Social complexity and transitive inference in corvids

ALAN B. BOND*, ALAN C. KAMIL* & RUSSELL P. BALDA†

*School of Biological Sciences, University of Nebraska, Lincoln
†Department of Biological Sciences, Northern Arizona University

The social complexity hypothesis asserts that animals living in large social groups should display enhanced cognitive abilities along predictable dimensions. To test this concept, we compared highly social pinyon jays, *Gymnorhinus cyanocephalus*, with relatively nonsocial western scrub-jays, *Aphelocoma californica*, on two complex cognitive tasks relevant to the ability to track and assess social relationships. Pinyon jays learned to track multiple dyadic relationships more rapidly and more accurately than scrub-jays and appeared to display a more robust and accurate mechanism of transitive inference. These results provide a clear demonstration of the association between social complexity and cognition in animals.

© 2003 Published by Elsevier Science Ltd on behalf of The Association for the Study of Animal Behaviour.

Animals that live in large, stable social groups face substantial cognitive demands. They must recognize other individuals in their group, continuously track the position, social behaviour and foraging success of other individuals, classify group members by age, sex, genetic relationship, reproductive status and dominance rank, and update this information as circumstances change (Cheney & Seyfarth 1990). Given these demands, it has been hypothesized that highly social animals should show commensurate increases in cognitive abilities related to social living (Kummer et al. 1997). Indeed, social complexity may have provided the primary selective impetus for the evolution of intelligence in primates (Jolly 1966; Humphrey 1976; Byrne & Whiten 1988).

The primary literature addressing this hypothesis has focused on data from field studies of Old World primates, constructing evolutionary accounts of the functional significance of their cognitive abilities (reviewed in Byrne & Whiten 1997). But these studies provide little in the way of direct tests of the evolutionary hypothesis, largely because comparable data from species differing in sociality are difficult to obtain with field methodologies. There have also been very extensive experimental studies on the cognitive capabilities of animals (reviewed in Shettleworth 1998). But for the most part, the tasks used have not been designed with regard to the cognitive demands of social complexity, and the species studied have not been selected with sociality in mind. As

Correspondence: A. B. Bond, School of Biological Sciences, 348 Manter Hall, UNL, Lincoln, NE 68588-0118, U.S.A. (email: abond@unl.edu). R. P. Balda is at the Department of Biological Sciences, Box 5640, Northern Arizona University, Flagstaff, AZ 86011, U.S.A.

a result, the social complexity hypothesis remains relatively untested.

One promising approach to testing the social complexity hypothesis is to use controlled experimental techniques to compare cognitive abilities in closely related species that differ strongly in their social organization (Endler 1986; Kamil 1988). These need not be primates: if the hypothesis is correct, enhanced cognitive abilities should be apparent in any highly social vertebrate (Balda et al. 1996; Holekamp et al. 1999). In this paper, we present the first explicit comparative test of the social complexity hypothesis conducted under well-controlled, experimental conditions, using two closely related species of birds with contrasting degrees of sociality.

Pinyon jays, *Gymnorhinus cyanocephalus*, are probably the most highly social North American corvids, living in stable groups of 50–500 individuals, foraging as a permanent flock and breeding colonially (Marzluff & Balda 1992). In contrast, the typical social unit of closely related western scrub-jays, *Aphelocoma californica*, is a pair of birds and their young of the year (Madge & Burn 1994). Although pinyon jays are more dependent on stored food during winter, the two species are quite comparable in most other aspects of their biology. They are sympatric over large parts of Arizona and Colorado, U.S.A., using similar resources (Madge & Burn 1994), and *Aphelocoma* and *Gymnorhinus* appear to be the most closely related of the six corvid genera endemic to the New World (Monteros & Cracraft 1997).

The social complexity hypothesis predicts that despite the close degree of relatedness and substantial ecological similarities between these species, the differing demands

of their social lifestyles should have selected for differences in relevant cognitive abilities. To test this hypothesis, we explored a subset of the cognitive abilities that social animals presumably use to track and assess relationships among group members. Understanding the web of relationships within the group should, in the first instance, require simultaneous tracking of a large number of dyadic interactions. We would thus expect highly social animals to learn multiple dyadic relationships more readily and be able to keep track of more of them. This aspect of social cognition was tested in experiment 1.

Second, it seems unlikely that an individual in a large social group could ever observe interactions between all possible pairs of group members. Social animals must base some of their relationship judgements on transitive inferences, concluding that if A is dominant to B and B is dominant to C, then A would probably also dominate C, even if they have never seen A and C interact (Smith 1988; Hogue et al. 1996). This suggests that social and nonsocial species should differ in their ability to make transitive inferences. This aspect of social cognition was tested in experiment 2.

We tested scrub-jays and pinyon jays using well-established operant procedures, in which the ordering of a set of arbitrary stimuli can be inferred from a series of dyadic comparisons (Fersen et al. 1991; Treichler & van Tilburg 1996). In the first experiment, the birds were trained on a series of discriminations between successive pairs of seven coloured stimuli, simultaneously tracking six dyadic relationships. Species differences were assessed in terms of the rate of task acquisition and asymptotic levels of performance. In the second experiment, we tested both species with novel stimulus pairs to analyse their transitive reasoning abilities.

GENERAL METHODS

Subjects

Five scrub-jays and five pinyon jays were captured as adults in live traps in Colorado and northern Arizona and were subsequently housed in individual cages in environmentally controlled rooms (22°C; 14:10 h light:dark cycle). They were maintained at 85–90% of their free-feeding weights by controlled feedings of turkey starter, sunflower seeds, parrot pellets, mealworms, pine nuts and a vitamin supplement, and were given unlimited access to grit and water. All subjects were initially naïve to operant procedures but had previously been used in several different open-room tests of spatial orientation.

Apparatus

Experiments were conducted in operant boxes equipped with three 2.5-cm diameter pecking keys spaced at 7.6-cm intervals across the front wall of the chamber, 22.5 cm above the floor. A perch was placed in front of the key array and adjusted so that the centres of the keys were approximately at eye level for each species. During trials, coloured light from a computer monitor was projected on the keys from behind through diffusion filters. The stimulus colours were red, green, blue, magenta, yellow, cyan, orange, or white, drawn from the standard VGA palette and chosen to be maximally discriminable to the human eye. The chamber was diffusely lit, and ambient white noise was provided to mask external sounds. Rewards were the most preferred food items for each species, pieces of pine nut for pinyon jays and halves of mealworm for scrub-jays. These were delivered into a food cup centred at floor level at the rear of the chamber, turning on an adjacent 'food light'.

Trial Sequence

At the start of each trial, the centre key was illuminated in white. If the key was not pecked within 15 s, it was turned off, and the trial was reinitiated after a 3-s delay. When the centre key was pecked, it was turned off, and both side keys were illuminated, presenting a choice between two of the seven nonwhite colours. Upon three successive pecks to one of the side keys, the entire panel was darkened. If the correct key had been pecked, a reward was delivered. After 10 s, the food light was turned off, and another trial began 20 s later. If the wrong key was pecked, the next trial was presented after a 30-s delay. If the bird failed to peck either key three times within 60 s, the panel was darkened, and the trial was repeated after a 30-s delay.

Initial Training

The birds were first habituated to the chamber and then shaped to peck the side keys when either was illuminated in white. They were subsequently given 6 days of experience pecking the side keys for reward, 36 trials per day, 18 to each side in random order. The central white key was then added for 6 days as an automatic precursor, and the response requirement for the side keys was increased to three pecks. After 6 days at this stage, responses to the initial white key were made obligatory, and the birds were transferred to experiment 1, which tested their ability to learn a set of dyadic relationships that defined an implicit ordinal sequence.

EXPERIMENT 1

The ability to track dyadic relationships was assessed by training each bird on a series of discriminations between successive pairs of seven implicitly ordered colours. The correct choice was defined by the particular pair of colours presented: choosing B was rewarded when B was presented with C, but not when B was presented with A. This contextual aspect of the task was introduced gradually, but all six pairs were ultimately fully intermixed within sessions, requiring simultaneous tracking of six dyadic relationships.

Methods

The seven nonwhite stimuli were assigned a unique ordering for each bird that was balanced across birds

within species, ensuring that no adjacent pairing occurred more than once and no stimulus occurred more than twice in the same ordinal position. The design thus ensured that the relative discriminabilities of the stimulus pairs could not confound the main effects of stimulus order either between or within species. For any given colour pair, only responses to the first, higher-ranked colour were rewarded, so the series of paired discriminations defined an implicit hierarchy: A>B>C>D>E>F>G. Birds were given daily sessions of 36 trials, with the position of the correct stimulus randomized between left and right keys on each trial.

In the first phase, all trials in a session consisted of a single colour pair. Thus, subjects were trained first on pair A/B, then on B/C, then on C/D. When a bird reached a criterion of 32 or more correct responses in each of three successive sessions, it was advanced to the next pair. As each bird completed the last pair (F/G), it was advanced to the second phase, an equal mixture of two adjacent colour pairs (e.g. B/C and C/D).

While the first phase was essentially a partial reversal task, the second phase entailed greater cognitive complexity, in that one of the three stimuli was present in every trial, but whether responses to it were rewarded depended on which other choice stimulus was present. For example, when B/C and C/D were being presented, C was correct when paired with D, but incorrect when paired with B. When each subject reached criterion on a set of three successive stimuli, testing with the next set began. For example, criterial performance during a B/C and C/D session would be followed the next day with a session of C/D and D/E trials.

In the third phase, the birds were required to track all six colour pairs presented in equal numbers and in randomized order each day. This task was continued for 100 sessions for the pinyon jays. Because the scrub-jays did not appear to reach asymptotic performance levels within 100 sessions, they were given an additional 100 training sessions.

Results

Although the two species learned the first (A/B) pairing at comparable rates (paired Student's t test: $t_{4,8} = -1.12$, NS), repeated measures analysis of variance (ANOVA) on the results of the first phase of acquisition (Fig. 1) indicated that pinyon jays learned the remaining dyads significantly faster than scrub-jays. The main effect of species was significant (ANOVA: $F_{1,8} = 8.79$, $P<0.02$), as was the effect of colour pair (ANOVA: $F_{5,40} = 5.05$, $P<0.002$), presumably reflecting the greater error rates displayed to the C/D and E/F pairs, but the species-by-pair interaction was not significant. During the second acquisition phase, pinyon jays again required much less experience than scrub-jays to reach criterion (Fig. 2). Pinyon jays averaged 266 trials to criterion per triad, while scrub-jays needed 684 trials per triad to attain the same performance level. Repeated measures ANOVA on the results of the second phase showed a significant main effect of species (ANOVA: $F_{1,8} = 13.18$, $P<0.01$), but no significant effect of triad and no significant interaction.
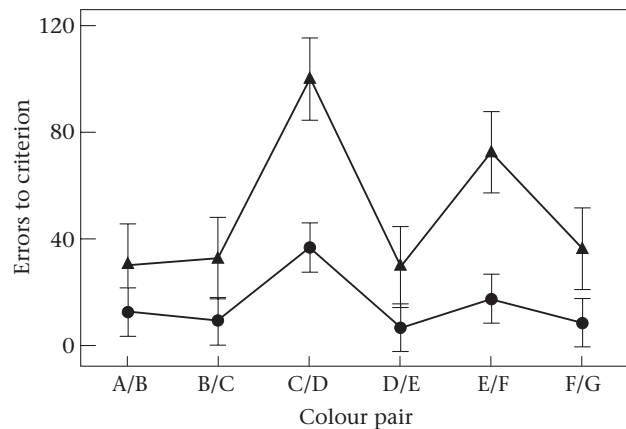

Figure 1. Acquisition of discriminative performance on single colour pairs. Data points indicate mean errors to criterial performance across birds. Error bars bracket the pooled within-subjects standard error estimate for each species. ●: Pinyon jays; ▲: scrub-jays.
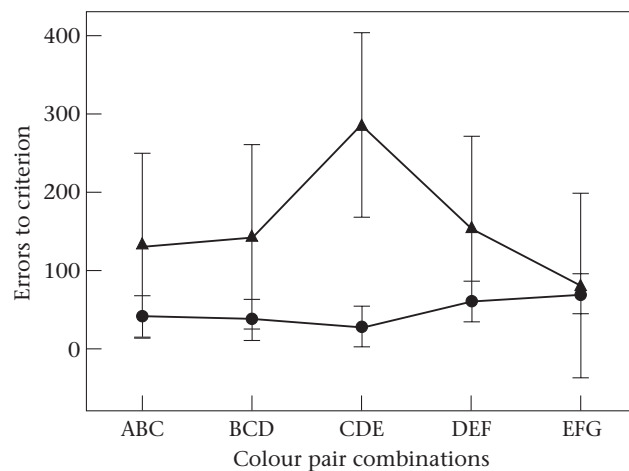

Figure 2. Acquisition of discriminative performance on mixtures of two successive colour pairs. Data points indicate mean errors to criterial performance across birds. Error bars bracket the pooled within-subjects standard error estimate for each species. ●: Pinyon jays; ▲: scrub-jays.

At the beginning of the third phase, accuracy was comparable for the two species, and well above chance levels (Fig. 3). However, pinyon jays improved more rapidly than scrub-jays and attained a higher asymptotic level of performance. Repeated measures ANOVA on data from the first 100 sessions, grouped into blocks of four consecutive days, showed a significant main effect of species (ANOVA: $F_{1,8} = 27.53$, $P<0.001$) and block (ANOVA: $F_{25,193} = 12.28$, $P<0.0001$), as well as a significant species × block interaction (ANOVA: $F_{25,193} = 3.56$, $P<0.0001$). Asymptotic levels of performance still differed between species, even after the scrub-jays had been given 100 sessions of additional training (paired Student's t test: $t_7 = 3.74$, $P<0.008$).

Discussion

The two species differed strikingly in their performance on this implicit ranking task, indicating a consistent
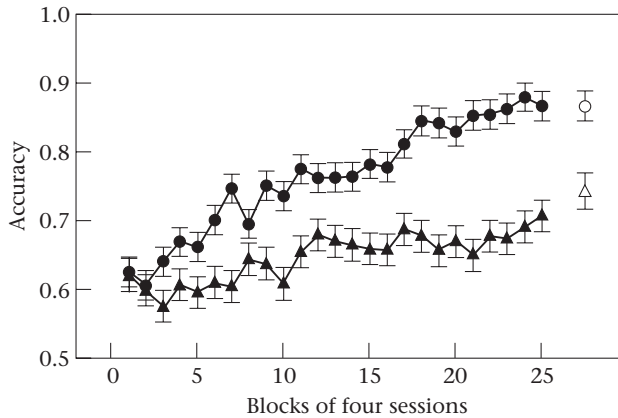
Figure 3. Acquisition of discriminative performance on mixtures of six successive colour pairs. Each data point indicates the mean accuracy across birds for a contiguous block of four 36-trial sessions. Error bars bracket the pooled within-subjects standard error estimate for each species. Circles are data from pinyon jays; triangles are data from scrub-jays. Filled symbols plot progress over the initial 100 sessions of training; open symbols indicate performance after the scrub-jays had been given an additional 100 training sessions.

species difference in cognitive ability. Pinyon jays learned the dyadic relationships more rapidly, with fewer errors, and reached a higher asymptotic level of performance than scrub-jays. Moreover, the pattern of errors during acquisition clearly differed between the species. In the first phase (Fig. 1), learning the initial A/B pair was an acquisition problem, which the two species solved equally well. Subsequent pairings involved a partial reversal, in that the rewarded item was previously incorrect, and pinyon jay error rates were significantly lower, particularly for the C/D and E/F combinations. This pattern suggests that pinyon jays adapt more readily to reversals of previous reward outcomes.

A similar effect was apparent during the second phase, learning pairs of dyads (Fig. 2). The conspicuous peak in errors to criterion for scrub-jays on the C/D, D/E treatment was wholly absent for pinyon jays. Training on the C/D, D/E triad, unlike earlier stages of this phase, involved simultaneous reversals or partial reversals for all three choice stimuli. This again suggests that interference from prior reward history was more problematic for scrub-jays.

When the task was expanded to include all six colour pairs in every session (Fig. 3), both species began at a similar level, well above chance. But while the pinyon jays rapidly improved, the scrub-jays had a great deal of difficulty responding accurately when all six dyads were presented in each session. While pinyon jays reached 80% correct on all pairs in an average of 68 sessions, only four of five scrub-jays managed to reach 75% correct even after 200 sessions. It is unlikely that these differences were due to some methodological detail that favoured the pinyon jays, as there were no initial species differences in any of the three training phases. In an earlier operant study of colour non-matching-to-sample (Olson et al. 1995), scrub-jays took longer to learn the task than pinyon jays, but they ultimately reached similar asymp-

totic levels of performance. Although methodological differences make comparisons to other studies difficult, the acquisition performance of scrub-jays on the implicit sequence task appeared to be significantly better than that of pigeons, *Columba livia* (Fersen et al. 1991), while that of pinyon jays was similar to but somewhat slower than that of rhesus macaques, *Macaca mulatta* (Treichler & van Tilburg 1996).

EXPERIMENT 2

In this experiment, we tested for transitive inference by presenting occasional probe trials of novel, nonadjacent colour pairs intermixed among familiar stimulus combinations. Although the birds had had previous experience with both B and D, for example, these stimuli had never been presented as alternative choices on the same trial. A consistent choice of B in this novel pairing would, therefore, indicate a transitive inference from the implicit ordinal sequence. A frequently cited indication of successful transitive reasoning is the symbolic distance effect (Moyer & Bayer 1976), an increase in the accuracy of transitive choices as the distance between items along the implicit sequence increases (i.e. B/E is responded to more accurately than B/D; Bryant & Trabasso 1971; Fersen et al. 1991; Steirn et al. 1995).

The literature suggests that transitive responding during testing with novel pairs may differ depending on how the implicit ordinal sequence is represented in memory. In an 'associative' (Terrace & McGonigle 1994) representation, the animal chooses between stimuli based on their relative associative strength. Associative strength accrues gradually during training, over many trials, through direct or indirect association with reward. Several algebraic models of this cognitive mechanism have been proposed (Fersen et al. 1991; Delius & Siemann 1998; Zentall & Sherburne 1998). The models differ substantially in detail, but when we evaluated each of them across a range of potential parameter values, we found that they generally produced higher associative strength at the beginning of an implicit sequence and greater variance in strength at the end.

This implies that using an associative representational system will yield transitive responding, but that error rates will be higher at the end of the implicit sequence. Associative representation, thus, should be characterized by a 'first-item accuracy effect' (Fersen et al. 1991), a decline in performance as a function of the position in the implicit sequence of the higher-ranked member of a dyad (i.e. B/C is responded to more accurately than C/D or D/E). Such first-item effects have been found in a number of studies with pigeons, which appear to form associative representations of sequentially presented colour dyads (Fersen et al. 1991; Steirn et al. 1995; Wynne 1997; Delius & Siemann 1998).

In a 'relational' (D'Amato 1991) or 'linear' (D'Amato & Columbo 1990) representation, in contrast, the animal interprets the outcomes of individual stimulus pairings as reflections of an underlying linear or spatial array and associates particular stimuli directly with their ordinal positions in the array. Animals using this system should

show minimal first-item accuracy effects. However, if determining relative position involves counting down from the top of the sequence, relational representations will produce longer choice latencies for items late in the list, a first-item effect on response times. Such latency effects have been reported in both humans and capuchin monkeys, *Cebus apella* (D'Amato & Columbo 1988, 1990; D'Amato 1991). Distinguishing between these two representational systems was the primary focus of experiment 2.

Methods

The same subjects, apparatus and trial contingencies were used as in experiment 1, with the exception of the inclusion of probe trials of nonadjacent pairs. For transitive inference testing, the birds were given 40 daily sessions of 36 trials each. During each session, they were presented with familiar, adjacent pairs on 33 trials and novel, nonadjacent pairs on three probe trials. The 120 probe trials presented to each bird included 20 of each of the central, nonadjacent pairs in the implicit sequence (B/D, B/E, B/F, C/E, C/F and D/F). We did not test pairs including A or G, because responses to A had always been rewarded and those to G never rewarded during training (the 'end-anchor effect': Bryant & Trabasso 1971). To avoid biasing choice on subsequent presentations, all choices during probe trials were rewarded (Steirn et al. 1995). Otherwise, the same reinforcement contingencies were used. All of the pinyon jays completed the transitive inference tests, but one of the scrub-jays had to be removed from the experiment for health reasons.

Results

Both species showed significantly higher accuracy on nonadjacent probe trials than expected by chance. For pinyon jays, the proportion of correct choices on transitive inference trials was 0.86 (paired Student's *t* test: $t_4 = 17.27$, $P < 0.001$); for scrub-jays, it was 0.77 ($t_3 = 3.98$, $P < 0.03$). The species difference was not significant ($t_7 = 1.45$, NS).

Analysis of the symbolic distance effect also indicated that pinyon jays and scrub-jays were equally capable of making transitive inferences. Because longer symbolic distances are inevitably associated with higher-ranking first items in a linear array, symbolic distance is fully confounded with first-item effects when all possible pairs are included in the analysis (as in Fersen et al. 1991). To avoid such problems in interpretation, we analysed for symbolic distance effects separately in stimulus pairs involving the second-ranked item (B/C, B/D, B/E and B/F) and in those involving the third-ranked item (C/D, C/E and C/F).

Repeated measures ANOVA revealed a significant main effect of symbolic distance in scrub-jays (ANOVA: $F_{3,9} = 4.05$, $P < 0.05$) but not in pinyon jays (ANOVA: $F_{3,12} = 1.38$, NS) for pairs involving the second-ranked item (Fig. 4a). The pinyon jays may simply have reached a performance ceiling, as two of them made no errors for
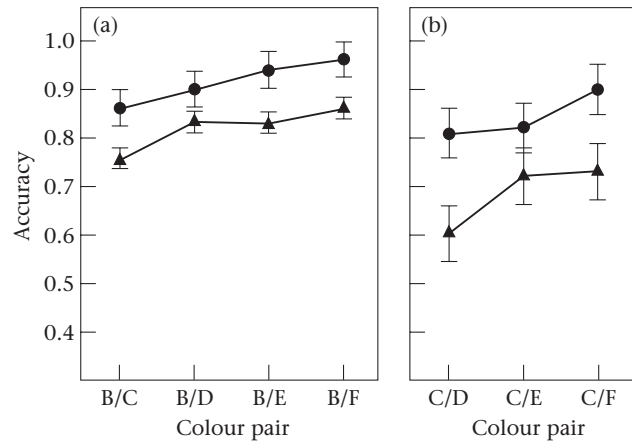

Figure 4. Performance accuracy of pinyon jays (●) and scrub-jays (▲) during transitive inference testing, analysed for symbolic distance effects. (a) Mean accuracy across subjects within species is plotted for colour pairs at successively greater distances on the implicit sequence, using only stimulus pairs that included the second-ranked item. (b) Mean accuracy across subjects is plotted for stimulus pairs including the third-ranked item. In each case, error bars bracket the pooled within-subjects standard error estimate for each species.

any nonadjacent stimulus pairs in this data set. The results for pairs involving the third-ranked item showed a similar tendency (Fig. 4b), but the effect was not significant in either species (ANOVA: pinyon jays: $F_{2,8} = 1.50$; scrub-jays: $F_{2,6} = 0.97$, both NS), owing to both higher variance and fewer degrees of freedom for this comparison.

In contrast, the position of stimuli in the implicit sequence had substantially different effects in the two species, differences that were evident in performance on trials with both familiar, adjacent pairs and novel, nonadjacent pairs. Our analysis of first-item effects in adjacent stimulus pairs, excluding the terminal dyads, compared pairs B/C, C/D, D/E and E/F. Analysis of nonadjacent pairs used dyads that were two steps apart on the implicit ranking scale (B/D, C/E and D/F). As there were large between-species differences in variance, these and subsequent analyses were carried out separately within each species.

For adjacent stimulus pairs (Fig. 5a), there was a strong, significant first-item effect on choice accuracy in scrub-jays (ANOVA: $F_{3,9} = 8.64$, $P < 0.006$), but not in pinyon jays (ANOVA: $F_{3,12} < 1$, NS). In the most extreme case, a scrub-jay that showed 92% accuracy in responding to B/C dropped to well below chance (37%) for E/F. Pinyon jay performance, in contrast, was almost invariably above 70% correct for all individuals and all stimulus pairs. Accuracy analysis for nonadjacent probe trials yielded a similar pattern. Response accuracy declined across first-item rankings for scrub-jays (Fig. 5b; ANOVA: $F_{2,6} = 4.97$, $P = 0.05$), but not for pinyon jays (ANOVA: $F_{2,8} = 2.17$, NS). (The apparent decline in performance for pinyon jays (Fig. 5b) was the result of a single subject that had anomalously low accuracies in responding to D/F probes.)

We then analysed the effects of sequential position on response time to adjacent colour pairs (Fig. 6a). Both
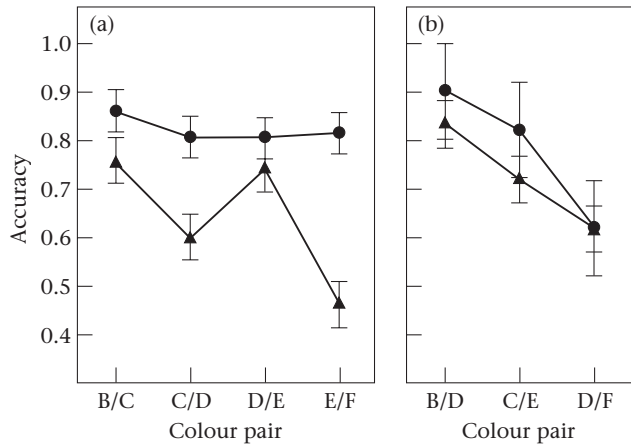
Figure 5. Performance accuracy of pinyon jays (●) and scrub-jays (▲) during transitive inference testing, analysed for first-item effects. (a) Mean accuracy across subjects within species is plotted for adjacent stimuli in which the rank of the higher-ranked member of the pair decreases from left to right along the abscissa. (b) Mean accuracy across subjects is similarly plotted for probe trials two steps apart. In each case, error bars bracket the pooled within-subjects standard error estimate for each species.
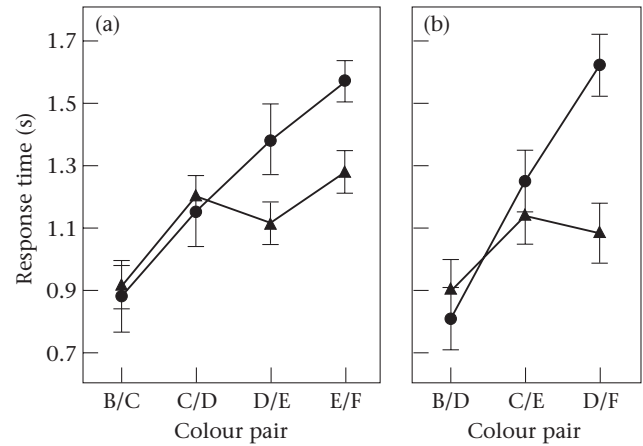

Figure 6. Mean response time of pinyon jays (●) and scrub-jays (▲) during transitive inference testing, analysed for first-item effects. (a) Response time across subjects within species is plotted for adjacent stimuli in which the rank of the higher-ranked member of the pair decreases from left to right along the abscissa. (b) Response time across subjects is similarly plotted for probe trials two steps apart. In each case, error bars bracket the pooled within-subjects standard error estimate for each species.

species responded more slowly to adjacent pairs near the end of the sequence (ANOVA: pinyon jays: $F_{3,12}=6.91$, $P<0.006$; scrub-jays: $F_{3,9}=5.35$, $P<0.025$). But the mean response time for pinyon jays was significantly longer than that for scrub-jays for the last two dyads (D/E and E/F; paired Student's t test: $t_7 \geq 2.86$, $P<0.03$). The same pattern was observed during nonadjacent probe trials (Fig. 6b), where response time increased as a function of sequential position for pinyon jays (ANOVA: $F_{2,8}=17.15$, $P<0.002$), but not for scrub-jays ($F_{2,6}=1.71$, NS).

Discussion

Both scrub-jays and pinyon jays showed high accuracy in novel, nonadjacent probe trials, demonstrating an ability to make transitive inferences. There were, however, striking differences in their patterns of responding as a function of the position of the stimulus pair in the implicit ordinal sequence. Pinyon jays showed no effect of position of the highest-ranked stimulus in the pair on choice accuracy, but were slower in responding to pairs that were lower in the sequence. Scrub-jays showed a clear first-item effect on accuracy, but only minimal effects on latency. This suggests that the species used different cognitive methods for representing the implicit stimulus sequence. Relational models predict effects on latency but not accuracy, the pattern shown by pinyon jays. Associative models predict effects on accuracy but not latency, the pattern shown by scrub-jays.

This conclusion is given added weight by the species differences in acquisition during experiment 1. When an associative representation is being established, the change in associative strength produced by each trial tends to be small. Accurate sequential categorization using an associative mechanism may, therefore, require

prolonged experience with individual pairs. In contrast, when a relational representation is constructed, the results of individual trials are inserted into a pre-existing structure, potentially requiring a smaller number of trials to attain competent performance. That pinyon jays learned the original set of dyadic discriminations with fewer errors and reached a higher asymptotic level of performance is, thus, consistent with the interpretation that the two species were using different cognitive mechanisms.

This finding is open to further exploration through additional experiments testing for species differences in list linkage. Treichler & van Tilburg (1996) trained macaques on two separate transitive series (e.g. A>B, B>C, C>D, D>E versus F>G, G>H, H>I, I>J), then linked the two series together by training the last item of the first series with the first item of the second (E>F or J>A). The monkeys were subsequently able to choose accurately between pairs drawn from across the two lists. Our interpretation of the results of experiment 2 predicts that pinyon jays would show much greater between-list accuracy in such an experiment than would scrub-jays.

GENERAL DISCUSSION

The results of both experiments are fully consistent with the social complexity hypothesis. Pinyon jays learned multiple dyadic relationships that defined an implicit ordinal sequence more rapidly than scrub-jays and attained higher asymptotic levels of choice accuracy. They also appeared to use a more sophisticated cognitive representation of the sequence, one that was less vulnerable to error. Because pinyon jays live in very large, stable social groups, their ability to infer the dominance status

of other individuals is critically dependent on being able to make accurate transitive inferences. Social complexity thus entails a predisposition to interpret dyadic contrasts as direct evidence of an underlying ordinal sequence, a linear representation of stimulus relationships that facilitates tracking a large number of dyads simultaneously and increases the accuracy of inferences about the relationship between novel items. Less social species, such as scrub-jays, can also make transitive inferences, but their capabilities are more limited, since they have not been subject to the same regimen of directional selection (Delius & Siemann 1998). Relatively nonsocial animals may require more extensive training, and the resulting cognitive structure produces less accurate transitive inferences, particularly in the later portions of the sequence. The ontogeny of these cognitive structures would certainly be well worth exploring. Pinyon jays mature in a more complex social environment than scrub-jays, and the concomitant experiential differences may contribute to the observed species differences in transitive inference in adult birds. Tests with hand-raised birds could be used to address this issue.

The primate literature provides tentative support for our inference of an association between the mechanism of transitive inference and social complexity. Monkeys that live in structured, long-lasting social groups, such as rhesus macaques (Treichler & van Tilburg 1996; Chen et al. 1997) or capuchins (D'Amato & Columbo 1990) show strong evidence of relational representation, including explicit encoding of ordinal position and high accuracy in transitive inference across linked lists. In contrast, squirrel monkeys, *Saimiri sciureus* (Robinson & Janson 1987) commonly live in large, fluid aggregations with little fixed structure, and their transitive inference performance is far less impressive and more suggestive of associative encoding (McGonigle & Chalmers 1992; Harris & McGonigle 1994). The pattern of cognitive differences between pinyon and scrub-jays may represent a more widespread phenomenon than has previously been appreciated.

Because any given pair of species always differ along more than one dimension, however, there are tenable alternative interpretations of the evolutionary significance of the differences we observed. For example, several authors have suggested that relational representation may be tied to spatial perception (D'Amato 1991; Davis 1992; Terrace & McGonigle 1994), and there is evidence that training rats with stimulus pairs in an invariant spatial arrangement increases the probability of subsequent transitive choices (Roberts & Phelps 1994). Pinyon jays cache large quantities of pine seeds during the late summer and autumn for retrieval throughout the winter (Marzluff & Balda 1992). Like other caching species, pinyon jays have performed better on open-room tests of spatial cognition than scrub-jays, which cache relatively little (Balda & Kamil 1989; Kamil et al. 1994). This suggests that the superior performance of pinyon jays in these experiments could be a reflection of the need to recall and evaluate locations of dispersed food caches rather than a consequence of the demands of social living.

These alternative hypotheses can be differentiated by further comparative research, as they make contrasting predictions about other species pairs. This is an example of how comparative studies of behavioural evolution are most powerful when used iteratively, applying similar experimental techniques to multiple sets of carefully selected, related species (Felsenstein 1985; Endler 1986; Kamil 1988).

To test the social complexity hypothesis directly against the alternative spatial hypothesis, for example, one could compare Clark's nutcrackers, *Nucifraga columbiana*, and European jackdaws, *Corvus monedula*, two related Old World corvids. The jackdaws are highly social, while nutcrackers are relatively solitary, but in this case the spatial cognition difference is reversed: solitary nutcrackers make extensive use of caching and spatial memory (Balda & Kamil 1989; Kamil et al. 1994; Olson et al. 1995), while social jackdaws apparently do not cache at all (Simmons 1968; Henty 1975).

The inferior performance of scrub-jays on this task also opens the question of possible compensatory abilities in other spheres. Are there cognitive tasks for which we would predict that scrub-jays might actually do better than their more social counterparts? One possibility derives from the fact that scrub-jays have a considerably broader niche than pinyon jays, occurring in oak woodlands and savannahs, as well as in pinyon–juniper forest (Bardwell et al. 2001). They also accept a broader range of foodstuffs, including more food of animal origin (Madge & Burn 1994). We might, thus, predict that scrub-jays would show superiority in cognitive tests related to ecological generality, such as exploratory behaviour (Haemig 1989), environmental sampling (Shettleworth et al. 1988), or neophobia (Greenberg 1984, 1990). Clayton's (Clayton et al. 2001; Griffiths & Clayton 2001) studies of episodic memory in scrub-jays also suggest that their generalist ecology may require them to encode more information about particular food-related events than a specialist species such as a pinyon jay or nutcracker.

The evolutionary origin of human intellectual capabilities is one of the most challenging issues in behavioural biology. Often, the question is treated as one that can be approached only through observational studies of higher primates. But some of the factors leading to the evolution of human intelligence must be general, having effects on the cognitive abilities and organization of other vertebrate species. The experiments and results reported here demonstrate the possibility of testing hypotheses about the evolution of intellectual abilities with nonprimate species.

Acknowledgments

References

Balda, R. P. & Kamil, A. C. 1989. A comparative study of cache recovery by three corvid species. *Animal Behaviour*, **38**, 486–495.

Balda, R. P., Kamil, A. C. & Bednekoff, P. A. 1996. Predicting cognitive capacities from natural histories: examples from four corvid species. *Current Ornithology*, **13**, 33–66.

Bardwell, E., Benkman, C. W. & Gould, W. R. 2001. Adaptive geographic variation in western scrub-jays. *Ecology*, **82**, 2617–2627.

Bryant, P. E. & Trabasso, T. 1971. Transitive inferences and memory in young children. *Nature*, **232**, 456–458.

Byrne, R. W. & Whiten, A. 1988. *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans*. Oxford: Clarendon Press.

Byrne, R. W. & Whiten, A. 1997. *Machiavellian Intelligence II: Extensions and Evaluations*. Cambridge: Cambridge University Press.

Chen, S., Swartz, K. B. & Terrace, H. S. 1997. Knowledge of the ordinal position of list items in rhesus monkeys. *Psychological Science*, **8**, 80–86.

Cheney, D. L. & Seyfarth, R. L. 1990. *How Monkeys See the World*. Chicago: University of Chicago Press.

Clayton, N. S., Griffiths, D. P., Emery, N. J. & Dickinson, A. 2001. Elements of episodic-like memory in animals. *Philosophical Transactions of the Royal Society of London, Series B*, **356**, 1483–1491.

D'Amato, M. R. 1991. Comparative cognition: processing of serial order and serial pattern. In: *Current Topics in Animal Learning: Brain, Emotion and Cognition* (Ed. by L. Dachowski & C. F. Flaherty), pp. 165–185. Hillsdale, New Jersey: L. Erlbaum.

D'Amato, M. R. & Columbo, M. 1988. Representation of serial order in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, **14**, 131–139.

D'Amato, M. R. & Columbo, M. 1990. The symbolic distance effect in monkeys (*Cebus apella*). *Animal Learning & Behavior*, **18**, 133–140.

Davis, H. 1992. Logical transitivity in animals. In: *Cognitive Aspects of Stimulus Control* (Ed. by W. K. Honig & J. G. Fetterman), pp. 405–429. Hillsdale, New Jersey: L. Erlbaum.

Delius, J. D. & Siemann, M. 1998. Transitive responding in animals and humans: exaptation rather than adaptation? *Behavioural Processes*, **42**, 107–137.

Endler, J. A. 1986. *Natural Selection in the Wild*. Princeton, New Jersey: Princeton University Press.

Felsenstein, J. 1985. Phylogenies and the comparative method. *American Naturalist*, **125**, 1–15.

von Fersen, L., Wynne, C. D. L., Delius, J. D. & Staddon, J. E. R. 1991. Transitive inference formation in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, **17**, 334–341.

Greenberg, R. 1984. Differences in feeding neophobia in tropical migrant warblers *Dendroica castanea* and *D. pennsylvanica*. *Journal of Comparative Psychology*, **98**, 131–136.

Greenberg, R. 1990. Ecological plasticity, neophobia and resource use in birds. *Studies in Avian Biology*, **13**, 431–437.

Griffiths, D. P. & Clayton, N. S. 2001. Testing episodic memory in animals: a new approach. *Physiology & Behavior*, **73**, 755–762.

Haemig, P. D. 1989. A comparative experimental study of exploratory behaviour in Santa Cruz Island and mainland California scrub jays *Aphelocoma coerulescens*. *Bird Behaviour*, **8**, 38–42.

Harris, M. R. & McGonigle, B. O. 1994. A model of transitive choice. *Quarterly Journal of Experimental Psychology B*, **47**, 319–348.

Henty, C. J. 1975. Feeding and food-hiding responses of jackdaws and magpies. *British Birds*, **68**, 463–466.

Hogue, M. E., Beaugrand, J. P. & Laguë, P. C. 1996. Coherent use of information by hens observing their former dominant defeating or being defeated by a stranger. *Behavioural Processes*, **38**, 241–252.

Holekamp, K. E., Boydston, E. E., Szykman, M., Graham, I., Nutt, K. J., Birch, S., Piskiel, A. & Singh, M. 1999. Vocal recognition in the spotted hyaena and its possible implications regarding the evolution of intelligence. *Animal Behaviour*, **58**, 383–395.

Humphrey, N. K. 1976. The social function of intellect. In: *Growing Points in Ethology* (Ed. by P. P. G. Bateson & R. A. Hinde), pp. 303–317. Cambridge: Cambridge University Press.

Jolly, A. 1966. Lemur social behavior and primate intelligence. *Science*, **153**, 501–506.

Kamil, A. C. 1988. A synthetic approach to the study of animal intelligence. In: *Comparative Perspectives in Modern Psychology: Nebraska Symposium on Motivation. Vol. 35* (Ed. by D. W. Leger), pp. 230–257. Lincoln: University of Nebraska Press.

Kamil, A. C., Balda, R. P. & Olson, D. J. 1994. The effects of requiring different response strategies following caching in Clark's nutcrackers (*Nucifraga columbiana*). *Animal Learning & Behavior*, **22**, 373–378.

Kummer, H., Daston, L., Gigerenzer, G. & Silk, J. 1997. The social intelligence hypothesis. In: *Human by Nature: Between Biology and the Social Sciences* (Ed. by P. Weingart, P. Richerson, S. D. Mitchell & S. Maasen), pp. 157–179. Hillsdale, New Jersey: L. Erlbaum.

McGonigle, B. & Chalmers, M. 1992. Monkeys are rational! *Quarterly Journal of Experimental Psychology B*, **45**, 189–228.

Madge, S. & Burn, H. 1994. *Crows and Jays: a Guide to the Crows, Jays and Magpies of the World*. New York: Houghton Mifflin.

Marzluff, J. M. & Balda, R. P. 1992. *The Pinyon Jay: Behavioral Ecology of a Colonial and Cooperative Corvid*. London: T. & A. D. Poyser.

de los Monteros, A. E. & Cracraft, J. 1997. Intergeneric relationships of the New World jays inferred from cytochrome *b* gene sequences. *Condor*, **99**, 490–502.

Moyer, R. S. & Bayer, R. H. 1976. Mental comparison and the symbolic distance effect. *Cognitive Psychology*, **8**, 228–246.

Olson, D. J., Kamil, A. C., Balda, R. P. & Nims, P. J. 1995. Performance of four seed-caching corvid species in operant tests of nonspatial and spatial memory. *Journal of Comparative Psychology*, **109**, 173–181.

Roberts, W. A. & Phelps, M. T. 1994. Transitive inference in rats: a test of the spatial coding hypothesis. *Psychological Science*, **5**, 368–374.

Robinson, J. G. & Janson, C. H. 1987. Capuchins, squirrel monkeys, and atelines: socioecological convergence with Old World primates. In: *Primate Societies* (Ed. by B. B. Smuts, D. L. Cheney, R. M. Seyfarth, R. W. Wrangham & T. T. Struhsaker), pp. 69–82. Chicago: University of Chicago Press.

Shettleworth, S. J. 1998. *Cognition, Evolution and Behavior*. Oxford: Oxford University Press.

Shettleworth, S. J., Krebs, J. R., Stephens, D. W. & Gibbon, J. 1988. Tracking a fluctuating environment: a study of sampling. *Animal Behaviour*, **36**, 87–105.

Simmons, K. E. L. 1968. Food-hiding by rooks and other crows. *British Birds*, **61**, 228–229.

Smith, P. K. 1988. The cognitive demands of children's social interaction with peers. In: *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans* (Ed. by R. W. Byrne & A. Whiten), pp. 94–109. Oxford: Clarendon.

Steirn, J. N., Weaver, J. E. & Zentall, T. R. 1995. Transitive inference in pigeons: simplified procedures and a test of value transfer theory. *Animal Learning & Behavior*, **23**, 76–82.

Terrace, H. S. & McGonigle, B. 1994. Memory and representation of serial order by children, monkeys, and pigeons. *Current Directions in Psychological Science*, **3**, 180–185.

Treichler, F. R. & van Tilburg, D. 1996. Concurrent conditional discrimination tests of transitive inference by macaque monkeys:

list linking. *Journal of Experimental Psychology: Animal Behavior Processes*, **22**, 105–117.

Wynne, C. D. L. 1997. Pigeon transitive inference: tests of simple accounts of a complex performance. *Behavioural Processes*, **39**, 95–112.

Zentall, T. R. & Sherburne, L. M. 1998. The transfer of value in simultaneous discriminations: implications for cognitive and social processes. In: *Brain and Values: Is a Biological Science of Values Possible?* (Ed. by K. H. Pribram), pp. 323–336. Hillsdale, New Jersey: L. Erlbaum.