

*TRANSITIVE RESPONDING IN HOODED CROWS REQUIRES LINEARLY
ORDERED STIMULI*

OLGA F. LAZAREVA, ANNA A. SMIRNOVA, MARIA S. BAGOZKAJA, ZOYA A. ZORINA,
VLADIMIR V. RAYEVSKY, AND EDWARD A. WASSERMAN

INSTITUTE OF HIGHER NERVOUS ACTIVITY, MOSCOW STATE UNIVERSITY, AND UNIVERSITY OF IOWA

Eight crows were taught to discriminate overlapping pairs of visual stimuli (A+ B−, B+ C−, C+ D−, and D+ E−). For 4 birds, the stimuli were colored cards with a circle of the same color on the reverse side whose diameter decreased from A to E (ordered feedback group). These circles were made available for comparison to potentially help the crows order the stimuli along a physical dimension. For the other 4 birds, the circles corresponding to the colored cards had the same diameter (constant feedback group). In later testing, a novel choice pair (BD) was presented. Reinforcement history involving stimuli B and D was controlled so that the reinforcement/nonreinforcement ratios for the latter would be greater than for the former. If, during the BD test, the crows chose between stimuli according to these reinforcement/nonreinforcement ratios, then they should prefer D; if they chose according to the diameter of the feedback stimuli, then they should prefer B. In the ordered feedback group, the crows strongly preferred B over D; in the constant feedback group, the crows' choice did not differ significantly from chance. These results, plus simulations using associative models, suggest that the orderability of the postchoice feedback stimuli is important for crows' transitive responding.

Key Words: transitive inference, associative learning, cognition, visual discrimination, hooded crows

A relation between premises is said to be transitive if the relation Rl that links Stimuli b and c and Stimuli c and d also links Stimuli b and d . A transitively competent individual should be able to deduce that if $b Rl c$ and $c Rl d$, then $b Rl d$. Moreover, such an individual should reach this conclusion if and only if Rl

supports a linear order between the stimuli. In other words, from the premises b is bigger than c and c is bigger than d , it follows that b is bigger than d . But, what follows from the premises b stands near to c and c stands near to d ? Because the spatial arrangement of the stimuli is not fully specified, it is impossible to infer the relation between b and d . The ability to distinguish transitive relations from those that do not allow correct inferences is thought to be an important characteristic of transitive inference (TI) formation in humans, at least in verbal tests (Evans, Newstead, & Byrne, 1993; Wright, 2001).

Olga Lazareva, formerly at the Institute of Higher Nervous Activity, Moscow, is now at the University of Iowa. Anna Smirnova, Maria Bagozkaja, and Zoya Zorina are at Moscow State University, Vladimir Rayevsky is at the Institute of Higher Nervous Activity, and Edward Wasserman is at the University of Iowa.

The research was supported by the collaborative grant from NATO to Z.A. Zorina and J. D. Delius, by the NCCR Neural Plasticity and Repair, and by 7IP No 62645 (Swiss National Scientific Foundation). The results were partially presented at UFAW Symposium, "Consciousness, Cognition, and Animal Welfare" (London, 2000), at the Fourth European Meeting for the Experimental Analysis of Behaviour (Amiens, 2000), and at the conference of Comparative Cognition Society (Florida, 2002). The experiment was a part of the doctoral thesis research of Olga Lazareva while she was a Ph.D. student at the Institute of Higher Nervous Activity (Moscow, Russia). We thank J. D. Delius, M. Siemann, M. J. Acerbo, M. Cleveland, B. Gibson, and I. A. Reyer for their helpful comments on an early version of the manuscript. We are grateful to Clive Wynne, Tom Zentall, and Tom Critchfield who provided valuable theoretical feedback on the manuscript.

Correspondence concerning this article should be addressed to Olga Lazareva, E11 Seashore Hall, Department of Psychology, University of Iowa, Iowa City, Iowa 52242 (e-mail: olga-lazareva@uiowa.edu).

Transitive inference could be beneficial for animals as well. Suppose that a monkey enters a new group and must establish its own social position. One way to achieve this objective is to fight with every other group member and to chart the success or failure of these encounters. A more economical way is to observe the relations among the members of the group and, after a few strategic encounters, to deduce one's own position in the existing hierarchy. Observe that an individual who fails to distinguish transitive relations from nontransitive relations can easily reach the wrong conclusions about an existing dominance hierarchy that will lead to a large number of unprofitable encounters. Field observations suggest that monkeys and apes ac-

tually use this kind of transitive information when they enter into a new society (Altmann, 1962; Kummer, 1982). This ability was documented in the laboratory for two bird species (Hogue, Beaugrand, & Lague, 1996; Paz-y-Mino, Bond, Kamil, & Balda, in press).

A method for presenting semiverbal TI tasks to very young children was developed by Bryant and Trabasso (1971); it was later adapted into a fully nonverbal method for animals by many researchers. The method normally involves presenting a series of overlapping pairs of stimuli: A+ B-, B+ C-, C+ D-, and D+ E- (where the letters stand for different stimuli and the plus and minus symbols indicate that choices of the corresponding stimuli are either reinforced or nonreinforced, respectively). Bryant and Trabasso used pairs of wooden sticks of different colors extending an *equal* distance above the top of a box. Only after a child had reported which stick was longer or shorter did the investigators show the child the *full* lengths of both sticks by taking them out of the box. The lengths of sticks, therefore, provided visual feedback *after* the choice was made (in some experiments, the children received verbal, rather than visual, postchoice feedback). Therefore, as in the verbal task, the choice items in this task could be ordered along a physical dimension, which supported a transitive relation between them.

To test for transitive inference, the novel BD testing pair was presented. This pair involved stimuli that were never before presented together; additionally, the choice of the stimuli B and D had been reinforced in one pair and nonreinforced in another pair. A transitively competent individual was expected to remember that $B > C$ and $C > D$, to infer that $B > D$, and to select B over D. Bryant and Trabasso (1971) found that 4-year-old children were quite able to select the transitively correct B stimulus when the discrimination task was presented in semiverbal form.

Several points must be emphasized about the design devised by Bryant and Trabasso (1971). First, at least five stimuli have to be presented; otherwise, there will be no new pair that does not contain end anchor stimuli, the choice of which was always reinforced or nonreinforced. Second, the stimuli have to be orderable along a dimension that supports a transitive relation between them. Third, any

dimensional comparison of the discriminative stimuli can be allowed only *after* the choice; otherwise, the successful choice between the stimuli can be achieved by direct dimensional comparison of the stimuli, without using associative information about the relations between different pairs. In the Bryant and Trabasso experiments, for example, the lengths of sticks provided visual feedback after the color choice was completed.

Of the three design principles described above, only the first has been followed in all of the prior animal studies. In some studies, the relation between the stimuli was explicit, just as in the Bryant and Trabasso (1971) experiments (McGonigle & Chalmers, 1977; Rapp, Kansky, & Eichenbaum, 1996; Roberts & Phelps, 1994; Treichler & Van Tilburg, 1996). In others (Bond, Kamil, & Balda, 2003; Boysen, Berntson, Shreyer, & Quigley, 1993; Davis, 1992b; Siemann, Delius, & Wright, 1996; Steirn, Weaver, & Zentall, 1995; von Fersen, Wynne, Delius, & Staddon, 1991), the stimuli were arbitrary, and the researchers believed that their subjects could establish such a transitive relation solely on the basis of the reinforcement or nonreinforcement of the stimuli. Note, on the one hand, that the nonordered TI task might be represented as equivalent to the verbal premises, "B is reinforced, C is not reinforced" and "C is reinforced, D is not reinforced"; in a verbal task, the transitive-like conclusion "B is reinforced, D is not reinforced" may well be incorrect because the relation "reinforced" is not transitive. On the other hand, a discrimination procedure such as this may establish a preference for one stimulus over the other and such *preferences* may be transitive. In fact, value transfer theory, as well as other associative theories, suggests that the relation becomes transitive during training, as an ordered series of associative values arises. We will return to this idea later.

Still, many animals have been reported to be able to select B in the new BD pair, even in experiments in which the discriminative stimuli were not orderable with respect to some physical characteristic. Unfortunately, only a few species have been studied in the same experimental conditions both with and without transitive relations. Roberts and Phelps (1994) trained rats using wooden tunnels marked with different odors as stimuli;

the tunnels were linearly arranged and took the same position throughout training. In the test, the rats selected B over D at a high level. When, however, the tunnels were circularly arranged during training, the rats selected B and D at a chance level, suggesting that the spatial information provided by the ordered placement of the tunnels was crucial for transitive responding.

It remains unclear whether the explicit orderability of the stimuli is important for all variants of the TI task or only when spatial cues are used. For example, in experiments by Davis (1992b) and Bunsey and Eichenbaum (1996), rats were found to respond transitively, although there was no transitive relation among the stimuli (different odors). New experiments that use stimuli from a nonspatial domain may help clarify whether the ordering of stimuli along a physical dimension may control organisms' transitive behavior.

The mechanisms underlying transitive responding are in dispute. Some researchers have assumed that an organism integrates the independently presented premises into an ordered series of internal representations and that these representations are spatial in nature (Davis, 1992a; Gillan, 1981; Rapp et al., 1996; Zorina, Kalinina, & Markina, 1996). Experimental evidence has been interpreted to support this point of view. The first finding is the "end-anchor" effect: higher performance in pairs that contain the first or the last stimulus (A or E) in the series compared to pairs containing the middle stimuli (B, C, or D). According to the spatial representation hypothesis, comparisons including the end items should be faster and easier to learn because they enjoy the position of the largest or the smallest object in the series. The second finding is the "symbolic distance" effect: the smaller the separation between the items in a testing pair, the longer the response time and the higher the probability of an error. Spatial representation hypothesis implies that the closer the representations of the items in a series, the more similar they are to one another, and hence the more difficult the discrimination between them should be.

Other theorists have proposed that the same pattern of results can be explained by associative models. Such models consider the organism's choice of the testing stimuli to be the result of the difference in their relative

reinforcement history (see Couvillon & Bitterman, 1992; Siemann & Delius, 1998; Wynne, 1995). Value transfer theory, the first in this class of models, proposed that the associative values of the training stimuli produce an ordered series, $A > B > C > D > E$, due to bidirectional transfer of associative value between reinforced and nonreinforced stimuli (von Fersen et al., 1991). Later studies revealed, however, that an ordered series of associative values arose even when no transfer of values across stimuli was postulated. In other words, the common outcome of training with the four premise pairs, $A+ B-$, $B+ C-$, $C+ D-$, and $D+ E-$, was found to be the formation of an ordered series of the associative values of those stimuli (see Siemann & Delius, 1998; Wynne, 1995, 1998 for more details).

Note that, in all current associative models, the ordered series of values arises from the training procedure, in which the premise pairs are presented in a given sequence and with a given frequency. Thus we should be able to manipulate the outcome of a TI test by changing the frequency of the premise pairs. If we were to increase the frequency of $D+ E-$ presentation, then we might expect that, eventually, the value of D would become greater than the value of B and, consequently, an organism should prefer Stimulus D. Such a test could provide strong support for an associative model of transitive responding, because the spatial representation hypothesis is unable to predict this outcome under any conditions.

The spatial representation hypothesis, however, suggests that the use of orderable training stimuli is the prime reason for the emergence of an ordered series of stimulus values during training. As argued by Markovits and Dumas (1992), "what makes a relationship transitive is . . . the implicit placement of the items along at least an ordinal scale" (p. 311). (See Wright, 2001, pp. 386–387, for a similar argument.) Therefore, if we were to show that the use of orderable stimuli facilitates transitive responding, then these results would speak in favor of the spatial representation hypothesis, especially because stimulus orderability is deemed to be irrelevant to associative models.

The present study examined transitive responding in hooded crows (*Corvus cornix L.*), a member of the *Corvidae* family. Corvids are

characterized by a large telencephalon and relatively high brain complexity (Rehkämper, Frahm, & Mann, 2001; Stingelin, 1958) as well as by an ability to solve various types of complex cognitive tasks (Balda, Kamil, Bednekoff, & Hile, 1997; Heinrich, 2000; Jones, Antoniadis, Shettleworth, & Kamil, 2002; Köhler, 1950; Krushinsky (1977/1990); Smirnova, Lazareva, & Zorina, 2000; Wilson, Mackintosh, & Boakes, 1985). Indeed, some of these tasks, such as tool making or complex numerical abilities, have been documented primarily in corvids and primates (Weir, Chappell, & Kacelnik, 2002; Zorina, 1997).

The present study compared transitive responding in crows in ordered and nonordered transitive inference tasks. In the ordered feedback group of the current experiment, 4 crows were shown colored choice stimuli followed by postchoice feedback stimuli of the same color, but of different size. This method is functionally equivalent to Bryant and Trabasso's (1971) task, in which children were given visual feedback after they had reported which of two sticks was shorter or longer. In the constant feedback group of the current experiment, 4 birds were trained with feedback stimuli that were of the same size; those feedback stimuli could not help to order the choice stimuli along a common dimension. In both groups, discrimination training was followed by extended D+ E- training, which led to a richer reinforcement history for Stimulus D than Stimulus B. Under these circumstances, the spatial representation hypothesis predicts transitive responding in the ordered feedback group, but not in the constant feedback group, where the choice stimuli cannot be ordered along any dimension; associative theories do not predict transitive responding in either group. In addition, we conducted post hoc simulations using several associative models including the value transfer model, to verify that associative models did, indeed, predict no response to "logically correct" Stimulus B in testing pair BD.

STUDY 1

METHOD

Subjects

Eight hooded crows (*Corvus cornix L.*), each more than 1 year old, were studied. The

crows were caught in the wild when they were about 6 months old and housed in outdoor aviaries in small groups of 2 or 3 birds. All of the birds were experimentally naive. Before the start of the experiment, the birds were food deprived for 2 days. Throughout the experiment, the birds had free access to water and food, except when they refused to work during training. In these cases, food without animal protein was given for 1 or 2 days.

Apparatus

A mesh wire cage (65 cm by 50 cm by 45 cm; mesh dimensions 4 cm by 4 cm) and a plastic tray (20 cm by 30 cm) with a handle (30 cm) were used. The left panel of Figure 1 shows a schematic representation of the experimental apparatus. Two cups (each 3.7 cm high and 5.0 cm in diameter) were placed on the tray (13 cm from the base of the tray and 1.5 cm from the side of the tray) with two plastic blocks (4 cm by 4 cm by 5 cm) behind each of them. Two transparent plastic pockets (7.1 cm by 7.1 cm) were attached to each of the blocks. During the experiment, color cardboard cards that served as the choice stimuli (see below) were put into the pockets. The pockets were attached so that the cards could serve as cup lids.

The tray was prepared for the trial out of the crow's view. An opaque plastic screen (70 cm by 40 cm) stood between the experimental cage and the experimenter, so that when the crow made its choice, it could not see the experimenter and vice versa.

Stimuli

Colored cards (7 cm by 7 cm) were used to represent the terms of the series A, B, C, D, and E. The colored upper surfaces of the cards (red, yellow, green, blue, or black) served as the choice stimuli. A circle of the same color was drawn on the underside of each card. These circles served as the post-choice feedback stimuli. The diameter of the feedback circles decreased from red (6.5 cm) to black (4.5 cm; right panel of Figure 1) in the ordered feedback group. In the constant feedback group, the diameter of all feedback circles was 6.5 cm. Additionally, two white cards were used for pretraining (see below).

Experimental Design

The stimuli and experimental design are shown in the right panel of Figure 1. The

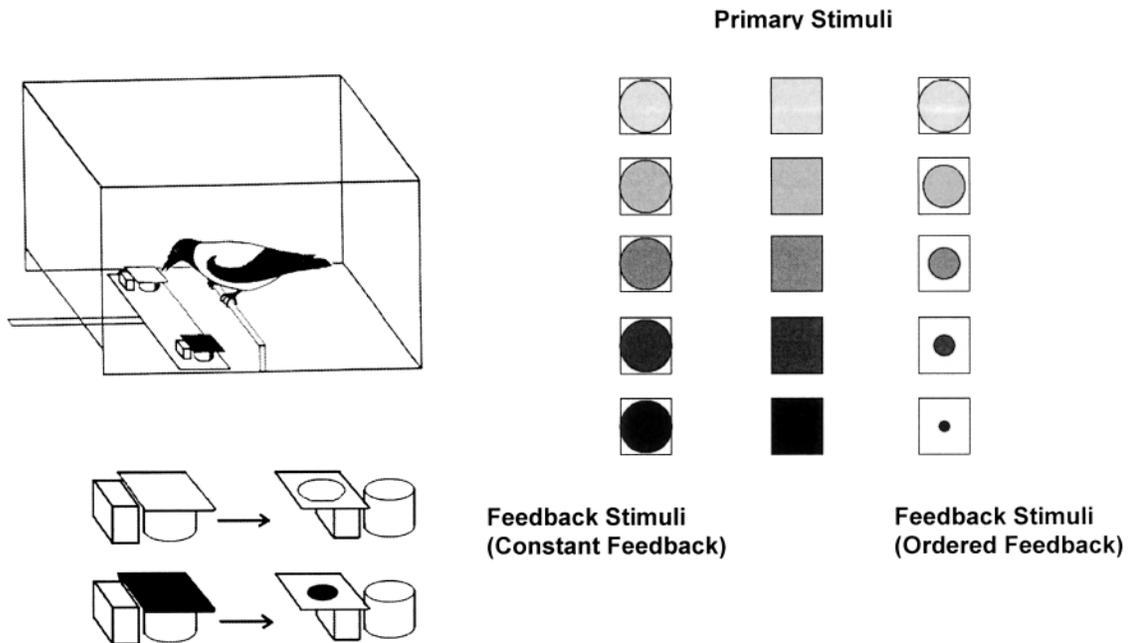


Fig. 1. Experimental setting (left) and design (right). Left top—bird in the experimental cage; left bottom—two cups with stimulus lids before and after opening. Right—choice and feedback stimuli for the constant and ordered feedback groups.

birds were trained on a multiple discrimination task involving overlapping pairs of visual choice stimuli: $A+ B-$, $B+ C-$, $C+ D-$, and $D+ E-$ (where the letters stand for choice stimuli of different colors, and the plus and minus symbols indicate that choices of the corresponding stimuli were either reinforced or nonreinforced, respectively).

In order to minimize possible color preferences, the crows were randomly assigned to two groups. The stimuli A, B, C, D, and E were represented by red, yellow, green, blue, black, respectively, for half of the crows within each group (ordered feedback: Zosia and Kotia; constant feedback: Korsar and Solomon), and by black, blue, green, yellow, and red for the other half of the crows (ordered feedback: Dascha and Zelenaja; constant feedback: Carolina and Kondrat). The postchoice feedback stimuli had the same color as the corresponding choice stimuli in both groups. Note that for 2 birds in the ordered feedback group, the postchoice feedback stimulus for the reinforced choice stimulus always had a larger diameter and the transitive series could be represented as $A > B > C > D > E$. For the other 2 birds in the ordered feedback

group, this postchoice feedback stimulus always had a smaller diameter and the relation in transitive series was reversed: $A < B < C < D < E$. In the constant feedback group, the postchoice feedback stimuli were of equal diameter (6.5 cm); therefore, these constant-size postchoice feedback stimuli could not help the pigeons order the choice stimuli along any dimension.

General Procedure

A trial began when the tray with the cups covered by the cards was slid into the crow's cage. One of the cups contained two mealworms as reinforcement. The bird could turn over either card and, if the choice was correct, then it received mealworms. After a correct choice, the birds usually also turned over the card over the incorrect cup; if not, then the experimenter turned over this card for the bird. Consequently, a bird could see *both* feedback stimuli and compare their diameters. After an incorrect choice, the tray was immediately angled so that the crow could not reach the correct cup. The experimenter then turned over the correct card and left the tray in the bird's field of vision for 3 to 5 s

before withdrawing it in order to give the bird an opportunity to see the postchoice feedback stimuli.

During Phases 1 through 8 (see below), if the bird made three repeated choices of the left or the right choice stimulus, then the trial was repeated until a correct response was made (correction procedure). During the testing phase, this correction procedure was not employed. The number of correction trials was recorded, but only the first response on each trial was scored in the data analysis.

Training in Phases 1 through 7 continued until a criterion of 80% correct or better over 30 consecutive trials was reached ($p < 0.001$; binomial probability test). Phases 8 and 9 consisted of a fixed number of trials. Experimental sessions were conducted 6 days each week and involved 40 or 60 trials (correction trials were not included in these totals). The exact number of trials per day depended on a bird's performance on each particular day.

Preliminary Training

The crows were trained to turn over the white cards to find food reinforcement. During preliminary training, reinforcement was placed into both cups; therefore, any choice was reinforced. Preliminary training continued until the birds turned over the white cards immediately after the tray was inserted into the cage.

Training

Phase 1. Only the single choice pair, A+ B−, was presented. The left–right location of reinforced responding was counterbalanced within a block of 10 trials, under the restriction that S+ did not appear in the right or left location on more than two successive trials. The birds were required to meet the 80% criterion during 30 consecutive trials.

Phase 2. The pair B+ C− was presented; the counterbalancing and the criterion remained the same as during Phase 1.

Phase 3. Two pairs, A+ B− and B+ C−, were presented during each session. The counterbalancing occurred in blocks of 16 trials, and both choice pairs were presented in a pseudorandom sequence so that no pair was presented on more than two successive trials. The crows were trained until they met the 80% criterion to each pair during 32 consecutive trials.

Phase 4. A pair C+ D− was presented; the counterbalancing and the criterion remained the same as during Phase 1.

Phase 5. Three pairs, A+ B−, B+ C−, and C+ D−, were presented during each session. The counterbalancing occurred in blocks of 16 sessions, and each choice pair was presented in a pseudorandom sequence so that no pair was presented on more than two successive trials. The crows were trained until they met the 80% criterion to each pair during 32 consecutive trials.

Phase 6. The pair D+ E− was presented; the counterbalancing and the criterion remained the same as during Phase 1.

Phase 7. Four pairs, A+ B−, B+ C−, C+ D−, and D+ E−, were presented during each session. The counterbalancing occurred in blocks of 16 trials, and both choice pairs were presented in a pseudorandom sequence so that no pair was presented on more than two successive trials. The crows were trained until they met the 80% criterion to each pair during 32 consecutive trials.

Bias Reversal Phase

It was conceivable that upon completion of all training phases the reinforcement history of stimulus B was richer when compared to stimulus D. If so, then the possible use of linearly ordered series of the stimuli could be confounded by the difference in their associative values. To disentangle the influence of the reinforcement history from the orderability of the stimuli along a dimension, we evaluated the reinforcement history before the test on a bird-by-bird basis.

To estimate the possible influence of the reinforcement history of testing pair BD, the number of reinforcements (N_r) and nonreinforcements (N_n) for selection of each choice stimulus during all of the training phases (including correction trials) was calculated. The reinforcement/nonreinforcement ratio $R = N_r/N_n$ was used to compare the value of the B and D choice stimuli. If the ratio for D was larger than the ratio for B, then the testing phase was begun; if not, then the crows were exposed to a bias reversal phase.

The bias reversal phase involved the D+E− choice pair, with number of presentations adjusted on the basis of the reinforcement/nonreinforcement ratios calculated earlier. This

Table 1

Number of correct and incorrect choices and reinforcement/nonreinforcement ratio both before and after the bias reversal phase in the ordered feedback group. N_r and N_n stand for number of reinforcements and nonreinforcements, correspondingly; R stands for reinforcement/nonreinforcement ratio.

	Stimulus	Bird			
		Zosia	Kotia	Dasha	Zelenaia
Before bias reversal					
N_r	B	124	114	96	127
	D	54	256	145	98
N_n	B	58	83	122	96
	D	52	151	50	39
R	B	2.1	1.4	0.8	1.3
	D	1.0	1.7	2.9	2.5
After bias reversal					
N_r	B	131			
	D	153			
N_n	B	71			
	D	60			
R	B	1.8			
	D	2.6			

bias reversal phase was conducted in order to augment the reinforcement/nonreinforcement ratio of D, and it was designed to continue until the reinforcement ratio of D was at least 1.1 times greater than that of B. After that, the birds were required to meet a criterion of 80% correct with each of the three remaining training pairs during 32 consecutive trials.

For instance, for the crow Zosia the choice of Stimulus B was reinforced 124 times and not reinforced 58 times, whereas the choice of Stimulus D was reinforced 54 times and not reinforced 52 times, yielding R values of 2.1 and 1.0, respectively. Therefore, the crow could prefer B over D on the basis of unequal histories of reinforcement. The bias reversal phase added 102 trials with D+ E- and 10 trials with each of the three remaining pairs: A+ B-, B+ C-, and C+ D-. This plan would have changed the ratio of D to 3.0 if the crow did not make any mistakes during that phase. Actually, the crow's accuracy was less than perfect (88.2%), yielding a final D ratio of 2.6. Consequently, if, during the subsequent BD test, this crow chose between stimuli according to these reinforcement/nonreinforcement ratios, then it should prefer D; if the crow chose according to the diameter of the feedback stimuli, then it should prefer B.

Testing

The testing phase consisted of the presentation of all four training pairs and the new testing BD pair. On BD testing trials, the selection of either choice stimulus was reinforced with a probability of 0.5. To prevent the birds from comparing the diameters of the postchoice feedback circles during the testing pair trials, only one lid could be opened and the other remained closed. The testing pair was presented 40 times out of total of 160 trials.

RESULTS

For the ordered feedback group, training with all four pairs of choice stimuli took an average of 727 trials, with a minimum of 524 trials and a maximum of 882 trials. For the constant feedback group, training with all four pairs of choice stimuli took an average of 781 trials (range, 462 to 1,050). There were no obvious differences between birds with different color assignments, so their results were pooled.

Table 1 shows the number of reinforced and unreinforced choices of Stimulus B and Stimulus D by each bird in the ordered feedback group before and after the bias reversal phase. The table also shows the final ratio, R , of reinforced and unreinforced choices. After

Table 2

Number of correct and incorrect choices and reinforcement/nonreinforcement ratio both before and after the bias reversal phase in the constant feedback group. N_r and N_n stand for number of reinforcements and nonreinforcements, correspondingly; R stands for reinforcement/nonreinforcement ratio.

	Stimulus	Bird			
		Korsar	Solomon	Carolina	Kondrat
Before bias reversal					
N_r	B	159	82	273	343
	D	101	35	269	115
N_n	B	72	70	117	61
	D	217	89	86	71
R	B	2.2	1.2	2.3	5.6
	D	0.5	0.4	3.1	1.6
After bias reversal					
N_r	B	170	96		364
	D	591	224		622
N_n	B	72	76		64
	D	217	93		89
R	B	2.4	1.3		5.7
	D	2.7	2.4		7.0

training, R for Stimulus D was greater than R for Stimulus B for Kotia, Dasha, and Zelenaiia, so they proceeded directly to the test. The 4th crow, Zosia, was exposed to the bias reversal phase. Originally, R for Stimuli B and D were 2.1 and 1.0, respectively, whereas after the bias reversal phase the values were 1.8 and 2.6, respectively.

Table 2 shows the number of reinforced and unreinforced choices of Stimulus B and Stimulus D by each bird in the constant feedback group before and after the bias reversal phase as well as the final reinforcement/nonreinforcement ratios. Carolina's R values for Stimuli B and D after training was complete were 2.3 and 3.1, respectively, so testing commenced immediately after training. Three crows—Korsar, Solomon, and Kondrat—had to be given the bias reversal phase. Thereafter, R was higher for Stimulus D than for Stimulus B.

Figure 2 depicts discrimination accuracy during testing for the four training pairs and preference among the stimuli in the testing pair (BD). (Henceforth, "choice" is used to refer to performance involving the training pairs, and "preference" is used to refer to performance involving the testing pairs.) For the constant feedback group (upper panel), training-pair accuracy was above chance (binomial probability test, $p < 0.05$) in all cases except pairs BC and CD for Korsar and pairs

AB and CD for Kondrat. No crow in this group strongly preferred Stimulus B over D. One (Korsar) preferred Stimulus D over Stimulus B (80.0%), and the other 3 birds' performance was near chance (Solomon, Kondrat, and Carolina). For the ordered feedback group (bottom panel), training-pair accuracy was above chance (binomial test, $p < 0.05$) in all cases except pair DE for Kotia and pairs CD and DE for Zelenaiia. All crows in this group strongly preferred Stimulus B over Stimulus D ($M = 83.1\%$).

Analyses of variance (ANOVA) were conducted with group (ordered, constant) and pairs (A+ B-, B+ C-, C+ D-, D+ E-, BD) as factors and with accuracy as the dependent variable. There were significant main effects of group, $F(1, 7) = 9.19$, $p < 0.05$, and pair, $F(4, 28) = 11.81$, $p < 0.05$, as well as a significant pair \times group interaction, $F(4, 28) = 15.48$, $p < 0.05$. A follow-up Tukey test found that, in the ordered feedback group, preference involving testing pair BD did not differ in magnitude from that of any of the four training pairs. In the constant feedback group, however, preference for Stimulus B was significantly lower than that for any of the reinforced stimuli in the training pairs. Most importantly, in the BD test trials, preference for Stimulus B was significantly higher in the ordered feedback group than in the constant feedback group.

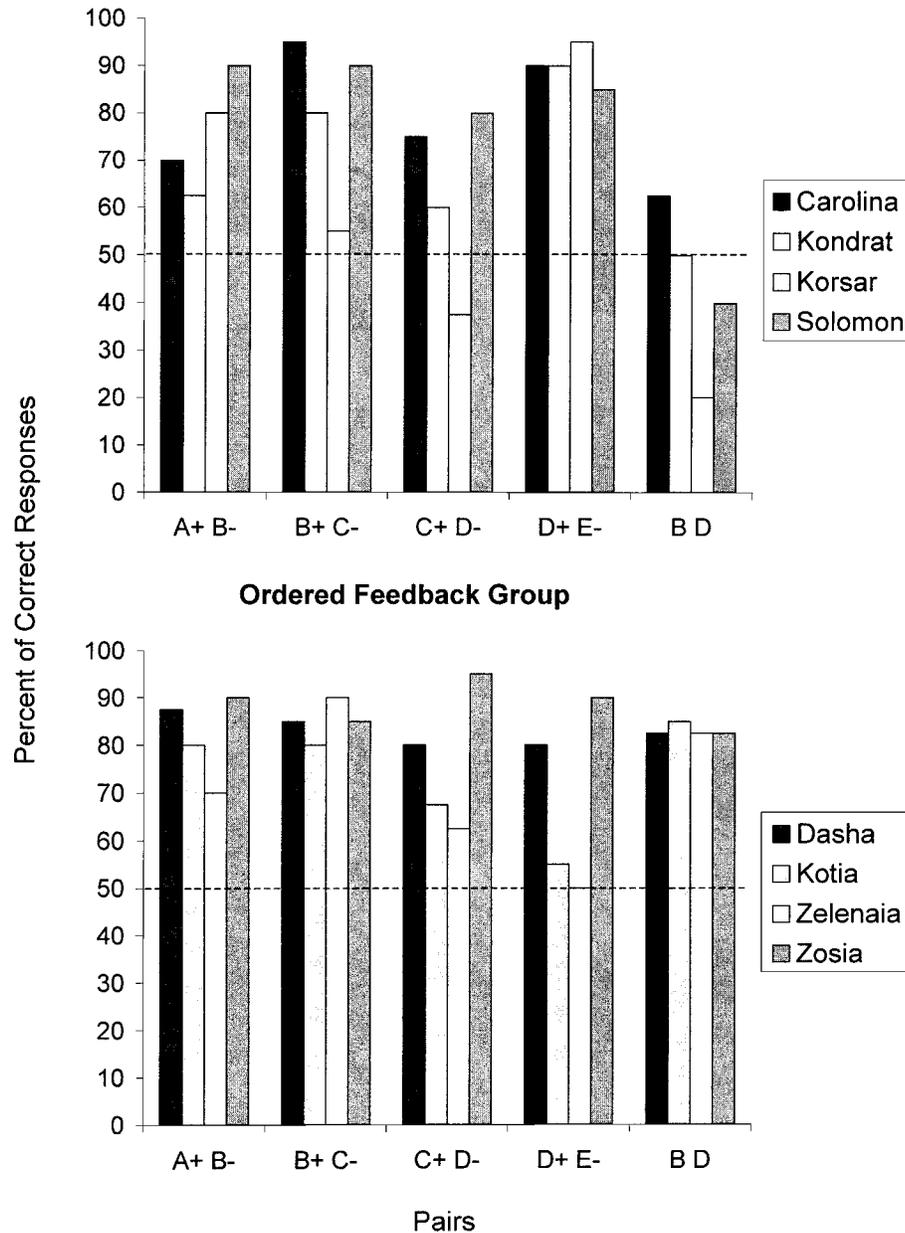


Fig. 2. Choice accuracy in the training pairs and preference for Stimulus B in the testing BD pair in the constant feedback group (upper panel) and the ordered feedback group (lower panel).

DISCUSSION

Strong transitive responding (i.e., preference of Stimulus B in new pair BD) emerged only in the ordered feedback group. In the constant feedback group, crows either selected the Stimuli B and D at chance levels or preferred Stimulus D, which was more often reinforced during training. These results are

consistent with the spatial representation hypothesis (Davis, 1992a; Gillan, 1981; Rapp et al., 1996; Zorina et al., 1996), and support the notion that orderability of the stimuli along a physical dimension such as size may be a necessary and sufficient condition for non-verbal transitive responding in crows.

It is important to stress that, in the present

experiment, a crow could not *directly* compare the sizes of the choice stimuli before a selection was made; the choice stimuli themselves were equal-sized squares of different colors. To establish an ordered series, a crow would need to notice that each choice stimulus preceded a feedback circle of the same color but of a different diameter, and that the diameters of the feedback circles could be dimensionally ordered.

One might note that 3 out of 4 birds in the constant feedback group were exposed to the bias reversal phase, whereas only 1 out of 4 birds required bias reversal in the ordered feedback group. Could the difference in transitive responding between groups be the result of mere exposure to the bias reversal phase? If that were true, then the crows that went through bias reversal in both groups should demonstrate lower transitive responding than the crows that proceeded directly to the test. In the constant feedback group, mean preference for Stimulus B during BD test trials was 36.7% for the crows exposed to the bias reversal phase and 62.5% for the crow that went directly to the test. In the ordered feedback group, preference for Stimulus B was 82.5% for the crow exposed to the bias reversal phase and 83.3% for the crows that went directly to the test. It seems unlikely, therefore, that mere exposure to the bias reversal phase could affect BD performance independently of the other factors. Future studies could evaluate this experimentally by including a BD test both before and after the bias reversal phase.

Another possible concern involves the low accuracy on training trials in the test for some of the birds in the constant feedback group (for example, Korsar, see Figure 2). Could the low performance to test pair BD be explained by low performance to the training pairs? Again, we found that poor performance to some training pairs during the test in the ordered feedback group (for example, Zelenaiia) did not prevent the crow from selecting Stimulus B of the BD pair at a very high level of accuracy. Moreover, the birds that were exposed to the bias reversal phase were retrained with the training pairs before they went into the test (see Method section). A decrease in accuracy involving the training stimuli when novel testing stimuli are intro-

duced has been documented before for other bird species (Pepperberg, 1987).

Although the ratio, R , of reinforced to unreinforced choices was used as an index of the associative strength of the choice stimuli, the question arises whether the manipulation of this ratio is adequate to change the predictions made by associative models for the BD test. An attempt was made, therefore, to simulate the present results by using associative models that previously have been found to yield reasonably good accounts of transitive responding (see Siemann & Delius, 1998; Wynne, 1995, 1998).

STUDY 2

METHOD

Preliminary simulations disclosed that elementary associative models (i.e., those that do not include a configural value component, such as the Luce, Bush-Mosteller, or Rescorla-Wagner models) could not provide a satisfactory fit to our data. (These simulations are available, upon request, from the corresponding author.) Hence subsequent simulations focused on two configural models, one based on the Rescorla-Wagner equations (Wynne, 1995) and the other based on the Luce equations (Siemann & Delius, 1998).

Also examined was a value transfer model that has been proposed as a possible account for nonverbal transitive responding (von Fersen *et al.*, 1991). The value transfer model does not specify how associative values are accrued; it simply states that the accrued values can be transferred between simultaneously presented stimuli. Therefore, the Siemann-Delius and Wynne models were modified to incorporate the value transfer mechanism. Short descriptions of each model, together with the equations, are provided in the Appendix.

The data of each bird were fitted individually, using the full sequence of trials presented during training (including correction trials) and employing the least-square error technique. Only the training data were fitted. The obtained associative values of the stimuli were used to calculate choice probability for testing pair BD according to the choice function used by the model (see Appendix).

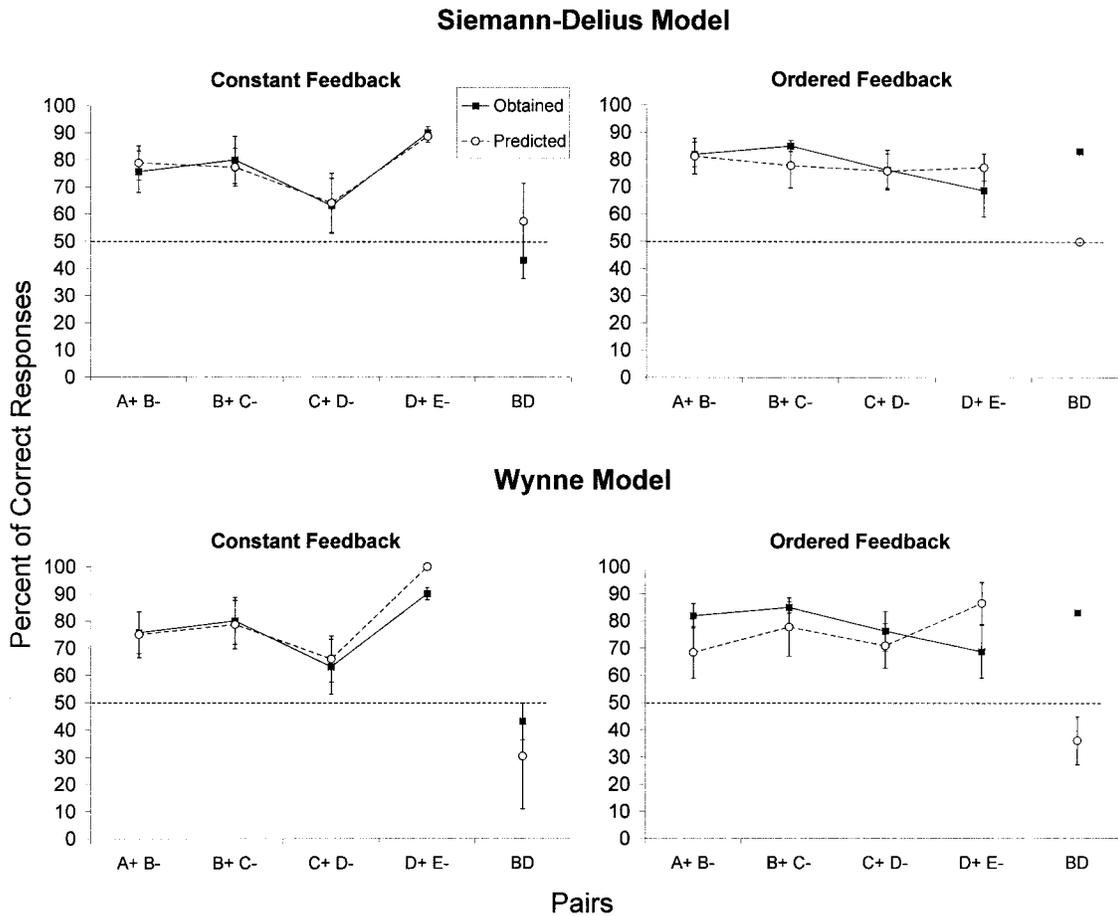


Fig. 3. The obtained percentage of correct choices in the training pairs and transitive responses in the testing BD pair compared with simulations using the Wynne and Siemann-Delius models.

RESULTS AND DISCUSSION

The results of the simulations are presented in Figures 3 and 4. The individual fits are given in Tables 3 and 4; Table 5 provides the values of the model parameters for the best-fitting result. Figure 3 (left panels) shows that both models provided a good fit for training-pair accuracy in the constant feedback group, although the Wynne model slightly overestimated accuracy for the D+ E- pair. The models predicted mean preference between test Stimuli B and D to be at or below chance. Table 3 documents that the Siemann-Delius model erroneously predicted preference of Stimulus B over D for 1 crow (Solomon), but accurately predicted the other crows' preference. The Wynne model either predicted too high (Carolina) or too low (Kondrat, Korsar, and Solomon) levels of transitive responding.

Both models also provided a reasonable fit for training-pair accuracy in the ordered feedback group (Figure 4, right panels). Importantly, the models again predicted mean preference involving test Stimulus B to be at or below chance when in fact all of the birds in this group strongly preferred Stimulus B over D (83.1% on average). Table 3 shows that the Siemann-Delius model predicted indifference for all birds, whereas the Wynne model predicted preference for Stimulus D for 2 crows (Dasha and Zosia) and indifference for the other 2 crows.

Figure 4 (left panels) shows that the Siemann-Delius and Wynne models, when modified to incorporate the value-transfer mechanism, both produced excellent fits of training-pair accuracy for the constant feedback group. The predictions of the modified

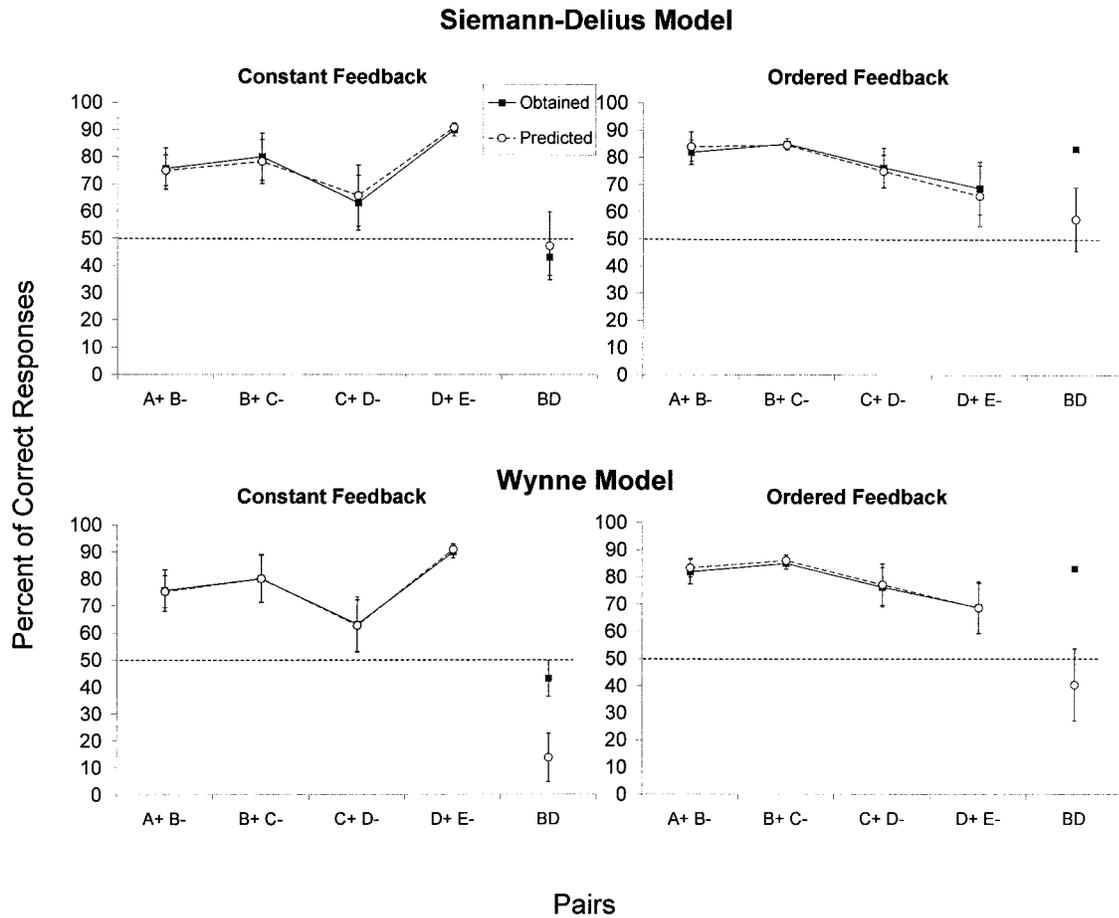


Fig. 4. The obtained percentage of correct choices in the training pairs and transitive responses in the testing BD pair compared with simulations using the modified Wynne and Siemann-Delius models that incorporated a value transfer mechanism.

Siemann-Delius model for preference involving the BD pair matched the obtained results, whereas the modified Wynne model predicted less preference for Stimulus B than was actually observed. Table 4 shows that the modified Siemann-Delius model erroneously predicted high transitive responding for 1 out of 4 crows (Carolina). The modified Wynne model predicted strong preference of Stimulus D over B for all 4 crows, instead of the chance responding that was observed.

Both models still predicted indifferent B-D preference in the ordered feedback group (Figure 4, right panel). Table 4 shows that the modified Siemann-Delius model predicted preference of Stimulus B over D for 2 crows (Dasha and Zelenaiia), chance responding for 1 crow (Kotia), and preference of Stimulus D

over B for 1 crow (Zosia). The modified Wynne model predicted preference of Stimulus B over D for 1 crow (Zelenaiia), chance responding for 1 crow (Dasha), and preference of Stimulus D over B for the remaining 2 crows (Kotia and Zosia). All birds, however, preferred Stimulus B over Stimulus D at a statistically significant level. Thus neither configural models alone nor configural models that incorporated the value transfer mechanism provided a satisfactory account for the present data.

Both the Siemann-Delius and Wynne associative models predicted performance on BD testing trials that was in clear disagreement with the obtained data. Yet in previous studies those models have been used to simulate successfully transitive responding in both pi-

Table 3
 Simulated accuracies for Siemann-Delius and Wynne models.

Bird		A+ B-	B+ C-	C+ D-	D+ E-	BD
Constant feedback group						
Carolina	Obtained	70.00	95.00	75.00	90.00	62.50
	Siemann-Delius	83.46	88.45	69.78	86.26	50.00
	Wynne	62.28	94.63	78.39	100.00	85.55
Kondrat	Obtained	62.50	80.00	60.00	90.00	50.00
	Siemann-Delius	62.28	79.86	60.19	89.73	48.66
	Wynne	59.98	76.44	63.51	99.99	28.99
Korsar	Obtained	80.00	55.00	37.50	95.00	20.00
	Siemann-Delius	77.13	56.80	37.07	94.39	33.16
	Wynne	82.01	54.51	42.94	100.00	0.05
Solomon	Obtained	90.00	90.00	80.00	85.00	40.00
	Siemann-Delius	92.62	84.16	89.48	84.47	97.84
	Wynne	95.41	89.05	79.00	100.00	6.90
Ordered feedback group						
Dasha	Obtained	87.50	85.00	80.00	80.00	82.50
	Siemann-Delius	88.55	85.20	77.11	84.83	49.98
	Wynne	83.57	97.03	78.18	99.99	21.47
Kotia	Obtained	80.00	80.00	67.50	55.00	85.00
	Siemann-Delius	64.29	59.69	68.66	73.26	50.00
	Wynne	52.73	57.99	57.90	73.11	50.40
Zelenaia	Obtained	70.00	90.00	63.50	50.00	82.50
	Siemann-Delius	77.83	69.59	64.41	64.90	50.00
	Wynne	51.00	60.50	56.93	73.11	52.50
Zosia	Obtained	90.00	85.00	95.00	90.00	82.50
	Siemann-Delius	94.08	96.87	93.15	85.83	50.00
	Wynne	86.33	95.51	90.50	100.00	20.07

geons and people (Delius & Siemann, 1998; Siemann & Delius, 1998; Wynne, 1995, 1997, 1998). What might be responsible for the disparity?

The associative value of a stimulus in the Siemann-Delius and Wynne models has two components—an elemental value and a configural value (see Appendix). The elemental value is updated whenever a given stimulus is presented, whereas the configural value is updated whenever a given stimulus is presented in a given pair. Testing pair BD had never been presented before, so the response to this pair is determined solely by the elemental values of Stimuli B and D. Thus, to predict transitive responding to testing pair BD, the elemental value of Stimulus B must be higher than the elemental value of Stimulus D.

The present experiment employed a bias reversal procedure that consisted of massive presentation of the training pair D+ E-. From a model's point of view, this procedure greatly increased the elemental value of Stimulus D and left unchanged the elemental value of Stimulus B, thereby assuring the prediction of choice of Stimulus D in the BD

testing pair. Unfortunately, transitive performance was not tested *before* the bias reversal phase. Such a test could show whether changes in predicted BD performance are due to the bias reversal procedure or to peculiarities in the individual training history of each bird.

It was also found that including a value transfer mechanism in the associative models improved their fit of the obtained training data, but did not systematically improve their predictions of BD performance. Value transfer was first suggested as a hypothetical process that might help to explain transitive responding in pigeons (von Fersen et al., 1991). Later, the transfer of associative value between stimuli in a simultaneous discrimination was experimentally documented (Clement & Zentall, 2000; Zentall & Sherburne, 1994). Nevertheless, some experimental data suggest that value transfer might not contribute importantly to transitive responding in pigeons (Siemann, Delius, Dombrowski, & Daniel, 1996; Steirn et al., 1995; Weaver, Steirn, & Zentall, 1997). Our simulations agree with these results: Table 5 shows that both associative models converged on a small

Table 4

Simulated accuracies for the modified Siemann-Delius and Wynne models that incorporated a value transfer mechanism.

Bird		A+ B-	B+ C-	C+ D-	D+ E-	BD
Constant feedback group						
Carolina	Obtained	70.00	95.00	75.00	90.00	62.50
	Siemann-Delius	69.32	93.90	75.24	91.09	80.63
	Wynne	70.63	92.33	75.33	88.81	3.30
Kondrat	Obtained	62.50	80.00	60.00	90.00	50.00
	Siemann-Delius	62.45	79.82	60.77	90.23	37.09
	Wynne	57.28	79.60	59.55	90.75	29.81
Korsar	Obtained	80.00	55.00	37.50	95.00	20.00
	Siemann-Delius	79.95	55.55	37.19	95.89	21.79
	Wynne	78.90	55.70	35.46	97.48	1.80
Solomon	Obtained	90.00	90.00	80.00	85.00	40.00
	Siemann-Delius	87.93	83.71	89.22	87.15	49.91
	Wynne	93.93	92.63	80.51	86.76	19.96
Ordered feedback group						
Dasha	Obtained	87.50	85.00	80.00	80.00	82.50
	Siemann-Delius	87.73	84.23	79.22	80.20	62.40
	Wynne	84.20	87.01	81.90	80.04	50.98
Kotia	Obtained	80.00	80.00	67.50	55.00	85.00
	Siemann-Delius	76.11	82.49	68.30	48.17	52.27
	Wynne	84.74	81.25	69.91	52.70	21.40
Zelenaia	Obtained	70.00	90.00	62.50	50.00	82.50
	Siemann-Delius	74.37	84.68	63.01	45.90	85.45
	Wynne	74.04	90.67	60.74	53.17	72.92
Zosia	Obtained	90.00	85.00	95.00	90.00	82.50
	Siemann-Delius	97.61	86.84	89.37	89.48	29.09
	Wynne	90.36	85.29	96.17	88.50	15.65

proportion of transferred value that did not importantly change the predicted results. We thus conclude that the associative models based predominately on the reinforcement history that accrued during training cannot predict transitive responding in the ordered feedback group and the absence of transitive responding in the constant feedback group.

GENERAL DISCUSSION

In our experiment, crows selected Stimulus B of novel testing pair BD only when the choice stimuli in the transitive series were associated with explicitly ordered postchoice feedback stimuli (ordered feedback group). Without such ordered feedback stimuli (constant feedback group), crows either selected Stimuli B and D at chance levels or preferred Stimulus D, which was more often reinforced during training. In short, the orderability of the choice stimuli along a physical dimension like size may be both a necessary and sufficient condition for nonverbal transitive responding in crows.

The present study is not the first attempt to control the reinforcement history of Stimuli B and D. Zorina *et al.* (1996) compared the number of reinforced choices for Stimuli B and D and found that crows responded transitively even when the choice of Stimulus B was reinforced less often than the choice of Stimulus D during training. However, experience with nonreinforced choices of the stimuli may be just as important as experience with reinforced choices; thus, the reinforcement/nonreinforcement ratio may be a better index of associative strength than the number of reinforced and nonreinforced choices. Unfortunately, it is not possible to calculate this ratio for Zorina *et al.* because raw numbers of reinforced and unreinforced choices were not reported.

In a series of studies, Zentall and his colleagues (Steirn *et al.*, 1995; Weaver *et al.*, 1997) attempted to control their pigeons' reinforcement history by changing the probability of reinforcement of the different training stimuli. For example, Weaver *et al.* trained a group of pigeons to discriminate

Table 5

Best-fitting parameters for the original and modified Siemann-Delius and Wynne models. LSD stands for least-square difference between the proportion of correct responses predicted by the model and obtained in the experiment; β , ε , α , γ , A_+ , and A_- are the model parameters (see Appendix).

	Constant feedback group				Ordered feedback group			
	Carolina	Kondrat	Korsar	Solomon	Dasha	Kotia	Zelenaia	Zosia
Siemann-Delius								
β_+	0.008099	0.004500	0.007000	0.000001	0.000100	0.006309	0.011000	0.000010
β_-	0.009000	0.004000	0.043000	0.180000	0.069999	0.000001	0.000010	0.240000
ε	0.0699	0.0599	0.9230	0.9999	0.0010	0.0001	0.0001	0.0010
LSD	0.02652	0.00002	0.00120	0.01312	0.00328	0.09940	0.07034	0.01784
Wynne								
α	24	13	26	29	13	1	1	35
β	0.0156	0.0240	0.0080	0.6590	0.0092	0.9999	0.00001	0.023
γ	0.1	0.2	0.3	0.2	0.6	0.9	0.9	0.2
LSD	0.01713	0.01314	0.00592	0.02462	0.05640	0.16482	0.17960	0.02418
Siemann-Delius with value transfer								
β_+	0.020	0.005	0.015	0.001	0.010	0.059	0.020	0.020
β_-	0.079	0.020	0.050	0.099	0.079	0.099	0.50	0.020
ε	0.800	0.400	0.900	0.020	0.500	0.900	0.900	0.001
A_+	0.0070	0.0010	0.0090	0.0001	0.0060	0.0599	0.0110	0.0040
A_-	0.006	0.005	0.009	0.001	0.009	0.015	0.010	0.008
LSD	0.00029	0.00007	0.00004	0.01449	0.00013	0.00686	0.00644	0.00932
Wynne with value transfer								
α	6	10	4	17	7	7	9	11
β	0.010	0.030	0.005	0.040	0.070	0.060	0.080	0.099
γ	0.80	0.20	0.50	0.16	0.20	0.20	0.11	0.41
A_+	0.0200	0.0090	0.0200	0.0500	0.0999	0.0699	0.0300	0.0699
A_-	0.0400	0.0090	0.0020	0.0500	0.0080	0.0699	0.0200	0.0379
LSD	0.00091	0.00281	0.00119	0.00257	0.00186	0.00351	0.00299	0.00038

pairs $A \pm B -$, $B + C \pm$, $C \pm D -$, and $D + E \pm$ (where \pm sign denotes the stimulus that choice was reinforced in 50% of the trials). This procedure was intended to eliminate the possible effects of value transfer. Normally, Stimulus B is paired with always reinforced Stimulus A, and Stimulus D is paired with never reinforced Stimulus E. Therefore, it is expected that the cumulative transferred value will be greater for Stimulus B than for Stimulus D, thus explaining the preference of Stimulus B in the BD testing pair. More importantly, Weaver et al. also reported the mean numbers of correct and incorrect choices for all of the training stimuli. Using these numbers, we found that, for the group described above, the mean ratio of Stimulus B was equal to 3.59 and the mean ratio of Stimulus D was equal to 10.98—but the birds nevertheless selected Stimulus B over Stimulus D on 78% of the trials.

Although these findings cast doubt on re-

inforcement history as the basis of transitive responding in animals, they do not rule it out completely. Associative models based on a history of reinforcement heavily depend on the exact sequence of reinforced and nonreinforced responses in training (Couvillon & Bitterman, 1992; Wynne, 1995, 1998). Global ratios of the number of reinforced and unreinforced choices cannot substitute for formal simulations. It may be noteworthy that, in several instances, the models examined here predicted transitive responding even when the ratio of reinforced to unreinforced training choices suggested otherwise (see, e.g., data for Carolina in Tables 2 and 3). It would be interesting to see whether pigeons would respond transitively even when the simulations using the exact sequence of the trials in training predicted the opposite result. So far, such simulations have agreed with pigeons' data (see Wynne, 1995, 1997, 1998).

It might be tempting to conclude that crows solve the TI problem by using a learning mechanism different from pigeons—a mechanism that required the presence of a transitive relation between stimuli in a series. Thus pigeons may be able to respond transitively even when the stimuli are unorderable, whereas crows may require ordered feedback stimuli. Such a conclusion would be premature in the absence of data that document pigeons' preference for Stimulus B over Stimulus D when real-time simulations predict the opposite trend and when the stimuli are unorderable.

CONCLUDING COMMENTS

The failure of several associative models to account for the present data is not particularly surprising because none of these models take into account the ordering of the stimuli along a physical dimension (see Siemann & Delius, 1998; Wynne, 1995, for reviews), which appeared to play a vital role in our study. Perhaps nonverbal transitive inference can be based on a history of past reinforcement when this history is the only cue available, but also can be based on other cues when they are available? We suspect that there is more to the story than that. Markovits and Dumas (1992) suggested that, if there is no scale underlying the relation “B is reinforced, C is not reinforced” and “C is reinforced, D is not reinforced,” then organisms' performance in such experiments could not *by definition* be based on transitive inference. In response, Wynne, von Fersen, and Staddon (1992) proposed that the mere fact that their pigeons were able to select B over D in the test is sufficient for showing that they exhibited transitive inference.

It is necessary to ask: What, exactly, are transitive inference and transitivity? The term transitive inference appears to require that two conditions be met: (a) there is a transitive relation among the stimuli, and (b) the organism performs an inferential operation. Although the importance of the second condition is normally recognized, the importance of the first condition is not.

In verbal syllogisms, transitivity is the property of a *relation*. It is the transitive relation that makes transitive inference possible and valid. For instance, given the nature of the

relation *to like*, no logical conclusion follows from the premises “the boy likes the girl” and “the girl likes the dog.” The transitivity of a relation does not change with the training schedule; no amount of exposure to the first and second premises renders the conclusion “the boy likes the dog” valid. According to existing associative theories of nonverbal transitive responding, however, transitive inference is possible and valid solely because of reinforcement history. In fact, in associative accounts, a transitive inference *becomes* valid as training proceeds. When the stimuli are not orderable and untrained, there is no reason to prefer Stimulus B over Stimulus D. During training, an ordered series of stimuli emerges because the stimuli are presented in a given sequence and with a given frequency. In other words, transitivity is treated as a property of *reinforcement history*.

Although the sensitivity to different probabilities of reinforcement that led to transitive responding in the nonordered transitive task is interesting per se, we suspect that it might be largely unrelated to transitive inference in the ordered task. This does not mean that “true” transitive inference can never be revealed in a nonverbal task. Our data as well as data obtained by Roberts and Phelps (1994) demonstrated that animals' transitive response could be controlled by transitive *relations*, not by the reinforcement history of unorderable stimuli.

We do not imply, however, that an associative approach is not useful for modeling nonverbal transitive inference in ordered tasks. Dynamic, reinforcement-based models have been shown to capture different effects of TI performance (end-anchor, serial position, and symbolic distance effects) and the outcomes of different orders of training (Wynne, 1995, 1998). It might be possible to develop a reinforcement-based model that would incorporate the relation between the stimuli—after all, this relation also needs to be learned as training proceeds. Future research will see if such an attempt can be successful. In conclusion, we suggest that, for the sake of clarity, performance in the nonordered transitive task should be called “transitive responding,” as proposed by Delius and Siemann (1998). Nonordered tasks omit an essential defining feature—the transitive relation among the

stimuli in a series—and should not be confused with ordered transitive tasks.

REFERENCES

- Altmann, S. A. (1962). A field study of the sociobiology of rhesus monkeys, *Macaca mulatta*. *Annals of the New York Academy of Sciences*, 102, 338–435.
- Balda, R. P., Kamil, A. C., Bednekoff, P. A., & Hile, A. G. (1997). Species differences in spatial memory performance on a three-dimensional task. *Ethology*, 103, 47–55.
- Bond, A. B., Kamil, A. C., & Balda, R. P. (2003). Social complexity and transitive inference in corvids. *Animal Behaviour*, 65, 479–487.
- Boysen, S. T., Berntson, G. G., Shreyer, T. A., & Quigley, K. (1993). Processing of ordinality and transitivity by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 107, 208–215.
- Bryant, P., & Trabasso, T. (1971). Transitive inferences and memory in young children. *Nature*, 232, 456–458.
- Bunsey, M., & Eichenbaum, H. (1996). Conservation of hippocampal memory function in rats and humans. *Nature*, 379, 255–257.
- Clement, T. S., & Zentall, T. R. (2000). Determinants of value transfer and contrast in simultaneous discriminations by pigeons. *Animal Learning and Behavior*, 28, 195–200.
- Couvillon, P. A., & Bitterman, M. E. (1992). A conventional conditioning analysis of “transitive inference” in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 18, 308–310.
- Davis, H. (1992a). Logical transitivity in animals. In W. K. Honig & J. G. Fetterman (Eds.), *Cognitive aspects of stimulus control* (pp. 405–429). Hillsdale, NJ: Erlbaum.
- Davis, H. (1992b). Transitive inference in rats. *Journal of Comparative Psychology*, 106, 342–349.
- Delius, J. D., & Siemann, M. (1998). Transitive inference in animals and humans: Exaptation rather than adaptation? *Behavioural Processes*, 42, 107–137.
- Evans, J. St. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human reasoning: The psychology of deduction*. Hillsdale, NJ: Erlbaum.
- Gillan, D. J. (1981). Reasoning in the chimpanzee: II. Transitive inferences. *Journal of Experimental Psychology: Animal Behavior Processes*, 7, 150–164.
- Heinrich, B. (2000). Testing insight in ravens. In C. Heyes & L. Huber (Eds.), *The evolution of cognition* (pp. 289–305). Cambridge, MA: MIT Press.
- Hogue, M.-E., Beaugrand, J. P., & Lague, P. C. (1996). Coherent use of information by hens observing their former dominants defeating or being defeated by a stranger. *Behavioural Processes*, 38, 241–252.
- Jones, J. E., Antoniadis, E., Shettleworth, S. J., & Kamil, A. C. (2002). A comparative study of geometric rule learning by nutcrackers (*Nucifraga columbiana*), pigeons (*Columba livia*) and jackdaws (*Corvus monedula*). *Journal of Comparative Psychology*, 116, 350–356.
- Köhler, O. (1950). The ability of birds to “count.” *Bulletin of Animal Behavior*, 9, 41–45.
- Krushinsky, L. V. (1990). *Experimental studies of elementary reasoning: Evolutionary, physiological, and genetic aspects of behavior*. (E. Tobach & I. Poletaeva, Trans.) New Delhi, India: Oxonian Press. (Original work published in 1977).
- Kummer, H. (1982). Social knowledge in free-ranging primates. In D. R. Griffin (Ed.), *Animal mind–human mind: Report of the Dahlem workshop on animal mind–human mind, Berlin 1981* (pp. 113–130). Berlin: Springer-Verlag.
- Markovits, H., & Dumas, C. (1992). Can pigeons really make transitive inferences? *Journal of Experimental Psychology: Animal Behavior Processes*, 18, 311–312.
- McGonigle, B., & Chalmers, M. (1977). Are monkeys logical? *Nature*, 267, 694–696.
- Paz-y-Mino, G., Bond, A. B., Kamil, A. C., & Balda, R. P. (in press). Pinyon jays use transitive inference to predict social dominance. *Nature*.
- Pepperberg, I. (1987). Acquisition of the same/different concept by an African Grey parrot (*Psittacus erithacus*): Learning with respect to color, shape, and material. *Animal Learning and Behavior*, 15, 423–432.
- Rapp, P. R., Kansky, M. T., & Eichenbaum, H. (1996). Learning and memory for hierarchical relationships in the monkey: Effects of aging. *Behavioral Neuroscience*, 110, 887–897.
- Rehkämper, G., Frahm, H. D., & Mann, M. D. (2001). Evolutionary constraints of large telencephala. In G. Roth & M. F. Wulliman (Eds.), *Brain evolution and cognition* (pp. 265–296). New York: Wiley.
- Roberts, W. A., & Phelps, M. T. (1994). Transitive inference in rats: A test of the spatial coding hypothesis. *Psychological Science*, 5, 368–374.
- Siemann, M., & Delius, J. D. (1998). Algebraic learning and neural network models for transitive and non-transitive responding. *European Journal of Cognitive Psychology*, 10, 307–334.
- Siemann, M., Delius, J. D., & Wright, A. A. (1996). Transitive responding in pigeons: Influences of stimulus frequency and reinforcement history. *Behavioural Processes*, 37, 185–195.
- Smirnova, A. A., Lazareva, O. F., & Zorina, Z. A. (2000). Use of number by crows: Investigation by matching and oddity learning. *Journal of the Experimental Analysis of Behavior*, 73, 163–176.
- Steirn, J. N., Weaver, I. E., & Zentall, T. R. (1995). Transitive inference in pigeons: Simplified procedures and a test of value transfer theory. *Animal Learning and Behavior*, 23, 76–82.
- Stingelin, W. (1958). *Vergleichende morphologische Untersuchungen am Vorderhirn der Vögel auf cytologischer und cytoarchitectonischer Grundlage* [Morphological study of cytology and cytoarchitecture of bird forebrain]. Basel, Switzerland: Helbing, Lichtenhahn.
- 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.
- 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.
- Weaver, J. E., Steirn, J. N., & Zentall, T. R. (1997). Transitive inference in pigeons: Control for different value transfer. *Psychonomic Bulletin and Review*, 4, 113–117.
- Weir, A. A. S., Chappell, J., & Kacelnik, A. (2002). Shaping of hooks in New Caledonian crows. *Science*, 297, 981.

- Wilson, B., Mackintosh, N. J., & Boakes, R. (1985). Transfer of relational rules in matching and oddity learning by pigeons and corvids. *Quarterly Journal of Experimental Psychology*, *37B*, 313–332.
- Wright, B. C. (2001). Reconceptualizing the transitive inference ability: A framework for existing and future research. *Developmental Review*, *21*, 375–422.
- Wynne, C. D. L. (1995). Reinforcement accounts for transitive inference performance. *Animal Learning and Behavior*, *23*, 207–217.
- Wynne, C. D. L. (1997). Pigeon transitive inference: Tests of simple accounts of a complex performance. *Behavioral Processes*, *39*, 95–112.
- Wynne, C. D. L. (1998). A minimal model of transitive inference. In C. D. L. Wynne & J. E. R. Staddon (Eds.), *Models of action: Mechanisms for adaptive behavior* (pp. 269–307). Mahwah, NJ: Erlbaum.
- Wynne, C. D. L., von Fersen, L., & Staddon, J. E. R. (1992). Pigeons' inferences are transitive and the outcome of elementary conditioning principles: A response. *Journal of Experimental Psychology: Animal Behavior Processes*, *18*, 313–315.
- Zentall, T. R., & Sherburne, L. M. (1994). Transfer of value from S+ to S- in a simultaneous discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, *20*, 1–8.
- Zorina, Z. A. (1997). Reasoning in birds. *Physiology and General Biology Reviews*, *11*, 1–47.
- Zorina, Z. A., Kalinina, T. S., & Markina, N. V. (1996). Capacity of birds for transitive inference: The solution of the Gillan test choice by corvids and pigeons. *Neuroscience and Biobehavioral Reviews*, *26*, 454–459.

Received March 21, 2003
Final acceptance June 17, 2004

APPENDIX

A detailed description of the focal models, their advantages and their defects, can be found in the publications of Wynne (1995, 1998) and Siemann and Delius (1998).

Siemann-Delius (or eta-kappa) model. In this model, the total associative value of each stimulus is a weighted average of its elemental and configural values. Assuming that pair X+Y- is presented, the values of the stimuli are updated according to the following equations:

For elemental stimuli

$$V(X)_{i+1} = V(X)_i + \beta_+ V(X)_i p_{X/XY} \varepsilon \quad (\text{A1})$$

$$V(Y)_{i+1} = V(Y)_i - \beta_- V(Y)_i (1 - p_{X/XY}) \varepsilon \quad (\text{A2})$$

where ε is a weighting parameter for elemental and configural values.

For configural values

$$V(X/XY)_{i+1} = V(X/XY)_i + \beta_+ V(X/XY)_i p_{X/XY} \kappa \quad (\text{A3})$$

$$V(Y/XY)_{i+1} = V(Y/XY)_i - \beta_- V(Y/XY)_i (1 - p_{X/XY}) \kappa \quad (\text{A4})$$

where $\kappa = 1 - \varepsilon$, $V(X/XY)$ and $V(Y/XY)$ are configural values for Stimulus X when paired with Y and for Stimulus Y when paired with X, respectively.

The probability of selecting Stimulus X in a given pair is

$$p_{X/XY} = \frac{V(X) V(X/XY)}{V(X) V(X/XY) + V(Y) V(Y/XY)}. \quad (\text{A5})$$

For the new testing pair that has never been presented before, all configural values are zero. Thus the probability of selecting Stimulus X in a new pair XA is

$$p_{X/XA} = \frac{V(X)}{V(X) + V(A)}. \quad (\text{A6})$$

Siemann-Delius model with value transfer modification. Here we report the results of simulations that incorporate value transfer only for elemental values of the stimuli. The model allows for different proportions of positive and negative value transfer by using two different parameters.

After the elementary values of the stimuli were updated according to equations A1 and A2, the model further modifies those values

$$V(Y)_{i+1} = V(Y)_i + A_+ V(X)_i \quad (\text{A7})$$

$$V(X)_{i+1} = V(X)_i - A_- V(Y)_i \quad (\text{A8})$$

where A_+ and A_- are the parameters for positive and negative value transfer, correspondingly.

All other values and probabilities are calculated as in the original Siemann-Delius model.

Wynne model. Here stimulus value is updated according to Rescorla-Wagner equations: for elemental values

$$V(X)_{i+1} = V(X)_i + \beta(1 - [V(X)_i + V(Z)_i]) \times p_{X/XY} \quad (\text{A9})$$

$$V(Y)_{i+1} = V(Y)_i + \beta(0 - [V(Y)_i + V(Z)_i]) \times (1 - p_{X/XY}) \quad (\text{A10})$$

and for configural values

$$\begin{aligned} V(X/XY)_{i+1} &= V(X/XY)_i \\ &+ \beta[1 - V(X/XY)_i]p_{X/XY} \end{aligned} \quad (\text{A11})$$

$$\begin{aligned} V(Y/XY)_{i+1} &= V(Y/XY)_i \\ &+ \beta[0 - V(Y/XY)_i](1 - p_{X/XY}) \end{aligned} \quad (\text{A12})$$

where $V(X/XY)$ and $V(Y/XY)$ are configural values for Stimulus X when paired with Y and for Stimulus Y when paired with X , respectively.

The probability of selecting X from a given pair is

$$p_{X/XY} = \frac{1}{1 + e^{-\alpha(2r-1)}} \quad (\text{A13})$$

where r is equal to

$$\begin{aligned} r &= [V(X) + V(Z) + \gamma V(X/XY)] \\ &\div [V(X) + V(Y) + 2V(Z) \\ &+ \gamma V(X/XY) + \gamma V(Y/XY)] \end{aligned} \quad (\text{A14})$$

where γ is a new parameter weighting configural stimulus values.

The probability of selecting Stimulus X in a new, first-presented pair (e.g. XA) is

$$p_{X/XA} = \frac{1}{1 + e^{-\alpha(2r-1)}} \quad (\text{A15})$$

where r is equal to

$$r = \frac{V(X) + V(Z)}{V(X) + V(A) + 2V(Z)} \quad (\text{A16})$$

because configural values for both X and A are equal to zero.

Wynne model with value transfer. As before, only elemental values of the stimuli were subjected to value transfer, according to equations A7 and A8. All other calculations remain the same as in the original Wynne model.