Reinforcement accounts for transitive inference performance

C. D. L. WYNNE
University of Western Australia, Perth, Australia

Transitive inference is the ability, given that $A > B$ and $B > C$, to infer that $A > C$. Pigeons, rats, chimpanzees, squirrel monkeys, and humans as young as 4 years have all been shown capable of this. In this paper, simple associative learning models are explored as accounts of nonverbal transitive inference performance. A Bush-Mosteller–based model can account for transitive inference under limited conditions. A Rescorla-Wagner–based model can account for transitive inference under all conditions in the literature, but cannot account for some additional nontransitive tests. A final configural model can also account for these nontransitive data. The ability of this model to account for transitive inference formation in humans is also considered.

The transitive inference task is a form of syllogism familiar to the ancient Greeks and introduced into psychology by Cyril Burt (1911; 1919a, 1919b). Of Burt's tests for assessing the reasoning powers of 7-year-old children, two of the three he considered best for a short test were of the following form:

Tom runs faster than Jim: Jack runs slower than Jim. Who is the slowest—Jim, Jack, or Tom?

(Burt, 1919a, p. 73)

Burt maintained an interest in this task because it correlated well with other measures of intelligence. Piaget was interested in Burt's syllogisms for the light that they could shed on the development of basic intellectual faculties (Piaget, 1928). For Piaget, success on transitive inference syllogisms is part of a general ability to seriate stimuli, as well as evidence for the concrete operational stage of intellectual development (Flavell, 1963).

The ability to seriate sets of stimuli, given only partial information about the relationships between them, has far wider generality than does the simple completion of syllogisms. For social animals to be able to estimate their rank in relation to many conspecifics without entering potentially dangerous interactions with every other member of the group must often be adaptive (Cheney & Seyfarth, 1986). Similarly, the ranking of food-item pref-

erences would clearly be simplified by an ability to seriate food items in a transitive manner. Thus, from a functional point of view, the ability to solve transitive inferences may have evolved in a wider variety of species than Burt or Piaget would have expected—and a burgeoning literature testifies that this is indeed the case (e.g., Boysen, Berntson, Shreyer, & Quigley, 1993 [chimpanzees]; Davis, 1992 [rats]; Fersen, Wynne, Delius, & Staddon, 1991 [pigeons]; Gillan, 1981 [chimpanzees]; McGonigle & Chalmers, 1977, 1986, 1992 [squirrel monkeys]; Steirn, Weaver, & Zentall, 1995 [pigeons]; Wynne, 1994 [pigeons]). Given that verbal reasoning still seems outside the reach of all except the human species, we may therefore predict that the mechanisms underlying these kinds of performances can be much simpler than hitherto assumed. This paper demonstrates how a complex cognitive ability can be the result of familiar associative mechanisms. This argument will be developed with the use of data from pigeon subjects, but the logic applies equally well to other data sets from other species, including humans. Few attempts have been made to study nonverbal inferential abilities in humans, but it is quite possible that human transitive-inference performance is also generated by a simple associative mechanism. Alternatively, there may be uniquely human aspects of the solution of inference tasks by *Homo sapiens*; however, only by testing the sufficiency of simple models will it be possible to identify whether this is the case. In this way the study of animals may further the understanding of the basic processes underlying human cognition.

A relation R between two objects x and y, xRy, is transitive if, given aRb and bRc, it correctly follows that aRc. Examples are "slower than," and "cleverer than." In a nonverbal transitive inference task, pairs of stimuli are presented; the choice of one item is rewarded, and the choice of the other is not. The pattern of reward on the several pairs "overlaps" so that a series can be inferred.

At least five stimuli (four pairs) must be used in order to avoid end anchor effects. Choice on the first and last items in the series (the end anchors) is theoretically uninteresting because they are only ever rewarded and nonrewarded, respectively. The minimal task configuration, known as the "five-term series," can be represented thus:

$$A+ \quad B-$$
$$B+ \quad C-$$
$$C+ \quad D-$$
$$D+ \quad E-,$$

where a "+" after the stimulus indicates a reward if the stimulus is chosen, and a "−" indicates nonreward for a choice. The letters of the alphabet are used for convenience only; there is no physical order to the stimuli. Different colors (McGonigle & Chalmers, 1977, 1992; Gillan, 1981, with primates), odors (Davis, 1992, with rats), or visual patterns (Boysen et al., 1993, with chimpanzees; Fersen et al., 1991; Wynne, 1994, with pigeons) have been used. Training proceeds as a standard operant simultaneous discrimination. In tests for transitivity, the stimuli B and D are presented together. Choice of the stimulus B implies that the subjects have performed an operation equivalent to seriating the stimuli and making a transitive inference. With the correct controls, the stimuli B and D will have been equally often presented in rewarded and nonrewarded pairs (B was rewarded in the presence of C, and nonrewarded in the presence of A; and D was similarly rewarded with E, and nonrewarded with C), and thus a molar reinforcement account (see, e.g., Humphreys, 1975) cannot explain success on the transitivity test pair.

Two points need to be emphasized in this procedure. First, at least five stimuli must be presented. If less than five stimuli are used, no test pair can be presented that does not contain at least one of the end anchor stimuli. Second, it is important that the stimuli not be orderable along any physical dimension. If the stimuli could be ordered according to their lengths (say), then successful choice on test trials could be achieved by direct comparison of the lengths of the stimuli, without the need to infer anything about the relationships between the various training pairs.

This procedure has been used, and successful BD test performance obtained, in at least some conditions in all the studies cited above. Nonverbal studies performed with human subjects have also found transitive choice in test (Chalmers & McGonigle, 1984; Siemann & Delius, 1993; Werner, Köppl, & Delius, 1992). Transitive choice has often been demonstrated in similar studies in humans where a verbal relation between the stimuli (such as "bigger than") has been used (see Halford, 1993, for a review).

In addition to transitive test pair choice, two other phenomena have been noted repeatedly in these studies. First, when response accuracies on the various training pairs are plotted against the training pairs in the order of the implied series, a U-shaped pattern is observed. Per-

formance is best on the training pairs at the ends of the series, and it is less good on training pairs in the middle of the series (this can be seen in Figure 1). By analogy with serial list learning, this is known as the *serial position effect*, although the effect does not depend on the order in which the pairs are trained, only on their ordering within the implied series of stimuli (Wynne, 1994). A second effect reported in many studies is the *symbolic distance effect* (see, e.g., Woocher, Glass, & Holyoak, 1978). As the number of stimuli intervening in the series between the two stimuli in a test pair increases, there is a progressive increase in response accuracies and/or decrease in reaction times.

Successful transitive choice, the symbolic distance effect, and the serial position effect have been taken as evidence of complex seriation processes (e.g., McGonigle & Chalmers, 1986; Trabasso, Riley, & Wilson, 1975). Theories developed to account for human transitive inference performance, with their heavy reliance on verbal constructs, are ipso facto unsuitable for nonverbal subjects. The purpose of this paper is to develop and test the simplest model sufficient to account for all the details of transitive inference performance in animals, and the adequacy of the resulting model as an account of human performance will also be considered. First, we will consider the utility of the simplest model of the effects of reinforcement; but because this will prove unsatisfactory for explaining transitive inference under all training regimens, we will progressively add complexity to the model until we have the simplest model capable of explaining all the data to hand.

MODEL 1
Bush-Mosteller

Subjects in a nonverbal transitive inference experiment are confronted with different stimuli, some of which are rewarded when responded to, and some not. The assumption that reward increases the probability of a subsequent response and that nonreward decreases the response probability goes back at least to Thorndike (1898). Bush and Mosteller (1955) assumed that reward increments the probability of a subsequent response by a quantity dependent on the difference between the present probability of that response and the asymptote. Nonreward decrements this value in a similar way. For our purposes, this can be expressed as follows:

$$V(X)_{i+1} = V(X)_i + U\beta * (1 - V(X)_i) \text{ on reward;}$$

$$V(X)_{i+1} = V(X)_i - D\beta * V(X)_i \quad \text{on nonreward,} \quad (1)$$

where $V(X)_i$ is the value of stimulus X on trial i; $U\beta$ is a rate parameter determining the effect of a reward, and $D\beta$ is a rate parameter for the effect of nonreward. The model makes three assumptions: (1) Each stimulus acquires a unique value; (2) changes to this value occur only after choice of that stimulus; and (3) choice on each trial (training or test) is governed by the relative values of the stimuli presented. Couvillon and Bitterman (1992) pro-
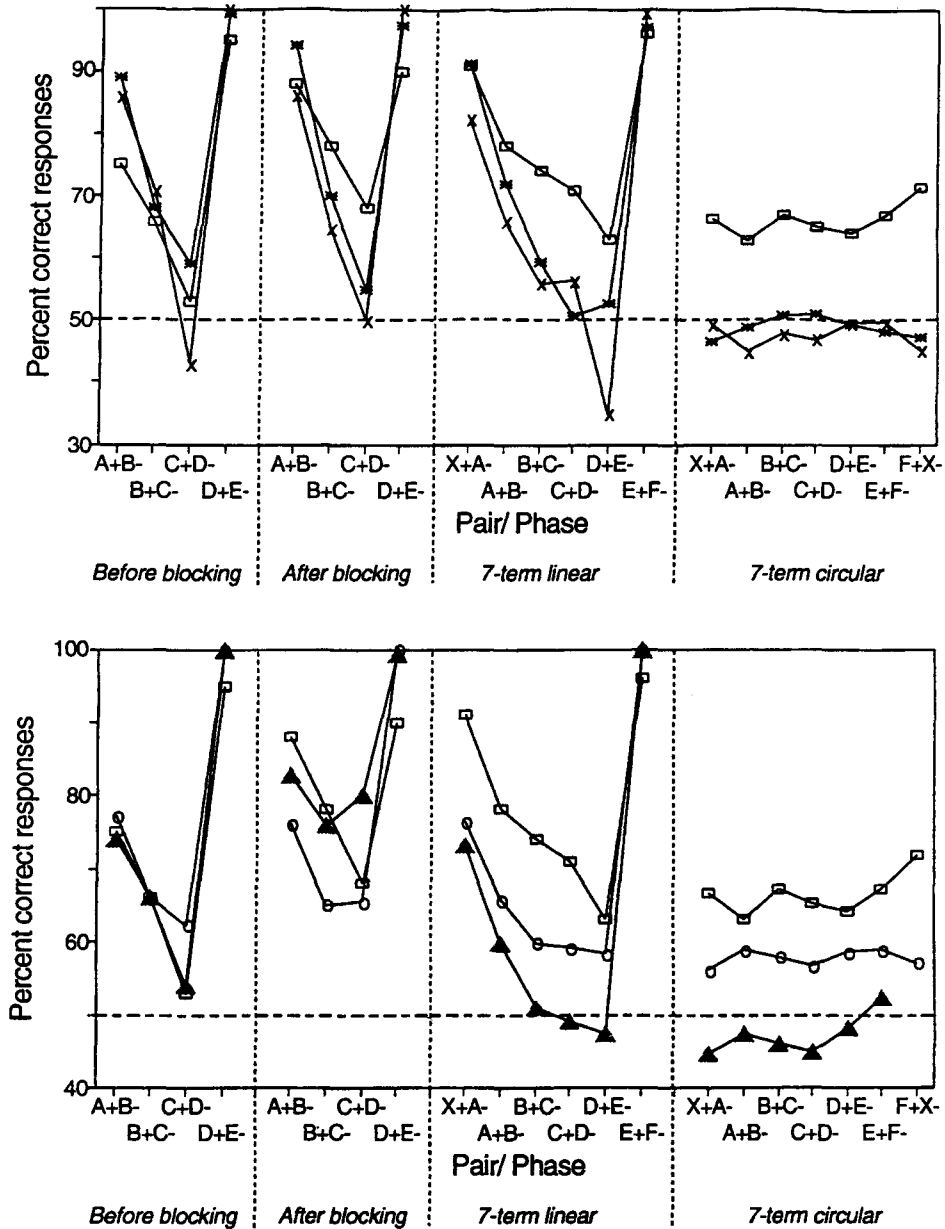
Figure 1. Percentage correct responding of Fersen, Wynne, Delius, and Staddon's (1991) pigeon subjects on training pairs in different phases of the study (open squares), compared with the results of simulations described in the text. "Before blocking" and "After blocking" refer to performance on the five-term series immediately before and after the blocking phase. "7-term linear" refers to the extended series of training, and "7-term circular" refers to that series after the ends of the series had been trained together. Top panels: simple Bush-Mosteller model (asterisks); Couvillon & Bitterman Bush-Mosteller model (crosses). Bottom panels: Model 2 (filled triangles); Model 3 (open circles).

posed that a model based on these assumptions could form transitive inferences. Simulations have shown that it is not critical how these three assumptions are implemented. In addition to the Bush-Mosteller operator defined above, some success was also obtained with simulations of Horner and Staddon's (1987) ratio invariance and constant difference models and with Luce's (1959) constant multiplier model (see Wynne, Fersen, & Stad-

don, 1992). For the present we will stay with Bush and Mosteller's linear operator, because it is analytically the simplest of these models and it forms the basis of many more recent accounts of associative learning.

On each training and test trial the values of the two stimuli are compared, and the stimulus with the most value is chosen. Again, it appears not to be critical how the stimulus values are compared—several simple pos-

sibilities lead to successful choice. Fersen et al. (1991) proposed a subtractive rule, ($p(X|XY) = V(X) - V(Y)$), but this is not always bounded between zero and one. The same drawback eliminates the ratio rule: ($p(X|XY) = V(X)/V(Y)$). A choice rule which is constrained between zero and one has been proposed by Luce (1959): ($p(X|XY) = V(X)/(V(X) + V(Y))$). For each of these choice rules, expected performance on the BD test pair can be predicted from observed performance on the $B + C-$ and $C + D-$ training pairs. When Wynne (1994) performed this analysis on a set of data from pigeons which showed wide variation in BD test performance, he found that Luce's choice rule gave the best correlation between predicted and obtained performance on BD. The model proposed by Couvillon and Bitterman (1992) takes the output from Luce's choice rule and adapts it thus:

$$r = V(X)/(V(X) + V(Y)).$$

If $r \geq 0.5$,

$$p(X|XY) = 0.5 + s(2r - 1)^k, \qquad (2A)$$

and if $r < 0.5$,

$$p(X|XY) = 0.5 - s(1 - 2r)^k, \qquad (2B)$$

where s and k are parameters fixed on the basis of previous research at 0.883 and 0.75, respectively.

Five- and Seven-Term Series Problems

Fersen et al. (1991) trained pigeons on the five-term series (diagrammed above); on a seven-term series (created by adding two more stimuli to the five-term series); and on a "circular" series of seven stimuli. The circular series was created by training the two end stimuli of the seven-term series together with their customary reward relations reversed (i.e., the always rewarded stimulus was now punished, and the never rewarded stimulus was now rewarded). The training and test pairs were presented mainly in randomized orders. This study provides a good initial test of the model, for three reasons: (1) Especially complete details of training and testing are available in Fersen (1989); as will become apparent, the predictions of the model are very sensitive to the details of training procedures. (2) The study includes additional tests which uncovered several patterns in the test pair results. (3) All subjects received the same amount and pattern of training on each stimulus pair; without this third condition, it would not be legitimate to average the performance of individual subjects for the purposes of comparison with the model.

The Bush-Mosteller model with and without the adaptation of response probabilities proposed by Couvillon and Bitterman (1992) (Equation 2 above) was exposed to the sequence of 27,000 stimulus pair presentations reported in Fersen et al. (1991; Fersen, 1989). Performance on training pairs was measured as the proportion of correct responses on the trials reported by Fersen et al. Performance on test pairs was estimated from the average stimulus values (the V) over the course of the

test phases. The small number of presentations of each test pair (typically, 12), combined with the stochastic nature of the response rule, meant that the variance in number of correct responses predicted from the model would be too high if assessed only on the 12 presentations of each test pair. All results are the average of 100 runs of the simulation. The parameters of the models did not change over the course of the study.

Training pairs. The open squares in Figure 1 show performance of the pigeons on the training pairs in each phase of each experiment reported in Fersen et al. (1991, Figures 2 and 4). Asterisks show the performance of the simple Bush-Mosteller model with best-fitting parameter values $U\beta = 0.001$; $D\beta = 0.050$. The performance of the Couvillon and Bitterman Bush-Mosteller model with best-fitting parameters ($U\beta = 0.01$; $D\beta = 0.16$) is shown by crosses.

During the first phase of training (Fersen et al., 1991, Experiment 1), a five-term series was trained as diagrammed above. Aside from a relatively short phase, during which the stimulus pairs were presented in blocks of eight trials of each type, the stimulus pairs were presented in pseudorandom orders. Performance was assessed just before and just after the blocking phase (first two panels of Figure 1).

Both models capture the typical U-shaped pattern of response accuracies on the training pairs (serial position effect). Indeed, the difference in performance between the best and worst solved training pairs is more pronounced in the performance of the models than in the observed levels of performance. This may be due to the models' more limited correction trial procedure. The models received only one correction trial for each incorrect response, whereas the pigeons continued receiving correction trials until they made a correct response. Each additional correction trial would improve the performance on a poorly solved pair.

That the last pair in the series is the best solved stimulus pair is a consequence of Luce's choice rule. Since the last stimulus in the series is never reinforced, its value will drop toward zero (for $D\beta > 0$). When the value of one stimulus is zero, the probability of choosing the other stimulus must always be one. Similarly, since the first stimulus can only gain value (for $U\beta > 0$), its value must be larger than that of any other stimulus, and performance on the first pair in the series must also be excellent. Performance on the other training pairs will generally be less good than performance on the end pairs, but the relative performance on these central pairs depends on the values of $U\beta$ and $D\beta$ and on the amount of exposure to each training pair.

The pigeons show an improvement in performance across the blocking phase—this the models capture less well.

Thereafter, the five-term series was extended to seven terms by the addition of a stimulus at each end of the series in the pairs $X + A-$ and $E + F-$ (Fersen et al., 1991, Experiment 2)—marked "7-term linear" in the figure. Here, performance of the pigeons on the six training

pairs re-forms again into an asymmetrical U shape. Now $D + E-$ (previously the best solved pairing) has become the worst solved stimulus pair. The results of the simulations capture this re-formation of the response accuracies, and also show a U shape, asymmetrical in the same direction. Predicted levels of performance on all pairs except the end pairs $(X + A-$ and $E + F-)$ are now clearly below observed levels.

Finally, in Experiment 3, the seven-term series was "closed" by training together the stimuli that had previously always been rewarded and nonrewarded respectively $(X$ and $F)$ with their reinforcement relations reversed; $F + X-$. Here the pigeons show roughly similar performance on all seven training pairs. The models also now predict similar performance on all training pairs, but both predict *indifference* on all pairs, whereas the pigeons actually showed correct choice significantly above chance.

Given the assumption of independent stimulus values, performance on the series is dependent on a ranking of those values. Clearly there is no ranking of values that could lead to successful performance after the two ends of the series have been trained together. The last panel of Figure 1 shows that although the pigeons' performance was poorer after series closure, and although differences in performance on the various stimulus pairs were no longer apparent, nonetheless the pigeons responded above chance on all training pairs. No value for the parameters in these models can fit this result. Gillan (1981), working with a single chimpanzee, has also reported that training pair performance remained above chance after training a circular series of stimulus pairs, although test pair performance fell to chance levels. Recently, Davis (1992) has shown that BD choice performance in rats was disrupted by training together the two ends of the series, although his method did not permit assessment of training pair performance at the end of training.

To summarize: Both Bush-Mosteller–based models can account for the serial position effect obtained on the training pairs of a linear series (though with an underprediction of performance on the central training pairs, possibly due to the correction procedure). These models also fit the absence of such an effect when the series is closed by training the ends together. However, they fail to predict the (slightly) above chance performance found in pigeons and a chimpanzee when the series was closed.

Test pairs. The open squares in Figure 2 show the test pair results from the linear seven-term series (Fersen et al., 1991, Experiment 2). The use of a seven-term, rather than the more usual five-term, series makes possible the analysis of two effects in the test pair results uncontaminated by end effects (a five-term series contains only one test pair [BD] which contains neither end stimulus). First, it can be seen that as the number of stimuli intervening in the series between the two test stimuli increases, there is a progressive increase in the percentage of correct responses (symbolic distance effect). Second, test pairs containing the stimulus A are in general solved better than those containing stimulus B, which in turn are solved better than those containing stimulus C. This
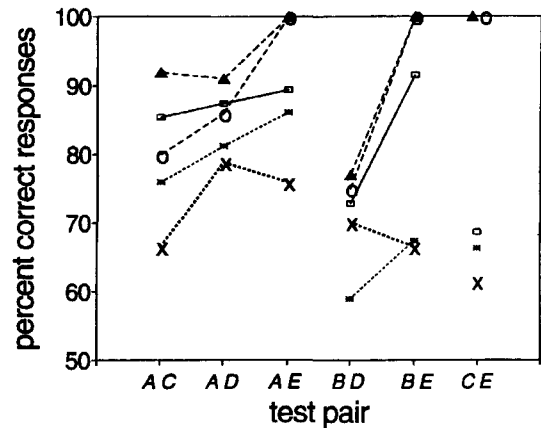

Figure 2. Test pair performance of the pigeon subjects (open squares), and of the simulations described in the text: Simple Bush-Mosteller model (asterisks); Couvillon & Bitterman Bush-Mosteller model (crosses); Model 2 (filled triangles); Model 3 (open circles).

is the *effect of first item* (Fersen et al., 1991). No other nonverbal study has reported sufficient test pair results to be able to assess these two effects, although the observation that BD test pair performance on the five-term series is sometimes better than the average of $B + C-$ and $C + D-$ training pair performance has been taken as evidence for the symbolic distance effect (see, e.g., McGonigle & Chalmers, 1986, 1992).

The predictions from the simple Bush-Mosteller model are shown as asterisks in the figure, and those from the Couvillon and Bitterman Bush-Mosteller model are shown as crosses. The simple Bush-Mosteller model predicts a clear symbolic distance effect; the Couvillon and Bitterman model fails to predict this effect. Both models underpredict performance on all pairs.

The symbolic distance effect is a direct consequence of the fact that, for successful performance on all training pairs, the stimulus values must become ordered. This value ordering implies that the farther apart two stimuli are in the series (and thus the greater the symbolic distance), the greater (on the average) will be the probability of making a correct choice.

In conclusion, the Bush-Mosteller models account for the gross features of training and test pair performance in Fersen et al.'s (1991) pigeon study. The two models do not differ greatly in their predictions. We will defer further discussion of the aspects of these data with which these models have some difficulty (particularly performance on the closed series of seven terms), and proceed to an experiment which clearly shows the inadequacy of this type of model.

Five-Term Series: Effects of Training Order

The trial-by-trial nature of the reinforcement models means that their predictions depend on the order in which training pairs are presented. Many different patterns of training have been used. In Fersen et al.'s (1991) study, the stimulus pairs were presented predominantly in ran-

domized orders. However, in most studies, at least during the early phases of training, the stimulus pairs have been presented only one at a time and in either the order of the implied series or the reverse order. Thus subjects are exposed to pair $A + B-$ until they reach some criterion, then $B + C-$, followed by $C + D-$, and finally $D + E-$. I shall refer to this as *forward series order*. Alternatively subjects may be trained in *reverse series order*: first $D + E-$ to criterion, then $C + D-$, followed by $B + C-$ and finally $A + B-$. Note that the difference between forward and reverse series training is not just a counterbalancing of the order of stimulus presentation ($A - B+$, followed by $B - C+$, $C - D+$, and $D - E+$ is reverse series training). The critical difference between the two forms of ordered training lies in the nature of the reversals of reward and nonreward experienced with each new stimulus pair. On forward series training, with the presentation of a new training pair, the previously nonrewarded stimulus becomes the rewarded stimulus (and the previously rewarded stimulus is no longer present). On backward series training, the previously rewarded stimulus becomes nonrewarded, and the previously nonrewarded stimulus disappears.

In studies of nonhuman primates, training pairs have been presented in reverse series order during the initial phases of training. Thus Boysen et al., (1993, chimpanzees), Gillan, (1981, chimpanzees), and McGonigle and Chalmers, (1977, squirrel monkeys) all used this method initially, followed by training of all pairs intermixed, and obtained transitive choice in test.

Many studies with human subjects have also used forward or backward training patterns. In comparisons of forward, backward, and randomized training paradigms, Kallio (1982), DeBoyssen-Bardies and O'Regan (1973), and Halford and Kelly (1984) have found that under certain conditions, randomized training is less effective than ordered training; differences between forward and backward training, however, have not been reported (see Wynne, in press, for a review).

Steirn et al. (1995) reported that pigeons failed to learn the task unless forward or backward series training was used. Davis (1992) also found that an ordered pattern of training pair presentation was necessary before rats could learn the five-term series. Both these studies, like the human studies cited above, reported no difference in performance on forward and backward series training.

In pigeons, training with the stimulus pairs in either forward or reverse series order has been found to yield relatively rapid learning of all four training pairs. Steirn et al. (1995) trained pigeon subjects for three 96-trial sessions on each training pair and found that this amount of training was more than sufficient for a high level of performance on each training pair. Subsequent choice on BD test trials was highly transitive. This total of 1,248 trials for all four training pairs contrasts with the performance of pigeons trained on the stimulus pairs predominantly in randomized orders (Fersen et al.'s, 1991, subjects required over 5,000 trials [plus correction

trials] to learn a five-term series; Wynne's, 1994, subjects required over 17,000 trials).

Thus, although within-study and within-subject controls are still lacking, there is a growing body of evidence that the transitive inference task is easier to train (more rapidly learned to a higher level of accuracy) if the training pairs are presented singly in each training session and in either the order of the series or its reverse.

Figure 3 shows BD test pair results from Steirn et al. (1995), alongside predictions from the simple Bush-Mosteller model and the Couvillon and Bitterman Bush-Mosteller model. Since Steirn et al. also used pigeon subjects with similar operant contingencies, the same parameter values were used for the simple Bush-Mosteller model as in the simulation of the Fersen et al. (1991) study. However, a systematic search of the parameter space of the Couvillon and Bitterman model failed to find any pair of values for $U\beta$ and $D\beta$ which could fit both Fersen et al.'s study and any part of Steirn et al.'s results. Therefore, the parameter values used for the results of the Couvillon and Bitterman model shown in Figure 3 are $U\beta = 0.00001$, $D\beta = 0.10$. Both models successfully reproduce the results of the *forward* group, but perform only at chance level after training on the *backward* condition. Pigeon performance on BD is similar for both groups (as is the performance of humans under similar training conditions; Wynne, in press). Although it has not been possible to prove analytically that these models cannot in principle account for transitive choice after backward series training, a thorough search of the parameter space has failed to uncover any values of $U\beta$ and $D\beta$ for which either of the Bush-Mosteller models could successfully predict choice of B over D.

Why should the model only be able to learn the transitive inference task if the stimulus pairs are presented in forward series order, rather than the reverse? The answer lies in the parameters chosen to fit the Fersen et al. (1991) and the Steirn et al. (1995) forward group data sets. The parameters that produce a good fit to these results assume that nonreward has a far greater impact on performance than reward has. Thus, when a forward group is trained, first stimulus B loses some value in $A + B-$ training. Then, when $B + C-$ training commences, stimulus C loses value until its value is ranked below that of stimulus B (which is already lower than that of stimulus A). This process iterates on the $C + D-$ and $D + E-$ training trials, until at the end of training the stimulus values are ranked with stimulus A having the most value, and stimulus E, the least. The same process operates more slowly when the stimuli are presented in a random order. When the training pairs are presented in reverse series order, first stimulus E loses value on $D + E-$ trials. On $C + D-$ trials, stimulus D will lose value. Stimulus D will stabilize with approximately the same amount of value as that of stimulus E. This process iterates on $B + C-$ and $A + B-$ trials, so that stimulus B will end with the same amount of value as that of stimuli C, D, and E. Only stimulus A maintains a higher level than do the other four stimuli.
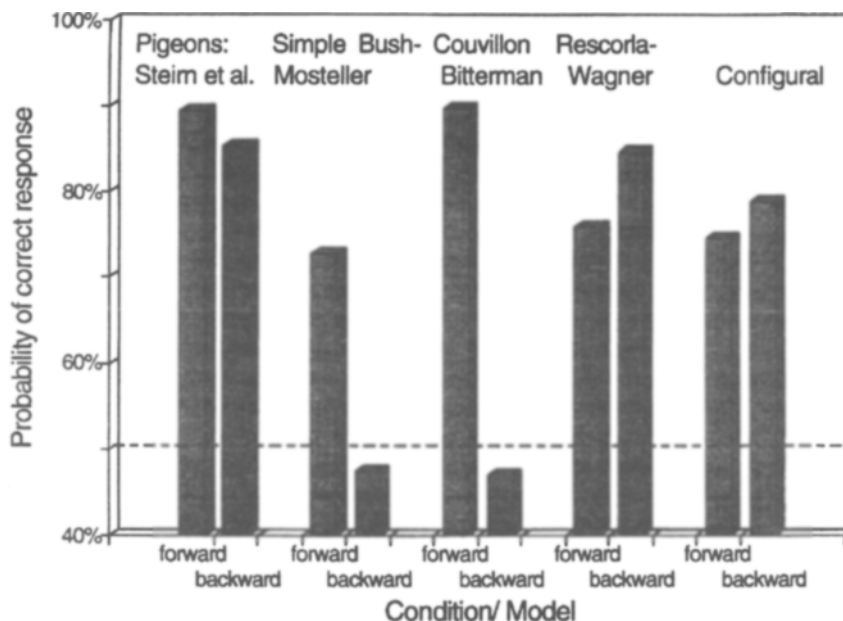
Figure 3. Performance on the BD test pair of Steirn, Weaver, and Zentall's (1995) pigeon subjects after forward and backward series training, and predicted performance levels from the four models described in the text.

MODEL 2
Rescorla-Wagner

The Bush and Mosteller assumption of completely independent stimulus values was challenged in 1972 by Rescorla and Wagner. These authors retained the Bush-Mosteller linear operator, but assumed that stimuli compete for a limited amount of stimulus value. The increment in the strength of a stimulus is due, not just to the difference between *that* stimulus' value and asymptote, but between the combined value of all stimuli present and asymptote. The assumption appeared necessary on the basis of results from experiments in classical conditioning (particularly those of Kamin, 1969), and it has also proven useful in accounting for data from operant tasks (see, e.g., Daly & Daly, 1982, 1984). Wagner and Rescorla (1972) proposed that their model could be applied to operant discrimination tasks by assuming that every time a stimulus (say) X is presented, the increment in that stimulus' value depends on the difference between asymptote and the combination of X's value and the value of whatever is present whenever any stimulus is chosen (Z: this includes the context and whatever else the different stimuli have in common). Thus, the update rules for the stimulus values become

$$V(X)_{i+1} = V(X)_i + \beta * (1 - [V(X)_i + V(Z)_i]) \text{ on reward,}$$

$$V(X)_{i+1} = V(X)_i - \beta * [V(X)_i + V(Z)_i] \text{ on nonreward. (3)}$$

Note that there is now only one growth parameter, β.

Values in this formation are no longer constrained between zero and one, and therefore a scaling parameter

(α) must be added to Luce's choice rule to derive response probabilities.

$$r = \frac{V(X) + V(Z)}{V(X) + V(Y) + 2V(Z)} \quad (4)$$

$$p(X|XY) = \frac{1}{1 + e^{-\alpha(2r-1)}} \quad (5)$$

With the two growth parameters reduced to one, the number of free parameters remains at two.

Five- and Seven-Term Series Problems

Performance of the Rescorla-Wagner–based model on the training pairs of the five- and seven-term series is included in Figure 1 (bottom panel). The parameter values used were $\beta = 0.20$ and $\alpha = 11$. It can be seen that the model produces a very similar fit to these data as that found with the Bush-Mosteller models. In particular, the poorer than observed performance on the central pairs of the seven-term linear series and the chance performance on the seven-term circular series is predicted by both models.

The test pair performance of this model is included in Figure 2. The fit shown by this model is somewhat better than was found with the Bush-Mosteller models (particularly on pairs BD and BE), though it is less accurate on the CE test pair.

Five-Term Series: Effects of Training Order

Having shown that the model produces a good fit to the data of Fersen et al. (1991), we ask how well it can

cope when the stimulus pairs are trained in forward and reverse series order.

Figure 3 compares the predictions for BD test performance of the two models under consideration here with the pigeon data of Steirn et al. (1995). It can be seen that, given the same parameter values as those used to fit the Fersen et al. (1991) data, the Rescorla-Wagner model predicts choice of B over D, no matter the direction in which the series was trained. Thus this model clearly gives a better fit to Steirn et al.'s data than do the Bush-Mosteller models. In addition, the Rescorla-Wagner model does not make the assumption that nonreward and reward have differential impacts on learning.

At this point, our original aim has been achieved. The Rescorla-Wagner–based model with two fixed parameters is sufficient to account for choice of B over D after training of a five- (or longer) term series—transitive inference performance. The model accounts for transitive choice after randomized or ordered presentation of the training pairs, and for the serial position and symbolic distance effects. There remains, however, an aspect of the performance of Fersen et al.'s (1991) pigeons (also observed in a chimpanzee by Gillan, 1981), for which the model cannot account. After the linear series was trained into a circular one by training together the two end-anchor stimuli with their reward relationships reversed (i.e., $F+X-$), it was found that performance on test pairs fell to chance levels, but performance on the training pairs remained above chance (see Figure 1, rightmost panel). That test pair performance fell to chance levels is logical (if Tom is faster than Jim, Jim faster than Jack, and Jack faster than Tom, it cannot be said who is the fastest), and it is predicted by both types of model considered thus far. However, neither model can account for the better-than-chance performance on the training pairs. We therefore consider next a model that can also account for this.

MODEL 3
Configural Cues

The last panel in Figure 1 shows the results from the final phase of training in Fersen et al. (1991). In this final phase the seven stimuli were trained in seven stimulus pairs which formed a logical circle ($X+A-$, $A+B-$, $B+C-$, $C+D-$, $D+E-$, $E+F-$, and $F+X-$). Subjects no longer performed above chance on any test pairs, and the typical U-shaped serial position effect in the training pair performances disappeared; but 2 (out of 4) subjects did perform above chance on all training pairs. Both the Bush-Mosteller and the Rescorla-Wagner-based models were correctly able to predict the absence of transitive choice on test pairs, but neither model could account for the above-chance performance of some subjects on training pairs.

Up till now, we have been assuming that a stimulus remains distinct whatever context it is presented in. Thus stimulus B (say) gains value when rewarded on $B+C-$ trials, and it loses value when punished on $A+B-$ trials—

stimulus B is always the same thing, no matter the context in which it is presented. Couvillon and Bitterman (1992) proposed that the introduction of configural stimulus values (after Rescorla, 1972, 1973, and Whitlow & Wagner, 1972) could enable a model to learn a circular series. Therefore, the stimulus B that gains value when rewarded on $B+C-$ trials is not entirely the same thing as the stimulus B that loses value on $A+B-$ trials. We will denote the value of the former $\langle B|BC\rangle$ and the latter $\langle B|AB\rangle$. These configural stimulus values are completely independent of each other, and they are distinct from the original context-independent elemental stimulus values, A, B, C, etc. If we consider the transitive inference task solely with these configural stimulus values, we can see that it makes no difference now that the task involves presenting the same stimuli in different pairings; for a purely configural account, each time a stimulus appears in a new context it is a completely new stimulus. For this reason, a model based on configural stimulus values can learn any pattern of training pairs. Even the circular pattern of training presented in the last phase of the Fersen et al. (1991) study poses no more difficulties to this model than a set of unrelated discrimination tasks. Indeed successful performance on a task similar to this one, the transverse patterning problem ($A+B-$, $B+C-$, and $C+A-$), has been used as evidence for the production of configural stimulus values of this type (Alvarado & Rudy, 1992). A purely configural model, however, can never solve test pairs. These pairs are, by definition, combinations of stimuli that have not been trained together, and therefore configural stimulus values cannot have developed for them. Therefore, this model maintains the existence of the elemental stimulus values from its predecessors, but adds in configural stimulus values to a degree controlled by a new parameter, γ.

For comparability with its predecessors, as much as possible from the previous model was left unchanged. The update rules for the stimulus values are as those for the Rescorla-Wagner model, but now they are also applied (independently) to configural values. Thus, the elemental values update as before:

$$V(X)_{i+1} = V(X)_i + \beta * (1 - [V(X)_i + V(Z)_i]) \text{ on reward,}$$

$$V(X)_{i+1} = V(X)_i - \beta * [V(X)_i + V(Z)_i] \text{ on nonreward, } (3)$$

With the same principle working on the configural values (and sharing the same β parameter),

$$V(\langle X|XY\rangle)_{i+1} = V(\langle X|XY\rangle)_i + \beta * (1 - V(\langle X|XY\rangle)_i)$$

$$\text{on reward,}$$

$$V(\langle X|XY\rangle)_{i+1} = V(\langle X|XY\rangle)_i - \beta * [V(\langle X|XY\rangle)_i]$$

$$\text{on nonreward. } (6)$$

The values are combined according to Luce's choice rule, but with the configural stimulus values weighted by a quantity γ ($\gamma > 0$):

$$r = \frac{V(X) + V(Z) + \gamma V(\langle X|XY\rangle)}{V(X) + V(Y) + 2V(Z) + \gamma V(\langle X|XY\rangle) + \gamma V(\langle Y|XY\rangle)}. (7)$$

The r values are scaled as before:

$$p(X|XY) = \frac{1}{1 + e^{-\alpha(2r-1)}}. \qquad (5)$$

It can be seen that this model has only one more parameter, γ, than the previous two models. When γ is sufficiently small, the behavior of this model becomes indistinguishable from that of the Rescorla-Wagner model. When γ is very large, the model gains the ability to solve any set of training pairs, but loses the ability to solve any test pairs.

Alvarado and Rudy (1992) have demonstrated (in rats) that configural stimulus values develop even when the task demands do not require them if the task contains some stimuli that are sometimes rewarded and sometimes not. We will assume, therefore, that the parameter weighting the configural stimulus values, γ, is constant throughout training.

Five- and Seven-Term Series Problems

In order to constrain the degrees of freedom in this model, and to test solely the effect of the configural stimulus values, the two parameters that this model has in common with Model 2 (Rescorla-Wagner) were fixed here at the values found optimal for that model: $\beta = 0.20$, $\alpha = 11$. Only the parameter γ was free to vary.

Figure 1 includes the mean results from a simulation of this model with the β and α parameters as for the Rescorla-Wagner model, and with $\gamma = 0.10$. In the first two panels of the figure, before and after blocking on the five-term series, the new model's performance is very similar to that of the previous two models and the pigeons. The third panel, performance on the seven-term linear series, shows an improvement relative to the other two models. Where both the Bush-Mosteller and Rescorla-Wagner models performed around chance levels on the "central" stimulus pairs, $B+C-$, $C+D-$, and $D+E-$, the configural model produces better performance, more similar to that of the pigeon subjects. The final phase of this figure, as expected, shows the most dramatic improvement over the other two models. On the seven-term circular condition, the Configural model is capable of producing above chance performance, with no serial position effect. Performance levels even closer to those shown by the pigeon subjects were obtained in an unconstrained search of the parameter space at $\beta = 0.15$, $\alpha = 13$, $\gamma = 0.06$. Thus this model is a highly successful fit to these pigeon data, which are similar in form to those obtained from a chimpanzee by Gillan (1981).

Figure 2 shows performance of the three models on test pairs: The configural model fits the data about as well as the Rescorla-Wagner model, and somewhat better than the Bush-Mosteller models.

Thus this model, considering the combined effects of elemental stimulus values with configural stimulus values, produces a good fit to all aspects of the transitive inference study of Fersen et al. (1991), including the final nontransitive phase of training in which the stimulus pairs formed a closed, circular series. Now we may consider how this model copes with forward and backward series training.

Five-Term Series: Effects of Training Order

Figure 3 includes predicted choice of stimulus B on nondifferentially reinforced BD test trials after training in forward and backward series order. The results of the Bush-Mosteller and Rescorla-Wagner models are compared with the predictions of the configural model with the parameter values from the preceding simulation of the Fersen et al. data ($\beta = 0.20$, $\alpha = 11$, $\gamma = 0.10$). Predicted performance on BD test trials after 1,200 trials of training forward or backward through the stimulus series is transitive at levels over 70%.

Thus we may terminate our exploration of models at this point, having found in the configural model with Rescorla-Wagner update rule an account for transitive choice after training both on randomized orders and on the training pairs in series order. The model also predicts above-chance performance on training pairs on a nontransitive circular series of stimuli.

DISCUSSION

Relation to Other Theories

Fersen et al. (1991) outlined an account of transitive inference performance which has some points of contact with the models presented here. According to value transfer theory, stimuli gain value through their histories of reward and nonreward, and choice on training and test trials is governed by the relative values of the stimuli presented. The model differs from those presented here both in the manner in which stimuli gain value, and in how the values are compared to produce choice. Value transfer theory assumes that stimuli gain a certain amount of value because they are themselves rewarded on some occasions, but that they receive in addition a proportion of the value of the rewarded stimulus with which they are presented. Thus, stimulus B, for example, receives a certain amount of value when rewarded on $B+C-$ trials, but also a (smaller) additional amount of value when presented on $A+B-$ trials—a portion of stimulus A's value. Choice is governed by the difference in value between the two stimuli presented.

Models 1 and 2 here (and Model 3 insofar as it subsumes in part Model 2) propose a mechanism for an effect similar to value transfer. According to these models, stimulus B (say) ends up with more value than stimulus C, not because of a direct transfer of value from stimulus A to stimulus B, but through an indirect mechanism. Stimulus A (being always rewarded) will gain value more rapidly than will stimulus B (which is only partially rewarded). As stimulus A gains value, the probability of a correct choice on $A+B-$ trials will increase rapidly—this protects stimulus B from the loss of value that would occur if it were chosen incorrectly on these trials. On $B+C-$ trials, stimulus B will gain value and thereby protect stimulus C from erroneous choice and

consequent loss of value. However, stimulus B's value gain will necessarily be slower than that of stimulus A, and therefore erroneous responses to stimulus C on $B+C-$ trials will be more common than erroneous responses to stimulus B on $B+C-$ trials, leading to stimulus C losing more value than stimulus B. Thus, these models predict a kind of "protection from value loss" which a higher ranked stimulus bestows on the stimulus with which it is presented during training, rather than the direct "value transfer" postulated by Fersen et al. (1991).

Most accounts of transitive inference performance assume that subjects order representations of the stimuli into a mental line (Banks, 1977; Breslow, 1981; Sternberg, 1980; Trabasso & Riley, 1975; Trabasso et al., 1975). Subjects may order training pair information into an ordered semantic framework or a quasispatial mental representation. Thus stimulus A becomes associated with one end of the mental line and stimulus E marks the other end of the mental line. The ends of the series are located first, because they are always positive or negative. The other stimuli gradually become ordered between the ends. Once the pairs have been trained better, choice on training and test trials is governed by a search along the mental line. The symbolic distance effect is given by the fact that, on the average, the farther apart two items are in the series, the faster the one or the other of them will be found in an ordered search of the mental representation. The serial position effect arises because the ends of the series are unambiguous and training pairs containing them therefore generate fewer errors.

Without an explicit theory of how mental lines are constructed, it is impossible to predict performance under different conditions. What does a mental line predict about performance on a circular training series (as in Fersen et al., Experiment 3), for example? Can a mental line become a mental circle?

The models presented here can be considered as possible implementations of the processes underlying a mental line. The ordering of a mental line is provided by stimulus values which become ordered during training. This ordering is produced by the update rules proposed here. The function of scanning the mental line is fulfilled by the choice rules presented above (e.g., Equations 3 and 4).

Possibly the authors of mental line models of transitive inference performance will see the functions of creating and scanning a mental line performed by quite different mechanisms. I hope that this paper will motivate them to postulate these mechanisms explicitly.

Inter- and Intraspecies Comparisons

Most studies of transitive inference have reported results in terms of the presence or absence of three things. (1) transitive choice on BD tests; (2) the serial position effect; and (3) the symbolic distance effect. These effects have been observed in all species tested, including humans as young as 4 years old, under a range of testing conditions. The models developed here can readily account for the symbolic distance and serial position ef-

fects, as well as predict transitive choice—thus they are potential explanations of transitive inference under a wide range of conditions and in a range of species (Wynne, in press).

A unique feature of inferential reasoning in humans is that the task can be presented verbally. This fact has motivated accounts of transitive inference in terms of linguistic representations. However, the similarities between the results from humans under verbal and nonverbal conditions on the one hand, and the similarities between human and nonhuman results on the other, suggest that linguistic representations may not be fundamental to task solution in humans (Wynne, in press).

Conclusions

The models presented here were not designed to solve transitive inference tasks. They are simply theories of conditioning, derived from familiar assumptions about the effects of reward and punishment. Therefore, this is not a novel theory designed post hoc to fit a new set of data, but rather a demonstration that the principles developed in one realm are applicable in a completely different one. More recent models of associative conditioning (e.g., Pearce & Hall, 1980) maintain the basic assumptions of competitive stimulus strengths from the Rescorla-Wagner formulation and thus would also predict successful performance here.

To show that transitive inference performance can be generated from simple associative models is to demonstrate that a complex conditional inference process can be modeled by processes which themselves are neither conditional nor inferential. This result has wide implications—it is an instantiation of the principle that task complexity is no proof that the underlying processes involved have to be complex (Braitenberg, 1984).

Models of this type have the advantage over other theories of transitive inference that they do not simply assume training pair performance and predict test pair choice on that basis. Rather, they make trial-by-trial predictions for choice on training and test pairs. Since training pair performance is not always observed at uniformly high levels, this makes these models potentially more realistic and also expands the database against which they can be tested.

The question of whether there are aspects of inferential reasoning in humans or other species that cannot be explained with models of the type developed here has to be an empirical one. I hope that this paper will motivate students of comparative cognition to develop more ingenious tests to clarify this issue.

REFERENCES

ALVARADO, M., & RUDY, J. W. (1992). Some properties of configural learning: An investigation of the transverse-patterning problem. *Journal of Experimental Psychology: Animal Behavior Processes*, **18**, 145-153.

BANKS, W. P. (1977). Encoding and processing of symbolic information in comparative judgments. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 11, pp. 101-159). New York: Academic Press.

BOYSEN, S. T., BERNTSON, G. G., SHREYER, T. A., & QUIGLEY, K. S. (1993). Processing of ordinality and transitivity by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, **107**, 1-8.

BRAITENBERG, V. (1984). *Vehicles: Experiments in synthetic psychology*. Cambridge, MA: MIT Press.

BRESLOW, L. (1981). Re-evaluation of the literature on the development of transitive inference. *Psychological Bulletin*, **89**, 325-351.

BURT, C. (1911). Experimental tests of higher mental processes and their relation to general intelligence. *Journal of Experimental Pedagogy*, **1**, 93-112.

BURT, C. (1919a). The development of reasoning in school children: I. *Journal of Experimental Pedagogy*, **5**, 68-77.

BURT, C. (1919b). The development of reasoning in school children: II. *Journal of Experimental Pedagogy*, **5**, 121-127.

BUSH, R. R., & MOSTELLER, F. (1955). *Stochastic models for learning*. New York: Wiley.

CHALMERS, M., & McGONIGLE, B. O. (1984). Are children any more logical than monkeys on the five-term series problem? *Journal of Experimental Child Psychology*, **37**, 355-377.

CHENEY, D. L., & SEYFARTH, R. M. (1986). The recognition of social alliances by vervet monkeys. *Animal Behavior*, **34**, 1722-1731.

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.

DALY, H. B., & DALY, J. T. (1982). A mathematical model of reward and aversive nonreward: Its application in over 30 appetitive learning situations. *Journal of Experimental Psychology: General*, **111**, 441-480.

DALY, H. B., & DALY, J. T. (1984). DMOD—A mathematical model of reward and aversive nonreward in appetitive learning situations: Program and instruction manual. *Behavior Research Methods, Instruments, & Computers*, **16**, 38-52.

DAVIS, H. (1992). Transitive inference in rats (*Rattus norvegicus*). *Journal of Comparative Psychology*, **106**, 342-349.

DE BOYSSON-BARDIES, B., & O'REGAN, K. (1973). What children do in spite of adults' hypotheses. *Nature*, **246**, 531-534.

FERSEN, L. VON (1989). *Kognitive Prozesse bei Tauben (Columba livia)* [Cognitive processes in pigeons]. Pfaffenweiler, Germany: Centaurus-Verlagsgesellschaft.

FERSEN, L. VON, 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.

FLAVELL, J. (1963). *The developmental psychology of Jean Piaget*. Princeton, NJ: Van Nostrand.

GILLAN, D. J. (1981). Reasoning in the chimpanzee: II. Transitive inference. *Journal of Experimental Psychology: Animal Behavior Processes*, **7**, 150-164.

HALFORD, G. (1993). *Children's understanding: The development of mental models*. Hillsdale, NJ: Erlbaum.

HALFORD, G., & KELLY, M. E. (1984). On the basis of early transitivity judgments. *Journal of Experimental Child Psychology*, **38**, 42-63.

HORNER, J. M., & STADDON, J. E. R. (1987). Probabilistic choice: A simple invariance. *Behavioral Processes*, **15**, 59-92.

HUMPHREYS, M. S. (1975). The derivation of endpoint and distance effects in linear orderings from frequency information. *Journal of Verbal Learning & Verbal Behavior*, **14**, 496-505.

KALLIO, K. D. (1982). Developmental change on a five-term transitive inference. *Journal of Experimental Child Psychology*, **33**, 142-164.

KAMIN, L. J. (1969). Selective association and conditioning. In N. J. Mackintosh & W. K. Honig (Eds.), *Fundamental issues in associative conditioning* (pp. 42-64). Halifax, NS: Dalhousie University Press.

LUCE, R. D. (1959). *Individual choice behavior: A theoretical analysis*. New York: Wiley.

McGONIGLE, B. O., & CHALMERS, M. (1977). Are monkeys logical? *Nature*, **267**, 694-696.

McGONIGLE, B. O., & CHALMERS, M. (1986). Representation and strategies during inference. In T. Myers, K. Brown, & B. McGonigle (Eds.), *Reasoning and discourse processes* (pp. 141-164). London: Academic Press.

McGONIGLE, B. O., & CHALMERS, M. (1992). Monkeys are rational! *Quarterly Journal of Experimental Psychology*, **45B**, 189-228.

PEARCE, J. M., & HALL, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, **87**, 532-552.

PIAGET, J. (1928). *Judgement and reasoning in the child*. New York: Harcourt, Brace.

RESCORLA, R. A. (1972). "Configural" conditioning in discrete-trial bar pressing. *Journal of Comparative & Physiological Psychology*, **79**, 307-317.

RESCORLA, R. A. (1973). Evidence for "unique stimulus" account of configural conditioning. *Journal of Comparative & Physiological Psychology*, **85**, 331-338.

SIEMANN, M., & DELIUS, J. D. (1993). Implicit deductive reasoning in humans. *Naturwissenschaften*, **80**, 364-366.

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**, 76-78.

STERNBERG, R. J. (1980). Representation and process in linear syllogistic reasoning. *Journal of Experimental Psychology: General*, **109**, 119-159.

THORNDIKE, E. L. (1898). Animal intelligence: An experimental study of the association processes in animals. *Psychological Review Monographs*, **2** (Whole No. 8).

TRABASSO, T., & RILEY, C. A. (1975). On the construction and use of representations involving linear order. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 381-410). Hillsdale, NJ: Erlbaum.

TRABASSO, T., RILEY, C. A., & WILSON, E. G. (1975). The representation of linear order and spatial strategies in reasoning. In R. Falmagne (Ed.), *Reasoning, representation and process* (pp. 201-229). Hillsdale, NJ: Erlbaum.

WAGNER, A. R., & RESCORLA, R. A. (1972). Inhibition in Pavlovian conditioning: Application of a theory. In R. A. Boakes & M. S. Halliday (Eds.), *Inhibition and learning* (pp. 301-306). London: Academic Press.

WERNER, U. B., KÖPPL, U., & DELIUS, J. D. (1992). Transitive Inferenz bei nicht-verbaler Aufgabendarbeitung [Transitive inference in nonverbal task presentation]. *Zeitschrift für experimentelle und angewandte Psychologie*, **39**, 662-683.

WHITLOW, J. W., JR., & WAGNER, A. R. (1972). Negative patterning in classical conditioning: Summation of response tendencies to isolable and configural components. *Psychonomic Science*, **27**, 299-301.

WOOCHER, F. D., GLASS, A. L., & HOLYOAK, K. J. (1978). Positional discriminability in linear orderings. *Memory & Cognition*, **6**, 165-173.

WYNNE, C. D. L. (1994). *Pigeon transitive inference: Tests of simple accounts of a complex performance*. Manuscript submitted for publication.

WYNNE, C. D. L. (in press). A minimal model of transitive inference. In C. D. L. Wynne & J. E. R. Staddon (Eds.), *Models of action*. Hillsdale, NJ: Erlbaum.

WYNNE, C. D. L., FERSEN, L. VON, & STADDON, J. E. R. (1992). Pigeons' inferences are transitive and are the outcome of elementary conditioning principles: A response. *Journal of Experimental Psychology: Animal Behavior Principles*, **18**, 313-315.