Journal of Comparative Psychology
1993, Vol. 107, No. 2, 208–215


0735-7036/93/$3.00

# Processing of Ordinality and Transitivity by Chimpanzees (*Pan troglodytes*)

Sarah T. Boysen, Gary G. Berntson, Traci A. Shreyer, and Karen S. Quigley

Three chimpanzees (*Pan troglodytes*) were trained to discriminate among pairs of boxes in an ABCDE-ordered series. The 2nd member of each pair was reinforced, until all 4 training pairs were learned. During novel tests the nonadjacent BD pair was presented, and all 3 animals reliably selected D. In Experiment 2, numerals 1–5 served as stimuli. One chimpanzee reliably selected the larger numeral 4 during testing with a nonadjacent pair (2–4), and 2 chimps showed no preference. In a 2nd phase, the same chimp demonstrated proficiency at reversing the task, reliably selecting the smaller of the 2–4 pair. In Experiment 4, after additional training, a 2nd test, which included novel test pairs composed of numbers that had not been used during training, was completed. Two of 3 animals were 100% correct on Trial 1 for all novel pairs. The results suggest that chimpanzees with experience in number concepts may recognize the ordinal character of numbers.



The study of logical processes in human cognition has been of great historical interest to a variety of subdisciplines within psychology (e.g., Burt, 1919; James, 1890). The investigation of related processes in nonhuman species has held a similar interest for investigators in comparative cognition (Matsuzawa, 1985; Premack, 1976; Roitblat, Bever, & Terrace, 1984). Among other processes recent studies have addressed the capacity for solving analogy problems in the chimpanzee (Gillan, 1981) and the ability to use transitive inference in monkeys (McGonigle & Chalmers, 1977) and chimpanzees (Gillan, 1981).

Transitive inference is an inferential judgment of the ordinal relation between two elements, derived from premises that specify the relation of each of two elements to a third (Halford, 1984; Kingma & Zumbo, 1987). As typically presented to children, the task consists of presentation or training of a series of object pairs, such as sticks of different colors and lengths, with pairs that serve as premises (e.g., A > B, B > C, C > D, and D > E). Under test conditions children are asked to identify the correct member of a novel, nonadjacent pair, such as BD. Because no nonadjacent pairs are trained, subjects must derive the relation between the two objects from their relative position within the ordered series. It has been proposed that under certain experimental conditions task solution depends on the use of transitive inference. Since the observations of Piaget, Inhelder, and Szeminska (1960), it has been widely assumed that young children who have not reached the concrete operational stage (approximately age 7) were unable to draw transitive inferences. Although the original inferential problem was first reported by Alfred Binet and predated Piaget's studies (Elkind, 1974), the question of age-related inferential capacities remains controversial (Chapman & Lindenberger, 1988; Halford & Kelly, 1984). Bryant and Trabasso (1971) asserted that children as young as 4 can demonstrate the use of transitive inference. They argued that previous failures were not due to children's inability to reason but, rather, their inability to remember the premises. Critics of Trabasso and his colleagues have proposed that his version of the task could be solved by nontransitive strategies (de Boysson-Bardies & O'Regan, 1973; Youniss & Furth, 1973). Currently, the debate continues and includes the nature of the representation of an ordered series, the age at which children can solve such tasks, and the cognitive processes that might subserve inferential abilities (Breslow, 1981; Chapman & Lindenberger, 1988; Halford & Kelly, 1984; Riley & Trabasso, 1974; Trabasso, 1975).

The development of tasks with minimal linguistic demands designed for younger children has provided models for testing nonhuman species, including squirrel monkeys (McGonigle & Chalmers, 1977) and chimpanzees (Gillan, 1981). McGonigle and Chalmers (1977) trained squirrel monkeys to select pairs of containers that differed in color and weight. Their results indicated that the monkeys' responses were transitive and virtually identical to profiles reported earlier by Bryant and Trabasso (1971) for children. They considered several alternative strategies to account for the monkeys' performance, although they noted

Sarah T. Boysen and Gary G. Berntson, Primate Cognition Project, Ohio State University and Yerkes Regional Primate Research Center; Traci A. Shreyer and Karen S. Quigley, Ohio State University.

This investigation was supported in part by National Institute of Mental Health Grants MH 44022 and MH 45460 to Sarah T. Boysen and Gary G. Berntson and National Science Foundation Grant BNS-9022355 to Sarah T. Boysen. Support was provided in part by National Institute of Health Grant RR-00165 from the National Center for Research Resources to the Yerkes Regional Primate Research. The Yerkes Center and the Ohio State University Laboratory Animal Center are fully accredited by the American Association for Accreditation of Laboratory Animal Care.

We gratefully acknowledge the contribution of Michelle Hannan and the continued cooperation of Jack Hanna, Director, Lynn W. Kramer, Director of Research, and Don Winstel, Curator, of the Columbus Zoo.

Correspondence concerning this article should be addressed to Sarah T. Boysen, Primate Cognition Project, Room 48, Townshend Hall, Ohio State University, Columbus, Ohio 43210-1222. Electronic mail may be sent to boysen.1@osu.edu.

that the animals nonetheless had to coordinate two separate pieces of information to solve the task. They concluded that some kind of inference was necessary for the animals to respond correctly during tests of transitivity (McGonigle & Chalmers, 1977).

In a similar study, Gillan (1981) explored transitive inference in young non–language-trained chimpanzees by using a series of colored boxes. Boxes were presented serially in pairs, with a single food item placed in one box of each pair (A–B+, B–C+, etc.). The chimps were tested on the novel BD pair after learning all four training pairs. One chimpanzee consistently chose the D box, a second animal was inconsistent, and a third was initially inconsistent but chose D during a second test (Gillan, 1981). In additional tasks, numerous manipulations, such as eliminating the high end point of the series, were introduced. From these results Gillan concluded that integration theories (i.e., subjects integrate information about pairs of stimuli into an ordered series) provided better theoretical support for the chimpanzee data than a nonintegration position and that language was not a necessary prerequisite for using transitive inference.

In light of these findings, we undertook this study to explore further processes of transitivity with a nonhuman primate species, the chimpanzee. The transitive inference paradigm, as applied by Gillan (1981), suggests an approach for evaluating ordinality in chimpanzees with previous experience with number concepts.

## General Method

### Subjects

Three juvenile, captive-born chimpanzees (*Pan troglodytes*) served as subjects (2 males, 8 and 8½ years of age, and 1 female, 6½ years of age). Kermit and Darrell were peer-raised in a laboratory nursery until ages 3 and 3½ years, and Sheba was human cross-fostered from 4 months until she joined the project at age 2½ years. All animals had been immersed in daily interactional training with their teachers and caregivers in a variety of conceptual tasks for a period of approximately 5 years.

All 3 animals had also received considerable training in counting arrays of edibles and objects, as well as number comprehension (Boysen, 1992a; Boysen & Berntson, 1989). Each of the animals, however, evidenced considerable differences in their numerical capabilities (Boysen, 1992a, 1992b). At the beginning of the study, Sheba had the most extensive experience, with a numbers repertoire that ranged from 0 to 6. She also had had additional testing for summation with objects and Arabic symbols (Boysen, 1992a; Boysen & Berntson, 1989) and preliminary training in a subtraction task. Darrell and Kermit, who each had the same training history, also showed individual differences in numerical abilities. At the beginning of the study, Darrell consistently performed correctly in number-related tasks that required him to label arrays with an Arabic numeral repertoire of 0–4 and was receiving concurrent training with the number 5. Kermit, however, despite 3 years of comparable training (with the same human teacher) persistently evidenced limited abilities with numbers. He had great difficulty in consistently labelling even small arrays of 1–3 items. Kermit's number repertoire included 0–4, with somewhat better performance with numbers 1 and 4. This suggested that he had some understanding of the smallest and largest arrays that might be presented to him and showed some reliability in associating the Arabic numeral with each of these. However, his discrimination of arrays of 2 or 3 was inconsistent, and continuing attempts at remediation yielded little change in his performance. Kermit had also demonstrated difficulty in other tasks that had been designed as diagnostic tools for children with Attention Deficit Disorder (i.e., vigilance tasks with continuous performance components; O'Dougherty, Berntson, Boysen, Wright, & Teske, 1988), which suggests that attentional problems may have interfered with his ability to perform well under some test conditions. He had, however, demonstrated superior performance over the other animals in tasks related to color discriminations. With the exception of the vigilance tasks, the individual differences observed among the animals had never resulted in appreciable performance differences in other non–number-related tasks that the animals had completed.

### Procedure

The chimps were tested daily, 5 days a week. Sheba, the younger female, was tested unrestrained in an open testing area, which was the usual training and testing context for her. The 2 males were tested individually in their shared home cage. Both had extensive experience in taking turns, such that one animal completed his testing session while the other played in the adjacent outdoor area and then changed places with him (Boysen, 1992b).

In all three experiments, pairs of stimuli (small painted boxes or plastic placards with Arabic numerals) were presented on the left and right sides of a small tray (15 × 23 cm) on each trial. The animals were required to indicate their choice by pointing to or touching the stimulus item. The animals' motor responses were readily discriminable and, together with the physical distance between the two choice stimuli, permitted an unambiguous evaluation of their selection. During training, the experimenter sat directly across from the animals as the stimuli were presented and provided immediate verbal feedback on each trial. The chimpanzees received a food reward and social praise when correct. When an error was made, a correction procedure was used: The animals were permitted to make a second choice, which was followed by the same reinforcement procedure. During blind testing, the animals were separated from the experimenter by a partition, and their choices were simultaneously monitored and recorded by a videocamera positioned in front of the cage. The animals' responses were viewed on a videomonitor several meters away by a second experimenter, who also recorded responses on data sheets.

## Experiment 1

### Method

The first study was a replication of Gillan (1981), with the procedures modeled directly after his approach. Five wooden boxes (9 × 20 cm) with solid lids were each painted a different color (black, yellow, white, silver, and red) and represented an ordered series ABCDE. Between trials the boxes were placed behind a partition, and a single food reinforcer (a small candy) was placed inside the higher order box. Boxes were always presented as adjacent pairs during each trial, with only one box of the pair baited. For example, when Pair AB was presented, Box A never contained food, whereas Box B always did; when Pair BC was presented, Box B never contained food, and Box C always did, and so on. The rule therefore was to always select the box of the pair that was closer to the high end of the series. The pairs were trained in order, beginning with Pair AB, to a criterion of 90% or better for two consecutive 16-trial sessions. Once criterion with AB was met, Pair BC was introduced and so on, until all four adjacent pairs had been trained to criterion. Next, mixed-pair sessions were conducted, beginning with sessions in which Pairs AB and BC were presented. The animals now had to meet criterion on each pair within a session, for two consecutive sessions, before the next pair was introduced into the mixed sessions.

After overall criterion performance on sessions in which all four training pairs were presented, blind tests were completed. During blind testing, the nonadjacent pair BD was presented in probe trials, with selection of D considered a correct choice. Test trials (2–5 per session) were introduced randomly among 12–16 training trials. During testing, the experimenter sat behind a partition that was positioned so that the animals' hands (and thus their choice of boxes) could be readily observed, but the experimenter could not be seen by the animals. No visual or verbal contact was made with the animals before their response. After the animals made a choice, they were reinforced for response with food and social praise.



### *Results*

All 3 animals selected Box D from the nonadjacent BD pair more often during novel testing than would be predicted by chance (Table 1), with individual performances statistically significant: For Sheba, $\chi^2(1, N = 12) = 5.33, p < .02$; for Darrell and Kermit, $\chi^2(1, N = 8) = 8.00, p < .01$. Their performance on the novel BD comparisons, in fact, was not significantly different from that on the concurrent training pairs, $\chi^2(1, N = 139) = 1.05$, *ns*. To minimize the potential problem of differential learning with the choice boxes throughout acquisition, the animals were trained to a fixed criterion for each pair. To examine this issue further, the ratios of reinforcement for the BD elements were then obtained by dividing the total number of reinforced trials when an element was the correct choice (including correction trials) by the total number of trials in which the element was chosen (Couvillon & Bitterman, 1992). The resulting ratios for B and D were not appreciably different (for B, $M = .87$, and for D, $M = .92$).

### *Discussion*

These data provide evidence in support of the findings of Gillan (1981), that chimpanzees are capable of determining the correct choice between two nonadjacent items in an ordered series, under novel test conditions. The results cannot be accounted for by differential reinforcement histories during acquisition because (a) all stimuli were trained to equivalent criterion, (b) the reinforcement ratios for B and D were not appreciably different, and (c) the total number of reinforcements, in fact, favored B (number of reinforcements for B, $M = 203.7$; for D, $M = 129.7$). In addition, because it was necessary that performance on all pairs reach criterion during random presentations within mixed-pair sessions before novel testing, the data do not appear to be accounted for by any type of recency effect.

## Experiment 2

To examine possible inferential processing in the chimpanzee further, a second experiment was proposed. To examine the animals' understanding of ordinality that may have emerged during their training on number concepts (Boysen & Berntson, 1989), the same transitive inference paradigm (Gillan, 1981) was used in Experiment 2, with Arabic number stimuli. Two of the three animals had demonstrated facility with number symbols in a variety of number-related tasks, which have been reported elsewhere (Boysen, 1992a, 1993; Boysen & Berntson, 1989). However, no task had yet provided evidence for the animals' understanding of the relative ordinal position of numbers within the counting sequence. The transitive inference task seemed ideal for possible demonstration that the animals had a grasp of increasing or decreasing numerosity, and thus Arabic number symbols replaced the colored box series as stimuli in Experiment 2.

### *Method*

The same 3 chimpanzees served as subjects for Experiment 2. The training procedures for the second experiment were identical to those described for Experiment 1, with the exception that the Arabic numeral series 1–5 served as stimuli. Arabic numerals, black on a silver background, were affixed to 7.5 × 12.5 cm Plexiglas placards and presented on a test tray as described for Experiment 1. Subjects were to select the larger of the two numbers presented, beginning with 1 versus 2. When correct, the animals were reinforced with nondiscrete edibles (yogurt or fruit juice), to avoid any task-specific association of the number symbols with some absolute number of reinforcers. After 90% criterion for two sessions, 2 versus 3 was introduced, and after criterion with the additional number pairs (3 vs. 4 and 4 vs. 5), training began with mixed pairs (1 vs. 2 and 2 vs. 3 within the same session). As in Experiment 1, during mixed-pair training, the animals were required to reach criterion with all pairs within a session, for two successive sessions, before the next number pair was introduced. Thus, the final phase of acquisition consisted of training sessions in which all four number pairs were randomly presented, and criterion performance with each pair for two successive sessions was required before novel testing.

After criterion, novel tests were completed in which the nonadjacent pair 2–4 was presented, embedded among a series of randomly ordered training trials with the four previously trained pairs.

Table 1
*Novel Test Performance With Nonadjacent Pair BD*

| | Training trials | | | Test trials | | | |
|---|---|---|---|---|---|---|---|
| Subject | No. correct | Total | % correct | No. correct | Total | % correct | *p* |
| Sheba | 41 | 47 | 87 | 12 | 2 | 83 | .02 |
| Darrell | 27 | 32 | 84 | 8 | 8 | 100 | .01 |
| Kermit | 27 | 32 | 84 | 8 | 8 | 100 | .01 |

*Note.* *p* values indicate the significance of chi-square tests to compare training and test trials.

During blind testing, all animals were tested individually in their home cages. The experimenter sat behind a partition placed against the front of the cage, which permitted presentation of the stimuli in the same manner as training. However, the experimenter was not able to see the animals nor could the animals see the experimenter's face. In turn, the experimenter monitored the animals' responses by a videocamera positioned in front of the cage, which permitted a clear picture of the test stimuli and the animal's hands. After each correct response, the experimenter reinforced the animals with yogurt or fruit juice.

### *Results*

Significant differences emerged across the 3 animals during training with the number pairs, with Sheba demonstrating much more rapid acquisition than either Kermit or Darrell. The total number of trials over all phases of acquisition for Sheba was 550, compared with 2,905 for Darrell and 2,787 for Kermit. The larger values for Kermit and Darrell were due largely to the final phase of mixed-pair training, in which all number pairs were presented.

During the initial novel test with numbers, Sheba correctly selected the number 4 when the nonadjacent 2–4 pair was presented among training pairs, $\chi^2(1, N = 22) = 11.63, p < .001$ (see Table 2). Both Kermit and Darrell failed the novel test, showing no significant preferences when presented with the nonadjacent novel pair 2–4 (Table 2). The total number of reinforcements during training did not differ for the elements 2 or 4 (number of reinforcements for 2, $M = 512$; for 4, $M = 515$). Although the reinforcement ratios did favor somewhat the selection of 4 (ratio for 2, $M = .80$; for 4, $M = .87$), these differences were minimal and, in fact, were not significantly different for the animal who performed successfully on the novel test: For Sheba, the ratio for 2 was .84 and for 4, .91, $\chi^2(1, N = 139) = 2.82$, *ns*.

### *Discussion*

The results from Experiment 2 indicate that 1 chimpanzee (Sheba) was able to correctly assess the larger of two numbers from an ordered series (1–5) when the nonadjacent number pair (2–4) was presented for novel testing. These data also suggest that the number sequence was represented as an ordered series by Sheba and that her responses may reflect a comparison that invokes the use of a transitive process for the correct choice of the larger number. Reinforcement ratios for the training pairs do not support the hypothesis that differential reinforcement history could account for her selections.

Results of initial blind testing with the other two animals were nonsignificant. Both Kermit and Darrell failed novel tests in which the nonadjacent pair 2–4 was presented, despite their having met the criterion demands of the training phase. Recall, however, that protracted training was necessary for both animals to meet criterion before administration of the blind tests. Darrell required over 2,400 trials in the final acquisition phase (mixed-pair training of all four number pairs), and Kermit completed over 1,800 trials, compared with only 160 trials of the same phase for Sheba to meet criterion. The similarly poor performance of Kermit and Darrell was particularly interesting, given their individual training history with number concepts. It would not have been surprising, with Kermit's seemingly poor grasp of number concepts, if he had not successfully integrated the individual discriminations represented by the training pairs into an ordered sequence. Darrell, however, had evidenced a consistent ability to select the correct Arabic symbol when presented with arrays of 1–4 candies or wooden shapes, as well as successful completion of number comprehension training with 1–3 (Boysen & Berntson, 1989). Like Sheba, his number skills were suggestive of a conceptual appreciation for some rudimentary principles of counting. In light of Darrell's training history, his failure to correctly choose the larger number of the nonadjacent novel pair was initially somewhat surprising.

Two significant tasks, functional counting and symbolic counting, in which Sheba demonstrated the spontaneous ability to sum small arrays of objects or Arabic symbols placed in different locations around the laboratory, may have contributed to elaboration of a more flexible "sense of number" (Davis & Pérusse, 1988, p. 568) in this animal (Boysen, 1992a; Boysen & Berntson, 1989). Results from the novel number test may therefore reflect the contributions of this additional training to Sheba's ability to represent the numbers 1–5 as an ordered series.

## Experiment 3

A third task was proposed in which the animals were required to assess the number pairs and now select the smaller of the two numerals. This represented a reversal of Experiment 2, in an effort to explore whether or not the unidirectionality of the typical transitive inference approach (toward increasing numerosity, as in Experiment 2) was contributing in some manner toward successful task performance.

Table 2
*Novel Test Performance with Nonadjacent Numbers 2 and 4 (Larger Than)*

| | Training trials | | | Test trials | | | |
|---|---|---|---|---|---|---|---|
| Subject | No. correct | Total | % correct | No. correct | Total | % correct | *p* |
| Sheba | 54 | 64 | 84 | 19 | 22 | 86 | .001 |
| Darrell | 47 | 64 | 73 | 12 | 22 | 56 | |
| Kermit | 44 | 64 | 69 | 14 | 22 | 64 | |

*Note.* *p* values indicate the significance of chi-square tests to compare training and test trials.

### Method

Sheba, now 7 years of age, served as the subject for Experiment 3. Given that Kermit and Darrell had failed the novel test in Experiment 2, they did not participate.

Training and testing procedures were the same, with the exception that Sheba was now required to select the smaller number represented by adjacent numeral pairs. Number pairs were presented in descending order, beginning with 5 versus 4, then 3 versus 4, and so forth. Other procedures were as described in the earlier experiments. After overall criterion performance with all four number pairs presented in mixed-pair sessions, blind tests were conducted under the same conditions as Experiment 2. During blind testing, the nonadjacent pair 2–4 was presented, randomly interspersed among 12–16 training trials per session.

### Results

Sheba selected the smaller number of the nonadjacent pair 2–4 in 28 of 30 opportunities (93%) over four test sessions, $\chi^2(1, N = 30) = 22.53, p < .001$, and maintained significant performance with the training pairs throughout blind testing (71 of 80 trials; 89%), $\chi^2(1, N = 80) = 48.05, p < .001$. Sheba's performance on the test session was equivalent to that of the training pairs. Reinforcement ratios for the 2 and 4 stimuli during training were again similar, and thus differential reinforcement does not appear to account for Sheba's selection of the smaller numeral.

### Discussion

The results of Experiment 3 further support the suggestion that Sheba was able to represent the numerals 1–5 as an ordered series and was therefore able to select the smaller of two nonadjacent numbers (2–4) presented during blind tests. It also suggests an understanding of the ordinal characteristics of the number line. Her ability to reverse the discrimination trained during Experiment 2 and to demonstrate recognition of the concept *smaller than* provides additional support for the suggestion that Sheba has a flexible sense of number and can readily ascribe the relative ordinal position of numbers within an ascending or descending number series.

## Experiment 4

Experiment 4 was designed to expand the range of test stimuli used in Experiment 2 to include novel numbers. At the beginning of the study, Sheba and Darrell were undergoing concurrent training on several number-related tasks. These included a receptive task (perceived numbers), during which the animals were required to create the correct size of array from a collection of spools, in response to a displayed number from 1 to 5, and fractions, during which one-half or one-fourth segments of fruits (e.g., bananas or apples) were presented for labeling with fraction symbols ½ or ¼ (Boysen, 1991). During the extended period of training and testing required for Experiment 1–3, Sheba and Darrell had also received concurrent training on productive labeling tasks with new numbers, increasing their counting repertoires (for Sheba, 0–8, and for Darrell, 0–6, with some experience with the number 7). It was hypothesized that the increased counting repertoire used to label arrays, as well as the greater breadth of numerical skills acquired through the introduction of the perceived numbers task and fractions, might provide the animals with a better understanding of relative ordinal positions of each number symbol.

### Method

All 3 chimps were subjects for Experiment 4.

*Preliminary training procedure.* After completion of Experiment 3, both Sheba and Darrell had received additional productive counting experience with numbers 0–7, as well as introduction to and training with fractions ¼ and ½. To equate for Sheba's experience with the descending number series, Darrell was trained to discriminate the smaller of each pair of the numerals 1–5. Training for Darrell was identical to Sheba's for Experiment 3, with training to a criterion of 90% correct performance for two successive sessions. Kermit's overall performance had been poor and very inconsistent. Thus, to avoid additional frustration, Kermit was not given further interim training but was included in Experiment 4 as a control. (For further elaboration of individual differences in the chimpanzees' abilities and training approaches [Boysen, 1992b], the reader is encouraged to see Davis & Balfour, 1992, on scientist–animal interactions).

*Procedure.* All three chimpanzees completed refresher training on the original number-pair discrimination procedure, in ascending order (numbers 1–5, beginning with 1 vs. 2). As 24 months had elapsed since the earliest training for Experiment 2, the retraining phase provided a measure of the chimps' current performance with *larger than* discriminations, which could be compared with novel tests with additional, untrained pairs of numbers. Criterion for the retraining was achieved within 4 weeks.

After completion of retraining, all 3 animals completed blind tests that included trials with training pairs of numbers between 1 and 5, the nonadjacent probe 2–4, and six number pairs designated as novel–novel probes. The novel–novel probes were number pairs that had not been used in any of the previous experiments and had never been presented for discrimination in either the larger than or smaller than tasks. The probes included: 0 vs. 1, 0 vs. 7, 5 vs. 7, 6 vs. 7, 6 vs. 4, and 6 vs. 5. Thus, some novel–novel probes represented pairs of numbers that were completely novel to the task (e.g., 6 vs. 7, 0 vs. 7), whereas other probes included one novel and one familiar element. The test session consisted of 34 trials, including 20 training trials, interspersed with 12 novel–novel probes (2 trials of each), and two trials of 2 versus 4. The animals' performance on the second novel test is presented in Table 3.

### Results

As seen in Table 3, all 3 animals achieved at least 90% performance in selecting the larger number represented by the training pairs, and all 3 were 100% accurate in selecting the larger number represented by the nonadjacent pair 2–4 on those two test trials. Two of the chimpanzees, Sheba and Darrell, also correctly selected the larger of two numbers represented by the novel–novel probes. As illustrated in Table 3, Sheba was correct for all 12 novel–novel trials, $\chi^2(1, N = 12) = 12.00, p < .001$, and Darrell was correct in 10 of 12 trials, $\chi^2(1, N = 12) = 5.33, p = .02$. Kermit's performance did not reach significance (8 of 12 trials). Of par-

Table 3
*Test Performance With Novel–Novel Probes*

| | Training trials | | | Novel probes | | | | Trial 1 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Subject | No. correct | Total | % correct | No. correct | Total | % correct | $p^a$ | No. correct | Total | % correct | $p^b$ |
| Sheba | 18 | 20 | 90 | 12 | 12 | 100 | .001 | 6 | 6 | 100 | .001 |
| Darrell | 20 | 20 | 100 | 10 | 12 | 83 | .02 | 6 | 6 | 100 | .001 |
| Kermit | 18 | 20 | 90 | 8 | 12 | 67 | | 4 | 6 | 67 | |

[a] *p* values indicate the significance of chi-square tests to compare training and novel probe trials. [b] *p* values indicate the significance of chi-square tests to compare training trials and Trial 1 responses.

ticular interest is the fact that both Sheba and Darrell achieved 100% ($p < .001$) performance on the first trial presentations of the novel–novel probes. Thus, the animals responded correctly before any reinforcement for the novel–novel probes.

Reinforcement ratios over all training trials for the 2–4 elements did not differ, either over all animals (for 2, $M = .86$; for 4, $M = .88$) or for the 2 animals who performed successfully on the novel–novel tests for Sheba and Darrell, for 2, $M = .84$; for 4, $M = .86$). In fact, over the three sessions just before the blind novel–novel test, which reflected the animals' most recent reinforcement history, the overall reinforcement ratios favored selection of 2 (in mixed-pairs trials in the last three sessions, for 2, $M = .95$; for 4, $M = .86$).

## General Discussion

The results of the three experiments provide evidence for the capacity of chimpanzees to demonstrate an understanding of ordinality and extend the transitivity findings of Gillan (1981) to tasks that use symbolic number representations. In Experiment 1, which replicated Gillan's (1981) Experiment 1A, all 3 chimpanzee subjects were able to correctly select Stimulus D when presented with the nonadjacent BD pair under novel test conditions. Our findings also compare favorably with similar experimental results from tests of transitivity with children (e.g., Bryant & Trabasso, 1971).

In Experiment 2, a series of Arabic numerals (1–5) were used as comparison stimuli instead of the colored box series of Experiment 1. Given that all 3 chimpanzees had eventually reached criterion during acquisition of the number pair discriminations and thus demonstrated the ability to select the larger number of the four adjacent number pairs, Sheba's ability to correctly select the larger number during novel tests with the nonadjacent pair 2–4 merits additional consideration. Sheba was also able to reverse the discrimination series in Experiment 3 and correctly select the smaller member of the 2–4 nonadjacent pair during testing.

As a subject, Sheba had not distinguished herself in any significant way on prior tasks (with the exception of the functional and symbolic counting tasks noted later) and therefore could not be described as more intelligent by either subjective or objective criteria. However, given both her younger age and smaller female body size, she was able to participate in several additional number-related tasks (Boysen, 1992a; Boysen & Berntson, 1989) that could not be readily adapted for testing with Kermit and Darrell. However, precisely how Sheba's previous experiences in number-related tasks may have contributed to these additional capabilities cannot yet be fully characterized.

Kermit's and Darrell's performance during training with adjacent pairs in Experiment 1 and 2 may indicate a grasp of the ordinal relation between adjacent stimulus elements. However, testing with the novel pair 2–4 in Experiment 2 indicated that neither animal, at that time, was able to reliably represent the entire series 1–5 as an ordered sequence in a fashion that permitted interrogations about the larger element of a novel nonadjacent pair. Given the significant differences in numerical competence between Kermit and Darrell at the start of the study, the results of Experiment 2 also suggested that prior number skills may not have been sufficient to permit a global representation of an ascending number series to emerge. In contrast, Sheba already had a good working knowledge of numbers 1–5 and continued to receive additional training on subsequent numbers 6 and 7 in other number-related tasks throughout the study period. It is possible that this more extensive numbers repertoire provided her with a conceptual frame of reference for the number series used in Experiment 2 and 3 and a better working knowledge of the ordinal sequence 1–5. For Sheba, number symbols were demonstrably representational (Boysen & Berntson, 1989) and linked to specific quantities. The other chimpanzees had not had the opportunity to acquire or demonstrate similar representational status with numbers at that time, and appeared unable to represent coherently the series of number symbols in the appropriate order.

In Experiment 4, training with descending number pairs was completed with Darrell, as well as the refresher training on ascending pairs with all 3 animals. The results are notable for several reasons. Two of the 3 animals tested chose the larger of the two numerical stimuli, even when the number pairs presented were completely novel. That is, the specific number pairs had never been used during training for other ordinality tests. It is particularly noteworthy that Trial 1 performance for both Sheba and Darrell was 100% for these novel–novel test probes. These findings suggest that both animals were able to represent an ordered series of numerical stimuli from 0 to 7 and correctly interrogate such a representation when confronted with novel number combinations.

Numerous alternative hypotheses have been raised in the animal and developmental literature to account for similar data in nonhuman primates, pigeons, and children (Breslow, 1981; Couvillon & Bitterman, 1992; Gillan, 1981; McGonigle & Chalmers, 1977, 1984; von Fersen, Wynne, Delius, &

Staddon, 1991). These include alternatives to a more symbolic or cognitive interpretation of transitivity, such as value transfer theory (von Fersen et al., 1991) or a simple conditioning model (Couvillon & Bitterman, 1992). However, in our studies, training was always completed to a specified criterion to minimize potential bias of differential reinforcement and thus to control for any history of associative strength. Moreover, the prior reinforcement ratios, as proposed by Couvillon and Bitterman (1992), fail to account for our results.

The most interesting findings from Experiment 4 are the results with the novel–novel probes, because the number stimuli used in these test trials had never been reinforced for selection in relation to another number. Thus, these stimuli had the weakest reinforcement history. Nonetheless, both animals were correct in their selection of the larger number for each novel–novel probe, on Trial 1. Such choices did not simply represent a preference (or aversion) for new numbers. For example, the number 1 was chosen correctly when paired with 0, even though selection of 1 had never been reinforced throughout the 2½ years of the complete study. Similarly, numbers 6 and 7 were differentially selected over smaller numbers in the series, such as 5, which had always been reinforced in the acquisition phases of the experiment. Clearly, reinforcement history alone does not support the animals' performance.

This series of experiments also speaks to a broader goal of exploring the chimpanzees' potential for understanding ordinality. In the developmental literature, tests for both seriation and transitivity are frequently used for the assessment of implicit ordinal number comprehension (Kingma & Zumbo, 1987). *Ordinality* refers to an appreciation of the intrinsic order of numbers, such that one of the conditions, $a > b$, $a = b$, or $a < b$, is met. It was proposed by Davis and Pérusse (1988) that an understanding of ordinality is a pivotal concept that must be demonstrated in order to assert that a nonhuman organism has exhibited true counting. Ordinality was of particular interest to us because the chimps had been trained to label arrays with corresponding Arabic numerals and had also received explicit training on number comprehension (Boysen & Berntson, 1989). The results of this study suggest that the chimpanzee is capable of using some process of transitivity with both symbolic stimuli, in the form of number representations, as well as nonsymbolic stimuli, represented by an ordered series of boxes. These data also provide evidence that chimpanzees are capable of demonstrating an understanding of the relative positions of elements in a series and indicate that animals trained in number concepts may come to recognize the ordinal nature of the counting sequence.

## References

Boysen, S. T. (1991, November). *Parts is parts: Use of fraction symbols by a chimpanzee.* Poster presented at the annual meeting of the Psychonomic Society, San Francisco.

Boysen, S. T. (1992a). Counting as the chimpanzee views it. In W. K. Honig & J. G. Fetterman (Eds.), *Complex extended stimuli in animals* (pp. 367–383). Hillsdale, NJ: Erlbaum.

Boysen, S. T. (1992b). Pongid pedagogy: The contribution of human/chimpanzee interaction to the study of ape cognition. In H. Davis & D. Balfour (Eds.), *The inevitable bond: Examining scientist–animal interactions* (pp. 205–217). New York: Cambridge University Press.

Boysen, S. T. (1993). Counting in chimpanzees: Nonhuman principles and emergent properties of number. In S. T. Boysen & E. J. Capaldi (Eds.), *The development of numerical competence: Animal and human models* (pp. 39–59). Hillsdale, NJ: Erlbaum.

Boysen, S. T., & Berntson, G. G. (1989). Numerical competence in a chimpanzee (*Pan troglodytes*). *Journal of Comparative Psychology, 101,* 23–31.

Breslow, L. (1981). Reevaluation of the literature on the development of transitive inferences. *Psychological Bulletin, 89,* 325–351.

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

Burt, C. J. (1919). The development of reasoning in school children. *Journal of Experimental Pedagogy, 5,* 68–77.

Chapman, M., & Lindenberger, U. (1988). Functions, operations, and decalage in the development of transitivity. *Developmental Psychology, 24,* 542–551.

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., & Balfour, D. (Eds.). (1992). *The inevitable bond: Examining scientist–animal interactions.* New York: Cambridge University Press.

Davis, H., & Pérusse, R. (1988). Numerical competence in animals: Definitional issues, current evidence and a new research agenda. *Behavioral and Brain Sciences, 11,* 561–615.

de Boysson-Bardies, B., & O'Regan, K. (1973). What children do in spite of adults' hypotheses. *Nature, 246,* 531–534.

Elkind, D. (1974). *Children and adolescents: Interpretive essays on Jean Piaget.* London: Oxford University Press.

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

Halford, G. S. (1984). Can young children integrate premises in transitivity and serial order tasks? *Cognitive Psychology, 16,* 65–93.

Halford, G. S., & Kelly, M. E. (1984). On the basis of early transitivity judgments. *Journal of Experimental Child Psychology, 38,* 42–63.

James, W. (1890). *Principles of psychology.* New York: Holt.

Kingma, J., & Zumbo, B. (1987). Relationship between seriation, transitivity, and explicit ordinal number comprehension. *Perceptual and Motor Skills, 65,* 559–569.

Matsuzawa, T. (1985). Use of numbers by a chimpanzee. *Nature, 315,* 57–59.

McGonigle, B., & Chalmers, M. (1977). Are monkeys logical? *Nature, 267,* 694–696.

McGonigle, B., & Chalmers, M. (1984). The selective impact of question form and input mode on the symbolic distance effect in children. *Journal of Experimental Child Psychology, 37,* 524–554.

O'Dougherty, M., Berntson, G. G., Boysen, S. T., Wright, F. S., & Teske, D. (1988). Psychophysiological predictors of attentional dysfunction in children with congenital heart defects. *Psychophysiology, 25,* 305–315.

Piaget, J., Inhelder, B., & Szeminska, A. (1960). *The child's conception of geometry* (E. Lunzer, Trans.). New York: Basic Books. (Original work published 1948)

Premack, D. (1976). *Intelligence in ape and man.* Hillsdale, NJ: Erlbaum.

Riley, C. A., & Trabasso, T. (1974). Comparatives, logical structures and encoding in a transitive inference task. *Journal of Experimental Child Psychology, 17,* 187–203.

Roitblat, H., Bever, T., & Terrace, H. S. (1984). *Animal cognition.* Hillsdale, NJ: Erlbaum.

Trabasso, T. (1975). The role of memory as a system in making transitive inferences. In A. D. Pick (Ed.), *Minnesota Symposium on Child Psychology.* Minneapolis: University of Minnesota Press.

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.

Youniss, J., & Furth, H. J. (1973). Reasoning and Piaget. *Nature, 244,* 314–315.

Received May 24, 1991
Revision received July 10, 1992
Accepted July 13, 1992 ■