A new approach to comparing problem solving, flexibility and innovation

Alice M.I. Auersperg,1,* Gyula K. Gajdon4 and Auguste M.P. von Bayern2,3

1Department of Cognitive Biology; University of Vienna; Vienna, Austria; 2Department of Zoology; University of Oxford; Oxford, UK; 3Max-Planck-Institute for Ornithology; Seewiesen, Germany; 4Messerli Research Institute; University of Veterinary Medicine Vienna; Medical University Vienna; University of Vienna; Vienna, Austria

Keywords: problem solving, flexibility, tool use, birds, comparative cognition

Comparative cognition aims at unfolding the cognitive processes underlying animal behavior and their evolution, and is concerned with testing hypotheses about the evolution of the brain and intelligence in general. It is a developing field still challenged by conceptual and methodological issues. Systematic cross-species comparisons of cognitive abilities, taking both phylogeny and ecology into account are still scarce. One major reason for this is that it is very hard to find universally applicable paradigms that can be used to investi- gate the same cognitive ability or ‘general intelligence’ in several species. Many comparative paradigms have not paid sufficient attention to interspecific differences in anatomical, behavioral and perceptual features, besides psychological variables such as motivation, attentiveness or neophobia, thus potentially producing misrepresentative results. A new stance for future comparative research may be to establish behavioral and psychological profiles prior or alongside to comparing specific cognitive skills across species. Potentially revealing profiles could be obtained from examining species differences in how novel experimental (extractive foraging) tasks are explored and approached, how solutions are discovered and which ones are preferred, how flexibly multiple solutions are used and how much individual variation occurs, before proceeding to more detailed tests. Such new comparative approach is the Multi-Access-Box. It presents the animal with a novel problem that can be solved in several ways thus offering the possibility to examine species differences in all the above, and extract behavioral and perceptual determinants of their performance. Simultaneously, it is a suitable paradigm to collect comparative data about flexibility, innovativeness and problem solving ability, i.e., theoretical covariates of ‘general intelligence’, in a standardized manner.

Understanding the function, development and evolution of complex brain structures, remains a major scientific undertaking involving numerous disciplines and viewpoints.1 The contemporary field of ‘comparative cognition’ is concerned with how different species acquire, process, store and use information from their environment and how in turn the environment has shaped these cognitive processes in the course of evolution.2,3 By comparing cognitive adaptations to ecological and social environments in distantly and closely related species, comparative cognition may also permit the testing of hypotheses about convergent neural structure evolution in large brained animals.1,5

Yet, such systematic cross-species comparisons taking both phylogeny and ecology into account (such, for example, as the studies by Balda and Kamil;6 Balda et al.;7,8 Bednekoff et al.;9 Clayton and Krebs;10,11 Olson et al.;12 Tebbich et al.)5 are only in their early stages. Instead, comparisons of cognitive abilities across different species emerge in the literature typically either prompted by the discovery of hitherto uncharted abilities on certain taxa or through the expansion of an established paradigm to other species, often without consideration of the different ecological backgrounds. Thus, much of the published comparative work as yet consists of reports of specific cognitive abilities in different species in a yes/no tick list manner but without standardized testing procedures. Yet, unfolding the evolution of cognitive abilities through comparative research is not unjustly considered a notoriously complex and contentious topic. One major methodological issue is the difficulty of finding universally applicable paradigms that can be standardized and that are ecologically valid for the species to be compared.4 Additionally, the field is challenged by several other conceptual and methodo- logical difficulties.

Conceptually, the comparison of cognitive abilities across species is a delicate topic, since cognitive abilities may be functionally related though mechanistically distinctive in different species, particularly if their brains anatomies have diverged. Furthermore, there are vast interspecific differences in perception, manipulatory skills, motivation, etc that are methodologically relevant.13,14 Research on cognitive abilities is typically focused on the outward manifestation of potentially complex inner cognitive processes that are to date not directly observable and quantifi- able without MSR (which is so far impossible to administer in unrestrictedly moving animals) and that may vary considerably between species. They may consist of complex interplays of different brain regions that could differ between species. An even greater problematic applies to cross-species comparisons of ‘intelligence’, given that ‘intelligence’ ought to comprise flexible interaction and integration of not just one, but many cognitive abilities and associated processes. Accordingly, it is hard to pinpoint and target certain cognitive abilities in order to compare them across different species.
Still more delicate and often neither meaningful nor revealing are comparisons based on a linear scale of ‘intelligence’. Different species are adapted to different environments, and thus, at the most, have been selected for qualitatively different forms of ‘intelligences’, which they employ in specific situations. This has also been termed domain-specific intelligence meaning that cognitive abilities can be restricted to specific contexts in which they occur.

The notion of (domain-) general or ‘overall’ intelligence in contrast suggests that cognitive abilities could also be expanded to several domains and may allow species to behave flexibly, innovate, solve novel problems and thus cope with changes to their environment. Recent demonstrations that species differ in behavioral flexibility and propensity to adopt novel foraging behaviors, have revived discussions of overall animal intelligence. The predominant approach to investigate this has been to take innovation rate as proxy for intelligence and indeed correlation of overall brain size and/or size of the forebrain and rate of innovation were found in several animal groups including birds. Yet, these results should be interpreted with caution, given that that innovation rate was measured crudely that innovation does not necessarily involve intelligence (exploration patterns can be equally important-see discussion below) and that equating brain size with intelligence is too simplistic.

With these caveats in mind and relinquishing any ranking of intelligence, the question why and in how far species differ in their ability to innovate and solve novel problems (abilities commonly associated with general intelligence), and why these abilities correlate with brain size, is valid to ask. The unconfirmed hypothesis behind this, that also concerns the evolution of human intelligence is that innovativeness, problem solving ability and flexibility e.g., in extractive foraging, are selected for in certain, unpredictable or harsh environments, because such species would better adapt behaviorally and cope with the changing living conditions less flexible and innovative ones. This leads to the idea that increased flexibility, innovativeness and problem solving ability ultimately added up to what one might call a ‘general’ or ‘overall’ intelligence. Alternatively, this may be affected by a species’ mode of exploration and trial-and-error learning strategies; species that try more are clearly more likely to acquire new behavioral strategies through individual learning.

These questions can only be addressed by systematic comparative data on innovative problem solving and flexibility and a standardized paradigm ecologically valid for a range of species.

Comparing Problem Solving: Different Approaches

Apart from the correlation studies mentioned above, the ability to innovate and solve novel problems flexibly, proxies for flexible cognitive capacities and potentially for domain-general intelligence, have not been investigated comparatively. Few experimental studies have specifically focused on innovativeness, and flexible problem solving per se.

The majority of experiments in animal cognition, in both the physical and social domain, present animals with specific problems and investigate how the animals solve them in order to examine specific cognitive abilities/processes. These can be physical tasks, social situations or time related. Mostly, these tasks have been tailored to be ecologically valid for one particular species, or have focused on contexts/questions so specific that they could not be reproduced in other species easily (e.g., dolphins; food-storing scrub jays).

Other problem solving tasks, particularly in physical cognition, have secondarily become comparative paradigms and have been established as so-called ‘benchmark tests’ for examining the existence of certain cognitive abilities in different species, e.g., the ‘trap-tube’ for examining an understanding of causality in terms of surface continuity, Povinelli’s cane task and Heinrich’s ‘string-pulling task’ for testing responsiveness to connectivity etc. Most of these typically started off from a single experiment designed for testing a particular species and some subsequently applied to other species, sometimes without paying sufficient attention to species differences in morphological (hand or beak), behavioral (e.g., object exploration, affordance learning) and perceptual features (e.g., field of vision), in addition to psychological variables (such as motivational, emotional or attentional states, inhibitory control or neophobia/neophilia). Yet all these factors can potentially have a big impact masking cognitive skills actually present in a species and producing misrepresentative results.

Another methodological problem of comparative cognition is that paradigms are applied to many species, but with slightly converted methodologies (better fitting the newly compared species’ demands), hence at a cost of comparability. If the methodology is not standardised, it is hard to interpret the findings of comparative studies, because any detected differences between species could be owing to the different procedures. An odd handicap for comparative cognition in this context appears to be that modification or ‘improvement’ of an already used comparative paradigm, instead of merely replicating it in a new species, may increase the chances of a study to become published. Often however, a direct comparison would have been scientifically more valuable than yet another improvement to an existing experimental paradigm. Recently, this has been acknowledged and several research groups have begun to run comparative studies with the exact same methods.

The difficulty of comparative cognition therefore is to find comparative paradigms that are compatible with many different species (i.e., that are ecologically valid for all the species to be compared, and not influenced by potentially confounding species-specific variables), and that have a standardised methodology that can be applied to different species in exactly the same way.

Because of the potentially confounding impact of different methodologies, the same cognitive ability should be investigated with not just a single but several tests with slightly different angles in any given species. Recently, comparative studies carry out entire batteries in different species tests, e.g., those comparing cognitive abilities in the social and physical domain. Yet, what is missing is to have a battery of tests establishing species differences that might affect performance in different cognitive tests, such as object exploration, motivation, attentiveness and fear/neophobia.
The Multi Access Box Approach

A new such comparative approach is the Multi-Access-Box (MAB), recently published in PLoS One (see Fig. 1). It presents the animal with a novel problem that can be solved in four different ways, thus offering the possibility to examine species differences in how novel problems are perceived, explored and approached and in which order solution(s) are discovered. This provides several data that can be used for establishing a behavioral (e.g., object exploration,) and psychological profile (e.g., motivation, flexibility, impulsivity, persistence, inhibitory control) and hence extract behavioral and perceptual determinants of different species’ performance in the tasks. Simultaneously, it is a suitable paradigm to collect data about problem solving ability, innovativeness and flexibility, i.e., theoretical covariates of ‘general intelligence’, across species in a standardized manner.

The MAB approach comprises not just one but several solutions to an extractive foraging problem at the same time (food out of reach in the center of a transparent box), i.e. it consists of a battery of alternative tasks that all lead to the same goal. Two solutions (opening a window and pulling a string) could be discovered by haptic exploration (touching the box at particular sites), while the other two additionally required the handling of objects, either wooden sticks or marbles, as tools (inserting a ball or a stick tool into specific openings).

The other important feature of the MAB is that subjects were forced to continue exploring alternative solutions, once they had successfully discovered and consistently used one particular solution, by blocking the one in use. This creates an order system which allows to detect species differences in which tasks are approached and explored first and how, how many solutions are discovered and how fast, whether and how quickly the subjects switch between options or whether they focus or settle on particular ones, as well as which tasks are problematic and why. In this manner we can detect not only species differences in problem solving performance, but also learn about the various underlying non-cognitive factors that may affect it. Although designed for large scale comparisons of different closely- and distantly related species from different ecological backgrounds, the initial MAB study compared just two avian species from different families, a corvid, the New Caledonian crow (NCC; Corvus moneduloides) and a parrot, the kea (Nestor notabilis) (see Fig. 2). Both species are known for their large brains, their innovativeness and problem solving skills, but NCCs are naturally tool-using species while kea have not been observed to use tools in the wild, but are famous for their neophilia. Subjects were exposed to the transparent MAB with the food reward in the center, which could be extracted by the four different methods corresponding to the four walls of the box. Once a method was mastered, it was blocked and the bird’s performance...
in reaching criterion in any of the others was recorded until all four methods had been discovered.

Auersperg et al.\textsuperscript{13} found that one kea and one NCC detected all four solutions, demonstrating that the solutions offered were within both species’ capacity. The kea were much quicker in discovering multiple solutions than the habitually tool-using NCCs and showed more individual variation. The keas were also more flexible once openings were blocked, switching to other solutions much quicker than the NCC.

Innovation rate as well as performance in this paradigm were strongly impacted by differences in exploration technique and neophilia rather than by cognitive discrepancies. The highly neophilic kea explored the apparatus more in a haptic than in a visual manner. They found its functional properties, while manipulating the affordances\textsuperscript{31,52} of the MAB they perceived as most salient. In contrast to the kea and probably due to their more visually guided exploration technique, the NC crows had problems solving the window solution. The window mechanism could not be deduced by visual inspection alone (without knowledge of hinge-mechanisms), but could be readily discovered by haptic exploration. Another difference was that the NC crows tended to persist with the first option that worked, whereas the kea, owing to their higher level of neophilia, switched between solutions.

Differences in beak morphology also affected the birds’ performance: the kea had problems maneuvering the stick tool because of their beak curvature, whereas the crows with their straight beaks had a good grip of the tools. Yet, the NC crows used their straight beak more for pecking than tearing actions, which would have been advantageous in detecting the MAB apparatus’s affordances in case of the window option (grasping and pulling the window crank).

An important new tool that could be incorporated in the MAB procedure and that could be revealing in comparisons of flexible problem solving may be ‘reversal learning’.\textsuperscript{5,53-55} Species with different ecological backgrounds may have been selected for different strategies in trial and error learning and problem solving. In terms of energy pay-off it may, under certain circumstances, be advantageous to persist (e.g., in the case of NCCs fishing for particularly nutritious wood-boring beetle larvae in rotten wood as, which can take considerable time but has a high potential return).\textsuperscript{56} In other contexts it may save energy to give up if something does not work and try something else instead. Reversal learning tasks reflect differences in flexibility and are informative of how fast animals can adjust their behavior to new external feedback, let go of previously reinforced behavioral patterns, but at the same time they offer a measure of persistence.

\textbf{Figure 2.} Kea and New Caledonian crow using the stick (left) and the ball-shaped (right) tool on the Multi Access Box.
To illustrate how reversal learning could be implemented, we present some data not published in the original Auersperg et al. paper. Once a subject had discovered all solutions, we incorporated a reversal learning task for the two solutions incorporating a tool (see Fig. 2). For the two successful subjects, Uék, a NCC, and Kermit, a kea, the last used tool option was blocked and the previous tool option was reopened. The results are presented in Table 1. As can be seen the crow required a similar amount of trials to relearn the previous option as the kea, although of course data on more individuals would be desirable.

The outcomes of Auersperg et al. illustrate how even diminutive differences in non-cognitive behavioral components such as neophilia or morphology can mask and/or interfere with the cognitive differences in non-cognitive behavioral components such as (haptic or visual exploration mode) and affordance learning, particularly what major impact differences in object exploration can have on performance in artificial experimental tasks, and hence how this affects the comparability of two species in the same task.

In future comparative research, establishing behavioral and psychological profiles of the species to be compared ought to precede comparative tests of specific cognitive skills or general intelligence. This may help to identify problem solving tasks that are equivalently applicable to the target species and hence achieve a high degree of ‘comparability’ of the obtained data.

Table 1. Reversal sessions for the two subjects (Uék = NCC, Kermit = Kea), which reached in all solutions of the Multi Access box. Session number in which reversal took place. Number of trials to reach criterion (9/10 correct)

| Subject          | From Ball to Stick | From Stick to Ball | From Window to Stick |
|------------------|--------------------|--------------------|---------------------|
| Uék              | 15                 | 14                 | 13                  |
| Kermit           | 12                 | 11                 | -                   |

References

1. Kamel AC. A synthetic approach to the study of animal intelligence. In: Leger DW, ed. Comparative Perspectives in Modern Psychology: Nebraska Symposium on Motivation, Lincoln, NE: U. of Nebraska Press, 1998; 35, 230-257.
2. Shettleworth SJ. The evolution of comparative cognition: is the snark still a boojum? Behav Processes 2009; 80:210-7; PMID:18824222; http://dx.doi.org/10.1016/j.beproc.2008.09.001
3. Shettleworth S. Cognition Evolution and Behaviour. Oxford University Press, 2010.
4. Maclean EL, Matthews LJ, Hare BA, Nunn CL, Anderson RC, Aureli F, et al. How does cognition evolve? Phylogenetic comparative psychology. Anim Cogn 2011. In press. PMID:21927850; http://dx.doi.org/10.1007/s10071-011-0448-8
5. Tébbich S, Sterelny K, Troschke I. The tale of the finch: adaptive radiation and behavioural flexibility. Phalos Trans R Soc Lond B Biol Sci 2010; 365:1099-109; PMID:20194172; http://dx.doi.org/10.1098/rspb.2009.0291
6. Balda RP, Kamel AC. A comparative study of cache recovery by three corvid species. Anim Behav 1989; 38:486-95; http://dx.doi.org/10.1016/S0003-3472(89)80041-7
7. Balda RP, Kamel AC, Bednokoff PA. Predicting cognitive capacity from natural history: examples from four species of corvids. In Nolan V, Ketterson ED, eds. Current Ornithology. New York: Plenum Press, 1996:13, 33-66.
8. Balda RP, Kamel AC, Bednekoff PA, Hile AG. Species differences in spatial memory performance on a three-dimensional task. Ethology 1997; 103:47-55; http://dx.doi.org/10.1046/j.1439-0310.1997.00006.x
9. Bednekoff PA, Balda RP, Kamel AC, Hile AG. Long-term spatial memory in four seed-caching corvid species. Anim Behav 1997; 53:335-41; http://dx.doi.org/10.1006/anbe.1996.0395
10. Clayton NS, Krets JR. Memory for spatial and object-specific cues in food-storing and non-storing birds. J Comp Psychol A 1994; 174:371-9.
11. Clayton NS, Dally JM, Emery NJ. Social cognition by food caching corvids: The western scrub-jay as a natural psychologist. Phil Trans Roy Soc B 2007; 326:506.
32. Huber L, Gajdon GK. Technical intelligence in animals: the kea model. Anim Cogn 2006; 9:295-305; PMID: 16099237; http://dx.doi.org/10.1007/s10071-006-0033-8
33. Gajdon GK, Lichtnegger M, Huber L. Toddler like object exploration helps captive mountain parrots to innovate tool use behaviour. Submitted.
34. Bird CD, Emery NJ. Insightful problem solving and creative tool modification by captive non-tool-using rooks. Proc Natl Acad Sci U S A 2009; 106:10370-5; PMID:19478068; http://dx.doi.org/10.1073/pnas.090108106
35. Burkart A, Strasser MF. Trade-offs between social learning and individual innovativeness in common marmosets, Callithrix jacchus. Anim Behav 2009; 77: 1291-301; http://dx.doi.org/10.1016/j.anbehav.2009.02.006
36. Reader SM, Macdonald K. Environmental variability and primate behavioural flexibility. In: Reader SM, Laland KN, eds. Animal Innovation. Oxford: Oxford University Press, 2003: 83-116.
37. Povinelli DJ. Folk physics for apes: a chimpanzee's theory by how the world works. Oxford University Press, Oxford, 2000.
38. Hare B, Call J, Agnetta B, Tomasello M. Chimpanzees 'know what conspecifics do and do not see. Anim Behav 2001; 62:945-54; http://dx.doi.org/10.1006/anbe.1999.1377
39. Hare B. Can competitive paradigms increase the validity of social cognitive experiments on primates? Anim Cogn 2003: 580-602.
40. Emery NJ, Clayton NS. Effects of experience and social context on prospective caching strategies by scrub jays. Nature 2001; 414:443-6; PMID:11719804; http://dx.doi.org/10.1038/35106560
41. Raby CR, Alexia DM, Dickinson A, Clayton NS. Planning for the future by western scrub-jays. Nature 2007; 445:919-21; PMID:17314979; http://dx.doi.org/10.1038/nature05579
42. Wimpenny JH, Weir AAS, Clayton L, Rutz C, Kachelh A. Cognitive processes associated with sequential tool use in New Caledonian crows. PLoS One 2009; 4:e6471; PMID:19654861; http://dx.doi.org/10.1371/journal.pone.0006471
43. Kuczaj SA, Walker RT. How do dolphins solve problems? In Wasserman EA, Zentall TR, eds. Comparativa Cognition. Oxford University Press, 2006:580-602.
44. Visalberghi E, Limongelli L. Lack of comprehension of cause-effect relations in tool-using capuchin monkeys (Cebus apella). J Comp Psychol 1994; 108:15-22; PMID:8174341; http://dx.doi.org/10.1037/0735-7036.108.1.15
45. Heinrich B. An experimental investigation of insight in common ravens (Corvus corax). Auk 1995; 112:994-1003.
46. Herrmann E, Call J, Hernández-Lloreda MV, Hare B, Tomasello M. Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. Science 2007; 317:1360-6; PMID:17823346; http://dx.doi.org/10.1126/science.1146282
47. Liedke J, Werdenich D, Gajdon GK, Huber L, Wanker R. Big brains are not enough: performance of three parrot species in the trap-tube paradigm. Anim Cogn 2011; 14:143-9; PMID:20814805; http://dx.doi.org/10.1007/s10071-010-0347-4
48. Schloegl C, Dierks A, Gajdon GK, Huber L, Kotrschal K, Bugnyar T. What you see is what you get? Exclusion performances in ravens and keas. PLoS One 2009; 4: e6368; PMID:19654864; http://dx.doi.org/10.1371/journal.pone.0006368
49. Herrmann E, Hernández-Lloreda MV, Call J, Hare B, Tomasello M. The structure of individual differences in the cognitive abilities of children and chimpanzees. Psychol Sci 2010; 21:102-10; PMID:20424039; http://dx.doi.org/10.1127/j0077295