Report

Crows Spontaneously Exhibit Analogical Reasoning

Anna Smirnova,[1] Zoya Zorina,[1] Tanya Obozova,[1]
and Edward Wasserman[2,*]
[1]Department of Biology, Lomonosov Moscow State
University, Moscow 119899, Russia
[2]Department of Psychology, The University of Iowa, Iowa City,
IA 52242, USA

Summary

Analogical reasoning is vital to advanced cognition and behavioral adaptation. Many theorists deem analogical thinking to be uniquely human and to be foundational to categorization, creative problem solving, and scientific discovery [1]. Comparative psychologists have long been interested in the species generality of analogical reasoning, but they initially found it difficult to obtain empirical support for such thinking in nonhuman animals (for pioneering efforts, see [2, 3]). Researchers have since mustered considerable evidence and argument that *relational matching-to-sample* (RMTS) effectively captures the essence of *analogy*, in which the relevant logical arguments are presented visually [4]. In RMTS, choice of test pair BB would be correct if the sample pair were AA, whereas choice of test pair EF would be correct if the sample pair were CD. Critically, no items in the correct test pair physically match items in the sample pair, thus demanding that only relational sameness or differentness is available to support accurate choice responding. Initial evidence suggested that only humans and apes can successfully learn RMTS with pairs of sample and test items [4–7]; however, monkeys have subsequently done so [8–12]. Here, we report that crows too exhibit relational matching behavior. Even more importantly, crows *spontaneously* display relational responding without ever having been trained on RMTS; they had only been trained on *identity matching-to-sample* (IMTS). Such robust and uninstructed relational matching behavior represents the most convincing evidence yet of analogical reasoning in a nonprimate species, as apes alone [7] have spontaneously exhibited RMTS behavior after only IMTS training.

Results and Discussion

Our findings come from three separate behavioral assessments that followed pretraining and pretesting on identity matching-to-sample (IMTS; see the Supplemental Experimental Procedures available online), deploying behavioral methods that were earlier used by Smirnova, Lazareva, and Zorina [13]. In that prior IMTS period, two hooded crows were shown several different kinds of visual stimuli: single stimuli of varying colors or shapes and clusters of stimuli containing varying numbers of items. The birds were taught to match stimuli by color, shape, and number of items, but not size. After learning these various IMTS tasks, the crows reliably transferred discriminative responding to new stimuli from the *same* categories that were used in training

*Correspondence: ed-wasserman@uiowa.edu

(Table S2) as well as to stimuli from an altogether *different* category (Figure S2)—stimuli varying in size—suggesting that the birds had acquired a general rule that was based on physical identity [14].

The key behavioral assessments followed, in which novel pairs of items served as the sample and test stimuli [7]. The visual dimensions in these assessments were size, shape, and color (Figures 1, 2, and 3 show several illustrative, not actual, trial sequences). On *identity* trials, we arranged *differential* reinforcement (food was given only after correct choices) to promote continued discriminative responding; accurate choice responding here could be based on either physical or relational matches between the sample pair and the correct test pair. On the critical *relational* trials, we arranged *nondifferential* reinforcement (food was given after all choices) because we did not want to teach the crows the very behavior that we were assessing; accurate choice responding here could only be based on relational matches between the sample pair and the correct test pair. Assessment sessions were conducted 5 days a week and contained 48 trials: 36 identity trials and 12 relational trials. The sessions contained six blocks (each block containing six identity trials and two relational trials); the trial order was randomized within each of the six blocks. On all trials, we scored as "correct" those choices that accorded with relational matching. Each assessment phase lasted eight sessions.

The first assessment phase involved size. Figure 1 shows that on half of the trials the sample pair involved shapes of the same size, whereas on the other half of the trials the sample pair involved shapes of different sizes. With the different-sized sample and choice pairs, the large shape always appeared on the left and the small shape always appeared on the right in order to keep the number of sample-comparison combinations within manageable limits. The palette of simple black shapes from which the sample and test pairs could be selected contained a square, a rectangle, a circle, an oval, a wide triangle, and a narrow triangle. These shapes could be large or small; the overall area of the two test pairs was equated, so that this stimulus property could not control the crows' behavior. Identity matching trials were arranged in which one test pair presented the same shapes in the same sizes as the sample pair; each of the sample and test pairs involved the same shape, and only correct choices were reinforced on these trials. Relational matching trials were arranged in which neither of the test pairs matched the sample pair in shape, thereby eliminating control by physical identity; on these trials, either correct or incorrect choices were reinforced.

The second assessment phase involved shape. Figure 2 shows that on half of the trials the sample pair involved identical shapes, whereas on the other half of the trials the sample pair involved nonidentical shapes. In this assessment phase and the next, the items in each pair were arranged along the positive diagonal in order for the spatial arrangement to differ from the horizontal arrangement in the prior assessment phase and to keep the number of sample-comparison combinations within manageable limits. The palette of shapes contained equal-sized circles, squares, triangles, and crosses that could be colored red, yellow, blue, or green. Identity matching trials

Trial	Left Test	Sample	Right Test
1			
2			
3			
4			
5			
6			
7			
8			

Figure 1. Examples of Identity and Relational Trials across Eight Exemplary Trials for Size Stimuli

On three-fourths of the trials (1–3 and 5–7), the correct test stimulus was an identity match to the sample. On one-fourth of the trials (4 and 8, shaded gray), the correct test stimulus was a relational match to the sample.

Trial	Left Test	Sample	Right Test
1			
2			
3			
4			
5			
6			
7			
8			

Figure 2. Examples of Identity and Relational Trials across Eight Exemplary Trials for Shape Stimuli

On three-fourths of the trials (1–3 and 5–7), the correct test stimulus was an identity match to the sample in shape. On one-fourth of the trials (4 and 8, shaded gray), the correct test stimulus was a relational match to the sample.

were arranged in which one test pair presented the same shapes as the sample pair; only correct choices were reinforced on these trials. Relational matching trials were arranged in which neither of the test pairs matched the sample pair in shape or color, thereby eliminating control by physical identity; on these trials, either correct or incorrect choices were reinforced.

The third assessment phase involved color (crows have excellent color vision [15]). Figure 3 shows that on half of the trials the sample pair involved identical colors, whereas on the other half of the trials the sample pair involved nonidentical colors. The palette of colors contained equal-sized circles, squares, triangles, and crosses that could be shown in red, yellow, blue, or green. Identity matching trials were arranged in which one test pair presented the same colors as the sample pair; only correct choices were reinforced on these trials. Relational matching trials were arranged in which neither of the test pairs matched the sample pair in color or shape, thereby eliminating control by physical identity; on these trials, either correct or incorrect choices were reinforced.

In the size-assessment phase, 60 combinations of stimuli were used as samples on relational trials (they were never repeated within sessions and were presented a maximum of two times, in different locations, across sessions) and 12 combinations of stimuli were used as samples on identity trials (they were never repeated within sessions and were presented a maximum of 24 times across sessions). In the shape and color assessment phases, trial-unique combinations of stimuli were used as samples on all relational and identity trials.

Throughout all three assessment phases, the crows exhibited highly accurate choice responding on both identity and relational trials (fourth column in Table 1). We individually assessed each crow's behavior with a logistic regression examining trial type, stimulus dimension, and session. These

regression analyses revealed no significant changes in choice accuracy over the eight sessions in each phase [crow 1: $\chi^2(1) = 0.94$, p = 0.33; crow 2: $\chi^2(1) = 1.57$, p = 0.21] and no interactions of Session with the other variables (daily accuracy scores are reported in Table S1, and details of the statistical analyses are reported in the Supplemental Experimental Procedures); therefore, subsequent discussion does not consider this factor.

We next compared accuracy scores on the different kinds of trials against 50% (representing random choice between the test pairs). Averaged across all eight sessions, crow 1 responded at significantly above chance accuracy levels to all six different kinds of stimuli (B = 1.06, SE = 0.068, Z = 15.57, p < 0.0001), with accuracy ranging from 61.46% to 79.51% correct. There was a significant main effect of trial type [$\chi^2(1) = 3.99$, p = 0.0457], with slightly lower accuracy on relational trials (69.44%) than on identity trials (75.46%). There was also a marginally significant trial type × stimulus dimension interaction [$\chi^2(2) = 5.67$, p = 0.059]. Follow-up tests within each stimulus dimension disclosed that the interaction was due to a significant effect of trial type for the color dimension (B = −0.68, SE = 0.25, Z = 2.74, p = 0.0061), but not for the shape (B = 0.18, SE = 0.27, Z = 0.68, p = 0.49) or size dimensions (B = −0.36, SE = 0.27, Z = 1.33, p = 0.183), suggesting that accuracy on color relational trials was a bit poorer than on the remaining kinds of trials. Averaged across all eight sessions, crow 2 responded at significantly above chance accuracy levels to all six different kinds of stimuli (B = 1.13, SE = 0.067, Z = 16.34, p < 0.0001). Accuracy ranged from 71.18% to 78.82% correct, with accuracy being numerically (but not reliably) lower on identity trials (72.22%) than on relational trials (80.56%). The logistic regression yielded no other significant effects.

Because of the importance of spontaneity to interpreting the crows' relational matching-to-sample (RMTS) behavior, we

Trial	Left Test	Sample	Right Test
1			
2			
3			
4			
5			
6			
7			
8			

Figure 3. Examples of Identity and Relational Trials across Eight Exemplary Trials for Color Stimuli

On three-fourths of the trials (1–3 and 5–7), the correct test stimulus was an identity match to the sample in color. On one-fourth of the trials (4 and 8, shaded gray), the correct test stimulus was a relational match to the sample.

Table 1. Mean Percentage of Correct Choices by Crows 1 and 2

Dimension	Trial Type	Session 1	Sessions 1–8
Crow 1			
Size	Identity	77.78	79.51
	relational	75.00	72.92
Shape	identity	75.00	70.83
	relational	91.67	73.96
Color	identity	66.67	76.04
	relational	58.33	61.46
Crow 2			
Size	identity	75.00	76.39
	relational	91.67	76.04
Shape	identity	63.89	71.18
	relational	75.00	77.08
Color	identity	77.78	78.82
	relational	75.00	71.88

Accuracy scores are from identity (IMTS) and relational (RMTS) trials on all three dimensions in session 1 and across sessions 1–8. See also Table S1.

also analyzed choice behavior in session 1 of each assessment phase (third column in Table 1) using a similar logistic regression as in the prior analyses. In session 1, crow 1 responded at significantly above chance levels of accuracy ($B = 1.08$, SE = 0.20, $Z = 5.39$, p < 0.0001), but there were no significant effects of trial type or stimulus dimension; summed across all three dimensions, accuracy on identity trials averaged 73.15% correct, and accuracy on relational trials averaged 75.00% correct. Crow 2 also responded at significantly above chance accuracy levels in session 1 ($B = 1.11$, SE = 0.20, $Z = 5.53$, p < 0.0001), but there were no significant effects of trial type or stimulus dimension; summed across all three dimensions, accuracy on identity trials averaged 72.22% correct, and accuracy on relational trials averaged 80.56% correct.

We thus found that, when tested during the size-assessment phase immediately after initial IMTS training, our crows not only responded discriminatively on identity matching trials, but also did so on relational matching trials, all of these trials for the first time involving two-item sample and test stimuli. These results represent striking behavioral evidence that the crows spontaneously perceived the relation between relations without ever having been explicitly trained to do so.

That initial size assessment entailed sample and comparison stimuli that were horizontally ordered large (left) to small (right) with different-sized pairs. Could that incidental spatial cue have affected the accuracy of the crows' later choice behavior? Evidently not. The crows continued to respond at similarly high levels of accuracy on identity and relational trials during the shape and color assessment phases, during which the items in the sample and comparison pairs were diagonally arranged.

This documentation of RMTS behavior is particularly noteworthy because our crows exhibited discriminative relational matching (averaging 77.78% correct) that was just as robust as their identity matching (averaging 72.69%) in session 1 across all three assessment phases. Although physical identity could have guided the crows' choice behavior on IMTS trials, physical identity could not have done so on RMTS trials, as no physical matches were possible between the sample pairs and the correct test pairs. These results perhaps surprisingly suggest that physical identity contributed little or nothing to our crows' testing performance; relational processing seems to have been of prime importance to controlling the birds' choice behavior.

Our results thus constitute unprecedented behavioral evidence of analogical reasoning by a nonprimate animal. They therefore add to growing research undermining the influential claims of such famous philosophers as René Descartes and John Locke that only humans are capable of abstract thought. Relational reasoning—particularly appreciating the relation between relations, as in analogies—can no longer be deemed to be the unique pinnacle of human cognition.

It may be no accident that crows performed so impressively in our study; they stand out among birds in their highly developed neuroanatomy [16, 17]. More generally, mounting evidence indicates that although birds do not have a brain structure that is homologous to the mammalian prefrontal cortex, the avian nidopallium caudolaterale may effectively mediate complex cognitive functions, perhaps representing a case of convergent evolution [18].

It should nevertheless be appreciated that we are not claiming that our crows' spontaneous relational matching behavior arose entirely de novo. Indeed, we believe that their earlier IMTS training is likely to have enabled them to grasp a broadly applicable concept of sameness that could apply to novel two-item sample and test stimuli involving only relational sameness [19, 20]. Just how that remarkable transfer is accomplished represents an intriguing matter for future study. Nor are we claiming that crows will prove to be the only nonprimate animals that are capable of exhibiting such spontaneous relational matching behavior. Future research must be undertaken in which different species are given comparable pretraining experience to our crows. Until such systematic comparative research is conducted, it would be premature to offer speculative evolutionary accounts as to why crows appear to have excelled in solving this challenging cognitive task. It would also be premature to suggest that simply because crows

successfully master RMTS tasks, they process the stimuli and relations in the same way as do humans and other nonhuman animals.

Finally, we should underscore the importance of research on animal behavior to the role of language or symbol systems in abstract conceptualization. The first evidence of analogical reasoning in animals came from symbol-trained chimpanzees [2, 5]. Accordingly, the hypothesis was advanced that only if an organism had acquired an elaborate symbol system might it encode and process abstract analogical relations [5]. The results of more recent primate studies [6, 7, 12], as well as the findings from our present experiment with crows, suggest that extensive prior experience with the abstract concept of "sameness" per se—rather than a linguistic symbol for such a concept—is sufficient for animals to succeed in solving analogical reasoning tasks.

Experimental Procedures

Subjects
Two experimentally naive hooded crows (*Corvus corone*), at least 2 years old, served as subjects. Both were housed in the aviary of the Biology Department of Lomonosov Moscow State University, Russia. Throughout the experiment, the birds had free access to water. Mealworms were used as reinforcement, as they are crows' favorite food and are attractive at most food-deprivation levels. If the crows refused to work in the experiment, then they received food without animal protein for 1 or 2 days. Refusals to work meant that unfinished sessions were continued on the next day. All of this research was conducted in full compliance with Russian research regulations: specifically, the bioethical requirements of Directive 86 EC.

Apparatus
A wire mesh cage (70 cm × 35 cm × 35 cm; 4 cm × 4 cm mesh spacing) and a plastic tray (20 cm × 30 cm) with a handle (30 cm) were used for training and testing (Figure S1). Two cups (3.7 cm high and 5.0 cm in diameter) were placed on the tray; during training, one cup contained two mealworms and the other cup was empty, whereas during testing both cups contained mealworms. The cups were covered by the comparison stimuli. The sample stimulus was placed between the comparison stimuli. All of the stimuli were drawn on cardboard cards (7 cm × 7 cm). An opaque plastic screen (70 × 40 cm) was placed between the experimenter and the crow; neither the bird nor the experimenter could see one another, precluding a "Clever Hans" error. Before each trial, the tray was prepared out of the bird's sight. The Supplemental Information includes additional procedural details, controls for confounded variables, and Movie S1 (a reenactment showing crow 1 performing the task).

General Procedure
The crows were trained and tested in two-alternative simultaneous matching-to-sample tasks. During the experiment, a bird was placed into the experimental cage. A trial started when the tray—containing the sample stimulus card in the center and the two cups covered by the comparison stimulus cards on each side—was slid into the cage. So that the crow could be given the opportunity to get acquainted with all three stimuli, the tray was initially placed in front of the bird for 2 to 3 s as the first step of the trial; the crow could see the cards, but it could not uncover the cups. Then, the second step followed. The tray was moved more deeply into the cage; the crow uncovered one of the cups and, in case of the correct choice, it received food. If the bird did not choose either card within 2 min, then the tray was removed from the cage. Which sample stimulus would be presented was determined by a quasirandom schedule under the restrictions that (1) the same card could not be used as a sample more than two times in succession and (2) the correct stimulus could not appear in the right or left location more than two times in succession. Otherwise, the order of stimulus presentation and the pairing of sample and comparison stimuli were randomized on each trial.

The minimum intertrial interval (ITI) was about 1 min, essentially the time to prepare the tray with new stimuli and mealworms. The maximum ITI was 5 min. If the bird did not choose either card within 2 min, then the tray was removed from the cage. The experimenter gave the bird a 5 min break and repeated that trial. If the bird again did not choose either card, then the experimenter ended the session.

Supplemental Information

Supplemental Information includes Supplemental Experimental Procedures, two figures, two tables, and one movie and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2014.11.063.

Author Contributions

A.S. and Z.Z. designed and conducted the studies and interpreted the data. T.O. and E.W. helped conduct the statistical analyses and interpret the data and composed the paper in collaboration with the other coauthors.

Acknowledgments

References

1. Gentner, D. (1999). Analogy. In The MIT Encyclopedia of the Cognitive Sciences, R.A. Wilson and F.C. Keil, eds. (Cambridge: MIT Press), pp. 17–20.

2. Gillan, D.D., Premack, D., and Woodruff, G. (1981). Reasoning in the chimpanzee: I. Analogical reasoning. J. Exp. Psychol. Anim. Behav. Process. *7*, 1–17.

3. Oden, D.L., Thompson, R.K.R., and Premack, D. (2001). Can an ape reason analogically? Comprehension and production of analogical problems by Sarah, a chimpanzee (*Pan troglodytes*). In The Analogical Mind: Perspectives from Cognitive Science, D. Gentner, ed. (Cambridge: MIT Press).

4. Thompson, R.K.R., and Oden, D.L. (2000). Categorical perception and conceptual judgments by nonhuman primates: The paleological monkey and the analogical ape. Cogn. Sci. *24*, 363–396.

5. Premack, D. (1983). The codes of man and beast. Behav. Brain Sci. *6*, 125–137.

6. Thompson, R.K.R., Oden, D.L., and Boysen, S.T. (1997). Language-naive chimpanzees (*Pan troglodytes*) judge relations between relations in a conceptual matching-to-sample task. J. Exp. Psychol. Anim. Behav. Process. *23*, 31–43.

7. Vonk, J. (2003). Gorilla (*Gorilla gorilla gorilla*) and orangutan (*Pongo abelii*) understanding of first- and second-order relations. Anim. Cogn. *6*, 77–86.

8. Fagot, J., and Maugard, A. (2013). Analogical reasoning in baboons (*Papio papio*): flexible reencoding of the source relation depending on the target relation. Learn. Behav. *41*, 229–237.

9. Fagot, J., and Parron, C. (2010). Relational matching in baboons (*Papio papio*) with reduced grouping requirements. J. Exp. Psychol. Anim. Behav. Process. *36*, 184–193.

10. Fagot, J., and Thompson, R.K.R. (2011). Generalized relational matching by guinea baboons (*Papio papio*) in two-by-two-item analogy problems. Psychol. Sci. *22*, 1304–1309.

11. Flemming, T.M., Thompson, R.K.R., and Fagot, J. (2013). Baboons, like humans, solve analogy by categorical abstraction of relations. Anim. Cogn. *16*, 519–524.

12. Fagot, J., Wasserman, E.A., and Young, M.E. (2001). Discriminating the relation between relations: the role of entropy in abstract conceptualization by baboons (*Papio papio*) and humans (*Homo sapiens*). J. Exp. Psychol. Anim. Behav. Process. *27*, 316–328.

13. Smirnova, A.A., Lazareva, O.F., and Zorina, Z.A. (2000). Use of number by crows: investigation by matching and oddity learning. J. Exp. Anal. Behav. *73*, 163–176.

14. Katz, J.S., Wright, A.A., and Bodily, K.D. (2007). Issues in the comparative cognition of abstract-concept learning. Comp. Cogn. Behav. Rev. *2*, 79–92.

15. Håstad, O., Victorsson, J., and Ödeen, A. (2005). Differences in color vision make passerines less conspicuous in the eyes of their predators. Proc. Natl. Acad. Sci. USA *102*, 6391–6394.

16. Emery, N.J. (2006). Cognitive ornithology: the evolution of avian intelligence. Philos. Trans. R. Soc. Lond. B Biol. Sci. *361*, 23–43.

17. Portmann, A. (1947). Études sur la cérébralisation chez les oiseaux. II. Les indices intracérébraux. Alauda *15*, 1–15.

18. Güntürkün, O. (2012). The convergent evolution of neural substrates for cognition. Psychol. Res. *76*, 212–219.

19. Truppa, V., Piano Mortari, E., Garofoli, D., Privitera, S., and Visalberghi, E. (2011). Same/different concept learning by capuchin monkeys in matching-to-sample tasks. PLoS ONE *6*, e23809.

20. Wasserman, E.A. (2008). Development and evolution of cognition: one doth not fly into flying!. Behav. Brain Sci. *31*, 400–401.