

Invited Review

Measuring variation in cognition

Candy Rowe^a and Susan D. Healy^b

^aCentre for Behaviour & Evolution, Institute of Neuroscience, Newcastle University, Henry Wellcome Building, Framlington Place, Newcastle upon Tyne NE2 4HH, UK and ^bSchool of Biology, University of St Andrews, St Andrews KY16 9TH, UK

Received 21 February 2014; revised 15 April 2014; accepted 15 April 2014; Advance Access publication 3 June 2014.

Across a range of disciplines, researchers are becoming increasingly interested in studying the variation in cognitive abilities found within populations. Behavioral ecology is no exception: the pursuit to understand the evolution of cognition has led to a rapidly expanding literature that uses various tasks to measure individuals' cognitive abilities. While this is an exciting time, we are concerned that without being clearer as to the cognitive abilities under test it will be difficult to design appropriate experiments and the interpretation of the data may be unsound. The aim of this review is 3-fold: 1) to highlight problems with designing tasks for measuring individual variation in cognitive abilities and interpreting their outcomes; 2) to increase awareness that noncognitive factors can cause variation in performance among individuals; and 3) to question the theoretical basis for thinking that performance in any cognitive task should necessarily correlate with a measure of fitness. Our take-home message is that variability in performance in cognitive tasks does not necessarily demonstrate individual variation in cognitive ability, and that we need to both design more stringent cognitive tests and be more cautious in their interpretation.

Key words: cognition, evolution, individual differences, innovation, motivation, problem solving.

INTRODUCTION

In the pursuit of a fuller understanding of the evolution of mating systems, foraging, parental care, and so on, it is becoming increasingly common to see papers in *Behavioral Ecology* that deal with the mechanistic basis of behavior rather than with its functional cause. For example, investigating the role that physiology, and particularly, parasites, immune systems, and hormones play is steadily leading to a fuller understanding of why animals behave as they do (e.g., Schwagmeyer et al. 2012; Spencer et al. 2010; Simmons 2012). Tinbergen would be proud.

Given the success of integrating physiological mechanisms and function, it is not surprising to find behavioral ecologists turning their attention to another type of mechanism underpinning behavior: cognition (e.g., Boogert et al. 2011; Keagy et al. 2012; Stoddard and Kilner 2013; Bokony et al. 2014). The interest is being especially directed at measuring the individual differences in “cognitive abilities” upon which selection can work and in identifying fitness benefits associated with having “better” cognitive abilities (e.g., Keagy et al. 2009; Cole et al. 2012; Sih and Del Giudice 2012). We, too, consider these to be attractive and exciting questions to address: who does not want to understand how and why cognitive abilities vary and evolve? However, in our view this new wave of research does not take sufficient advantage of the wealth

of methodologies, understanding, and insights from many earlier studies on animal cognition conducted by our psychological colleagues. As a result we are concerned that these more recent experiments and analyses may not be as rigorous as is required and that, consequently, they will fail to answer the questions posed.

Our overarching aim is to highlight how invaluable the history of animal cognition testing is to behavioral ecologists. The field of animal cognition is a mature one in its own right but one whose historical proponents are interested in asking questions that differ to those posed by behavioral ecologists. Specifically, animal cognition researchers direct their focus toward determining the mechanisms that underpin cognitive abilities and, as a consequence, they have established robust methods for measuring a range of cognitive processes in a wide range of species (Pearce 2008; Shettleworth 2010). As examination of variation and selective benefits of cognitive abilities is gathering apace, this seems a useful juncture to recognize and draw on the wealth of knowledge collected in the laboratory by psychologists but also to recognize the pitfalls of assuming that one can take psychological tools “off the shelf” to measure variation in cognitive abilities.

Here, we attempt to create an awareness of what we can gain from donning the psychologist's lab coat, but also the shortcomings inherent in doing so. We want to provide a useful research framework for pointing behavioral ecologists investigating cognitive abilities in a direction we think likely to be efficient, effective, and thereby profitable. By so doing we do not want to deter behavioral ecologists from studying the cognitive abilities of their study species.

*Address correspondence to C. Rowe. E-mail: candy.rowe@ncl.ac.uk

Quite the opposite: we really want to stimulate research in this area. We consider that this will involve a far more rigorous approach, however, than that we currently observe, including: 1) defining more clearly the specific cognitive abilities that are being studied; 2) recognizing that noncognitive factors can affect performance in tasks designed to measure a cognitive ability; and 3) examining the assumption that fitness advantages will accrue from performance in a cognitive task. Each of these represents a significant challenge, but we hope that our review harnesses the new energy and enthusiasm in this pursuit of characterizing the cognitive phenotype.

BETTER COGNITION?

Many behavioral ecologists approach the evolution of cognitive abilities with the assumption that those animals that have “better cognition” will enjoy some kind of fitness advantage. Before going any further, we need to unpick and define 2 crucial words: “better” and “cognition.”

So, let us start with cognition. The definition of cognition that we prefer is both broad and commonly used, and can be described as “information processing.” This definition encompasses the acquisition of sensory information, the storage, retrieval, and use of that information for making behavioral decisions (Healy and Rowe 2010; Shettleworth 2010). Key to this definition of cognition is that cognition is not a unitary trait. Nor is it a black box: open any textbook on animal cognition, and you will find dozens of different cognitive abilities that have been identified and studied in detail (e.g., Pearce 2008; Shettleworth 2010). Psychologists now have an increasingly robust understanding of the mechanisms underlying information acquisition, learning, and memory.

However, in some contrast to the psychological tradition of explicit characterization of specific cognitive abilities, behavioral ecologists are currently tending to be both less explicit and less specific about the trait they wish to investigate. Perhaps because the behavioral ecologist’s approach is to understand behavior in terms of the fitness benefits to the whole organism rather than with regard to the mechanisms underlying that behavior, there is an enthusiasm for addressing cognitive abilities in a similar manner. For example, “problem solving” is defined by what an animal does behaviorally, rather than through the identification of any specific cognitive processes. To some, this distinction may appear unimportant in a functional context: if performing a behavior well results in fitness benefits, then the cognitive abilities underlying it, whatever they may be, should also be favored by selection. We think, however, that this approach can lead to problems.

Crucially, by not defining the cognitive ability carefully, we have no way of understanding how selection acts on that ability. An example of this comes from classic work on the spatial memory abilities of food storing and nonstoring parids (*Parus spp.*, the tits and the chickadees). After it was demonstrated that food-storing songbirds relied on memory for successful relocation of their stores, and that the area associated with spatial memory, the hippocampus, was larger in storers compared with nonstorers, it seemed obvious that storers would have superior spatial cognition compared with nonstorers (Krebs et al. 1989; Sherry et al. 1989). After multiple attempts, food storers were indeed shown to be able to remember even a single location for longer than nonstorers. Importantly, however, storers and nonstorers did not differ in 2 other components of spatial memory used to retrieve their stores: 1) storers were not more accurate in their spatial memory, and 2) storers could not remember more locations than nonstorers (Biegler et al. 2001;

see also Hampton and Shettleworth 1996). We say “importantly,” because better performance could have been due to any one (or all) of these mechanisms but was not. For food storers, then, all possible components of spatial cognition have not been enhanced by selection but rather selection has focused on just one aspect. When we consider, therefore, apparently more “complex” behavioral phenomena, such as “problem-solving” or “innovation” (e.g., Liker and Bókonyi 2009; Cole et al. 2011; Benson-Amram and Holekamp 2012; Thornton and Samson 2012), we should consider that these are unlikely to be underpinned by a single cognitive ability. And if we want to understand how these “complex” behaviors evolve, we need to be able to specify exactly what those abilities should be.

This leads us to our second word, “better”: what do researchers mean when they use the word “better” (e.g., Boogert et al. 2011; Keagy et al. 2011; Morand-Ferron and Quinn 2011; Sih and Del Giudice 2012)? As demonstrated in the example of spatial memory, there are a number of ways that an animal’s cognitive abilities might be considered to be better than those of another animal: the animal can learn more information, learn information faster or more accurately, or remember it for longer. There is no reason to think that every animal in a species learns about the world and remembers past experiences in the same way (Brooks and Endler 2007; Harris et al. 2008; Archard and Braithwaite 2011; Sih and Del Giudice 2012; van Heijningen et al. 2013). Take, for example, the speed with which an animal learns to associate a color cue with a food reward. If an animal learns quickly, it could be that it learns that a color predicts reward. Nothing more. However, if it is slower, it may be because it learns more about the visual cue, such as its size or shape, or spend time fitting the new experience into the context of previous experiences (e.g., through category formation). As to which of these is “the best” is open to argument, but is likely to depend upon an individual’s experiences during development, an animal’s current environment, and/or natural selection (e.g., Boogert et al. 2013; Salvanes et al. 2013). Therefore, we should be careful when considering both what we mean by “better” and when we would expect an animal, group, or species to have such better cognitive abilities.

Behavioral ecologists should find this an example of “teaching granny to suck eggs” since trade-offs are fundamental to our field: no behavioral ecologist would argue that it is always better to produce fewer larger offspring than a large number of small offspring. The same logic can, and should, be applied to cognitive abilities.

MEASURING VARIATION IN COGNITION

Having discussed the problem of defining cognition, we now want to consider its measurement. For all of our enthusiasm for the work of our psychological colleagues, there may be real problems in taking a cognitive task that has been used successfully in the lab by psychologists “off the shelf” in the hope that a few simple modifications will enable the measurement of variation in cognitive abilities across individuals. Importantly, psychologists aim to determine what it is that animals can do, a question that is best addressed using tasks designed to allow the measurement of cognitive performance under tightly controlled conditions. In particular, the intent is to minimize any effects of variation among the performance of their subjects. For example, psychologists work hard to ensure that their animals are kept in and trained under the same standardized conditions, that they reach a criterion level of performance before they are tested, that their age and gender are matched, and so on. Perhaps the most pertinent feature of these tasks for our purposes,

however, is the removal of what might best be described as ecological validity. Tasks may be presented on computer screens or in Skinner boxes, and cognitive abilities are tested using lights, tones, key presses, and similar. The essential feature is that that variation caused by factors extraneous to the question of interest (e.g., what can an animal remember, can an animal count?) is diminished as much as is humanly possible.

So, the behavioral ecologist interested in variation needs to adopt a similarly rigorous approach in order to reduce the possibility that variability in performance in a cognitive task is due to some other, extraneous factor (Healy and Rowe 2013). Herein lies a problem, however, perhaps best explained using an example. Leaving aside for the moment our concerns over whether learning something more quickly is better than learning it more slowly, let us imagine a scenario where we want to test the prediction that those birds that learn to solve a foraging task the soonest will have higher reproductive success. This is based on the assumption that the individuals that solve a task more quickly will have better access to food and be able to provision their young more effectively. So far, so good: we can measure individual reproductive success as the number or mass of the offspring at fledging, the rate of survival from laying to fledging, or some other measure that conveys a fitness advantage.

Now for the other side of the equation, our measurement of cognitive performance. The speed at which an animal learns to associate a conditioned stimulus (such as a light or tone) with an unconditioned stimulus (some form of reward or punishment) is a well-established paradigm for measuring learning in experimental psychology (e.g., Rescorla and Wagner 1972). We will take for our example, the speed at which our birds associate a color cue with a food reward. Suppose we initially train a group of birds to visit an array of 10 white bowls placed on the ground that contain food (note that we are, actually, describing a situation that is relatively “ecologically valid”). At this point the bowls are not covered so that the bird can see the food, but over successive trials, we progressively cover the bowls with white lids until the birds have to flip the lids to retrieve the food. Once the birds have learned to flip all 10 lids, we cover half of the bowls with a green lid. Into these bowls we place food. The remaining bowls we cover with a purple lid and these bowls remain empty. To ensure that the particular colors we have chosen for the lids do not influence the outcome of the task, we counterbalance the colors so that half of the birds have purple lids signaling food reward. The array is presented once a day at the same time, and the number of days taken for the bird to attack all 5 of the rewarded bowls before attacking any unrewarded bowls is recorded. The task appears to be successful: birds show significant variation in the number of trials it takes for them to reach the learning criterion and we now have a measure of their cognitive ability. Or do we?

Before we can reach this conclusion, there are number of factors we need to consider and exclude as alternative explanations (Healy and Rowe 2013). One of these factors is the animals’ past experience. For example, if birds vary in the proportion of green insects they eat or some individuals prefer to include purple fruits in their diets, then these individuals will bring associations between certain colors and rewards to the task. These biases can differ across individuals quantitatively (e.g., the strength of association that already exists between green and reward) or qualitatively (i.e., whether they are more likely to associate green or purple with a reward, depending on their past diet). These hidden biases to associate certain colors with rewards will affect the number of trials it takes them to reach the learning criterion. Indeed, if all the animals in a

population learn more readily to associate green with reward, color reversals can even introduce unwanted variation into the task, as the animals that need to learn that purple is rewarding are immediately at a disadvantage, and likely to be slower to associate green with reward. This example demonstrates that we need to carefully think about whether techniques that we use to compare treatment groups are relevant for testing at the level of the individual. One certainly needs to be cautious in attributing variation in performance in an associative learning task involving colors and rewards to differences in learning speed per se (Isden et al. 2013).

A factor that is even more obvious as a possible explanation for the variation in speed of acquisition is energetic state. Hungry animals, or those with higher energetic requirements, are likely to be more motivated to learn the discrimination. In addition, a fixed amount of food can be perceived as more rewarding when the animal is hungry than when it is satiated (e.g., Pompilio et al. 2006). Therefore, variation in the evaluation of the food reward can affect learning speed: the ability of animals to associate cues may be identical but variation arises because some perceive the reward (or the punishment of an empty bowl) to be greater.

Salience of the conditioned stimulus, the cue that the animal is learning about, is also highly likely to have an impact on an animal’s task performance. Salience generally means the degree to which a stimulus “stands out” to an animal and in psychological terms can determine the amount of attention that an animal pays to a stimulus during the learning process. It is important to recognize that salience does not depend just on the strength of the signal. In our example, lids of the same green color may not be equally salient for all test individuals: a whole raft of motivational, cognitive, and emotional factors may determine cue salience. For example, the salience of a color cue can depend on the range of colours that an animal has experienced in its past, and in particular, what it has already learned about colored stimuli in its environment. Perceptual processes can also be important: perceptual sensitivities are likely to vary, due to experience, aging, or sensory tuning. Therefore, variation in the speed with which birds learn this particular color–reward association would not necessarily be a reliable predictor of how well that animal is likely to learn another color–reward association or any other more general stimulus–reward association.

At risk of producing a seemingly endless list of killjoy alternative explanations for variation in the speed of acquisition of a task, we want to add one more that we think is especially worthy of consideration here. And this is that, even if we could be sure that our observed variation is highly unlikely to be due to past experience, cue salience, or motivation among birds, it does not exclude the possibility that the variation is due to individuals acquiring and using information in different ways. Imagine that, all other things being equal, the 2 birds perform differently in this task: one reaches criterion (uncovering all the rewarded bowls first) in just 5 trials, whereas the other takes 10 trials. The second bird appears an obvious dullard. But if the dullard takes those extra trials because, for example, it chooses to sample the unrewarded colored bowls so as to collect information about how they vary in their reward value, then at the end of the task, the dullard will in fact have acquired more information about its environment. This variation in what is learned during the discrimination task may be because of the animal’s past experiences or its developmental history or be due to personality differences (Sih and Del Giudice 2012). The point we want to make here is that a task may be learned or solved in different ways and that considering alternative explanations of variation in performance are crucial.

In our example, we describe concerns over the interpretation of variation observed in what appears to a simple learning task. Behavioral ecologists are not, however, restricting themselves to measuring the speed of acquisition but are also addressing variation in more “complex” cognitive abilities, such as “innovation” and “problem solving.” Just as in our simple learning task, however, performance in tasks aimed at measuring problem solving and innovation will also be affected by apparently extraneous cognitive abilities (e.g., memory, discrimination) and noncognitive factors (e.g., motivation, perception, motor skills). Indeed, as the exact cognitive abilities and processes underlying behaviors such as innovation and problem solving are not yet well described, measurement of variation in these behaviors may be even more susceptible to problems of interpretation.

MEASURING INTERINDIVIDUAL VARIATION IN COGNITION

We have attempted to demonstrate some of the pitfalls of simply applying experimental designs from psychology to measure the cognitive abilities of individuals. But these problems leave us with a real challenge as to how we best deal with them. And it will doubtless seem that we are advocating an impossible task to design experiments that completely eliminate these potential confounds.

Ideally now is the point at which the critics (us) provide our pet solution and advocates that everyone follows the proposed way forward. But we will come clean straight off: we do not have a definitive answer to this set of problems. The 2 things that we are sure of, however, are that: 1) researchers need to attempt to deal with these factors in their experimental designs in order to mitigate their effects; and 2) researchers need to be aware of the limitations of those experimental designs and interpret their data accordingly.

There might also be ways in which we can reduce the impact of these factors. Let us take the problem of cue preferences and biases that can affect animals' abilities to form learned associations (e.g., [ten Cate and Rowe 2007](#); [Halpin et al. 2008](#); [Guillette et al. forthcoming](#)). To our knowledge, no experimental psychology study is free from these biases and preferences: these exist in animals. The psychologists' challenge has been to reduce the impact of extraneous variables, which can be dealt with in part by counterbalancing and using cues that do not elicit strong biases. Counterbalancing can introduce its own problems as we have already discussed, so perhaps one could look to use cues that do not elicit not strong biases. In our example with birds trained to associate colors with reward, one might replace the colors with grayscale cues, which could reduce the effects of visual biases (e.g., [Alatalo and Mappes 1996](#); Smith et al. unpublished data).

When it comes to motivation, there again, is no simple solution. Since psychologists commonly motivate animals to perform in their experiments using food rewards, they have developed various methods to try to ensure that individuals are in a similar motivational state during experiments. And, indeed, behavioral ecologists do make valiant efforts to control for motivation, but in our view, they still fall short of being able to fully eliminate it as a contributing factor to their findings. Attempts to consider motivation in tasks involving food rewards include measuring individual body condition (e.g., [Cole et al. 2011](#); [Thornton and Samson 2012](#)), or deriving other behavioral measures associated with performing the task (e.g., [Thornton and Samson 2012](#); [Cauchard et al. 2013](#)). However, body mass may not accurately reflect an animal's condition or energetic requirements (e.g., [Chatelain et al. 2013](#)), and other behavioral

measures could also be influenced by confounding factors associated with personality or past experience. The behavioral ecologist has to face the fact that motivation is tricky; it was when [McFarland and others \(e.g., Houston and McFarland 1976\)](#) made heroic attempts to understand it in the 1970s and 80s, and we still do not fully understand its mechanistic basis. While we can try to reduce the effects of motivation in our experiments, we cannot assume that our experimental design, however elegant, has made this problem go away. We would argue that we should always be prepared to concede that motivation may have influenced the variation in performance we observe among the individuals we test.

Ensuring that we give these factors the respect that they deserve, both when designing our experiments and when we interpret our data, certainly means that findings will be more robust to the criticisms we have raised, and potentially lead to more significant advances. While we are unable to come up with a neat solution to eradicate the effects of noncognitive factors in our designs, this is not to say that this may be solved in the future, perhaps through the use of some novel task or methodology. We certainly see this as a major challenge for this type of study and hope that we can stimulate researchers to think harder about ways to solve this problem.

COGNITION AND FITNESS

So, having dealt with what we see as rather logistical concerns, we want to turn our attention to one more, rather conceptual, concern. That is, that we have to question whether or not it is realistic to expect that cognition should correlate with fitness. This is the basic tenet of many recent studies directed toward understanding the selection pressures acting on animal cognition (e.g., [Boogert et al. 2011](#); [Cole et al. 2012](#); [Keagy et al. 2012](#)). But given that cognition is not a unitary trait, that an animal's ability to complete a single task may require multiple cognitive processes, and that cognitive processes are likely to be used across many different contexts, we would argue that one should not necessarily predict that measures of cognition will correlate with fitness.

We return to our example of the food storers, where storing is associated with an improved ability to remember locations for longer. In this example, the cognitive ability is clear and specific: it is the length of time over which birds remember locations, which is greater in storers compared with nonstorers and not in the number of items that are remembered or in the accuracy with which spatial locations are remembered. In this example, one can see how selection could act so specifically: those birds that are better at remembering the locations of their stores through a long winter would be the ones that are more likely to survive and to reproduce (or to produce more offspring, a prediction that remains untested). There is also clear directional selection, as we have a good idea of what is “better” in this case: longer memories, as opposed to shorter ones, will help the food storer to relocate its stores and survive. The fact that the cognitive process is so specific to the foraging task, and that the direction in which selection should act is plain to see, would leave us unsurprised to find a link between this cognitive ability and fitness.

For broad cognitive processes, such as discrimination, learning, or memory, however, as we have already discussed, what is “better” is much less obvious and they do not map so clearly onto performance in a single behavioral task, as they are often employed in multiple contexts. Each of these processes will be valuable to an animal in a context-dependent fashion: being good at discriminating between objects will sometimes be very useful and other times

be pointless, learning some information very quickly will sometimes be vital but at other times be costly, while being able to remember some information for a long time may be useful but it will not be too useful to remember all information (e.g., Dunlap et al. 2009; Ferrari et al. 2010; Halpin et al. 2012; Gonzalo et al. 2013). Selection will act on the outcome of all these costs and benefits, leading to an optimal solution that maximizes lifetime fitness, not just performance in a single task.

This is perhaps best illustrated by example: a choosy female bowerbird needs to find bowers (often quite well hidden under bushes, sometimes separated by several kilometers) before assessing the features of the bower that will enable her to make a good choice of mate (without this assessment taking all day). In making that decision, she might consider the current range of males from which she has to choose, which may be limited and, as bowerbirds are relatively long lived (some 30 years), she may also incorporate memories of past decisions and their outcomes. All this before she chooses the male(s) with which to mate. A range of cognitive processes (e.g., learning, discrimination, memory) will underpin her behavior and decision making, and there may be different optimal combinations of cognitive traits that will lead her make a good choice (e.g., Kelley and Endler 2012). Furthermore, before she even gets to admire a single bower, our female needs cognitive abilities that get her through a long juvenile phase (potentially 10–15 years) during which she needs to be “good” at finding food, avoiding predators, and coping with the variety of situations life will throw at her. Before she lays her eggs, she must also consider her own state and the current environmental conditions. Our bowerbird female, then, needs a range of cognitive abilities to achieve and enhance her reproductive success but she is unlikely to call on all of them all of the time.

Even if we did design an experiment that robustly showed a correlation between variation in a cognitive trait and reproductive success, having eliminated all of the confounding factors that we have raised, could we then be confident that this cognitive trait increases fitness? To address this, we return to our food storers, for which we have already identified a valuable cognitive ability. We could, for example, take coal tits into the lab in autumn and measure variation in the duration over which they can remember spatial locations before releasing them to follow their overwinter survival and reproduction in the following spring. This appears to provide an explicit test of the prediction that improved spatial cognition would correlate with increased fitness. Even if we find that those birds with better spatial cognition produce more offspring the following spring, however, we need to realize that this is not sufficient confirmation of our prediction. Rather, we have identified a foraging benefit to improved spatial cognition during the winter months when times are hard and food is scarce but what of the rest of the year? Better spatial cognition means having a larger hippocampus but it appears that this enlargement is more costly to maintain at other times of year (Smulders et al. 1995). Our experiment would provide only a snapshot of the birds’ lifetime reproductive success, and as such, we have to be wary about concluding that those food-storing birds with better spatial cognition enjoy better overall lifetime reproductive success (see also Cole et al. 2012).

In our view, then, we should not expect to see selection for maximizing across cognitive abilities, that is, selection will not tend to favor “smart” animals. Rather we would expect to see selection favoring animals that are smart for their environment, that is, with a suite of cognitive abilities that is optimized for that environment.

To be a “smart” blue tit, desert ant, or saltwater crocodile is likely to require a different suite of cognitive abilities. It might also be that there is more than one way to be a “smart” blue tit, desert ant, or saltwater crocodile. Some of the strongest support for this suggestion comes from research on sex differences, especially in mammals: males typically outperform females on spatial tasks (e.g., Jones et al. 2003; Jones and Healy 2006). Although there is considerable debate as to the interpretation of these data, and what environmental factors might lead to such variation, no one suggests that female reproductive success suffers because of this. There are also some supportive data from species in which there are clear environmental differences between morphs. For example, laboratory experiments showed that 3-spined sticklebacks (*Gasterosteus aculeatus*) from ponds were more likely use visual landmarks to relocate a rewarded goal than were sticklebacks taken from fast-flowing rivers (Odling-Smee and Braithwaite 2003). Although these data may be explained by differences in attention to different cues rather than by differences in memory, they nonetheless support the idea that information use may differ even within species, depending on the environment. We are still some distance from having a sufficiently thorough natural history of cognitive abilities to enable us to predict with any precision which abilities should be favored in specific environments. Our food storers remain a surprisingly rare example.

CONCLUSION

We welcome the increasing diversity of species and contexts in which cognitive abilities are being examined. This will no doubt bring us closer to understanding how natural selection can act on cognitive abilities. However, behavioral ecologists need to recognize that measuring variation in cognition is challenging and that they would benefit from addressing the concerns that we raise here. The growing number of studies purporting to show variation in cognitive abilities in different species has lead us to question whether there are, as yet, compelling data that demonstrate individual differences in cognition. Cognitive processes need to be more clearly defined, which may mean moving away from measuring behaviors such a “problem solving” to identifiable cognitive processes. We also call for more care when considering how selection could be acting on cognition: more memory, faster learning, finer discriminations, may not always be better. Finally, we think that until experimental tests are more carefully designed and experimental output is more carefully interpreted, demonstrating that variation in cognitive abilities leads to a selective advantage is likely to remain elusive. Perhaps we are the ones who have to be smarter.

We would like to thank D. Shuker for provocative discussion, and J. Skelhorn, L. Guillette, and 2 anonymous reviewers for critical and insightful comments on an earlier draft.

Editor-in-Chief: Leigh Simmons

REFERENCES

- Alatalo RV, Mappes J. 1996. Tracking the evolution of warning signals. *Nature*. 382:708–710.
- Archard GA, Braithwaite VA. 2011. Variation in aggressive behaviour in the poeciliid fish *Brachyrhaphis episcopi*: population and sex differences. *Behav Processes*. 86:52–57.
- Benson-Amram S, Holekamp KE. 2012. Innovative problem solving by wild spotted hyenas. *Proc Biol Sci*. 279:4087–4095.

- Biegler R, McGregor A, Krebs JR, Healy SD. 2001. A larger hippocampus is associated with longer lasting spatial memory. *Proc Natl Acad Sci (USA)*. 98:6941–6944.
- Bókonyi V, Lendvai ÁZ, Vágási CI, Pátrás PL, Németh J, Vincze E, Papp S, Preiszner B, Seress G, Líker A. 2014. Necessity or capacity? Physiological state predicts problem-solving performance in house sparrows. *Behav Ecol*. 25:124–135.
- Boogert NJ, Fawcett TW, Lefebvre L. 2011. Mate choice for cognitive traits: a review of the evidence in nonhuman vertebrates. *Behav Ecol*. 22:447–459.
- Boogert NJ, Zimmer C, Spencer KA. 2013. Pre- and post-natal stress have opposing effects on social information use. *Biol Lett*. 9:20121088.
- Brooks R, Endler JA. 2007. Female guppies agree to differ: phenotypic and genetic variation in mate-choice behaviour and the consequences for sexual selection. *Evolution*. 55:1644–1655.
- Cauchard L, Boogert NJ, Lefebvre L, Dubois F, Doligez B. 2013. Problem-solving performance is correlated with reproductive success in a wild bird population. *Anim Behav*. 85:19–26.
- Chatelain M, Halpin CG, Rowe C. 2013. Ambient temperature influences birds' decisions to eat toxic prey. *Anim Behav*. 86:733–740.
- Cole EF, Morand-Ferron J, Hinks AE, Quinn JL. 2012. Cognitive ability influences reproductive life history variation in the wild. *Curr Biol*. 22:1808–1812.
- Cole EF, Cram D, Quinn JL. 2011. Individual variation in spontaneous problem-solving performance among wild great tits. *Anim Behav*. 81:491–498.
- Dunlap AS, McLinn CM, MacCormick HA, Scott ME, Kerr B. 2009. Why some memories do not last a lifetime: dynamic long-term retrieval in changing environments. *Behav Ecol*. 20:1096–1105.
- Gonzalo A, López P, Martín J. 2013. Adaptive forgetting in Iberian green frog tadpoles (*Pelophylax perezi*): learned irrelevance and latent inhibition may avoid predator misidentification. *J Comp Psychol*. 127:56–62.
- Guillette LM, Morgan KV, Hall ZJ, Bailey IE, Healy SD. Forthcoming. Food preference and copying behavior in zebra finches *Taeniopygia guttata*. *Behav Proc*.
- Ferrari MC, Brown GE, Bortolotti GR, Chivers DP. 2010. Linking predator risk and uncertainty to adaptive forgetting: a theoretical framework and empirical test using tadpoles. *Proc Biol Sci*. 277:2205–2210.
- Halpin CG, Skelhorn J, Rowe C. 2008. Being conspicuous and defended: selective benefits for the individual. *Behav Ecol*. 19:1012–1017.
- Halpin CG, Skelhorn J, Rowe C. 2012. The relationship between sympatric defended species depends upon predators' discriminatory behaviour. *PLoS One*. 7:e44895.
- Hampton RR, Shettleworth SJ. 1996. Hippocampal lesions impair memory for location but not color in passerine birds. *Behav Neurosci*. 110:831–835.
- Harris AP, D'Eath RB, Healy SD. 2008. Sex differences, or not, in spatial cognition in albino rats: acute stress is the key. *Anim Behav*. 76:1579–1589.
- Healy SD, Rowe C. 2010. Information processing: the ecology and evolution of cognitive abilities. In: Westneat DE, Fox CW, editors. *Evolutionary behavioural ecology*. Oxford: OUP.
- Healy SD, Rowe C. 2013. Costs and benefits of evolving a larger brain: doubts over the evidence that large brains lead to better cognition. *Anim Behav*. 86:E1–E3.
- Houston A, McFarland D. 1976. Measurement of motivational variables. *Anim Behav*. 24:459–475.
- 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. *Anim Behav*. 86:829–838.
- Jones CM, Braithwaite VA, Healy SD. 2003. The evolution of sex differences in spatial ability. *Behav Neurosci*. 117:403–411.
- Jones CM, Healy SD. 2006. Differences in cue use and spatial memory in men and women. *Proc Biol Sci*. 273:2241–2247.
- Keagy J, Savard J-F, Borgia G. 2011. Complex relationship between multiple measures of cognitive ability and male mating success in satin bowerbirds, *Ptilonorhynchus violaceus*. *Anim Behav*. 81:1063–1070.
- Keagy J, Savard J-F, Borgia G. 2012. Cognitive ability and the evolution of multiple behavioral display traits. *Behav Ecol*. 23:448–456.
- Keagy J, Savard J-F, Borgia G. 2009. Male satin bowerbird problem-solving ability predicts mating success. *Anim Behav*. 78:809–817.
- Kelley LA, Endler JA. 2012. Male great bowerbirds create forced perspective illusions with consistently different individual quality. *Proc Natl Acad Sci USA*. 109:20980–20985.
- Krebs JR, Sherry DE, Healy SD, Perry VH, Vaccarino AL. 1989. Hippocampal specialization of food-storing birds. *Proc Natl Acad Sci USA*. 86:1388–1392.
- Líker A, Bókonyi V. 2009. Larger groups are more successful in innovative problem solving in house sparrows. *Proc Natl Acad Sci USA*. 106:7893–7898.
- Morand-Ferron J, Quinn JL. 2011. Larger groups of passerines are more efficient problem solvers in the wild. *Proc Natl Acad Sci USA*. 108:15898–15903.
- Odling-Smee L, Braithwaite VA. 2003. The influence of habitat stability on landmark use during spatial learning in the three-spined stickleback. *Anim Behav*. 65:701–707.
- Pearce JM. 2008. *Animal learning and cognition, 3rd ed: an introduction*. Psychology Press.
- Pompilio L, Kacelnik A, Behmer ST. 2006. State-dependent learned valuation drives choice in an invertebrate. *Science*. 311:1613–1615.
- Rescorla RA, Wagner AR. 1972. 'A theory of Pavlovian conditioning: variations in the effectiveness and nonreinforcement'. In: Black AH, Prokasy WF, editors. *Classical conditioning: current research and theory*. New York: Appleton-Century-Crofts. p. 64–99.
- Salvanes AG, Moberg O, Ebbesson LO, Nilsen TO, Jensen KH, Braithwaite VA. 2013. Environmental enrichment promotes neural plasticity and cognitive ability in fish. *Proc Biol Sci*. 280:20131331.
- Schwagmeyer PL, Parker PG, Mock DW, Schwabl H. 2012. Alternative matings and the opportunity costs of paternal care in house sparrows. *Behav Ecol*. 23:1108–1114.
- Sherry DE, Vaccarino AL, Buckenham K, Herz RS. 1989. The hippocampal complex of food-storing birds. *Brain Behav Evol*. 34:308–317.
- Shettleworth SJ. 2010. *Cognition, evolution and behavior*. Oxford: Oxford University Press.
- Sih A, Del Giudice M. 2012. Linking behavioural syndromes and cognition: a behavioural ecology perspective. *Philos Trans R Soc Lond B Biol Sci*. 367:2762–2772.
- Simmons LW. 2012. Resource allocation trade-off between sperm quality and immunity in the field cricket, *Teleogryllus oceanicus*. *Behav Ecol*. 23:168–173.
- Smulders TV, Sasson AD, DeVoogd TJ. 1995. Seasonal variation in hippocampal volume in a food-storing bird, the black-capped chickadee. *J Neurobiol*. 27:15–25.
- Spencer KA, Heidinger BJ, D'Alba LB, Evans NP, Monaghan P. 2010. Then versus now: effect of developmental and current environmental conditions on incubation effort in birds. *Behav Ecol*. 21:999–1004.
- Stoddard MC, Kilner RM. 2013. The past, present and future of 'cuckoos versus reed warblers'. *Anim Behav*. 85:693–699.
- ten Cate C, Rowe C. 2007. Biases in signal evolution: learning makes a difference. *Trends Ecol Evol*. 22:380–387.
- Thornton A, Samson J. 2012. Innovative problem solving in wild meerkats. *Anim Behav*. 83:1459–1468.
- van Heijningen CA, Chen J, van Laatum I, van der Hulst B, ten Cate C. 2013. Rule learning by zebra finches in an artificial grammar learning task: which rule? *Anim Cogn*. 16:165–175.