Problem-Solving and Spontaneous Tool-Using Ability in European Brown Bears (*Ursus arctos arctos*)

Helen R. Chambers[*] and Sean J. O'Hara

University of Salford, School of Science, Engineering & Environment

*Corresponding author (Email: h.r.chambers@edu.salford.ac.uk)

Abstract – Little is known about the cognitive abilities of bears, despite possessing relatively large brains for their body size. One reason this group is perhaps overlooked is because they are thought to fail to conform to the "social brain" hypothesis, by being relatively solitary species but still possessing large brains. Here, to better understand the proposed benefits afforded by encephalization, the cognitive abilities of 17 captive European brown bears *Ursus arctos arctos* were tested. The aim was to determine whether bears possess problem-solving and object-manipulation abilities. Two experimental tests – a puzzle box and an object-manipulation set-up – were presented to bears at seven UK zoological parks. Generalised linear mixed models were used to determine which variables, specifically age, sex, motivational levels, behavioral diversity and persistence, influence cognitive performance. Results revealed evidence of trial-and-error learning; however, two juveniles appeared to acquire a latch association, suggesting some individuals have potential to adopt successful strategies and draw perceptive associations. Individual variation in motivation levels appears to be an important factor influencing cognitive performance. Overall, the bears failed to spontaneously use a tool but still managed to retrieve the food reward, instead using alternative techniques to solve the problem. Analyses revealed both age and sex to be negatively associated with time-to-solve in our sample, indicating the younger male bears solved the task more quickly. Results suggest social dynamics of group-living bears to be influencing cognitive performance, as the collective nature of testing resulted in increased competition over a high-value reward. These results are discussed herein. Brown bears are confirmed to be an excellent model species for testing the cognitive abilities of Ursids, as well as theories of cognitive evolution.

Keywords – Bears, Brain size, Problem-solving, Object-manipulation, Intelligence, Tool use

---

Encephalization (Jerison, 1985) is thought to afford cognitive advantages (Boddy et al., 2012; Sol, 2009; Weisbecker et al., 2015), such as behavioral flexibility in primates (Amici et al., 2018), innovation in mammalian carnivores (Benson-Amram et al., 2016), learning and predator evasion in fish (Kotrschal et al., 2013, 2015), invasion success in amphibians, reptiles and birds (Amiel et al., 2011; Sol et al., 2005), and self-control across a broad range of taxa (MacLean et al., 2014). This is thought to ultimately aid in increased fitness since individuals who can adjust their behavior in a novel or challenging situation, are less likely to experience decreases in fitness (Sol, 2009). Accordingly, research attention is now focused on the evolution of problem-solving abilities and whether species can be innovative under socioecological challenges (Benson-Amram et al., 2016; Benson-Amram et al., 2013; Borrego & Gaines, 2016; Drea & Carter, 2009; von Bayern et al., 2009; Wat et al., 2020).

Much of the cognitive literature investigating enhanced cognitive abilities has taxonomically favored primates, birds, and domestic dogs (Boesch, 2012; Emery & Clayton, 2004; Horschler et al., 2019; Kubinyi et al., 2007; Miklósi et al., 2004; Seed et al., 2009; Seed & Tomasello, 2010); however, carnivores

are now receiving more attention (see e.g., Benson-Amram et al., 2016, 2023; Borrego & Gaines, 2016; Daniels et al., 2019; Holekamp & Benson-Amram, 2017). Attention has primarily been focused on the social species of the order Carnivora, such as spotted hyenas *Crocuta crocuta* (Benson-Amram & Holekamp, 2012) and big cats (see Borrego, 2017), with one understudied family within the order being Ursids. Bears have unexpectedly large relative brain sizes, showing similar encephalization increases to Canidae (Finarelli & Flynn, 2009), despite living minimally social lives (Gittleman, 1999). This group has perhaps been overlooked as they fail to conform to the "social brain" hypothesis (Dunbar, 1998), by not being social-living species but still possessing encephalized brains. Testing cognition in bears has been limited (as highlighted by Vonk & Beran, 2012), meaning little is known about whether they too possess those cognitive abilities seen in large-brained social species. Moreover, it is not known why encephalized brains have evolved in these taxa nor their function or what specific socioecological challenge has spurred increased brain size. For instance, black bears *Ursus americanus* and brown bears *U. arctos* are classified as generalists, having high levels of foraging flexibility (Gittleman, 1986). Their flexible diets are perhaps important in driving increased brain size within this group (Chambers et al., 2021), as has found to be the case in primate species (Chambers et al., 2021; DeCasien et al., 2017; Powell et al., 2017).

While relatively limited, research examining the capabilities of bears can be found, with American black bears being the focal taxa of many studies. Particular attention has been paid to the visual, spatial and numerical abilities of bears (Bacon & Burghardt, 1976a, b, 1983; Dungl et al., 2008; Kelling et al., 2006; Perdue et al., 2009; Perdue et al., 2011; Tarou, 2004; Vonk & Beran, 2012; Vonk et al., 2012; Vonk & Leete, 2017). Reports indicate that bears may be capable of tool use (Deecke, 2012), a behavior thought to be indicative of higher cognitive function (Emery & Clayton, 2009; Seed & Byrne, 2010). Beck et al. (2012, p. 3330) defined tool-use as "*the external employment of an unattached or manipulable attached environmental object; to purposively alter the form, position, or condition of another object, another organism, or the user itself...*". Following this definition, six captive bears were found to be capable of tool use, manipulating inanimate objects so as to obtain a food reward (Waroff et al., 2017). However, sloth bears *Melursus ursinus* have failed in a similar scenario (Amici et al., 2019). Benson-Amram et al. (2016) used puzzle boxes to test mammalian carnivores' ability to problem solve, finding species in the family Ursidae, including grizzly bears *U. arctos horribilis*, polar bears *U. maritimus*, and American black bears, to be most successful at solving the problem and accessing the puzzle box. Such research highlights the cognitive potential of bears generally; however, it is clear further research is needed to examine *brown* bear cognitive capabilities.

The tendency for bear studies to focus on American black bears (see e.g., Johnson-Ulrich et al., 2016; Myers & Young, 2018; Vonk et al., 2012; Zamisch & Vonk, 2012) is perhaps due to both their prevalence and accessibility in the United States of America, meaning brown bears have seldom been evaluated. This is surprising from the standpoint that brown bears offer an excellent model system for testing the cognitive abilities of Ursids. They are numerous and the most widespread bear species found across Europe, Asia, and North America, occupying a diversity of habitats (Belant et al., 2010; Hilderbrand et al., 2018; Servheen et al., 1999). Furthermore, they exhibit variable foraging strategies within populations, with diets ranging from highly mixed (i.e., meat and vegetation) to exceedingly specialised ones (Costello et al., 2016; Lafferty et al., 2015; Mangipane et al., 2018, 2020). Such factors are thought to demonstrate how brown bears come to display high levels of behavioral plasticity (Van Daele et al., 2012) or behavioral diversity, especially in terms of foraging behavior, and makes them excellent candidates for measuring cognitive ability.

In efforts to further substantiate the assertion that an encephalized brain results in enhanced cognitive abilities, and to further elucidate whether studying innovative ability is a useful tool for testing such assumptions, the cognitive abilities of captive European brown bears *U. arctos arctos* are tested. Albeit seldom tested, research is starting to surface suggesting brown bears show enhanced cognitive abilities (see e.g., Benson-Amram et al., 2016) and studying animals' ability to be innovative and solve unique socioecological problems has become prevalent in animal cognition studies (Bandini & Harrison, 2020). Here, two cognitive trials are implemented to test the cognitive abilities of brown bears. These trials have been previously implemented (see e.g., Benson-Amram et al., 2016; Waroff et al., 2017) and proven useful

in testing cognitive ability. The first hypothesis investigated here is that: *brown bears possess problem-solving abilities*, with the testable predictions being that: *brown bears are able to gain access to a baited puzzle box using their problem-solving abilities* and *brown bears are able to use the latch to gain access to the puzzle box*. The hypothesis related to the second task is that: *brown bears possess object-manipulation abilities*, with the testable predictions being that: *brown bears are able to manipulate an object leading to facilitate the retrieval of an out of reach food reward* and *brown bears are capable of tool-use*. As captive bears have previously proven successful at similar tasks, we predict the bears should be able to solve both tasks, confirming their problem-solving and object-manipulation abilities. We also aimed to examine the influence of age, sex, motivation levels, behavioral diversity and persistence on time-to-solve scores, to better understand the variables influencing bears' cognitive competence.

Methods

Ethics Statement

This research received ethical approval from the University of Salford Research Ethics Committee (STR1819-64). The application was guided by the guidelines for the treatment of animals in behavioral research and teaching of the Association for the Study of Animal Behaviour (ASAB) (Buchanan et al., 2012). This research was granted a letter of support from the British and Irish Association of Zoos and Aquariums (BIAZA) Research Committee.

Study Sites and Subjects

Experimental trials were implemented at seven zoological parks in the United Kingdom (UK): Five Sisters Zoo, Wildwood Trust, Camperdown Wildlife Centre, Wildwood Escot, Welsh Mountain Zoo, Scottish Deer Centre and Port Lympne Reserve. Seventeen captive European brown bears were included in this study: eight adult females, one juvenile female, five adult males, and three juvenile males ($N = 17$). Ages ranged from 1- 23 yrs (mean = 10.35, ± 7.75). Weight ranged from 80 to 300 kg. Fourteen of the bears were captive born, and three were thought to be wild caught. None of the bears had previous experience with cognitive trials. However, all but two bears had previous experience with artificial enrichment, such as boomer balls (Table 1 and see supplementary materials for demographic information on the bears).

Experimental Apparatus

Trials utilized (i) a puzzle box, and (ii) an object manipulation set-up. The puzzle box was a small (30 cm x 30 cm x 30 cm) baited steel box (Figure 1), which had a simple latch on the front that required the bears to slide laterally for the door to open, allowing access the food reward inside. It was designed similarly to those used in previous studies that have proven useful in testing mammalian problem-solving ability (see Benson-Amram et al., 2013, 2016; Benson-Amram & Holekamp, 2012; Borrego & Gaines, 2016); however, it is worth noting, the latch was increased in size to accommodate the size and weight of the bears, which allowed the bears to move the latch using their mouth, nose or paws. The barred box design meant the bears could both see and smell the food reward inside. This task draws on the bear's ability to manipulate small objects, similar to foraging challenges bears typically face, such as retrieving hard to reach berries or extracting honey from bee nests.

Table 1

Summary of Information on the Bears Involved in the Study

Location	Bear (sex, age)	Relation	Number of bears in group	Cognitive testing experience?	Enrichment experience?	Origin
Five Sisters Zoo	Eso (F,7)	N/A	1	No	Yes	W
Wildwood Trust	Scruff (M,23)	Not related	2	No	No	C
	Fluff (M,23)	Not related	2	No	No	C
Camperdown	Brumm (M,6)	Siblings	3	No	Yes	C
	Maja (F,6)	Siblings	3	No	Yes	C
	Brumma (F,6)	Siblings	3	No	Yes	C
Wildwood Escot	Mish (M,2)	Siblings	2	No	Yes	W
	Lucy (F,2)	Siblings	2	No	Yes	W
Welsh Mountain Zoo	Athena (F,17)	Siblings	2	No	Yes	C
	Fivi (F,17)	Siblings	2	No	Yes	C
Scottish Deer Centre	Loki (M,9)	Offspring	2	No	Yes	C
	Nelly (F,17)	Dam	2	No	Yes	C
Port Lympne	Enciam (F,19)	Dam	5	No	Yes	C
	Julio (M,19)	Sire	5	No	Yes	C
	Neu (F,5)	Offspring + Siblings	5	No	Yes	C
	Rojo (M,1)	Offspring + Siblings	5	No	Yes	C
	Tornillo (M,1)	Offspring + Siblings	5	No	Yes	C

Note. For sex, F = female, M = male. For origin, C = captive born, W = wild caught.

The individual was considered "successful" in solving the problem if they gained access to the box using any technique; but a further distinction was made if they used the latch to open the door. The box was chained during the trials to reduce the likelihood of the box being pushed around and the latch inadvertently falling open. The ability to chain the box varied at the collections, which resulted in the trial data being split into two categories (i) box well chained with limited maneuverability, and (ii) box loosely chained with high maneuverability. The main distinction between the two categories was that the box could more easily be 'tipped' when loosely chained, potentially resulting in the latch sliding open, which was not typically possible with the low maneuverability trials. While variability in how tightly the box could be chained limits the comparability of trials conducted across collections, by categorizing them, this allowed us to examine those two categories to help better understand the presence of latch-use.

The object-manipulation set-up involved suspending a food reward out of reach and providing tree stumps for the bears to manipulate so as to retrieve the reward. This typically involved securing a rope between two adjacent trees, with a food item hung in the middle and stumps provided underneath (Figure 2). At certain zoological parks, it was logistically easier to run trials in indoor enclosures with food suspended from roof meshing. While practical considerations led to some inter-collection set-up variability, trials remained conceptually identical. Between one to three stumps were provided for the bears depending on age, weight and size. This task ultimately required the individual to work out that the food item could be accessed by repositioning the stump under the suspended item so as to reach it (Figure 3) – a behavior that could be considered tool-use (Waroff et al., 2017).

Figure 1

Bear Interacting with the Puzzle Box. Photo credit: Page, 2021

Figure 2

Object-Manipulation Set-Up: A Rope Secured Between Two Trees with a Food Item Hung in the Middle and Tree Stumps Provided Underneath. Photo credit: Chambers, 2021

Figure 3

Bear Succeeding in Acquiring the Food Reward Using the Stump in Stage 1 of the Object-Manipulation Task. Photo credit: Chambers, 2021

Variations of this set-up have been previously deployed in captivity (Amici et al., 2019; Waroff et al., 2017). Previous researchers have trained or provided cues for the bears; therefore, this task was divided into stages. Initially, the tree stump(s) were placed directly underneath the food reward, so the individual had only to approach and stand on the stump(s) so as to be successful. If successful, in the next trial, the stump(s) were placed on their side so the individual had to push flat the stump(s) and stand on it/them to be successful. If successful, in the next trial, the stump(s) were positioned flat, but they were positioned away from the food reward, so the individual had to actively manipulate and maneuver the stump(s) to reach the food reward and be successful. Stages one and two are illustrated in Figure 4; stage 2 was reached but not successfully passed and therefore stage three was not presented. The full set-up is further illustrated in Figure 5. The set-up presented only changed if stump use and manipulation was recorded. The individual was considered "successful" in solving the problem if they managed to retrieve the hung food reward using any technique; however, a further distinction was made if they used the stump(s) for elevation.
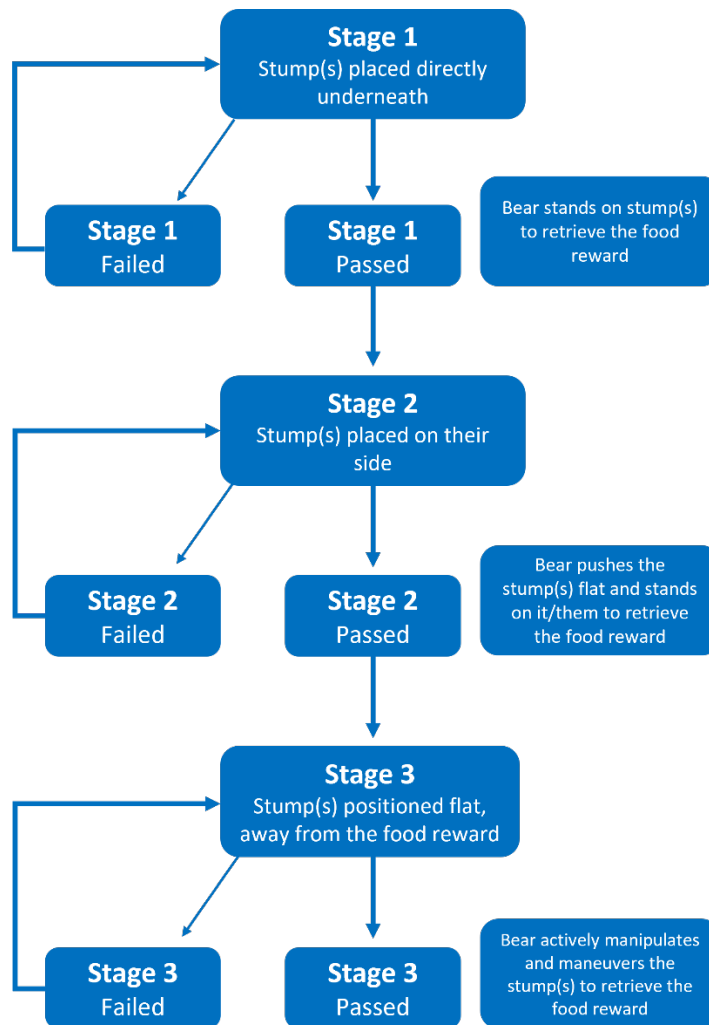
Figure 4

Object-Manipulation Stages One and Two. Photo credit: Chambers, 2021

Note. In stage one, tree stumps were placed beneath the food reward (left) and in stage two, stumps beneath the reward but first required maneuvering to lay flat/stabilise (right)

Figure 5

Object-Manipulation Stages and Criterion

```
                        ┌─────────────────────────┐
                        │        Stage 1          │
            ┌──────────▶│ Stump(s) placed directly │
            │           │      underneath          │
            │           └─────────────────────────┘
            │              │              │
            │              ▼              ▼         ┌──────────────────────┐
      ┌──────────┐   ┌──────────┐                  │ Bear stands on stump(s) │
      │ Stage 1  │   │ Stage 1  │                   │ to retrieve the food    │
      │ Failed   │   │ Passed   │                   │      reward             │
      └──────────┘   └──────────┘                  └──────────────────────┘
                           │
                           ▼
                        ┌─────────────────────────┐
                        │        Stage 2          │
            ┌──────────▶│ Stump(s) placed on their │
            │           │         side            │
            │           └─────────────────────────┘
            │              │              │
            │              ▼              ▼         ┌──────────────────────┐
      ┌──────────┐   ┌──────────┐                  │ Bear pushes the          │
      │ Stage 2  │   │ Stage 2  │                   │ stump(s) flat and stands │
      │ Failed   │   │ Passed   │                   │ on it/them to retrieve   │
      └──────────┘   └──────────┘                  │ the food reward          │
                           │                        └──────────────────────┘
                           ▼
                        ┌─────────────────────────┐
                        │        Stage 3          │
            ┌──────────▶│ Stump(s) positioned flat,│
            │           │  away from the food reward│
            │           └─────────────────────────┘
            │              │              │
            │              ▼              ▼         ┌──────────────────────┐
      ┌──────────┐   ┌──────────┐                  │ Bear actively manipulates│
      │ Stage 3  │   │ Stage 3  │                   │ and maneuvers the        │
      │ Failed   │   │ Passed   │                   │ stump(s) to retrieve the │
      └──────────┘   └──────────┘                  │ food reward              │
                                                    └──────────────────────┘
```

Experimental Procedure

Trials were implemented between June and October 2021. Separation of group-housed bears was rarely practical. This was of particular concern to many keepers who referenced that separation would likely lead to increases in behavioral anxiety indicators or stereotypical behaviors in the bears. As a result of this, trials were run with the bears collectively, meaning all bears had access to, and could interact with, the set-ups presented. Each bear was presented with at least three trials of each set-up; however, this was dependent on which bear engaged with the trials and whether one bear monopolised the trials, as there was no way to guarantee which bear was going to interact with the trial. Whilst this ultimately resulted in not all bears having the same exposure time with the trials, this was unavoidable and points to how the social dynamics of group-housed bears may influence engagement with both cognitive trials and behavioral enrichment activities.

Two to four trials were run per day, typically one in the morning and one in the afternoon, meaning trials typically coincided with feeding times. It is possible running trials in this way influenced bear motivation levels, as trials run in the morning were presented to bears who had not yet been fed, whereas bears had already been fed, having had access to food prior to the afternoon trials. This is of interest as

problem-solving abilities are often found to result from motivational differences rather than complex cognitive processes (van Horik & Madden, 2016). To ensure this was accounted for, it was noted how many bears had access to the set-up during each trial, whether they had already been fed and whether food was present at the same time. In terms of the food chosen to use as bait, preferences tests were not conducted, instead the choice was informed by keeper discussions, previously known bear preferences, and food availability. This usually was a 'high value' food item, such as monkey nuts, melon, or quail; however, there were a few occasions when the food item used was not as high value, such as corn, which could have influenced engagement and persistence. Keepers also acknowledged bear preferences change periodically (and seasonally) and thus, sometimes a presumed high value item was not well received. Such instances were few and most trials were baited with a high value item.

Trials commenced when the bear(s) had a direct line of sight to the set-up and trials ran for up to 30 minutes in duration or until the food reward had been retrieved. This varied slightly for trials run from September onwards, which were extended to last up to one hour in duration, to account for bears tested at this time of year having started to slow down for torpor, and not being as highly motivated to engage with the trials.

In total, trials involved 16 captive brown bears (nine females and seven males). One male bear did not engage with any of the trials. Similarly, one female only briefly engaged with one trial and motivation was very low. A second female failed to engage with the puzzle-box but engaged with the object-manipulation trials. Only bears that engaged with one, or both, of the trials were considered for analyses.

Data Extraction from Videotaped Trials

All trials were video recorded and behavioral data were extracted from the video recordings. For each individual their best three trials were scored for each set-up. Video recordings were scored by one observer. To validate that the footage was scored accurately, a randomly selected 25% of trial recordings were reviewed by an independent observer, blind to the hypotheses. Interobserver reliability was very high across all measures (Latency $R = .996$; Time-to-solve $R = .991$; Behavioral diversity $R = .90$; Persistence $R = .978$; Latch use/Stump use $R = 1.00$, Successful $R = 1.00$; using Spearman rank correlation for all measures).

In terms of measuring problem-solving ability for both set-ups, performance measures were used. Latency to approach (t^1) was recorded as the time taken (secs) to approach the set-up after first detecting it, as a measure of motivation to obtain the food reward. To also measure motivation, the following scoring system was additionally implemented: low (L), medium (M) and high (H). This was scored based on the time of day of the trial, whether the subject had already been fed, the number of bears with access, the availability of food alongside the trial and the overall activity levels of the bear prior to the trials. Time-to-solve (t^2) was recorded as the time taken (secs) to solve the test, after approaching the set-up, during which the individual was oriented on the set-up and focused on solving the task, until successful. If unsuccessful, this was scored N/A and excluded from further analyses. To score the range of behaviors seen during attempts, the same behavioral diversity score was used as Benson-Amram et al. (2016), including 12 different behaviors: *rub, foot on box, sniff, lick, dig, bite, pull box with mouth, push box with head, push box with paw, pull box with paw, stand on box, and tip box*. 'Flip box' was excluded as this was not possible and this was instead replaced with '*claw*', used to represent the behavior during which the bear used its front paw to '*claw*' at the box, either in an attempt to open the box or to pull the food item through the bars. Thus, each individual received a score from 0 to 13. This scoring system was used only for the puzzle box trials. For both tests it was noted whether the individual was successful (Yes or No); however, a further distinction was made in the puzzle box trials, in terms of the presence of latch use (Yes or No), and a further distinction was made in the object-manipulation trials, in terms of the presence of stump use (Yes or No). Number of attempts, with an attempt being any behavior used to try and retrieve the food reward, was used as a measure of persistence.

Statistical Analyses

Interobserver reliability analyses and chi-square tests were conducted in Minitab 21.1.0. (Minitab LLC, 2021). All further analyses were conducted in *R* 4.1.2 (R Core Team, 2021), using the packages 'lme4' (Bates et al., 2015), 'lmerTest' (Kuznetsova et al., 2017), 'car' (Fox & Weisberg, 2019), 'MuMIn' (Bartoń, 2010), 'r2glmm' (Byron, 2017) and several functions provided by Roger Mundry. To examine the presence of latch-use and determine whether the two puzzle box categories (i) box well chained and (ii) box loosely chained, influenced the presence of latch-use, a chi-square test was conducted. Generalised linear mixed models (GLMM) were implemented to estimate the effects of age, sex, behavioral diversity, persistence, number of successful trials, trial number, latency to approach, motivation score and the number of bears present on time-to-solve (Baayen, 2008). Five models were run, models one to three using the puzzle box data and models four and five using the object-manipulation data. In model one, behavioral diversity, persistence and their interaction were included as fixed effects. The interaction between behavioral diversity and persistence was included because it seemed there was likely to be a relationship between both the *number* and *types* of attempts. In model two, number of successful trials and trial number were included as fixed effects. The interaction between number of successful trials and trial number was included because it seemed possible for there to be a relationship between the number of *successful* trials and the trial *number*. In model three, latency to approach and motivation score were included as fixed effects. To avoid model convergence issues, the interaction between these two variables was not considered within this model. In model four, age, sex, and their interaction were included as fixed effects. The interaction between age and sex was included because we reasoned it be conceivable that they may interconnect. In model five, the number of bears present and the trial number were included as fixed effects. It seemed unlikely that these two variables would be interconnected and therefore the interaction between these two variables was not considered in this model. Time-to-solve (the task) was the response variable in the five models. In all models, subject was included as the as random effect, to control for the inclusion of multiple datapoints from one individual.

A series of preliminary models were implemented that included various combinations of the variables of interest to determine which variables best explained the data and to uncover any significant associations. AIC (Akaike Information Criterion) values of each model were then compared (Akaike, 1974). These preliminary models included variables such as behavioral diversity, persistence, age, sex, trial number, number of successful trials, time of trial, latency to approach, number of bears present, month, and motivation score. The five models (outlined above) were selected for use following this preliminary testing. Time of day and month were excluded from the final models as the inclusion of those variables did not improve the model fit or uncover any significant associations.

Prior to fitting the models, we inspected all predictors and responses for whether their distributions were roughly symmetric. As a consequence, all continuous variables (time-to-solve, age, behavioral diversity, persistence, number of successful trials, trial number, latency to approach and number of bears present) were log-transformed prior to analysis to satisfy the assumption of normality. Continuous predictors (age, behavioral diversity, persistence, number of successful trials, trial number, latency to approach and number of bears present) were z-transformed to make model interpretation easier (Schielzeth, 2010) and for easy model convergence. To handle the binary data, sex and motivation score were both dummy coded, with female and high (H) being the reference category, respectively. After fitting the models, we checked that the assumptions of the residuals - to be normally distributed and homogeneous - were fulfilled. No deviations from these assumptions were indicated.

Full-null model comparisons were conducted to test the influence of the fixed effects and their interactions (Forstmeier & Schielzeth, 2011), whereby the null models lacked the interaction effects but were otherwise identical to the full model. The effect of individual fixed effects was tested by means of the Satterthwaite approximation (Luke, 2017). To test for the presence of multicollinearity, variance inflation factor (VIF) scores were checked. These were produced using models lacking the interactions between fixed effects and collinearity appeared to be of no issue (maximum VIF = 3.53). Model stability was also assessed based on the levels of the estimated coefficients and standard deviations by excluding the levels

of the random effects one at a time (Nieuwenhuis et al., 2012). Both models appeared to be of acceptable stability, with the exception of the random effect (subject ID).

The sample for model one, two, and three encompassed 32 trial values, from 14 individuals, with 10 observations per estimated term. The sample for model four encompassed 30 trial values, from 13 individuals, with 9.5 observations per estimated term. The sample for model five encompassed 30 trial values, from 13 individuals, with 12 observations per estimated term.

Results

Hypothesis one posited that *brown bears possess problem-solving abilities*, which was evaluated using the following testable predictions: *brown bears are able to gain access to a baited puzzle box using their problem-solving abilities* and *brown bears are able to use the latch to gain access to the puzzle box*. These predictions and hypothesis were confirmed by the following results.

Out of 32 trials, latch use was recorded 17 times and alternative techniques were used in the remaining 15 trials (Table 2). There was a significant difference between the two maneuverability categories, with maneuverability of the puzzle box influencing the presence of latch-use ($\chi^2 = 14.191$, $df = 1$, $p = .001$).

In model one, the interaction between behavioral diversity and persistence failed to be significant (full-null model comparison: $\chi^2 = 1.353$, $df = 1$, $p = .245$) and consequently the model was re-run excluding the interaction terms. In this reduced model, behavioral diversity failed to be significant ($p = .054$, $R^2m = .722$, $R^2c = .852$) (R^2m = marginal R^2 value including just fixed effects, R^2c = conditional R^2 value including fixed and random effects). Persistence was positively correlated with time-to-solve and impacted success times ($p = .010$, $R^2m = .722$, $R^2c = .852$). More specifically, as persistence (the number of attempts) increased, time-to-solve also increased (Table 3; Figure 6).

Table 2

Results of the Puzzle Box Trials, Including How the Box Was Presented and Presence/Absence of Latch Use

Subject (ID)	Total *N* of successful trials[1]	Trials with latch use[2]	Trials with no latch use	Box well chained/ box loosely chained
Eso	3	0	3	0/3
Fluff	3	1	2	3/0
Brumm	2	0	2	0/2
Maja	1	0	1	0/1
Brumma	3	0	3	0/3
Mish	3	3	0	3/0
Lucy	1	1	0	1/0
Athena	3	2	1	3/0
Loki	1	1	0	1/0
Nelly	3	1	2	3/0
Enciam	2	1	1	2/0
Julio	1	1	0	1/0
Roja	3	3	0	3/0
Tornilla	3	3	0	3/0
TOTAL	**32**	**17**	**15**	**23/9**

Note. [1]When using up to three trials per bear, some individuals had more than three successful trials, [2]Whether accidental or intentional not distinguished.

Table 3

Results of GLMM Model One of Behavioral Diversity and Persistence on Log-Time-to-Solve

Term	estimate	SE	lower CI	upper CI	t	df	P	min	max
Intercept	1.47	.066	1.355	1.597	22.415	12.809	.000	1.405	1.507
Behavioral diversity[2]	0.199	.099	-0.013	0.392	2.001	31.999	.054	0.115	0.265
Persistence[3]	0.279	.101	0.074	0.485	2.754	31.407	**.010**	0.219	0.399

Note. [1]Results displayed are estimates and standard errors, together with confidence limits, results of tests, and the range of estimates obtained when dropping levels of random effects one at a time, [2] z-transformed to a mean of 0 and a standard deviation of 1; mean and SD of the original variable were .645 and .242, respectively and [3] z-transformed to a mean of 0 and a standard deviation of 1; mean and SD of the original variable were .794 and .539, respectively.

Figure 6

The Influence of Persistence and Behavioral Diversity on Time-to-Solve
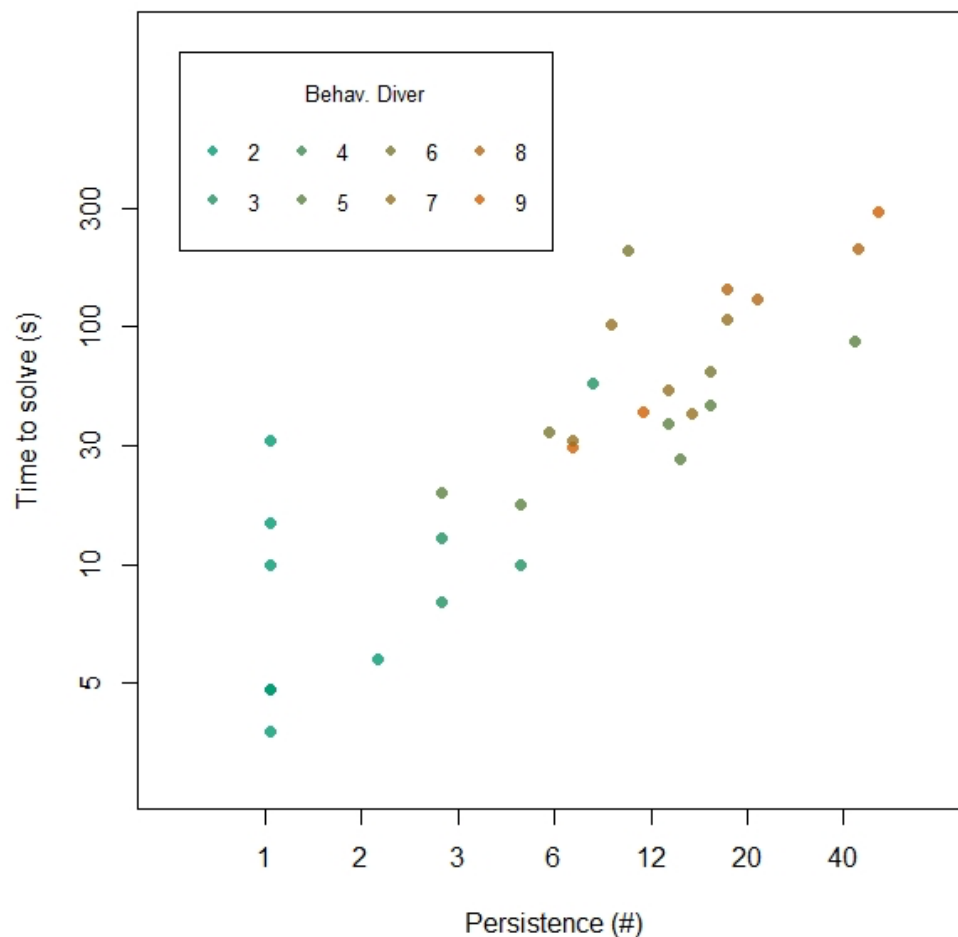
Note. Axes are in $\log_{10}$ scale. As time-to-solve increases both the number of attempts and the types of attempts increases (greater behavioral diversity displayed in orange dots). Data points represent all trial values ($N = 32$) from 14 individuals (mean number of trial values or data points per individual = 2.29, ± 0.91).

In model two, the interaction between the number of successful trials and trial number significantly improved the model and thus the full model was retained (full-null model comparison: $\chi2 = 6.233$, $df = 1$, $p = .013$). In this model, the number of successful trials appeared to influence success times, as this was significantly negatively correlated with time-to-solve ($p = .020$, $R^2m = .563$, $R^2c = .653$). Trial number

failed to be significant (p = .336, R^2m = .563, R^2c = .653). In addition, there appears to be an interaction between number of successful trials and trial number, indicating that time-to-solve is dependent both on *exposure to success* and the exact trial *number* (p = .009, R^2m = .563, R^2c = .653) (Table 4).

Table 4

Results of GLMM Model Two of Number of Successful Trials and Trial Number and their Interaction on Log-Time-to-Solve[1]

Term	estimate	SE	lower CI	upper CI	t	df	P	min	max
Intercept	1.625	.099	1.432	1.802	16.477	9.751	.000	1.554	1.734
# of successful trials[2]	-0.271	.11	-0.460	-0.038	-2.470	28.733	**.020**	-0.384	-0.212
Trial number[3]	-0.102	.103	-0.313	0.083	-0.983	21.899	.336	-0.165	-0.009
ST:TN	-0.234	.08	-0.383	-0.078	-2.919	17.991	**.009**	-0.317	-0.182

Note. [1]Results displayed are estimates and standard errors, together with confidence limits, results of tests, and the range of estimates obtained when dropping levels of random effects one at a time, [2]z-transformed to a mean of 0 and a standard deviation of 1; mean and SD of the original variable were 0.546 and 0.296, respectively, and [3]z-transformed to a mean of 0 and a standard deviation of 1; mean and SD of the original variable were 0.382 and 0.315, respectively.

In model three, latency to approach appeared to influence success times, as this was significantly positively associated with time-to-solve (p = .028, R^2m = .13, R^2c = .794). Specifically, as latency to approach increases, time-to-solve also increases. It is worth noting, this variable did fail to be significant when paired with time of day. Motivation score also failed to be significant (p = .166, p = .259, R^2m = .13, R^2c = .794). This was the result produced when using high (H) as the reference category (Table 5).

Table 5

Results of GLMM Model Three of Latency to Approach and Motivation Score on Log-Time-to-Solve[1]

Term	estimate	SE	lower CI	upper CI	t	df	P	min	max
Intercept	1.750	0.223	1.287	2.239	7.833	27.123	0.000	1.620	1.926
Latency[2]	0.205	0.088	0.015	0.385	2.327	27.592	**0.028**	0.100	0.253
Motivation score L[3]	-0.371	0.262	-0.919	0.132	-1.420	31.422	0.166	-0.587	-0.154
Motivation score M[3]	-0.211	0.182	-0.555	0.151	-1.159	21.728	0.259	-0.426	-0.011

Note. [1]Results displayed are estimates and standard errors, together with confidence limits, results of tests, and the range of estimates obtained when dropping levels of random effects one at a time, [2]z-transformed to a mean of 0 and a standard deviation of 1; mean and SD of the original variable were 1.284 and 0.436, respectively, and [3] coded as factor with high (H) being the reference category.

Hypothesis two posited that *brown bears possess object-manipulation abilities*, which was evaluated using the following testable predictions: *brown bears are able to manipulate an object leading to facilitate the retrieval of an out of reach food reward* and *brown bears are capable of tool-use*. These predictions and hypothesis two were not supported. Out of 30 trials, stump use was present 20 times and alternative techniques were used in the other ten trials (Tables 6 and 7).

Table 6

Results of the Object-Manipulation Trials, Including Number of Variations Presented and Presence/Absence of Stump Use

Subject (ID)	Object-manipulation stage reached	Stage with stump use present	Total N of successful trials[1]	Trials with stump use/no stump use
Eso	2	1	3	1/2
Brumm	1	1	3	2/1
Maja	1	1	1	0/1
Brumma	1	1	1	1/0
Mish	1	1	3	0/3
Lucy	1	1	1	0/1
Athena	2	1	3	2/1
Fivi	2	1	2	1/1
Nelly	1	1	2	2/0
Enciam	1	1	2	2/0
Julio	1	1	3	3/0
Roja	1	1	3	3/0
Tornilla	1	1	3	3/0
TOTAL	–	–	30	20/10

Note. [1] When using up to three trials per bear, some individuals had more than three successful trials.

Table 7

The Alternative Techniques Implemented by the Bears When Solving the Object-Manipulation Set-Up

Techniques used[1]	*N* times witnessed	Object-manipulation stage
Climb tree and loosen rope leading to the food dropping down to a reachable height[2]	2	2
Climb apparatus and shake rope leading to the food dropping to the floor	2	1
Climb apparatus and put pressure on the rope, leading to the food sliding to the bear	3	1
Climb and hang from apparatus so the rope is in reach, then pull rope close and grab the food	1	1
Pull down on apparatus attached to chain (instead of rope) bringing the food down to a reachable height	2	2

Note. [1]Techniques are described how they happened, it was not obvious whether they were intentional or accidental, [2]This individual even tried to suspend herself/hang from the rope, almost like she was going to pull herself along the rope to the food, but the rope did not hold (she tried this twice).

In model four, including the interaction between age and sex significantly improved the model and so the full model was retained (full-null model comparison: $\chi^2 = 6.579$, $df = 1$, $p = .010$). In this model, age appeared to influence success times, as age was significantly negatively correlated with time-to-solve ($p = .022$, $R^2m = .463$, $R^2c = .597$) or simply, with increasing age comes increasing time to success. Similarly, sex influenced time-to-solve, with sex significantly negatively correlated with time-to-solve ($p = .004$, $R^2m = .463$, $R^2c = .597$) (Figure 7). This was the result produced when using female as the reference category, indicating that males typically solved the trials faster in comparison to females. There also appears to be an

interaction between age and sex, indicating that time-to-solve is dependent on *both* age and sex ($p = .011$, $R^2m = .463$, $R^2c = .597$) (Table 8).

In model five, the number of bears present appeared to influence success times, as this was significantly negatively associated with time-to-solve ($p = .004$, $R^2m = .442$, $R^2c = .537$). Specifically, as the number of bears present during trials increases, the time taken to solve the task grows longer. Trial number failed to be significant ($p = .11$, $R^2m = .442$, $R^2c = .537$) (Table 9).

Figure 7

The Influence of Age and Sex on Bears' Time-To-Solve
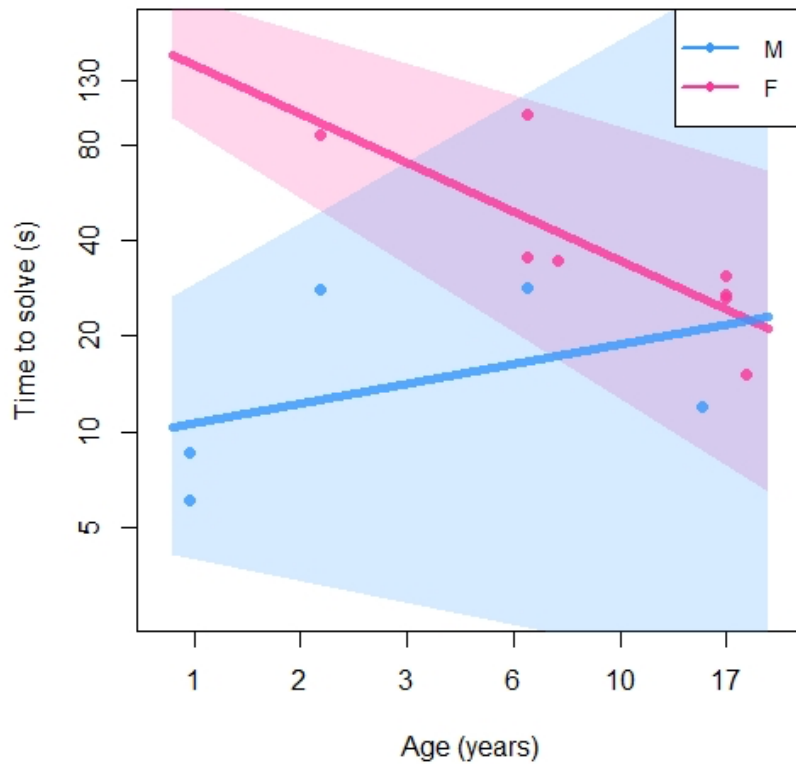
Note. Axes are in $\log_{10}$ scale. Time-to-solve is dependent on both sex and age, with younger males typically having improved scores. Colored lines are the model predictions per sex (M – male; F – female), shaded areas are the 95% confidence intervals. Data points represent mean values of all trial values ($N = 30$) per individual ($N = 13$) (mean number of trial values per individual = 2.31, ± 0.85).

Table 8

Results of GLMM Model Two of Age, Sex and their Interaction on Log-Time-to-Solve[1]

Term	estimate	SE	lower CI	upper CI	t	df	P	min	max
Intercept	1.690	.102	1.482	1.888	16.549	23.372	.000	1.634	1.720
Age[2]	-0.306	.124	-0.557	-0.058	-2.460	22.197	**.022**	-0.337	-0.259
SexM[3]	-0.478	.141	-0.740	-0.201	-3.398	16.420	**.004**	-0.572	-0.153
A:S	0.428	.151	0.131	0.729	2.825	17.700	**.011**	0.343	0.684

Note. [1]Results displayed are estimates and standard errors, together with confidence limits, results of tests, and the range of estimates obtained when dropping levels of random effects one at a time, [2]z-transformed to a mean of 0 and a standard deviation of 1; mean and SD of the original variable were .744 and .489, respectively and [3] dummy coded with female being the reference category.

Table 9

Results of GLMM Model Five of Number of Bears Present and Trial Number on Log-Time-to-Solve[1]

Term	estimate	SE	lower CI	upper CI	*t*	*df*	*P*	min	max
Intercept	1.332	.054	1.228	1.429	24.556	9.828	.000	1.284	1.375
# Of bears present[2]	-0.224	.056	-0.333	-0.117	-3.992	8.410	**.004**	-0.297	-0.178
Trial number[3]	-0.081	.049	-0.177	0.020	-1.649	29.977	.110	-0.126	-0.051

Note. [1]Results displayed are estimates and standard errors, together with confidence limits, results of tests, and the range of estimates obtained when dropping levels of random effects one at a time, [2]*z*-transformed to a mean of 0 and a standard deviation of 1; mean and SD of the original variable were .446 and .23, respectively, and [3]*z*-transformed to a mean of 0 and a standard deviation of 1; mean and SD of the original variable were .382 and .315, respectively.

Discussion

Our findings confirm brown bears to be competent problem-solvers. Despite this, we were unable to find evidence of spontaneous tool using among the bears in our study. While not always solving the tasks in the ways designed in the study, the bears persisted and typically still solved the problem, highlighting their behavioral flexibility and adaptability when faced with novel problems.

Puzzle Box

All 14 bears that engaged with the puzzle box had at least one successful trial and were successful in solving the problem. Hypothesis one is supported, with bears able to use their problem-solving abilities to gain access to the puzzle box, in some instances using the latch. This outcome concurs with Benson-Amram and colleagues (2016) who found bear species able to succeed at a similar task.

Data analyzed from the puzzle box trials revealed persistence to be associated with time-to-solve. Analyses further revealed a potential association between behavioral diversity and time-to-solve, albeit this fell short of statistical significance, perhaps due to small sample sizes and/or individual variation. This suggests that with increasing time-to-solve, both the *number* and *types* of attempts to gain access to the box also increased. The first suggestion is typically expected during cognitive trials since as engagement time increases, the *number* of attempts also increases, as the individual increasingly tries to retrieve the food reward. The second suggestion is more interesting, as it hints toward the fact that the *types* of behaviors used also increases with time-to-solve. Thus, with increasing time, the individual broadens its approach utilizing different techniques (i.e., elaboration), until one is successful. This result is indicative of trial-and-error learning and suggests bears, alongside other species (Galef & Laland, 2005; Heyes, 1994), use this technique when facing novel challenges. Similar results were found by Waroff and colleagues (2017) who also suggest bears use trial and error techniques when approaching new tasks, usually invoking physical force. This type of learning stands in contrast to insight learning, which is thought to require causal knowledge, means-end understanding and mental models (Heinrich, 2000; Heinrich & Bugnyar, 2005; Huber & Gajdon, 2006).

On several occasions the use of successful techniques was repeated during subsequent trials, suggesting the bears remembered previous exposures and adopted successful strategies. This was reinforced by the presence of an interaction between the number of successful trials and trial number, with the number of successful trials negatively associated with time-to-solve. That indicates that as the number of successful exposures increases, the bears learn successful strategies, perhaps first uncovered by their use of trial-and-error learning, which ultimately improves their performance over time. African lions *Panthera leo*, too, are capable of solving a novel problem, while learning and remembering the task solution in subsequent trials (Borrego & Dowling, 2016). Benson-Amram et al. (2016) when presenting a similar task to a wide range of carnivoran species, found successful individuals improved their performance with experience. Thus, when compared to other carnivoran species, the bears were similarly capable of gaining an understanding of the puzzle and how to open it.

Latency to approach appeared important in terms of influencing success times. It was positively associated with time-to-solve, indicating that as time to engage the task increased, the time taken to solve it also increased. This suggests that individual variation in motivational levels is an important factor influencing cognitive engagement and performance. Cooke et al. (2021) likewise found variation in problem-solving performance to be best explained by motivational differences. This concurs with the notion that motivation is an important driver of innovative behavior (Laland & Reader, 1999; Sol et al., 2012). Here, the result is most likely the consequence of motivational differences in terms of the food reward present, rather than the influence of neophobia, as bears typically did not show signs of fear toward the novel object. Therefore, the latency to approach result is most likely due to differing levels of interest to engage with the trials. This highlights how motivational differences are an important factor to consider when investigating the presence of cognitive abilities.

Frequent behaviors implemented in efforts to gain access to the food reward included tipping, shaking, and pounding the box, as well as clawing at the food item through the bars. Despite this, whilst the bears often interacted with the door and latch, they frequently failed to draw associations between that and retrieving the food reward. From observations of trials, instead it appears the bears became hyper-focused on the food reward, failing to fully assess the situation. Amici et al. (2019) similarly note how bears fail to "cognitively represent" certain situations. Despite this, we did witness two individuals appearing to draw the association between latch-use and access to the box. The bears, both male and juvenile in age, by the end of their trials were gaining access to the box in less than ten seconds, often with only one attempt and clear latch use (see supplementary materials for footage). While we are hesitant to make assumptions and broad generalisations of such behavior, this hints at evidence that bears can make such associations, particularly through asocial learning. Asocial learning, involving direct interaction with the inanimate environment (Laland, 2004), contrasts with social learning, in which individuals learn through the influence of the behavior or products of the behavior of conspecifics (van Schaik et al., 2016). Waroff and colleagues (2017) also suggest that when physical force is not successful, bears often display '*insight-like*' behavior, indicative of intelligent behavior.

Object-Manipulation

Of 15 individuals that engaged with the object-manipulation set-up, 13 successfully solved the problem. While successful, this result fails to offer support to hypothesis two since the bears failed to show evidence of object-manipulation ability. Bears did not make use of the supplied tree stumps as an intermediatory to access a food reward. Consequently, we find no evidence of tool use in brown bears.

Data analyzed from the object-manipulation set-up revealed that both age and sex were influencing time-to-solve, as it was revealed there was an interaction between age and sex. In this scenario, younger male bears more quickly solve the task. Some caution is warranted, however, since the sample sizes for this task are small (males = 5, females = 8). The data show wide confidence intervals for males and are funnel-shaped, suggestive of the sample not being adequate to provide an accurate representation of the population as a whole; instead, potentially being a product of the individuals examined here and the sampling techniques implemented.

For age association, the direction of the relationship is negative, indicating that younger bears typically were quicker at solving the problem. This result contrasts with a long-held assumption that with increasing age comes increased knowledge in terms of certain environmental and social events, seen for example in elephants (McComb et al., 2001, 2011). We instead find bears younger in age typically outperforming those older in age. This is likely due to the fact that the younger bears were generally more excitable, receptive, and willing to engage with the trials. Juvenile hyenas have also been found to display greater exploratory behaviors, while being more persistent and less neophobic than adults (Benson-Amram & Holekamp, 2012). The juvenile bears are also at the age where the most social learning occurs, transferred through the mother-offspring relationship (Breck et al., 2008), meaning cubs will typically be more curious in nature and receptive to the environment. For example, play behavior in cubs has been found to influence survival (Fagen & Fagen, 2004, 2009). This highly receptive period, coupled with the investigatory nature

of young bears (highlighted by Bacon, 1980), means they were more likely to engage with trials and better predisposed to learning. This increased their exposure to the trials, allowing them to refine their skills or techniques in solving the problem, which over trials ultimately reduced their time-to-solve. Likewise, with the puzzle box trials, it was juvenile individuals who picked up the latch association; with this association appearing to be consistent across both experimental tests.

Regarding the sex association, male bears typically had shorter time-to-solve scores, more quickly solving the problem. This could indicate that there may be sex differences in the cognitive abilities of bears (see, for example, Carazo et al., 2014; Jonasson, 2005; Jones et al., 2003). However, we believe this association is likely rather the result of collective testing and instead indicates that social dynamics are influencing cognitive performance. Specifically, because trials were run so that all bears had access to and could interact with the set-ups at the same time, this meant that those bears that were more dominant, i.e., males, would often monopolize the trials due to the food reward and competition present. This proposal is further reinforced by the findings that the number of bears present during trials was negatively associated with time-to-solve. This is likely the consequence of the collective nature of testing which promoted increased competition between bears over the reward being present. Although brown bears are considered obligatorily solitary (Gittleman, 1999), in captivity they are commonly kept in pairs or groups (mean group size here = 3), with this social environment likely influencing bear behavior, especially in terms of engagement with enrichment devices or high-value food rewards. While running trials individually is preferred, that may perhaps neglect the influence of social dynamics on cognitive processes. For example, while individual testing is possible in some zoological settings, this does not represent the wild environment in which socioecological challenges are presented, whereby individuals could be influenced by conspecifics in their surrounding social environment (Hansen et al., 2021).

In terms of the variations of the object-manipulation set-up presented, most trials were run with the stump already directly underneath the food reward. Bears often failed to identify the stump as an object to use and instead tried other techniques to solve the problem, such as climbing the tree/apparatus involved and shaking the food reward free. Ultimately, stump use was limited. While unexpected, the bears instead exhibited great resourcefulness in the techniques they used to solve the problem and almost all trials run were successful, even if tree stump use was not present. This apparent behavioral flexibility warrants further investigation. When stump use was present and the set-up was subsequently altered in further trials, the bears often then either failed to solve the problem or reverted to other techniques. Thus, the bears failed to manipulate an object to retrieve a food reward. Amici and colleagues (2019) concluded similarly. The result displays not a lack of ability, as bears have previously proven successful (Waroff et al., 2017) but rather that the scenario potentially fails to hold ecological relevance and neglects to draw on typical bear behaviors. We recommend that more ecologically relevant scenarios be designed to test this cognitive ability.

Conclusions

The evidence presented here supports the proposition that bears possess problem-solving abilities. Brown bears were able to use their problem-solving abilities so as to gain access to our puzzle box, in some instances using the latch. The bears failed, however, to spontaneously tool use, although 13 out of 15 individuals were successful in retrieving the food reward, instead using alternative techniques to solve the problem. Data analyes highlighted several variables influencing time-to-solve, including persistence, behavioral diversity, number of successful trials, latency to approach, age, sex and number of bears present. Since these results may be 1) an artefact of the trial conditions and/or 2) a product of specific individuals sampled within this study, further investigation is needed to confirm these findings before extrapolation to the wider population.

We further advocate the use of zoo-housed cognition studies for what they can reveal about the lived-experiences of individuals housed in zoological settings. The captive environment differs from the wild one, leading to a different (though not necessarily a lesser) existence for those individuals who represent the conservation concerns of their wild counterparts to the viewing public. In the UK there are

currently 32 brown bears, 12 polar bears, 12 American black bears, 3 Asiatic black bears *U. thibetanus*, 9 sun bears *Helarctos malayanus*, 16 Andean bears *Tremarctos ornatus*, 3 sloth bears, and 2 giant pandas *Ailuropoda melanoleuca* (N = 89 total). Such "ambassador" individuals (Rees, 2013), which are representative of their wild counterparts, warrant our fullest attention to ensure they live the most fulfilled lives we can offer them.

In conclusion, our study provides a demonstration that brown bears are an excellent model system for testing the cognitive abilities of Ursids and subsequently testing theories of cognitive evolution. The explanation for encephalized brains in Ursids has been somewhat of an enigma since bears appear as an outlier in prominent explanations for the evolution of encephalization (e.g., the social brain hypothesis) due to their relatively asocial nature. Selection for enhanced cognitive abilities, specifically in terms of their behavioral flexibility and ability to problem-solve and innovate, offers one explanation for their presence in these taxa.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716-723.

Amici, F., Call, J., Watzek, J., Brosnan, S., & Aureli, F. (2018). Social inhibition and behavioural flexibility when the context changes: A comparison across six primate species. *Scientific Reports*, *8*, 3067.

Amici, F., Holland, R., & Cacchione, T. (2019). Sloth bears (*Melursus ursinus*) fail to spontaneously solve a novel problem even if social cues and relevant experience are provided. *Journal of Comparative Psychology*, *133*, 373-379.

Amiel, J. J., Tingley, R., & Shine, R. (2011). Smart moves: Effects of relative brain size on establishment success of invasive amphibians and reptiles. *PLOS ONE*, *6*(4), e18277.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press.https://doi.org/10.1017/CBO9780511801686

Bacon, E. S. (1980). Curiosity in the American black bear. *Bears: Their Biology and Management*, *4*, 153-157.

Bacon, E. S., & Burghardt, G. M. (1976a). Ingestive behaviors of the American black bear. *Bears: Their Biology and Management*, *3*, 13-25.

Bacon, E. S., & Burghardt, G. M. (1976b). Learning and color discrimination in the American black bear. *Bears: Their Biology and Management*, *3*, 27-36.

Bacon, E. S., & Burghardt, G. M. (1983). Food preference testing of captive black bears. *Bears: Their Biology and Management*, *5*, 102-105.

Bandini, E., & Harrison, R. A. (2020). Innovation in chimpanzees. *Biological Reviews*, *95*, 1167-1197.

Bartoń, K. (2010). *MuMIn: multi-model inference* [Computer software].

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1 - 48.

Beck, B. B., Walkup, K., & Shumaker, R. (2012). Tool use and problem solving in animals. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 3330-3333). Springer.

Belant, J. L., Griffith, B., Zhang, Y., Follmann, E. H., & Adams, L. G. (2010). Population-level resource selection by sympatric brown and American black bears in Alaska. *Polar Biology*, *33*, 31-40.

Benson-Amram, S., Griebling, H.J., & Sluka, C.M. (2023). The current state of carnivore cognition. *Animal Cognition, 26*, 37-58.

Benson-Amram, S., Dantzer, B., Stricker, G., Swanson, E. M., & Holekamp, K. E. (2016). Brain size predicts problem-solving ability in mammalian carnivores. *PNAS*, *113*(9), 2532-2537.

Benson-Amram, S., & Holekamp, K. E. (2012). Innovative problem solving by wild spotted hyenas. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1744), 4087-4095.

Benson-Amram, S., Weldele, M. L., & Holekamp, K. E. (2013). A comparison of innovative problem-solving abilities between wild and captive spotted hyaenas, *Crocuta crocuta*. *Animal Behaviour*, *85*, 349-356.

Boddy, A. M., McGowen, M. R., Sherwood, C. C., Grossman, L. I., Goodman, M., & Wildman, D. E. (2012). Comparative analysis of encephalization in mammals reveals relaxed constraints on anthropoid primate and cetacean brain scaling. *Journal of Evolutionary Biology*, *25*, 981-994.

Boesch, C. (2012). The ecology and evolution of social behavior and cognition in primates. In J. Vonk & T. K. Shackelford (Eds.), *The Oxford handbook of comparative evolutionary psychology* (pp. 486-503). Oxford University Press.

Borrego, N. (2017). Big cats as a model system for the study of the evolution of intelligence. *Behavioural Processes*, *141*, 261-266.

Borrego, N., & Dowling, B. (2016). Lions (*Panthera leo*) solve, learn, and remember a novel resource acquisition problem. *Animal Cognition*, *19*, 1019-1025.

Borrego, N., & Gaines, M. (2016). Social carnivores outperform asocial carnivores on an innovative problem. *Animal Behaviour*, *114*, 21-26.

Breck, S. W., Williams, C. L., Beckmann, J. P., Matthews, S. M., Lackey, C. W., & Beecham, J. J. (2008). Using genetic relatedness to investigate the development of conflict behavior in black bears. *Journal of Mammalogy*, *89*, 428-434.

Buchanan, K., Perera, T., Carere, C., Carter, T., Hailey, A., Hubrecht, R., Jennings, D., Metcalfe, N., Pitcher, T., Péron, F., Sneddon, L., Sherwin, C., Talling, J., Thomas, R., & Thompson, M. (2012). Guidelines for the treatment of animals in behavioural research and teaching. *Animal Behaviour*, *83*, 301-309.

Byron, J. (2017). *r2glmm: computes R squared for mixed (multilevel) models* [Computer software].

Carazo, P., Noble, D. W. A., Chandrasoma, D., & Whiting, M. J. (2014). Sex and boldness explain individual differences in spatial learning in a lizard. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1782), 20133275.

Chambers, H. R., Heldstab, S. A., & O'Hara, S. J. (2021). Why big brains? A comparison of models for both primate and carnivore brain size evolution. *PLOS ONE*, *16*(12), e0261185.

Cooke, A. C., Davidson, G. L., van Oers, K., & Quinn, J. L. (2021). Motivation, accuracy and positive feedback through experience explain innovative problem solving and its repeatability. *Animal Behaviour*, *174*, 249-261.

Costello, C. M., Cain, S. L., Pils, S., Frattaroli, L., Haroldson, M. A., & van Manen, F. T. (2016). Diet and macronutrient optimization in wild ursids: A comparison of grizzly bears with sympatric and allopatric black bears. *PLOS ONE*, *11*(5), e0153702.

Daniels, S. E., Fanelli, R. E., Gilbert, A., & Benson-Amram, S. (2019). Behavioral flexibility of a generalist carnivore. *Animal Cognition*, *22*, 387-396.

DeCasien, A. R., Williams, S. A., & Higham, J. P. (2017). Primate brain size is predicted by diet but not sociality. *Nature Ecology and Evolution*, *1*, 0112.

Deecke, V. B. (2012). Tool-use in the brown bear (*Ursus arctos*). *Animal Cognition*, *15*, 725-730.

Drea, C. M., & Carter, A. N. (2009). Cooperative problem solving in a social carnivore. *Animal Behaviour*, *78*, 967-977.

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology*, *6*, 178-190.

Dungl, E., Schratter, D., & Huber, L. (2008). Discrimination of face-like patterns in the giant panda (*Ailuropoda melanoleuca*). *Journal of Comparative Psychology*, *122*, 335-343.

Emery, N. J., & Clayton, N. S. (2004). The mentality of crows: Convergent evolution of intelligence in corvids and apes. *Science*, *306*(5703), 1903-1907.

Emery, N. J., & Clayton, N. S. (2009). Tool use and physical cognition in birds and mammals. *Current Opinion in Neurobiology*, *19*, 27-33.

Fagen, R., & Fagen, J. (2004). Juvenile survival and benefits of play behaviour in brown bears, *Ursus arctos*. *Evolutionary Ecology Research*, *6*, 89-102.

Fagen, R., & Fagen, J. (2009). Play behaviour and multi-year juvenile survival in free-ranging brown bears, *Ursus arctos*. *Evolutionary Ecology Research*, *11*, 1053-1067.

Finarelli, J. A., & Flynn, J. J. (2009). Brain-size evolution and sociality in carnivora. *PNAS*, *106*(23), 9345-9349.

Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology*, *65*, 47-55.

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed). Sage.

Galef, B. G., & Laland, K. N. (2005). Social learning in animals: Empirical studies and theoretical models. *BioScience*, *55*, 489-499.

Gittleman, J. L. (1986). Carnivore brain size, behavioral ecology, and phylogeny. *Journal of Mammalogy*, *67*, 23-36.

Gittleman, J. L. (1999). Hanging bears from phylogenetic trees: Investigating patterns of macroevolution. *Ursus*, *11*, 29-39.

Hansen, J. E., Hertel, A. G., Frank, S. C., Kindberg, J., & Zedrosser, A. (2021). Social environment shapes female settlement decisions in a solitary carnivore. *Behavioral Ecology*, *33*(1), 137-146.

Heinrich, B. (2000). Testing insight in ravens. In C. M. Heyes & L. Huber (Eds.), *The evolution of cognition Vienna series in theoretical biology* (pp. 289-305). The MIT Press.

Heinrich, B., & Bugnyar, T. (2005). Testing problem solving in ravens: String-pulling to reach food. *Ethology*, *111*, 962-976.

Heyes, C. M. (1994). Social learning in animals: Categories and mechanisms. *Biological Reviews of the Cambridge Philosophical Society*, *69*, 207-231.

Hilderbrand, G. V., Gustine, D. D., Mangipane, B. A., Joly, K., Leacock, W., Mangipane, L. S., Erlenbach, J., Sorum, M. S., Cameron, M. D., Belant, J. L., & Cambier, T. (2018). Body size and lean mass of brown bears across and within four diverse ecosystems. *Journal of Zoology*, *305*, 53-62.

Holekamp, K. E., & Benson-Amram, S. (2017). The evolution of intelligence in mammalian carnivores. *Interface Focus*, *7*, 20160108.

Horschler, D. J., Hare, B., Call, J., Kaminski, J., Miklósi, Á., & MacLean, E. L. (2019). Absolute brain size predicts dog breed differences in executive function. *Animal Cognition*, *22*, 187-198.

Huber, L., & Gajdon, G. (2006). Technical intelligence in animals: The kea model. *Animal Cognition*, *9*, 295-305.

Jerison, H. J. (1985). Animal intelligence as encephalization. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *308*(1135), 21-35.

Johnson-Ulrich, Z., Vonk, J., Humbyrd, M., Crowley, M., Wojtkowski, E., Yates, F., & Allard, S. (2016). Picture object recognition in an American black bear (*Ursus americanus*). *Animal. Cognition*, *19*, 1237-1242.

Jonasson, Z. (2005). Meta-analysis of sex differences in rodent models of learning and memory: A review of behavioral and biological data. *Neuroscience & Biobehavioral Reviews*, *28*, 811-825.

Jones, C. M., Braithwaite, V. A., & Healy, S. D. (2003). The evolution of sex differences in spatial ability. *Behavioral Neuroscience*, *117*, 403-411.

Kelling, A. S., Snyder, R. J., Marr, M. J., Bloomsmith, M. A., Gardner, W., & Maple, T. L. (2006). Color vision in the giant panda (*Ailuropoda melanoleuca*). *Learning & Behavior*, *34*, 154-161.

Kotrschal, A., Buechel, S. D., Zala, S. M., Corral-Lopez, A., Penn, D. J., & Kolm, N. (2015). Brain size affects female but not male survival under predation threat. *Ecology Letters*, *18*, 646-652.

Kotrschal, A., Rogell, B., Bundsen, A., Svensson, B., Zajitschek, S., Brännström, I., Immler, S., Maklakov, A. A., & Kolm, N. (2013). The benefit of evolving a larger brain: Big-brained guppies perform better in a cognitive task. *Animal Behaviour*, *86*, e4-e6.

Kubinyi, E., Viranyi, Z., & Miklósi, Á. (2007). Comparative social cognition: From wolf and dog to humans. *Comparative Cognition and Behavior Reviews*, *2*, 26-46.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*, 1 - 26.

Lafferty, D. J. R., Belant, J. L., & Phillips, D. L. (2015). Testing the niche variation hypothesis with a measure of body condition. *Oikos*, *124*, 732-740.

Laland, K. N. (2004). Social learning strategies. *Animal Learning & Behavior*, *32*, 4-14.

Laland, K. N., & Reader, S. M. (1999). Foraging innovation in the guppy. *Animal Behaviour*, *57*, 331-340.

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*, *49*, 1494-1502.

MacLean, E. L., Hare, B., Nunn, C. L., Addessi, E., Amici, F., Anderson, R. C., Aureli, F., Baker, J. M., Bania, A. E., Barnard, A. M., Boogert, N. J., Brannon, E. M., Bray, E. E., Bray, J., Brent, L. J., Burkart, J. M., Call, J., Cantlon, J. F., Cheke, L. G., . . . Zhao, Y. (2014). The evolution of self-control. *PNAS*, *111*(20), E2140-2148.

Mangipane, L. S., Belant, J. L., Lafferty, D. J. R., Gustine, D. D., Hiller, T. L., Colvin, M. E., Mangipane, B. A., & Hilderbrand, G. V. (2018). Dietary plasticity in a nutrient-rich system does not influence brown bear (*Ursus arctos*) body condition or denning. *Polar Biology*, *41*, 763-772.

Mangipane, L. S., Lafferty, D. J. R., Joly, K., Sorum, M. S., Cameron, M. D., Belant, J. L., Hilderbrand, G. V., & Gustine, D. D. (2020). Dietary plasticity and the importance of salmon to brown bear (*Ursus arctos*) body size and condition in a low Arctic ecosystem. *Polar Biology*, *43*, 825 - 833.

McComb, K., Moss, C., Durant, S. M., Baker, L., & Sayialel, S. (2001). Matriarchs as repositories of social knowledge in African elephants. *Science*, *292*(5516), 491-494.

McComb, K., Shannon, G., Durant, S. M., Sayialel, K., Slotow, R., Poole, J., & Moss, C. (2011). Leadership in elephants: the adaptive value of age. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1722), 3270-3276.

Miklósi, Á., Topál, J., & Csányi, V. (2004). Comparative social cognition: What can dogs teach us? *Animal Behaviour*, *67*, 995-1004.

Myers, P. J., & Young, J. K. (2018). Consistent individual behavior: Evidence of personality in black bears. *Journal of Ethology.*, *36*, 117-124.

Nieuwenhuis, R., Grotenhuis, M., & Pelzer, B. (2012). influence.ME: Tools for detecting influential data in mixed effects models. *The R Journal*, *4*, 38-47.

Perdue, B. M., Snyder, R. J., Pratte, J., Marr, M. J., & Maple, T. L. (2009). Spatial memory recall in the giant panda (*Ailuropoda melanoleuca*). *Journal of Comparative Psychology*, *123*, 275-279.

Perdue, B. M., Snyder, R. J., Zhihe, Z., Marr, M. J., & Maple, T. L. (2011). Sex differences in spatial ability: A test of the range size hypothesis in the order Carnivora. *Biology Letters*, *7*, 380-383.

Powell, L. E., Isler, K., & Barton, R. A. (2017). Re-evaluating the link between brain size and behavioural ecology in primates. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1865), 20171765.

R Core Team. (2021). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing.

Rees, P. A. (2013). *Dictionary of zoo biology and animal management*. Wiley-Blackwell.

Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, *1*, 103-113.

Seed, A., & Byrne, R. (2010). Animal tool-use. *Current Biology*, *20*, R1032-1039.

Seed, A., Emery, N., & Clayton, N. (2009). Intelligence in corvids and apes: A case of convergent evolution? *Ethology*, *115*(5), 401-420.

Seed, A., & Tomasello, M. (2010). Primate cognition. *Topics in Cognitive Science*, *2*, 407-419.

Servheen, C., Herrero, S., Peyton, B., Pelletier, K., Moll, K., Moll, J., & Group, I. S. P. B. S. (1999). *Bears: Status survey and conservation action plan*. IUCN.

Sol, D. (2009). The cognitive-buffer hypothesis for the evolution of large brain. In R. Dukas & J. M. Ratcliffe (Eds.), *Cognitive ecology II* (pp. 111-136). The University of Chicago Press.

Sol, D., Duncan, R., Blackburn, T., Cassey, P., & Lefebvre, L. (2005). Big brains, enhanced cognition, and response of birds to novel environments. *PNAS*, *102*, 5460-5465.

Sol, D., Griffin, A. S., & Bartomeus, I. (2012). Consumer and motor innovation in the common myna: The role of motivation and emotional responses. *Animal Behaviour*, *83*, 179-188.

Tarou, L. R. (2004). *An examination of the role of associative learning and spatial memory in foraging in two species of bear (family: Ursidae) (Ailuropoda melanoleuca, Tremarctos ornatus).* Georgia Institute of Technology.

Van Daele, L. J., Barnes, V. G., & Belant, J. L. (2012). Ecological flexibility of brown bears on Kodiak Island, Alaska. *Ursus*, *23*, 21-29, 29.

van Horik, J. O., & Madden, J. R. (2016). A problem with problem solving: Motivational traits, but not cognition, predict success on novel operant foraging tasks. *Animal Behaviour*, *114*, 189-198.

van Schaik, C., Graber, S., Schuppli, C., & Burkart, J. (2016). The ecology of social learning in animals and its link with intelligence. The *Spanish Journal of Psychology*, *19*, E99.

von Bayern, A. M., Heathcote, R. J., Rutz, C., & Kacelnik, A. (2009). The role of experience in problem solving and innovative tool use in crows. *Current Biology*, *19*, 1965-1968.

Vonk, J., & Beran, M. J. (2012). Bears "count" too: Quantity estimation and comparison in black bears (*Ursus americanus*). *Animal Behaviour*, *84*, 231-238.

Vonk, J., Jett, S. E., & Mosteller, K. W. (2012). Concept formation in American black bears, *Ursus americanus*. *Animal Behaviour*, *84*953-964.

Vonk, J., & Leete, J. A. (2017). Carnivore concepts: Categorization in carnivores "bears" further study. *International Journal of Comparative Psychology*, *30*, 1-22.

Waroff, A. J., Fanucchi, L., Robbins, C. T., & Nelson, O. L. (2017). Tool use, problem-solving, and the display of stereotypic behaviors in the brown bear (*Ursus arctos*). *Journal of Veterinary Behavior*, *17*, 62-68.

Wat, K. K. Y., Banks, P. B., & McArthur, C. (2020). Linking animal personality to problem-solving performance in urban common brushtail possums. *Animal Behaviour*, *162*, 35-45.

Weisbecker, V., Blomberg, S., Goldizen, A. W., Brown, M., & Fisher, D. (2015). The evolution of relative brain size in marsupials is energetically constrained but not driven by behavioral complexity. *Brain, Behavior and Evolution*, *85*, 125-135.

Zamisch, V., & Vonk, J. (2012). Spatial memory in captive American black bears (*Ursus americanus*). *Journal of Comparative Psychology*, *126*, 372-387.