retained. The first component explained 90% of the variance and all four traits loaded equally onto it. Mean persistence, mean motor diversity, and mean activity all loaded moderately positively onto this component, whereas neophobia loaded moderately negatively onto it (Supplementary Table 2).

Factors predicting repeated innovation

We used each subject's score on the first component of our principal component analysis (hereafter proactivity scores) for inclusion in our final model in place of mean persistence, mean motor diversity, mean activity, or neophobia scores. We chose the term "proactivity" because high scores on the first component closely matched the proactive end of the commonly studied proactive–reactive axis of behavior (Sih et al. 2004). With only two predictor variables in our final model, proactivity scores and inhibition-C, we conducted no further model reduction. Using this model (Table 5), proactivity scores positively and significantly predicted repeated innovation scores (GLM, $N = 10$, $p = 0.02$, Fig. 5), whereas inhibition-C did not (GLM, $N = 10$, $p = 0.27$). Note that, with a sample size of 10, null results should be interpreted cautiously as the risk of Type II error is high.

We had intended to also look at a correlation between inhibition-C and inhibition-M, but we were unable to do so because inhibition-M was measured multiple times, whereas inhibition-C was only measured once. For comparison, we could have taken a mean inhibition-M score; however, we did not calculate mean inhibition-M scores due to low repeatability within subjects.

Table 4 Correlation matrix among subjects' scores on behavioral measures from Table 1

	1	3	4	8	10
1. Mean persistence	1.00	0.85**	0.81**	− 0.90**	0.02
3. Mean motor diversity		1.00	0.78**	− 0.67*	− 0.08
4. Mean activity			1.00	− 0.70*	− 0.08
8. Neophobia				1.00	− 0.20
10. Inhibition-C					1.00

*Indicates $p < 0.05$, **indicates $p < 0.01$, spearman's rank correlation. Table 4 differs from Table 2 in that it shows the correlations between measures at the subject level with means, rather than the trial level with raw scores

Table 5 Parameter estimates from the GLM

	Estimate	SE	95% CI	z	p
Intercept	− 0.74	0.91	–	− 0.81	0.42
Proactivity	0.61	0.26	0.10, 1.11	2.65	0.02*
Inhibition-C	0.14	0.13	− 0.11, 0.39	1.54	0.27

*Indicates $p < 0.05$

Discussion

We examined the effects of persistence, motivation, motor diversity, activity, efficiency, and two measures of inhibitory control on repeated innovation. We found no support for our hypothesis that inhibitory control would positively predict repeated innovation. Inhibition-C failed to predict repeated innovation, and inhibition-M did not vary significantly among hyenas. Opposite to our predictions, several of our non-cognitive traits strongly predicted repeated

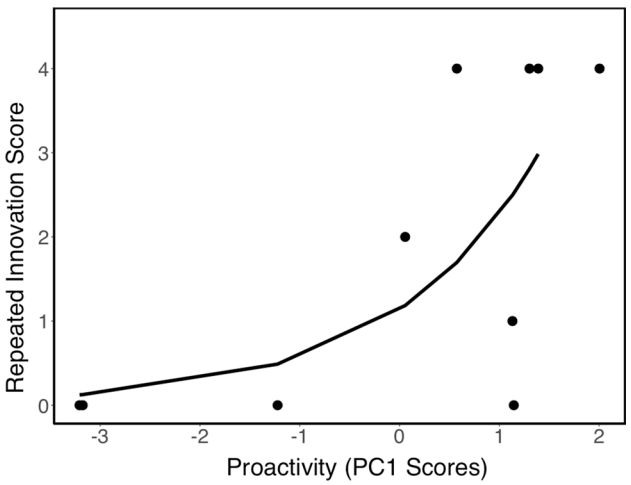

Fig. 5 Relationship between PC1 scores and the number of MAB solutions learned (repeated innovation score). Each point represents one subject. The black line is a Poisson GLM regression of the number of MAB solutions learned against PC1 scores. High PC1 scores represent high persistence, high motor diversity, high activity, and low neophobia

innovation through a composite proactivity score. High motor diversity, high persistence, high activity, and low neophobia were associated with higher repeated innovation scores. Not only had we predicted that most of these traits would have no relationship with repeated innovation, but we also predicted that activity would have a negative, not positive, relationship. Like inhibition-M, motivation and efficiency were not repeatable within subjects and therefore were not comparable to repeated innovation scores.

Inhibitory control and repeated innovation

The lack of a significant relationship between inhibitory control and repeated innovation was in contradiction to previous research on inhibitory control and problem-solving (Herrmann et al. 2007; Thornton and Samson 2012; Manrique et al. 2013; Müller et al. 2016; Burkart et al. 2017). Although we observed a great deal of variation in repeated innovation, inhibition-M was not repeatable within hyenas. This might suggest that inhibiting previous learning with the MAB did not affect the ability to learn a new solution. However, our sample size for inhibition-M was only 6 hyenas which may have been too small a sample in which to observe an effect. Inhibition-C was also not a significant predictor of MAB scores. The confidence interval for the estimated coefficient for inhibition-C overlapped zero which suggests the possibility that inhibition-C had no relationship with repeated innovation. On the other hand, with a sample size of only 10 subjects we also cannot rule out the high risk of type II error. Because of that risk, we also examined the effect size of the estimates for inhibition-C and proactivity. At the upper end of the confidence interval, an increase of 1 in inhibition-C scores corresponded with an increase of $e^{0.39} = 1.48$ in MAB scores, suggesting a potential relationship, but this effect size was quite small relative to the effect of proactivity in the GLM (Table 5). The few other studies that independently assessed individual's inhibitory control and problem-solving ability have also failed to find a simple positive relationship (Müller et al. 2016; Shaw 2017). Although we had also planned to examine the relationship between inhibition-M and inhibition-C in this study, because inhibition-M was not repeatable within hyenas, we were unable to correlate inhibition-M and inhibition-C. Overall, our results suggest that variation among individuals on either type of inhibitory control may not be related to variation in repeated innovation; at best, the effect of the relationship is small relative to effects of other behavioral traits. Our results highlight the need for further study into how individuals use inhibitory control and how that contributes to successful innovation and learning.

Non-cognitive factors influencing repeated innovation

In the current study, higher repeated innovation scores were predicted by a composite proactivity score that consisted of high persistence, high activity, high motor diversity, and low neophobia. Previous research with hyenas also found consistent individual differences in motor diversity and found that high motor diversity, high persistence and low neophobia predicted problem-solving success with a single-access puzzle box (Benson-Amram et al. 2013). Previous research with the MAB paradigm in keas and corvids produced results similar to those obtained in the current study. Birds that were less neophobic and more exploratory with the MAB were more likely to learn more solutions (Auersperg et al. 2011).

Not only were the results of the current study in contradiction to our predictions, but, persistence, activity, motor diversity, and neophobia were also both repeatable within subjects and strongly correlated with one another, suggesting that captive hyenas exhibit a "behavioral syndrome" within a problem-solving context. Sih et al. (2004) specifically define a behavioral syndrome as "a suite of correlated behaviors within a given behavioral context or across different contexts." The current behavioral syndrome appears to closely mirror a syndrome of personality on the proactive–reactive axis (Sih et al. 2004). Proactive individuals are less neophobic, more persistent, active and exploratory, whereas reactive individuals show the opposite pattern (Sih et al. 2004). Here, proactivity was positively associated with repeated innovation; hyenas that learned more MAB solutions were more persistent, more active during exploration, had higher motor diversity scores and had lower neophobia scores.

A growing body of literature links behavioral syndromes and personality traits to cognition (Carere and Locurto 2011; Sih and Del Giudice 2012; Griffin et al. 2015; Guillette et al. 2017). Current hypotheses that describe the adaptive function of proactive–reactive syndromes with regard to cognition do not suggest whether proactive or reactive individuals should be better at learning and problem-solving, but instead suggest that the proactive–reactive axis may reflect a trade-off between alternative learning styles in which proactive individuals are more innovative and faster at learning novel tasks whereas reactive individuals are slower and less innovative, but more flexible or accurate learners (Carere and Locurto 2011; Sih and Del Giudice 2012; Griffin et al. 2013; Ducatez et al. 2015). This trade-off is believed to have ecological relevance because it represents two alternative strategies for survival in the wild (Chittka et al. 2009; Sih and Del Giudice 2012). In the current study, no clear trade-off exists; proactive hyenas had higher repeated innovation scores and therefore also showed a high degree of flexibility by switching between four different solutions to the MAB. In

addition, proactive hyenas were not more efficient at opening the MAB than reactive hyenas who learned fewer solutions. Indeed, whether proactive or reactive individuals perform better or worse on a task may be heavily task-dependent or may not be apparent in captivity where the costs and benefits of one strategy over another are reduced or absent. Other studies have also shown inconsistencies in the relationship between proactivity-reactivity and cognitive trade-offs (Titulaer et al. 2012; Bousquet et al. 2015; Schuster et al. 2017). Of note, the proactive syndrome observed in our study was only measured in the context of the MAB which could account for the lack of a clear trade-off.

Conclusion

In sum, we found that a behavioral syndrome consisting of persistence, motor diversity, activity and neophobia predicted repeated innovation in a MAB paradigm with captive spotted hyenas. Contrary to our predictions, neither of our two measures of inhibitory control bore a clear relationship to repeated innovation scores. Although inhibitory control may be required for many problem-solving tasks, here the effect of inhibitory control was small or non-existent relative to the effect of proactivity scores. The finding that a proactive syndrome of correlated traits, including activity, motor diversity, persistence, and neophobia significantly predicted repeated innovation score was also contrary to our prediction that these traits would have either no effect or a negative effect. Overall, our results add support to the increasing body of literature that suggests that motor diversity, activity, persistence and neophobia have a strong relationship with innovative problem-solving and that these traits are often strongly interrelated. However, our results contradict hypotheses suggesting the direction of the relationship between proactive traits and innovation. Our results also shed some light on the murky relationship between behavioral traits, problem-solving, and cognition.

The relationship between behavioral traits and problem-solving is a burgeoning field (see Guillette et al. 2017), and though some consider these traits to be non-cognitive factors to be controlled when trying to measure variation in cognition (Rowe and Healy 2014; van Horik and Madden 2016), a relationship between behavioral traits and problem-solving does not necessarily negate cognitive explanations for problem-solving success (Griffin et al. 2015; Guillette et al. 2017). Here, repeated innovation required memory of previous solutions in order to reach learning criterion such that success could not be attributed solely to these non-cognitive behavioral traits. However, the strong influence of these non-cognitive behavioral traits does support the hypothesis that innovation and innovative problem-solving might not be exclusively cognitive constructs themselves. Although the proximate relationship between these traits and cognition remains unclear, ultimately, if an individual is highly innovative, it may be adaptive to perform behaviors that both encourage learning and increase opportunities for innovation, resulting in correlated evolution (e.g. Griffin 2016).

Acknowledgements We thank Courtney Frenchak at the Oak Creek Zoological Conservatory, Heather Genter and the other zookeepers at the Denver Zoo, and undergraduate research assistants Mike Kowalski and Paige Barnes for reliability coding.

Funding This work was supported by National Science Foundation Grants OISE 1556407, IOS 1755089, and DEB 1353110 to KEH, and by a National Science Foundation Graduate Research Fellowship to LJU. LJU was also supported by fellowships from the College of Natural Sciences, the Integrative Biology Department, and the program in Ecology, Evolutionary Biology, and Behavior at Michigan State University. This work was also supported in part by the BEACON Center for the Study of Evolution in Action, funded by National Science Foundation Grant OIA 0939454.

Compliance with ethical standards

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

Ethical approval The data collection procedure followed here was reviewed by Michigan State University Institutional Animal Care and Use Committee: AUF #04/16-050-00. All procedures were also reviewed and approved for each zoological institution's ethical guidelines.

Data availability The datasets generated and analyzed during the current study are available in the Knowledge Network for Biocomplexity (KNB) Repository, ID: knb.92181.2, (https://knb.ecoinformatics.org/#view/knb.92181.2).

References

Auersperg AMI, von Bayern AMP, Gajdon GK et al (2011) Flexibility in problem solving and tool use of kea and New Caledonian crows in a multi access box paradigm. PLoS ONE 6:e20231. https://doi.org/10.1371/journal.pone.0020231

Bates D, Maechler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67(1):1–48. https://doi.org/10.18637/jss.v067.i01

Benson-Amram S, Weldele ML, Holekamp KE (2013) A comparison of innovative problem–solving abilities between wild and captive spotted hyaenas, *Crocuta crocuta*. Anim Behav 85:349–356. https://doi.org/10.1016/j.anbehav.2012.11.003

Benson-Amram S, Dantzer B, Stricker G et al (2016) Brain size predicts problem–solving ability in mammalian carnivores. Proc Natl Acad Sci 113:2532–2537. https://doi.org/10.1073/pnas.1505913113

Borrego N, Dowling B (2016) Lions (*Panthera leo*) solve, learn, and remember a novel resource acquisition problem. Anim Cogn 19:1019–1025

Bousquet CAH, Petit O, Arrivé M et al (2015) Personality tests predict responses to a spatial-learning task in mallards, *Anas platyrhynchos*. Anim Behav 110:145–154

Bray EE, Sammel MD, Seyfarth RM et al (2017) Temperament and problem solving in a population of adolescent guide dogs. Anim Cogn 20(5):923–939. https://doi.org/10.1007/s10071-017-1112-8

Brucks D, Marshall-Pescini S, Wallis LJ et al (2017) Measures of dogs' inhibitory control abilities do not correlate across tasks. Front Psychol 8:849

Brust V, Wuerz Y, Krüger O (2013) Behavioural flexibility and personality in zebra finches. Ethology 119:559–569

Burkart JM, Schubiger MN, van Schaik CP (2017) The evolution of general intelligence. Behav Brain Sci 40:e195. https://doi.org/10.1017/S0140525X16000959

Carere C, Locurto C (2011) Interaction between animal personality and animal cognition. Curr Zool 57:491–498

Cattell RB (1966) The scree test for the number of factors. Multivar Behav Res 1:245–276

Chittka L, Skorupski P, Raine NE (2009) Speed–accuracy tradeoffs in animal decision making. Trends Ecol Evol 24:400–407

Chow PKY, Lea SEG, Leaver LA (2016) How practice makes perfect: the role of persistence, flexibility and learning in problem–solving efficiency. Anim Behav 112:273–283. https://doi.org/10.1016/j.anbehav.2015.11.014

Diquelou MC, Griffin AS, Sol D (2015) The role of motor diversity in foraging innovations: a cross-species comparison in urban birds. Behav Ecol 27:584–591

Ducatez S, Audet JN, Lefebvre L (2015) Problem–solving and learning in Carib grackles: individuals show a consistent speed–accuracy trade-off. Anim Cogn 18:485–496

Friard O, Gamba M (2016) BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. Methods Ecol Evol 7:1325–1330

Gosling SD, Hawk JE, Beer JS et al (1998) Personality dimensions in spotted hyenas (*Crocuta crocuta*). J Comp Psychol 112:107–118

Greenberg JR, Holekamp KE (2017) Human disturbance affects personality development in a wild carnivore. Anim Behav 132:303–312

Greggor AL, Thornton A, Clayton NS (2015) Neophobia is not only avoidance: improving neophobia tests by combining cognition and ecology. Curr Opin Behav Sci 6:82–89. https://doi.org/10.1016/j.cobeha.2015.10.007

Griffin AS (2016) Innovativeness as an emergent property: a new alignment of comparative and experimental research on animal innovation. Philos Trans R Soc Lond B Biol Sci 371:20150544. https://doi.org/10.1098/rstb.2015.0544

Griffin AS, Guez D (2014) Innovation and problem solving: a review of common mechanisms. Behav Process 109:121–134

Griffin AS, Guez D, Lermite F, Patience M (2013) Tracking changing environments: innovators are fast, but not flexible learners. PLoS ONE 8:e84907

Griffin AS, Guillette LM, Healy SD (2015) Cognition and personality: an analysis of an emerging field. Trends Ecol Evol 30:207–214. https://doi.org/10.1016/j.tree.2015.01.012

Guillette LM, Naguib M, Griffin AS (2017) Individual differences in cognition and personality. Behav Process 134:1. https://doi.org/10.1016/j.beproc.2016.12.001

Herrmann E, Call J, Hernàndez-Lloreda MV et al (2007) Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. Science 317:1360–1366

Holekamp KE, Sakai S, Lundrigan B (2007) The spotted hyena (*Crocuta crocuta*) as a model system for study of the evolution of intelligence. J Mammal 88:545–554

Huebner F, Fichtel C (2015) Innovation and behavioral flexibility in wild redfronted lemurs (*Eulemur rufifrons*). Anim Cogn 18:777–787. https://doi.org/10.1007/s10071-015-0844-6

Kabadayi C, Bobrowicz K, Osvath M (2018) The detour paradigm in animal cognition. Anim Cogn 21:21–35. https://doi.org/10.1007/s10071-017-1152-0

Kaiser HF (1960) The application of electronic computers to factor analysis. Educ Psychol Meas 20:141–151

Lefebvre L, Reader SM, Sol D (2004) Brains, innovations and evolution in birds and primates. Brain Behav Evol 63:233–246. https://doi.org/10.1159/000076784

Lüdecke D (2018) Sjstats: statistical functions for regression models. R package version 0.14.0. https://cran.r-project.org/package=sjstats

MacLean EL, Hare B, Nunn CL et al (2014) The evolution of self-control. Proc Natl Acad Sci U S A 111:E2140–E2148. https://doi.org/10.1073/pnas.1323533111

Manrique HM, Völter CJ, Call J (2013) Repeated innovation in great apes. Anim Behav 85:195–202. https://doi.org/10.1016/j.anbehav.2012.10.026

Mischel W, Shoda Y, Rodriguez M (1989) Delay of gratification in children. Science 244:933–938. https://doi.org/10.1126/science.2658056

Müller CA, Riemer S, Virányi Z et al (2016) Inhibitory control, but not prolonged object-related experience appears to affect physical problem–solving performance of pet dogs. PLoS ONE 11:e0147753

Peduzzi P, Concato J, Kemper E et al (1996) A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 49:1373–1379

R Core Team (2016) R: A language and environment for statistical computing. R foundation for statistical computing, Vienna. https://www.R-project.org/

Reader SM, Morand-Ferron J, Flynn E (2016) Animal and human innovation: novel problems and novel solutions. Philos Trans R Soc B Biol Sci 371:20150182. https://doi.org/10.1098/rstb.2015.0182

Rowe C, Healy SD (2014) Measuring variation in cognition. Behav Ecol 25:1287–1292. https://doi.org/10.1093/beheco/aru090

Schuster AC, Zimmermann U, Hauer C, Foerster K (2017) A behavioural syndrome, but less evidence for a relationship with cognitive traits in a spatial orientation context. Front Zool 14:19. https://doi.org/10.1186/s12983-017-0204-2

Shaw RC (2017) Testing cognition in the wild: factors affecting performance and individual consistency in two measures of avian cognition. Behav Process 134:31–36. https://doi.org/10.1016/j.beproc.2016.06.004

Sih A, Del Giudice M (2012) Linking behavioural syndromes and cognition: a behavioural ecology perspective. Philos Trans R Soc B Biol Sci 367:2762–2772. https://doi.org/10.1098/rstb.2012.0216

Sih A, Bell AM, Johnson JC, Ziemba RE (2004) Behavioral syndromes: an integrative overview. Q Rev Biol 79:241–277

Tabachnick BG, Fidell LS (1996) Analysis of covariance. Using MultivarStat 8:321–374

Taylor AH, Hunt GR, Medina FS, Gray RD (2009) Do New Caledonian crows solve physical problems through causal reasoning? Proc R Soc London B Biol Sci 276:247–254

Thornton A, Samson J (2012) Innovative problem solving in wild meerkats. Anim Behav 83:1459–1468

Titulaer M, van Oers K, Naguib M (2012) Personality affects learning performance in difficult tasks in a sex-dependent way. Anim Behav 83:723–730

van Horik JO, Madden JR (2016) A problem with problem solving: motivational traits, but not cognition, predict success on novel operant foraging tasks. Anim Behav 114:189–198. https://doi.org/10.1016/j.anbehav.2016.02.006

van Horik JO, Langley EJG, Whiteside MA, Madden JR (2017) Differential participation in cognitive tests is driven by personality, sex, body condition and experience. Behav Process 134:22–30. https://doi.org/10.1016/j.beproc.2016.07.001

Vittinghoff E, McCulloch CE (2007) Relaxing the rule of ten events per variable in logistic and cox regression. Am J Epidemiol 165:710–718

Wickham H (2009) ggplot2: elegant graphics for data analysis. Springer, New York

Wickham H, Francois R, Henry L, Müller K (2017) dplyr: a grammar of data manipulation. R package version 0.7.4. https://CRAN.R-project.org/package=dplyr

Yoshida KCS, Van Meter PE, Holekamp KE (2016) Variation among free-living spotted hyenas in three personality traits. Behaviour 153:1665–1722