CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	No evidence for cross-contextual consistency in spatial cognition or behavioral flexibility in a passerine
Authors	Troisi, Camille A.;Cooke, Amy C.;Davidson, Gabrielle L.;de la Hera, Iván;Reichert, Michael S.;Quinn, John L.
Publication date	2020-08-01
Original Citation	Troisi, C. A., Cooke, A. C., Davidson, G. L, de la Hera, I., Reichert, M. S., & Quinn, J. L. (2021). No evidence for cross-contextual consistency in spatial cognition or behavioral flexibility in a passerine. Animal Behavior and Cognition, 8(3), pp. 446-461. doi: 10.26451/abc.08.03.08.2021
Type of publication	Article (peer-reviewed)
Link to publisher's version	https://www.animalbehaviorandcognition.org/article.php?id=1282
Rights	© 2020 the authors. Attribution 3.0 Unported (CC BY 3.0) - https://creativecommons.org/licenses/by-sa/3.0/
Download date	2026-10-05 12:09:54
Item downloaded from	https://hdl.handle.net/10468/11817

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

ABC 2021, 8(3):446-461
Animal Behavior and Cognition
DOI: https://doi.org/10.26451/abc.08.03.08.2021
©Attribution 3.0 Unported (CC BY 3.0)

No Evidence for Cross-Contextual Consistency in Spatial Cognition or Behavioral Flexibility in a Passerine

Camille A. Troisi[1,2,*], Amy C. Cooke[1], Gabrielle L. Davidson[1,3], Iván de la Hera[1,4], Michael S. Reichert[1,4], and John L. Quinn[1,5]

[1] School of Biology, Earth and Environmental Sciences, University College Cork, Ireland
[2] Department of Experimental Psychology, Ghent University, Belgium
[3] Department of Psychology, University of Cambridge, UK
[4] Department of Integrative Biology, Oklahoma State University, USA
[5] Environmental Research Institute, University College Cork, Ireland

*Corresponding author (Email: camille.troisi@mail.mcgill.ca)

Citation – Troisi, C. A., Cooke, A. C., Davidson, G. L, de la Hera, I., Reichert, M. S., & Quinn, J. L. (2021). No evidence for cross-contextual consistency in spatial cognition or behavioral flexibility in a passerine. *Animal Behavior and Cognition, 8*(*3*), 446-461. https://doi.org/10.26451/abc.08.03.08.2021

Abstract – Although the evolution of cognitive differences among species has long been of interest in ecology, whether natural selection acts on cognitive processes within populations has only begun to receive similar attention. One of the key challenges is to understand how consistently cognitive traits within any one domain are expressed over time and across different contexts, as this has direct implications for the way in which selection might act on this variation. Animal studies typically measure a cognitive domain using only one task in one context and assume that this captures the likely expression of that domain in different contexts. This use of limited and restricted measures is not surprising because, from an ecologist's perspective, cognitive tasks are laborious to employ, and if the measure requires learning a particular aspect of the task (e.g., reward type, cue availability, scale of testing), then it is difficult to repeat the task as the learning is context specific. Thus, our knowledge of whether individual differences in cognitive abilities are consistent across contexts is limited, and current evidence suggests that consistency is weak. We tested up to 32 wild great tits (*Parus major*) to characterize the consistency of two cognitive abilities, each in two different contexts: 1) spatial cognition at two different spatial scales, and 2) behavioral flexibility as performance in a detour reaching task and reversal learning in a spatial task. We found no evidence of a correlation between individuals' performance in two measures of spatial cognition or two measures of behavioral flexibility. This suggests that cognitive performance is highly plastic and sensitive to differences across tasks, that variants of these well-known tasks may tap into different combinations of both cognitive and non-cognitive mechanisms, or that the tasks simply do not adequately measure each putative cognitive domain. Our results highlight the challenges of developing standardized cognitive assays to explain natural behavior and to understand the selective consequences of that variation.

Keywords – Cognition, Spatial cognition, Inhibitory control, Great tits, Consistency, Repeatability

The importance of evolutionary processes in explaining why individuals vary in their cognitive abilities is an emerging question in behavioral ecology. On the one hand, support for natural selection is accumulating in studies that show associations between correlates of fitness and cognitive measures (Ashton et al., 2018; Cauchard et al., 2013; Cole et al., 2012; Cole & Quinn, 2012; Keagy et al., 2009; Raine & Chittka, 2008; although see Isden et al., 2013; Sewall et al., 2013 for studies that do not find an association). On the other hand, it remains unclear whether selection on individual variation in cognitive traits will result in a meaningful response (Shaw & Schmelz, 2017). One reason for this uncertainty over

the evolutionary consequences of selection on cognition is that estimates of heritability are almost entirely lacking from natural populations (but see Langley et al., 2020; Quinn et al., 2016; reviewed in Croston et al., 2015), partly because generating reliable estimates of heritability demands large pedigrees (Quinn et al., 2006), and instead most researchers are forced to accept the phenotypic gambit, i.e., to assume that if a trait is repeatable, it is likely to be heritable, or that if two traits are phenotypically correlated, they are also genetically correlated (but see Quinn et al., 2016).

A recent meta-analysis has shown that repeatability among cognitive traits is highly variable, and that behaviors in certain types of cognitive tasks are less repeatable than others (Cauchoix et al., 2018), thus providing less reliable measures of cognition (R = .15-.28, compared to an average of .37 for other behavioral measures as in Bell et al., 2009). This is particularly the case for contextual repeatability, whereby two different tasks or the same task in different contexts aim to test the same cognitive trait, as opposed to temporal repeatability in which the exact same task is repeated over time (Cauchoix et al., 2018). Robust measures of individual differences in cognitive abilities are generally lacking, particularly within the same cognitive domain (i.e., statistically derived group of factors that capture the variance in a set of tasks; (Shaw & Schmelz, 2017; van Horik, Langley, Whiteside, Laker, & Madden, 2018). In fact, other studies examining the same domain-specific cognitive abilities found little evidence of contextual repeatability, particularly for associative learning (Boogert et al., 2011; Bray et al., 2014; Brucks et al., 2017; Guenther & Brust, 2017; Isden et al., 2013; Keagy et al., 2009; Morand-Ferron et al., 2011; Shaw et al., 2015; van Horik et al., 2019; van Horik, Langley, Whiteside, Laker, Beardsworth et al., 2018; van Horik, Langley, Whiteside, Laker, & Madden, 2018; Vernouillet et al., 2018). The gold standard would be to measure multiple traits across time and contexts in order to validate those measures of cognition (Völter et al., 2018; Vonk & Povinelli, 2011). In this study, we examined performance of two cognitive traits of key functional significance in behavioral ecology, spatial cognition across contexts (i.e., spatial scale) and behavioral flexibility across two types of cognitive tasks.

Spatial cognition, a combination of learning and memory relating to information about orientation and location, is a fundamental cognitive process that affects many aspects of an animal's ecology. For example, spatial cognition helps individuals to find food and mates, to monitor their territory boundaries, and at a larger scale, to navigate their migration routes (Healy & Hurly, 2004). Several processes are involved in different navigational methods. Animals can use view-matching strategies, such as landmark-matching or panorama-matching, where the apparent size or distance of landmarks, or the shape of the surroundings are important in matching a remembered view (Pritchard et al., 2018; Pritchard & Healy, 2018). Animals may also use strategies where the absolute distance or direction of a goal from a landmark is of importance (Pritchard et al., 2018; Pritchard & Healy, 2018). Moreover, these methods are not mutually exclusive and can be applied simultaneously in different environments or during different stages of navigation (Pritchard & Healy, 2018). In fact, evidence in the field, and in captivity, suggests that individuals rely on different information depending on the scale of their environment, or the size of their enclosure (Chiandetti et al., 2007; Healy & Hurly, 1998; Sovrano et al., 2005, 2006). Although spatial cognition is commonly referred to as a domain-specific cognitive trait (e.g., Herrmann et al., 2010), such results question whether spatial cognition should in fact be broken down into more specific mechanisms. In humans, spatial cognition abilities measured at different scales have been shown to be determined by some common processes for encoding, maintaining, and transforming spatial representation, as well as some unique processes not shared at different scales of space (Hegarty et al., 2006; Montello, 1993). Yet the majority of studies in nonhuman animals adopt tasks that are relatively small in scale (i.e., short distance between cues, relative to body size), and do not differ in context (i.e., the context is almost always in relation to foraging in a small space; e.g., Branch et al., 2019; Sewall et al., 2013; Sonnenberg et al., 2019). If animals use different cues depending on the spatial scale, we may expect performance in a spatial cognition task to differ across contexts. Nevertheless, the use of different navigational mechanisms does not preclude individual consistency across contexts, which would be suggestive of a meaningful trait that has the potential to be heritable.

Another cognitive trait that has received a lot of attention in cognitive ecology is behavioral flexibility, which allows individuals to adapt their behavior to changes in their environment (Brown &

Tait, 2014). In psychology, behavioral flexibility refers specifically to attentional shifting, rule switching, and response reversal (Brown & Tait, 2014). Recently, behavioral and cognitive ecologists have been criticized for adopting the term to broadly describe any flexible behavior, thus grouping behaviors that may be guided by different cognitive mechanisms (Audet & Lefebvre, 2017). As such, different assays of behavioral flexibility may involve different cognitive mechanisms, thereby limiting our ability to make broad inferences about what an animal's performance in a particular context might mean for adaptive responses in the wild. It is therefore important to test if there are correlations between performance on these putative measures of behavioral flexibility that will help us evaluate whether flexibility is indeed a general trait.

Two tasks have been frequently used to test how well animals respond to changes in their environment – the detour reach and reversal learning tasks. The detour reach task is thought to measure inhibitory control; i.e., an executive cognitive function that determines the ability to overcome a prepotent but disadvantageous response in favor of a more advantageous but less instinctive response. In this task, animals must avoid a transparent barrier by inhibiting a motor response to go the most direct route towards a reward and instead move around the barrier (Boogert et al., 2011; MacLean et al., 2014). Reversal learning tasks are thought to measure how flexibly animals can adjust to changes in learned contingencies, whereby a novel response is rewarded, and the previously rewarded response is not. While reversal learning can involve inhibitory control (Bari & Robbins, 2013; particularly on the first reversal as opposed to multiple reversals in which rules are formed), it also involves instrumental conditioning (extinction and relearning of reinforced stimuli; Brown & Tait, 2014).

Previous work has directly tested the association between individual performance in a detour reach and a reversal learning task, with mixed results (Anderson et al., 2017; Ashton et al., 2018; Boogert et al., 2011; Brucks et al., 2017; Shaw et al., 2015; van Horik, Langley, Whiteside, Laker, & Madden, 2018). Most work that specifically tests for correlations between performance in the detour reach and reversal learning task have used a cylinder task to measure detour reach performance, and a color association task to measure reversal learning (Anderson et al., 2017; Ashton et al., 2018; Boogert et al., 2011; Shaw et al., 2015; van Horik, Langley, Whiteside, Laker, & Madden, 2018). The measures used for both tasks vary between studies: whereas detour reach performance is mostly measured using a learning criterion where an individual is considered to have learned the detour reach task when it has completed 6/7 consecutive trials without touching the transparent cylinder (Anderson et al., 2017; Boogert et al., 2011; Shaw et al., 2015), one study used the number of pecks at the transparent cylinder during one trial as their measure (van Horik, Langley, Whiteside, Laker, & Madden, 2018), and another study used a learning criterion of three consecutive trials without touching the transparent cylinder (Ashton et al., 2018). Only the latter study reports a (positive) significant association between performance on a detour reach task, and performance on a reversal learning task. Furthermore, among the studies reported above, only Brucks et al. (2017) used a series of tasks to specifically target behavioral flexibility, rather than a battery of tasks to look at different aspects of cognition. Overall, there are a variety of methods used to study behavioral flexibility, and a wide variety of results. Moreover, all of the reversal learning tasks in such studies were based on object or color associations with no spatial dimension, so it remains unclear how reversal learning in a spatial context relates to detour reach performance. This is crucial, as performance on a detour reach task has a spatial dimension (van Horik et al., 2020), albeit at a local scale. Thus, although there is reason to expect that detour reaching and reversal learning correlate due to an overlap in a domain general cognitive mechanism (i.e., inhibitory control), other mechanisms may be exclusive to one task only, raising uncertainty as to whether this is necessarily the case.

The great tit (*Parus major*) is a model species for ecological and behavioral studies (e.g., Aplin et al., 2015; Cole et al., 2011; Dutour et al., 2020; Loukola et al., 2020; Morand-Ferron et al., 2011). Great tits adapt well to temporary captivity, allowing for their use in controlled experiments on individual differences. Here we investigated consistency in spatial cognition and behavioral flexibility across contexts. We measured spatial cognition at two different spatial scales: at a large scale within an experimental room (460 [W] x 310 [L] x 265 [H] cm), and at a smaller scale within the birds' individual home cages (46 [W] x 56 [L] x 57 [H] cm). We also measured behavioral flexibility across two different

tasks (also at different scales): reversal learning, using a spatial feeder array in the experimental room, and a detour apparatus in the home cage (MacLean et al., 2014). Our study had two objectives: first, to examine whether a spatial cognition task conducted in a home cage predicts measures of the same putative cognitive mechanisms at a larger scale. Second, to investigate whether behavioral flexibility measured by the detour reach task predicts the behavior of birds in a reversal learning task. If spatial cognition at different spatial scales is determined by the same cognitive mechanism, then we would expect our measure of spatial cognition in the home cage to be correlated to our measure of spatial cognition in the experimental room. By contrast, if spatial cognition at different scales requires different mechanisms, or attention to different cues, then we would not expect such a correlation. Similarly, if behavioral flexibility is determined by inhibitory control, then we would expect measures emerging from our two tasks to be correlated. However, if behavioral flexibility is a consolidated measure of different cognitive processes, or if experimental design has a strong effect on performance, then we expect measures emerging from those tasks to be uncorrelated with each other (Miyake et al., 2000).

Method

Subjects

Wild-caught great tits (N = 36) from the Bandon Valley, County Cork, Ireland, were brought into captivity, and later released at their capture site upon completing the experiments (O'Shea et al., 2018). Birds were captured using mist nets, and therefore, our sample is likely biased towards trappable individuals (Webster & Rutz, 2020). Each bird was fitted with a BTO ring for individual identification and a Passive Integrated Transponder (PIT) tag. Birds were housed individually, in 46 (W) x 56 (L) x 57 (H) cm plywood cages, with two perches each and with an internal light set from 7:30 to 18:00. Birds were in auditory contact with each other, but not in visual contact. Birds had *ad libitum* access to food and water. Food consisted of sunflower hearts, peanuts, mealworms, and waxworms. Out of the 36 birds brought into captivity, 28 birds learned the large-scale initial spatial cognition task, and 25 of those also learned the large-scale reversal spatial cognition task. Out of the 36 birds, 30 learned the small-scale spatial cognition task, and 32 completed all 10 trials of the detour reach task. Each bird was tested in isolation, in its home cage during the small-scale spatial cognition task and detour reach task, or in the experimental room during the large-scale initial and reversal spatial cognition tasks. Data collection took place between January and March, 2019, and the mean time that birds were in captivity was 12 days. In terms of task order, the large-scale initial spatial cognition task was always followed by the large-scale reversal spatial cognition task, which was always followed by the small-scale spatial cognition task. Training for the detour reach task started simultaneously with training for the large-scale initial spatial cognition task, so the order (and test intervals) of the detour reach task in relation to the large-scale initial and reversal spatial cognition tasks depended on how long it took each bird to train for and learn each task. The detour reach test always took place before the small-scale spatial cognition task.

Materials

Large-scale Spatial Cognition Task

The large-scale (initial and reversal) spatial cognition task took place in an experimental room of 460 (W) x 310 (L) x 265 (H) cm, with four feeders containing sunflower seeds placed in a square with sides of one meter, and a small plastic Christmas tree (150 cm high) placed in the center for birds to rest and hide. Feeders were equipped with RFID readers to remotely log each visit by detecting the individual's PIT tag.

Small-scale Spatial Cognition Task

The small-scale spatial cognition task took place in the birds' home cage (46 [W] x 56 [L] x 57 [H] cm). Individuals were given artificial food items designed to mimic seeds/insect prey enclosed in an outer shell (Ihalainen et al., 2007). In our study, these artificial food items consisted of a sunflower seed encased in a paper parcel (1.8 cm x 1.8 cm).

Detour Reach Task

The detour reach task also took place in the birds' home cage. For this task, birds were required to retrieve a waxworm from inside a transparent cylinder, requiring them to make a detour around the cylinder to obtain the reward (Boogert et al., 2011; MacLean et al., 2014). The cylinder (3 cm length, 3.5 cm diameter) was made from plastic sheeting, open at both ends, and glued onto a cardboard base (7 cm x 20 cm). We also added a small perch (8 cm wide, 8 cm high) parallel to the cylinder, to avoid any biases in approach direction. The task had three phases: habituation, training and testing, and a waxworm was used as a reward in all phases. During the habituation and training phases, the cylinder was opaque (black plastic), whereas it was transparent during the test.

Procedure

Large-scale Spatial Cognition Task

The large-scale spatial cognition task had four phases: habituation, training, initial learning test phase and reversal learning test phase. Birds were food deprived 1 hour before each trial, which consisted of a 1 hr "block" where the bird was placed in the experimental room, and was allowed to move freely within that experimental room and visit the feeders as many times as it wanted within that hour. Trials were randomized each day, to account for time-of-day effects on performance. The number of trials each bird had within a phase varied between birds, as this depended on when they reached the criterion for each phase. Each phase took place at least 12 hrs apart (i.e., typically the following day), and birds accessed the experimental room from a small opening in their home cage. In the habituation phase, food was accessible and visible in all four feeders. Birds had to eat 10 seeds within 1 hr (based on seed husk collected on the floor after the trial) before progressing to the training phase. In the training phase, an opaque paper sheet concealed seeds in each feeder, so the food was no longer visible by the birds except from the RFID reader platform, from which it was also accessible. Once a bird visited any of the four feeders a total of 10 times (based on logged visits from RFID reader), they were advanced to the initial learning phase of the testing trials. During the testing trials, all feeders remained wrapped in paper but only one (randomly assigned) feeder contained food. The criterion for having learned the feeder position was to have visited the correct feeder 8 times within a moving window of 10 visits (Guenther & Brust, 2017). This success rate is significantly different from the expectation if birds selected feeders at random (binomial test, $p < .001$). The number of visits to reach criterion was used as a measure of learning. Once the birds met the criterion, they were advanced to the reversal learning phase, in which a new feeder was allocated as the rewarded feeder. The same criterion was used as for the initial learning phase: 8 visits to the correct feeder within a moving window of 10 visits. We also calculated the time (in seconds) it took for each bird to reach criterion, which took into account only the time that birds spent in the trials, not the time in between trials.

In all phases, birds were given one hour to visit the feeders before being returned to their home cage. Data from the loggers were reviewed on the same day to determine whether birds had met the criterion. Because of time constraints and welfare issues, if the birds did not reach their criterion within five trials, they were released and excluded from the analysis. Out of the 36 birds brought into the aviary, three never reached the initial learning phase (they did not complete the training) and were released early,

and another four did not learn the initial cognition task within five trials and were released as well. Of the 29 birds that learned the initial cognition task and progressed to the reversal cognition task, one bird was released because it was exhibiting stress-related behavior during the reversal cognition task, and three were released because they did not learn the reversal cognition task within five trials. For the initial cognition experiment, most birds (n = 22 out of 29 birds) learned over multiple trials (> 1 block of 1 hr), whereas for the reversal cognition, half of them learned over multiple trials (n = 13 out of 25 birds).

These data were collected as part of another experiment (Cooke et al., in prep) in which birds were exposed to different levels of simulated predation risk (treatment), during both the initial learning and reversal learning part of the experiment. We found no evidence that treatment in the previous experiment (Cooke et al., in prep) affected behavior in the current one, so we analyzed all individuals together (but see Supplementary Material for analysis of birds with no perceived predation risk -control birds).

Small-Scale Spatial Cognition Task

Adapting from Ihalainen et al. (2007), all birds were trained to handle the artificial food in their home cages in four steps in which they had to consume the seeds before advancing to the next step: (i) five food items with the seed sticking out from each parcel; (ii) five food items with the seed inside each parcel, but with a hole in the middle showing the seed inside; (iii) five food items with the seed completely hidden inside each parcel; and finally, (iv) five food items with the seed completely hidden inside of each parcel with three rewarded (i.e., seed) and two unrewarded (i.e., made the reward inaccessible once the parcel was opened by wrapping the seed in duct tape). In this last step, unrewarded items were used so that the birds would learn that not all parcels had accessible food. The birds had to eat all items before the training progressed to the next phase or eat three rewarded food items and open two unrewarded food items in the last phase of training to proceed to the testing phase. In each training step, parcels were placed centrally in the home cage in a small dish.

The testing phase consisted of ten parcels placed in a small dish in each corner of the cage. Three of those locations contained only parcels that were unrewarded, while all parcels from the rewarded location contained seeds. The rewarded corner was allocated to each bird randomly. Every time the bird made 10 choices, irrespective of their location, each corner was rebaited to have 10 total parcels, so that the amount of parcels at each corner would not act as a cue for the bird.

During training and testing, birds had no *ad libitum* access to food, and had access to food only through the parcels. Individuals were not food deprived beforehand to allow for longer training and testing sessions. Water was still available *ad libitum*. Birds were trained and tested for a maximum of three hours consecutively. After three hours, their *ad libitum* food was replaced in their cages. During training and testing, if a bird had not eaten any food for the last 1.5 hrs, or less than 10 seeds for the last 2 hrs, then the training/testing was stopped, and *ad libitum* food was placed back in their cages.

Trials consisted of periods of 3h where the birds' *ad libitum* food was replaced by the parcels in each corner. Trials stopped either once the bird had learned where to find seeds, based on a criterion of 8 correct visits in a moving window of 10 visits, or if they were not eating enough (see above). The number of choices to reach criterion was used as a measure of learning. We also calculated the time (in seconds) it took for each bird to reach criterion, which considered only the time that birds spent in the trials, not the time in between trials. The number of trials each bird took part in for this experiment varied between birds, as this depended on when they reached the learning criterion. Out of the 32 birds that took part in this experiment (four of the 36 captured birds were released before the start of this experiment because of restless behavior and disinterest in the other tasks), two stopped participating in their second and third trial, when the experiment was stopped for these birds. Of the 30 birds that did learn, 17 learned within one trial, eight learned in two trials, three learned in three trials, and two birds learned in four trials (average: 1.67 trials). For the 13 birds that learned over more than one trial, six birds had the 8/10 successful visits (of the moving window) within one trial, while seven birds had the 8/10 successful visits (of the moving window) overlap several trials. Birds were given *ad libitum* food in between trials.

Detour Reach Task

The detour reach task had three phases: habituation, training, and testing. Before habituation and training, birds were food deprived for 1 hr. For habituation, training, and testing, birds had *ad libitum* access to water and had access to food only through the task.

Birds were first habituated to ensure they were not fearful towards the novel apparatus. For habituation, birds were required to eat a waxworm placed in front of the opaque cylinder. Once they had completed this task three times, birds advanced to the training phase of the experiment. During training, birds were required to eat a waxworm placed in the middle of the opaque apparatus, by reaching around the cylinder into the open end without touching the exterior of the tube. Training was repeated until birds ate the food without touching the outside of the cylinder four times, at which point birds could advance to the test phase. During the test phase, birds were presented the transparent apparatus with a waxworm placed in the middle. Birds were scored either a success (obtaining the worm without touching the tube) or a failure (touching the outside of the tube prior to obtaining the food) and the apparatus was then removed from the cage. This was repeated ten times in succession on the same day (except for one bird that was tested over two days due to restless behavior and a disinterest in the task). Performance on this task was quantified as the number of successes out of ten. The test phase always occurred at least the day after the training phase was completed: 18 birds were tested the day after their training, 15 birds two days later (one of those 15 birds started testing but later had to stop because of restless behavior and disinterest in the task, so is not included in the analysis) and one bird was tested three days later owing to time constraints (M = 1.5, SD = 0.56).

Statistical Analysis

All analyses were conducted in R (R Core Team, 2019). Data followed a Poisson distribution, and therefore Kendall's Tau correlations were conducted for non-normal distributions. We conducted a correlation between the performance on the large-scale initial spatial cognition task and the small-scale spatial cognition task, and between the performance on the detour reach task and the large-scale reversal spatial cognition task. Some birds learned on their first choice (birds could have reached the criterion on their first visit if that first visit was correct and, if out of the nine following visits, seven of them were to the rewarded food source). As we cannot say whether they learned or made a correct first choice by chance and then just continued to feed from a rewarding feeder (large-scale initial spatial cognition task n = 1; large-scale reversal spatial cognition task n = 3; small-scale spatial cognition task n = 2), we also provide the analysis only on the birds that did not learn on their first choice, in the supplementary material. Because there was variation in inter-trial intervals for the small-scale spatial cognition task (see Method), we also ran a correlation between the small-scale and large-scale initial spatial cognition tasks on birds that learned the small-scale spatial cognition task in just one trial (n = 17). This enabled us to examine whether the results held when there was no variation in inter-trial intervals. For the two spatial cognition tasks, we also compared overall learning speeds in terms of number of visits, and time, using generalized linear mixed-models with a Poisson error distribution and log function (Bates et al., 2015). Model assumptions were checked using DHARMa (Hartig, 2020). Because of evidence of overdispersion, we added an observation-level random effect (Harrison, 2014). Repeatability analysis was conducted for the spatial cognition tasks, using the *rptR* package (Stoffel et al., 2017). We provide both adjusted and non-adjusted repeatability, using a Poisson distribution, and the adjusted repeatability includes the following fixed factors: the experiment (large-scale initial spatial cognition task or small-scale spatial cognition task), as well as the age and sex of each individual. A summary of which birds took part in which experiment, as well as whether their data was excluded, is available in Table S1. In light of the STRANGE framework - a framework that helps identify potential biases in the sample, and the representativeness of the sample - we also provide the sex, age and weight for each of the 36 birds that were brought into temporary captivity (Webster & Rutz, 2020).

Data Availability

The dataset analyzed during the current study, and the R code used to analyze them are available on the Open Science Framework repository: https://osf.io/khf3m/ (DOI: 10.17605/OSF.IO/KHF3M).

Ethics

We performed the experiment in accordance with the Association for the Study of Animal Behaviour guidelines for the Treatment of Animals in Behavioral Research and Teaching, and the Animal Welfare Body of the University College Cork approved the study, under the number 2014/014 "The evolutionary and behavioral ecology of birds." This study was conducted under licenses from the Health Products Regulatory Authority (AE19130/P017), the National Park and Wildlife Services (C01/2019) and the British Trust for Ornithology, and permission from Coillte Forestry and private landowners.

Results

On average, birds that took part in both tasks (n = 26) made 30.65 (SE = 4.62) choices before learning the large-scale spatial cognition task, and 22.77 (SE = 3.30) choices before learning the small-scale spatial cognition task. There was little evidence of a difference in the learning speed between the two tasks (model estimate = 0.15; 95% CI [-0.31, 0.59]; p = .492), and in the time (seconds) it took to learn the two tasks (model estimate = -0.26; 95% CI [-0.70, 0.18]; p = .233). Birds took on average 135.35 (SE = 12.28) minutes to learn the large-scale spatial cognition task, while it took them 130.46 (SE = 21.97) minutes to learn the small-scale spatial cognition task. There was no evidence for consistent performance in the two tasks based on a correlation analysis (z = 0.99; tau = 0.14; p = .320; Figure 1A), and a repeatability analysis (Table S2). When examining only birds that learned the small-scale spatial task in one trial, we similarly found no evidence for consistent performance in the two tasks (z = 0.86; tau = 0.16; p = .391, n = 16).

On average, birds that took part in both tasks (n = 25) made 17.44 (SE = 2.97) choices before reversal learning, and made 4.04 (SE = 0.50) correct choices in the detour reach task. There was no evidence of a correlation between the detour reach and reversal learning (z = 0.59; tau = 0.09; p = .554; Figure 1B).

Discussion

Here we report a lack of correlation among any of the traits measured in our tasks, which we discuss in the context of consistency, cognitive domains and other potentially confounding factors like task design.

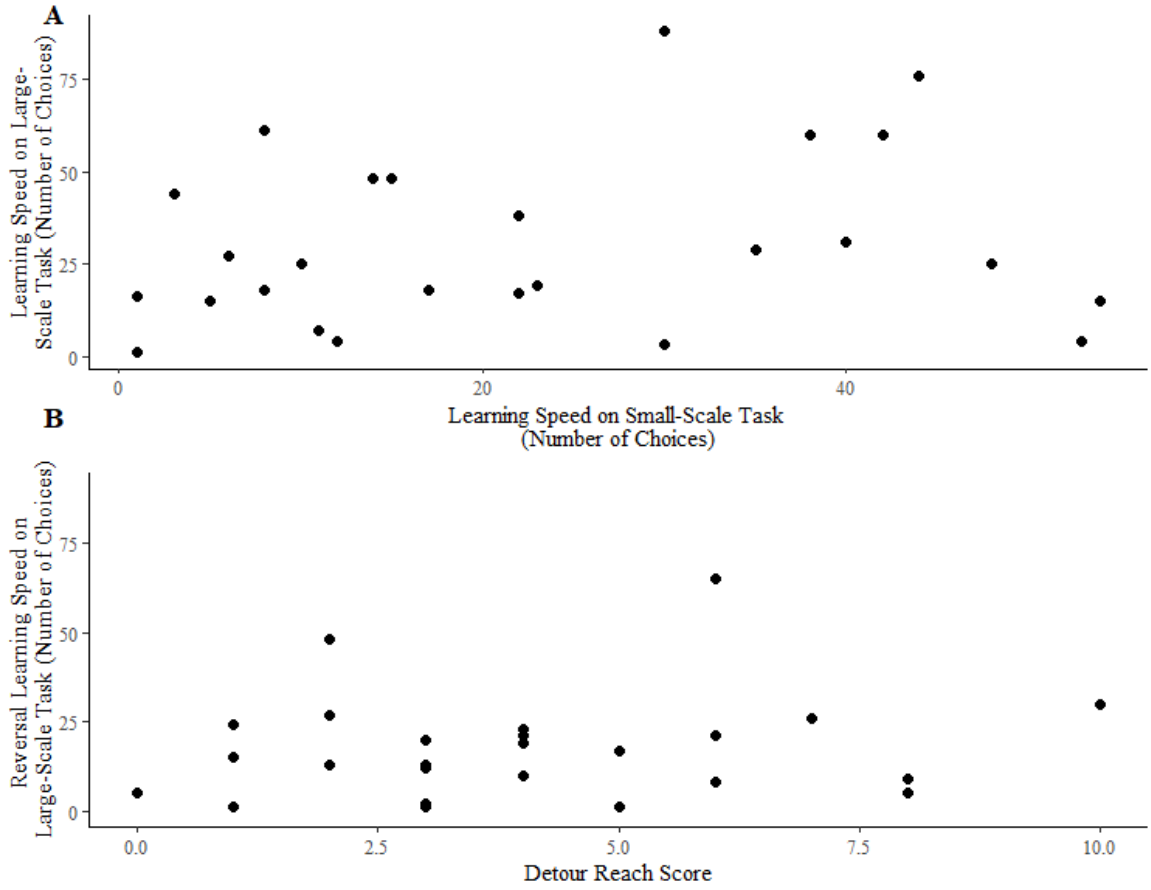
Spatial Cognition

We did not find a correlation between spatial cognition performance in the small-scale setting of the home cage and in the larger scale setting of the experimental room. As far as we are aware, this is the first attempt to directly compare performance on a spatial cognition task at different scales in non-human animals (but see Hegarty et al., 2006; Montello, 1993 for work in humans). Theoretically, both tasks we used should measure spatial cognition because they required individuals to use and remember specific location cues with food rewards, allowing them to return to that location more often than one would expect by chance (Olton, 1977). One possible explanation for the lack of correlation between cognitive tasks is that non-cognitive factors - for example motivation, inter-trial and inter-test intervals, personality,

stress, the external environment, and motor skills - influenced performance differently across tasks (Schubiger et al., 2020). Although we had no a priori reason to expect this might be the case, it must remain a possibility since we did not control for these effects. Inter-trial and inter-test intervals varied among individuals, and this has the potential to lead to confounding effects of short- and long-term memory. However, when analyzing data only from birds that learned the small-scale spatial cognition task in one trial (this task had the most variation in inter-trial intervals of the two tasks) and comparing it to the birds' performance on the large-scale initial spatial cognition task, we similarly found no evidence of an association between the performance on those two tasks. Another study with great tits also found no effects of the amount of time between an individual's visits on learning speed (Reichert et al., 2020). Hence, variation in inter-trial interval is unlikely to explain the lack of evidence of a correlation between spatial cognition performance in the small-scale and in the larger scale setting.

Figure 1

Relation Between Performance on Small-Scale and Large-Scale Spatial Cognition Tasks and Performance on Detour Reach and Reversal Cognition Tasks



Note. A) Small-scale and large-scale spatial cognition tasks. B) Detour reach and reversal cognition tasks

A more plausible explanation is that different kinds of cues or strategies were used to recognize locations in our two tasks and that this confounded any correlation in cognitive performance (Morgan et al., 2014). Work in chimpanzees and humans suggests that performance on spatial tasks loads on the same factor (Herrmann et al., 2010). However, work in captivity has shown that several species are more reliant on cues from the geometry of the room when they have to navigate in small enclosures, and absolute direction to landmarks when they have to navigate large enclosures (Chiandetti et al., 2007; Sovrano et

al., 2005, 2006). Visual information might also change at a faster pace in the home cage than in the exploration room, creating a potential bias against relying on optical-flow in the home cage (Pritchard et al., 2018). If these different perceptual abilities are either not correlated among individuals (Healy et al., 2009; Jones & Healy, 2006; Pike et al., 2018; Sovrano et al., 2003), or are independent of other processes involved with spatial learning, for example memory (Tello-Ramos et al., 2018), this could readily explain the lack of a correlation in our data (Rowe & Healy, 2014b). In this case, categorizing the sensory information available to the individuals could explain individuals' spatial cognition performances (Pritchard et al., 2017). The lower cost of making errors in a smaller environment could also explain why we find no evidence of a correlation (Zamisch & Vonk, 2012). If reward locations are closer together (as in the small-scale spatial cognition task), then making errors (i.e., visiting the wrong location), is not very costly in time and energy, compared to visiting different reward locations that are far apart. In a smaller environment, individuals could therefore visit every location in an efficient manner without the need to learn which location is rewarded (Zamisch & Vonk, 2012). However, in our experiment, we found no evidence that the small-scale spatial cognition task was more difficult to learn than the large-scale initial spatial cognition task. Our results are in contrast to findings (in humans) that, despite involving some different mechanisms, learning at different spatial scales was still found to be partly determined by a common process for encoding, maintaining, and transforming spatial representation (Hegarty et al., 2006; Montello, 1993). Measuring temporal repeatability of spatial cognitive tasks could help tease apart whether the lack of correlation found in this study is underpinned by differing cognitive traits or non-cognitive factors influencing performance. High temporal repeatability would indicate that the performance in each of the two tasks is influenced by different cognitive mechanisms, while low temporal repeatability would more likely indicate an effect of confounding factors. However, high temporal repeatability could also indicate that the same confounding factors have the same effects in both tasks.

Whatever the reason for the lack of correlation, subtle differences in experimental design (e.g., environments or reward type) may preclude meaningful comparison across studies (Thornton & Lukas, 2012). Furthermore, using only one task is unlikely to capture the expression of spatial cognition in different contexts (Pritchard et al., 2017). If individuals use different cognitive processes for spatial cognition in different environments, or under different conditions (e.g., season, stress), then it becomes crucial to measure spatial cognition using several tasks to better understand individual variation in the underlying processes. For example, a current focus of research in food caching species is to measure spatial cognition under standardized conditions, usually on a small scale, to infer performance in cache retrieval at a larger spatial scale, but it remains unclear whether observed associations relate to navigation through the environment, or cache retrieval at a fine spatial scale (Healy, 2019; Healy et al., 2005, 2009; Krebs et al., 1990; McGregor & Healy, 1999). Similarly, work on parasitic cowbirds in which the females, but not males, need to locate and remember potential host nests, have found that females out-perform males in some spatial cognition tasks, but not others, further suggesting that individual differences in spatial ability may depend on task design and scale of spatial location (Sherry & Guigueno, 2019). This limitation mirrors the more general problem in evolutionary ecological studies of cognition, where there is often a lack of a clear link between standardized cognitive tests (e.g., problem solving) and functional behavior (innovative foraging) under natural conditions (Morand-Ferron et al., 2016; Rowe & Healy, 2014a, b; Thornton et al., 2014).

Behavioral Flexibility

Behavioral flexibility is a loosely defined term (Audet & Lefebvre, 2017), and based on previous work, as well as our results, performance on a detour reaching task and a reversal spatial cognition task may not be underpinned by the same cognitive mechanism. However, both tasks measure an aspect of behavioral flexibility, and although they might measure different traits, they could fall under the same executive function umbrella, and may therefore be correlated. However, we did not find any evidence of an association between the performance of the birds on a detour reach and reversal spatial cognition task, despite the prediction that both measures of behavioral flexibility would be correlated because they

involve inhibitory control. This lack of correlation between detour reaching and reversal cognitive tasks is in keeping with previous work in wild and captive birds that have also used single reversals (song sparrows, *Melospiza melodia*: Anderson et al., 2017; Boogert et al., 2011; New Zealand robin, *Petroica longipes*: Shaw et al., 2015; pheasants, *Phasianus colchicus*: van Horik, Langley, Whiteside, Laker, & Madden, 2018, but is dissimilar to Ashton et al.'s (2018) findings in Australian magpies, *Cracticus tibicen dorsalis*). It is interesting to note that experimental designs and measures vary between those studies, and a standardized test would be greatly beneficial. However, standardizing tests across species might prove difficult, since, as our results suggest, scale and context matter for behavioral flexibility, and this is likely to be different for every species. Yet, our study increases the generality of this finding on the lack of correlations between performance on tasks measuring behavioral flexibility, because these previous studies focused on color discrimination reversal, rather than the spatial discrimination reversal we used here.

There are multiple possible explanations for why these putative measures of behavioral flexibility were not correlated. One is that these different tasks, in reality, measure different components of inhibitory control (namely stopping, delaying, or withholding motor responses in the first place; Bari & Robbins, 2013; Bray et al., 2014; Brucks et al., 2017; van Horik, Langley, Whiteside, Laker, Beardsworth et al., 2018; Vernouillet et al., 2018). Moreover, although our results suggest that components of inhibitory control do not fall under the same general domain umbrella of behavioral flexibility, the lack of correlation may also be explained by the reversal cognition task also requiring a spatial cognition component (Boogert et al., 2018; Brucks et al., 2017; Miyake et al., 2000). Our measure during the reversal spatial cognition task may involve a mix of behavioral flexibility, and spatial cognition, and the effect of spatial cognition on performance on the reversal spatial cognition task could mask the relation between performance on the reversal spatial cognition task and the detour reach task. However, we find this explanation unlikely because several studies found little consistency between initial learning and reversal learning in spatial tasks (e.g., Reichert et al., 2020), and the detour reach task also requires a spatial component (van Horik et al., 2020). Both detour reach and reversal learning probably tap into multiple cognitive domains, which makes a correlation between the two less likely. Discrepancies could also be explained by differences in motivation, which may have played a bigger role in the detour reach task as the food was a visible worm, compared to the reversal spatial cognition task where the food was a non-visible seed. Without measuring the degree to which individuals differ in their preference of one food type over the other, this interpretation is only speculative. Finally, although we acknowledge that the detour reach task has been recently criticized as a measure of inhibitory control and is subject to being influenced by various factors that are difficult to control and were not controlled here – for example perception, stress, prior experience (Kabadayi et al., 2018; van Horik, Langley, Whiteside, Laker, Beardsworth et al., 2018) - this is likely true of all cognitive tasks (Morand-Ferron & Quinn, 2015).

Conclusions

Our results across both task comparisons highlight that caution needs to be taken when making conclusions about learning speed or behavioral flexibility based on a single test, because performance may be highly sensitive to the context and type of task. On the one hand, if the lack of evidence for correlations reflects true cognitive differences related to either spatial cognition or behavioral flexibility, then this would point towards greater complexity in how cognitive processes that drive animal navigation and behavioral plasticity interact with the individual's environment to enable specific behavior. On the other hand, if the lack of correlation arises because of confounding effects that were not controlled for, or because one or both of the tasks within each domain does not measure spatial cognition or behavioral flexibility as expected, then this would point towards a common issue with experimental design used in cognitive tests. Either way, the context in which we measure cognition is essential to consider if we want to better understand causes and consequences of individual variation in cognition. Further investigation into the neurobiology related to performance in tasks which *a priori* measure the same cognitive

processes may facilitate progress in validating cognitive tasks (van Horik, Langley, Whiteside, Laker, Beardsworth et al., 2018), and distinguishing which of our interpretations are most valid. However, this approach may not often be feasible in non-model, wild animals. Pinning down the meaningful measures of individual differences in cognitive mechanisms remains a major challenge. Nevertheless, studies that aim to validate tasks, as we do here, are a step forward in understanding causes and consequences of individual variation in cognition.

Acknowledgements

We would like to thank Sam Bayley, Ciara Sexton, and Karen Cogan for capturing and releasing birds. We are also grateful to Allen Whitaker for help in coordinating the aviary building and Karen Cogan for the administrative support throughout the project.

Author Contributions

Conceptualization: CAT, JLQ, ACC, GLD; Data collection: CAT, ACC, I de la H, MSR; Analysis: CAT; Writing: CAT and JLQ, with contributions from all authors. This work was supported by funding from the European Research Council under the European Union's Horizon 2020 Programme (FP7/2007-2013)/ERC Consolidator Grant "Evoecocog" Project No. 617509 to JLQ, and by a Science Foundation Ireland ERC Support Grant 14/ERC/B3118 to JLQ

References

Anderson, R. C., Searcy, W. A., Peters, S., Hughes, M., DuBois, A. L., & Nowicki, S. (2017). Song learning and cognitive ability are not consistently related in a songbird. *Animal Cognition*, *20*(2), 309–320. https://doi.org/10.1007/s10071-016-1053-7

Aplin, L. M., Farine, D. R., Morand-Ferron, J., Cockburn, A., Thornton, A., & Sheldon, B. C. (2015). Experimentally induced innovations lead to persistent culture via conformity in wild birds. *Nature*, *518*(7540), 538–541. https://doi.org/10.1038/nature13998

Ashton, B. J., Ridley, A. R., Edwards, E. K., & Thornton, A. (2018). Cognitive performance is linked to group size and affects fitness in Australian magpies. *Nature*, *554*(7692), 364–367. https://doi.org/10.1038/nature25503

Audet, J.-N., & Lefebvre, L. (2017). What's flexible in behavioral flexibility? *Behavioral Ecology*, *28*(4), 943–947. https://doi.org/10.1093/beheco/arx007

Bari, A., & Robbins, T. W. (2013). Inhibition and impulsivity: Behavioral and neural basis of response control. *Progress in Neurobiology*, *108*, 44–79. https://doi.org/10.1016/j.pneurobio.2013.06.005

Bates, D., Machler, M., Bolker, B. M., & Walker, S. C. (2015). *Fitting linear mixed-effects models using lme4*.

Bell, A. M., Hankison, S. J., & Laskowski, K. L. (2009). The repeatability of behaviour: A meta-analysis. *Animal Behaviour*, *77*(4), 771–783. https://doi.org/10.1016/j.anbehav.2008.12.022

Boogert, N. J., Anderson, R. C., Peters, S., Searcy, W. A., & Nowicki, S. (2011). Song repertoire size in male song sparrows correlates with detour reaching, but not with other cognitive measures. *Animal Behaviour*, *81*(6), 1209–1216. https://doi.org/10.1016/j.anbehav.2011.03.004

Boogert, N. J., Madden, J. R., Morand-Ferron, J., & Thornton, A. (2018). Measuring and understanding individual differences in cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1756), 20170280. https://doi.org/10.1098/rstb.2017.0280

Branch, C. L., Pitera, A. M., Kozlovsky, D. Y., Bridge, E. S., & Pravosudov, V. V. (2019). Smart is the new sexy: Female mountain chickadees increase reproductive investment when mates to males with better spatial cognition. *Ecology Letters*, *22*(6), 897–903. https://doi.org/10.1111/ele.13249

Bray, E. E., MacLean, E. L., & Hare, B. A. (2014). Context specificity of inhibitory control in dogs. *Animal Cognition*, *17*(1), 15–31. https://doi.org/10.1007/s10071-013-0633-z

Brown, V. J., & Tait, D. S. (2014). Behavioral flexibility: Attentional shifting, rule switching, and response reversal. In I. P. Stolerman & L. H. Price (Eds.), *Encyclopedia of Psychopharmacology* (pp. 1–7). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-27772-6_340-2

Brucks, D., Marshall-Pescini, S., Wallis, L. J., Huber, L., & Range, F. (2017). Measures of dogs' inhibitory control abilities do not correlate across tasks. *Frontiers in Psychology*, *8*, 849. https://doi.org/10.3389/fpsyg.2017.00849

Cauchard, L., Boogert, N. J., Lefebvre, L., Dubois, F., & Doligez, B. (2013). Problem-solving performance is correlated with reproductive success in a wild bird population. *Animal Behaviour*, *85*(1), 19–26. https://doi.org/10.1016/j.anbehav.2012.10.005

Cauchoix, M., Chow, P. K. Y., van Horik, J. O., Atance, C. M., Barbeau, E. J., Barragan-Jason, G., Bize, P., Boussard, A., Buechel, S. D., Cabirol, A., Cauchard, L., Claidière, N., Dalesman, S., Devaud, J. M., Didic, M., Doligez, B., Fagot, J., Fichtel, C., Henke-von der Malsburg, J., … Morand-Ferron, J. (2018). The repeatability of cognitive performance: A meta-analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1756), 20170281. https://doi.org/10.1098/rstb.2017.0281

Chiandetti, C., Regolin, L., Sovrano, V. A., & Vallortigara, G. (2007). Spatial reorientation: The effects of space size on the encoding of landmark and geometry information. *Animal Cognition*, *10*(2), 159–168. https://doi.org/10.1007/s10071-006-0054-3

Cole, E. F., Cram, D. L., & Quinn, J. L. (2011). Individual variation in spontaneous problem-solving performance among wild great tits. *Animal Behaviour*, *81*(2), 491–498. https://doi.org/10.1016/j.anbehav.2010.11.025

Cole, E. F., Morand-Ferron, J., Hinks, A. E., & Quinn, J. L. (2012). Cognitive ability influences reproductive life history variation in the wild. *Current Biology*, *22*(19), 1808–1812. https://doi.org/10.1016/j.cub.2012.07.051

Cole, E. F., & Quinn, J. L. (2012). Personality and problem-solving performance explain competitive ability in the wild. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1731), 1168–1175. https://doi.org/10.1098/rspb.2011.1539

Cooke, A. C., Davidson, G. L., Reichert, M. S., & Quinn, J. L. (In prep). *Non-lethal effects of predators on prey in the context of spatial learning and cognitive flexibility*.

Croston, R., Branch, C. L., Kozlovsky, D. Y., Dukas, R., & Pravosudov, V. V. (2015). Heritability and the evolution of cognitive traits. *Behavioral Ecology*, *26*(6), 1447–1459. https://doi.org/10.1093/beheco/arv088

Dutour, M., Suzuki, T. N., & Wheatcroft, D. (2020). Great tit responses to the calls of an unfamiliar species suggest conserved perception of call ordering. *Behavioral Ecology and Sociobiology*, *74*(3), 37. https://doi.org/10.1007/s00265-020-2820-7

Guenther, A., & Brust, V. (2017). Individual consistency in multiple cognitive performance: Behavioural versus cognitive syndromes. *Animal Behaviour*, *130*, 119–131. https://doi.org/10.1016/j.anbehav.2017.06.011

Harrison, X. A. (2014). Using observation-level random effects to model overdispersion in count data in ecology and evolution. *PeerJ*, *2*, e616. https://doi.org/10.7717/peerj.616

Hartig, F. (2020). *DHARMa: Residual diagnostics for hierarchical (Multi-Level / Mixed)* (0.2.7) [Computer software]. https://CRAN.R-project.org/package=DHARMa

Healy, S. D. (2019). Animal cognition. *Integrative Zoology*, *14*(2), 128–131. https://doi.org/10.1111/1749-4877.12366

Healy, S. D., Bacon, I. E., Haggis, O., Harris, A. P., & Kelley, L. A. (2009). Explanations for variation in cognitive ability: Behavioural ecology meets comparative cognition. *Behavioural Processes*, *80*(3), 288–294. https://doi.org/10.1016/j.beproc.2008.10.002

Healy, S. D., Dekort, S., & Clayton, N. (2005). The hippocampus, spatial memory and food hoarding: A puzzle revisited. *Trends in Ecology & Evolution*, *20*(1), 17–22. https://doi.org/10.1016/j.tree.2004.10.006

Healy, S. D., & Hurly, T. A. (1998). Rufous hummingbirds' (*Selasphorus rufus*) memory for flowers: Patterns or actual spatial locations? *Journal of Experimental Psychology*, *24*(4), 296–404. https://doi.org/10.1037/0097-7403.24.4.396

Healy, S. D., & Hurly, T. A. (2004). Spatial learning and memory in birds. *Brain, Behavior and Evolution*, *63*(4), 211–220. https://doi.org/10.1159/000076782

Hegarty, M., Montello, D. R., Richardson, A. E., Ishikawa, T., & Lovelace, K. (2006). Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence*, *34*(2), 151–176. https://doi.org/10.1016/j.intell.2005.09.005

Herrmann, E., Hernández-Lloreda, M. V., Call, J., Hare, B., & Tomasello, M. (2010). The structure of individual differences in the cognitive abilities of children and chimpanzees. *Psychological Science*, *21*(1), 102–110. https://doi.org/10.1177/0956797609356511

Ihalainen, E., Lindström, L., & Mappes, J. (2007). Investigating Müllerian mimicry: Predator learning and variation in prey defences. *Journal of Evolutionary Biology*, *20*(2), 780–791. https://doi.org/10.1111/j.1420-9101.2006.01234.x

Isden, J., Panayi, C., Dingle, C., & Madden, J. (2013). Performance in cognitive and problem-solving tasks in male spotted bowerbirds does not correlate with mating success. *Animal Behaviour*, *86*(4), 829–838. https://doi.org/10.1016/j.anbehav.2013.07.024

Jones, C. M., & Healy, S. D. (2006). Differences in cue use and spatial memory in men and women. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1598), 2241–2247. https://doi.org/10.1098/rspb.2006.3572

Kabadayi, C., Bobrowicz, K., & Osvath, M. (2018). The detour paradigm in animal cognition. *Animal Cognition*, *21*(1), 21–35. https://doi.org/10.1007/s10071-017-1152-0

Keagy, J., Savard, J.-F., & Borgia, G. (2009). Male satin bowerbird problem-solving ability predicts mating success. *Animal Behaviour*, *78*(4), 809-817. https://doi.org/10.1016/j.anbehav.2009.07.011

Krebs, J. R., Healy, S. D., & Shettleworth, S. J. (1990). Spatial memory of paridae: Comparison of a storing and a non-storing species, the coal tit, Parus ater, and the great tit, P. major. *Animal Behaviour*, *39*(6), 1127–1137. https://doi.org/10.1016/S0003-3472(05)80785-7

Langley, E. J. G., Adams, G., Beardsworth, C. E., Dawson, D. A., Laker, P. R., van Horik, J. O., Whiteside, M. A., Wilson, A. J., & Madden, J. R. (2020). Heritability and correlations among learning and inhibitory control traits. *Behavioral Ecology*, araa029. https://doi.org/10.1093/beheco/araa029

Loukola, O. J., Adamik, P., Adriaensen, F., Barba, E., Doligez, B., Flensted-Jensen, E., Eeva, T., Kivelä, S. M., Laaksonen, T., Morosinotto, C., Mänd, R., Niemelä, P. T., Remeš, V., Samplonius, J. M., Sebastiano, M., Senar, J. C., Slagsvold, T., Sorace, A., Tschirren, B., … Forsman, J. T. (2020). The roles of temperature, nest predators and information parasites for geographical variation in egg covering behaviour of tits (*Paridae*). *Journal of Biogeography*, jbi.13830. https://doi.org/10.1111/jbi.13830

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. N., Burkart, J. M., Call, J., Cantlon, J. F., Cheke, L. G., … Zhao, Y. (2014). The evolution of self-control. *Proceedings of the National Academy of Sciences*, *111*(20), E2140–E2148. https://doi.org/10.1073/pnas.1323533111

McGregor, A., & Healy, S. D. (1999). Spatial accuracy in food-storing and nonstoring birds. *Animal Behaviour*, *58*(4), 727–734. https://doi.org/10.1006/anbe.1999.1190

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Montello, D. R. (1993). Scale and multiple psychologies of space. In A. U. Frank & I. Campari (Eds.), *Spatial information theory: A theoretical basis for GIS* (Vol. 716). Springer. https://doi.org/10.1007/3-540-57207-4_21

Morand-Ferron, J., Cole, E. F., & Quinn, J. L. (2016). Studying the evolutionary ecology of cognition in the wild: A review of practical and conceptual challenges: Evolutionary ecology of cognition in the wild. *Biological Reviews*, *91*(2), 367–389. https://doi.org/10.1111/brv.12174

Morand-Ferron, J., Cole, E. F., Rawles, J. E. C., & Quinn, J. L. (2011). Who are the innovators? A field experiment with 2 passerine species. *Behavioral Ecology*, *22*(6), 1241–1248. https://doi.org/10.1093/beheco/arr120

Morand-Ferron, J., & Quinn, J. L. (2015). The evolution of cognition in natural populations. *Trends in Cognitive Sciences*, *19*(5), 235–237. https://doi.org/10.1016/j.tics.2015.03.005

Morgan, K. V., Hurly, T. A., & Healy, S. D. (2014). Individual differences in decision making by foraging hummingbirds. *Behavioural Processes*, *109*, 195–200. https://doi.org/10.1016/j.beproc.2014.08.015

Olton, D. S. (1977). Spatial memory. *Scientific American*, *236*(6), 82–99. https://www.jstor.org/stable/24954072

O'Shea, W., O'Halloran, J., & Quinn, J. L. (2018). Breeding phenology, provisioning behaviour, and unusual patterns of life history variation across an anthropogenic heterogeneous landscape. *Oecologia*, *188*(4), 953–964. https://doi.org/10.1007/s00442-018-4155-x

Pike, T. W., Ramsey, M., & Wilkinson, A. (2018). Environmentally induced changes to brain morphology predict cognitive performance. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1756), 20170287. https://doi.org/10.1098/rstb.2017.0287

Pritchard, D. J., & Healy, S. D. (2018). Taking an insect-inspired approach to bird navigation. *Learning & Behavior*, *46*(1), 7–22. https://doi.org/10.3758/s13420-018-0314-5

Pritchard, D. J., Hurly, T. A., & Healy, S. D. (2018). Wild hummingbirds require a consistent view of landmarks to pinpoint a goal location. *Animal Behaviour*, *137*, 83–94. https://doi.org/10.1016/j.anbehav.2018.01.014

Pritchard, D. J., Tello Ramos, M. C., Muth, F., & Healy, S. D. (2017). Treating hummingbirds as feathered bees: A case of ethological cross-pollination. *Biology Letters*, *13*(12), 20170610. https://doi.org/10.1098/rsbl.2017.0610

Quinn, J. L., Charmantier, A., Garant, D., & Sheldon, B. C. (2006). Data depth, data completeness, and their influence on quantitative genetic estimation in two contrasting bird populations. *Journal of Evolutionary Biology*, *19*(3), 994–1002. https://doi.org/10.1111/j.1420-9101.2006.01081.x

Quinn, J. L., Cole, E. F., Reed, T. E., & Morand-Ferron, J. (2016). Environmental and genetic determinants of innovativeness in a natural population of birds. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1690), 20150184. https://doi.org/10.1098/rstb.2015.0184

R Core Team. (2019). *R: A Language and Environment for Statistical Computing* (3.6.1) [Computer software]. R Foundation for Statistical Computing.

Raine, N. E., & Chittka, L. (2008). The correlation of learning speed and natural foraging success in bumble-bees. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1636), 803–808. https://doi.org/10.1098/rspb.2007.1652

Reichert, M. S., Crofts, S. J., Davidson, G. L., Firth, J. A., Kulahci, I. G., & Quinn, J. L. (2020). Multiple factors affect discrimination learning performance, but not between-individual variation, in wild mixed-species flocks of birds. *Royal Society Open Science*, *7*(4), 192107. https://doi.org/10.1098/rsos.192107

Rowe, C., & Healy, S. D. (2014a). Measuring cognition will be difficult but worth it: A response to comments on Rowe and Healy. *Behavioral Ecology*, *25*(6), 1298–1298. https://doi.org/10.1093/beheco/aru190

Rowe, C., & Healy, S. D. (2014b). Measuring variation in cognition. *Behavioral Ecology*, *25*(6), 1287–1292. https://doi.org/10.1093/beheco/aru090

Schubiger, M. N., Fichtel, C., & Burkart, J. M. (2020). Validity of cognitive tests for non-human animals: Pitfalls and prospects. *Frontiers in Psychology*, *11*(1835). https://doi.org/10.3389/fpsyg.2020.01835

Sewall, K. B., Soha, J. A., Peters, S., & Nowicki, S. (2013). Potential trade-off between vocal ornamentation and spatial ability in a songbird. *Biology Letters*, *9*(4), 20130344. https://doi.org/10.1098/rsbl.2013.0344

Shaw, R. C., Boogert, N. J., Clayton, N. S., & Burns, K. C. (2015). Wild psychometrics: Evidence for 'general' cognitive performance in wild New Zealand robins, *Petroica longipes*. *Animal Behaviour*, *109*, 101–111. https://doi.org/10.1016/j.anbehav.2015.08.001

Shaw, R. C., & Schmelz, M. (2017). Cognitive test batteries in animal cognition research: Evaluating the past, present and future of comparative psychometrics. *Animal Cognition*, *20*(6), 1003–1018. https://doi.org/10.1007/s10071-017-1135-1

Sherry, D. F., & Guigueno, M. F. (2019). Cognition and the brain of brood parasitic cowbirds. *Integrative Zoology*, *14*(2), 145–157. https://doi.org/10.1111/1749-4877.12312

Sonnenberg, B. R., Branch, C. L., Pitera, A. M., Bridge, E., & Pravosudov, V. V. (2019). Natural selection and spatial cognition in wild food-caching mountain chickadees. *Current Biology*, *29*(4), 670-676.e3. https://doi.org/10.1016/j.cub.2019.01.006

Sovrano, V. A., Bisazza, A., & Vallortigara, G. (2003). Modularity as a fish (*Xenotoca eiseni*) views it: Conjoining geometric and nongeometric information for spatial reorientation. *Journal of Experimental Psychology: Animal Behavior Processes*, *29*(3), 199–210. https://doi.org/10.1037/0097-7403.29.3.199

Sovrano, V. A., Bisazza, A., & Vallortigara, G. (2005). Animals' use of landmarks and metric information to reorient: Effects of the size of the experimental space. *Cognition*, *97*(2), 121–133. https://doi.org/10.1016/j.cognition.2004.08.003

Sovrano, V. A., Bisazza, A., & Vallortigara, G. (2006). How fish do geometry in large and in small spaces. *Animal Cognition*, *10*(1), 47–54. https://doi.org/10.1007/s10071-006-0029-4

Stoffel, M. A., Nakagawa, S., & Schielzeth, H. (2017). rptR: Repeatability estimation and variance decomposition by generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *8*(11), 1639–1644. https://doi.org/10.1111/2041-210X.12797

Tello-Ramos, M. C., Branch, C. L., Pitera, A. M., Kozlovsky, D. Y., Bridge, E. S., & Pravosudov, V. V. (2018). Memory in wild mountain chickadees from different elevations: Comparing first-year birds with older survivors. *Animal Behaviour*, *137*, 149–160. https://doi.org/10.1016/j.anbehav.2017.12.019

Thornton, A., Isden, J., & Madden, J. R. (2014). Toward wild psychometrics: Linking individual cognitive differences to fitness. *Behavioral Ecology*, *25*(6), 1299–1301. https://doi.org/10.1093/beheco/aru095

Thornton, A., & Lukas, D. (2012). Individual variation in cognitive performance: Developmental and evolutionary perspectives. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1603), 2773–2783. https://doi.org/10.1098/rstb.2012.0214

van Horik, J. O., Beardsworth, C. E., Laker, P. R., Whiteside, M. A., & Madden, J. R. (2020). Response learning confounds assays of inhibitory control on detour tasks. *Animal Cognition*, *23*(1), 215–225. https://doi.org/10.1007/s10071-019-01330-w

van Horik, J. O., Langley, E. J. G., Whiteside, M. A., Laker, P. R., Beardsworth, C. E., & Madden, J. R. (2018). Do detour tasks provide accurate assays of inhibitory control? *Proceedings of the Royal Society B: Biological Sciences*, *285*(1875), 20180150. https://doi.org/10.1098/rspb.2018.0150

van Horik, J. O., Langley, E. J. G., Whiteside, M. A., Laker, P. R., & Madden, J. R. (2018). Intra-individual variation in performance on novel variants of similar tasks influences single factor explanations of general cognitive processes. *Royal Society Open Science*, *5*(7), 171919. https://doi.org/10.1098/rsos.171919

van Horik, J. O., Langley, E. J. G., Whiteside, M. A., & Madden, J. R. (2019). A single factor explanation for associative learning performance on colour discrimination problems in common pheasants (*Phasianus colchicus*). *Intelligence*, *74*, 53–61. https://doi.org/10.1016/j.intell.2018.07.001

Vernouillet, A. A. A., Stiles, L. R., Andrew McCausland, J., & Kelly, D. M. (2018). Individual performance across motoric self-regulation tasks are not correlated for pet dogs. *Learning & Behavior*, *46*(4), 522–536. https://doi.org/10.3758/s13420-018-0354-x

Völter, C. J., Tinklenberg, B., Call, J., & Seed, A. M. (2018). Comparative psychometrics: Establishing what differs is central to understanding what evolves. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1756), 20170283. https://doi.org/10.1098/rstb.2017.0283

Vonk, J., & Povinelli, D. (2011). Individual differences in long-term cognitive testing in a group of captive chimpanzees. *International Journal of Comparative Psychology*, *24*, 137–167. https://doi.org/10.5070/P4241020255

Webster, M. M., & Rutz, C. (2020). How STRANGE are your study animals? *Nature*, *582*, 337–340. https://doi.org/10.1038/d41586-020-01751-5

Zamisch, V., & Vonk, J. (2012). Spatial memory in captive American black bears (*Ursus americanus*). *Journal of Comparative Psychology*, *126*(4), 372–387. https://doi.org/10.1037/a0028081