



Published in final edited form as:

J Comp Psychol. 2009 February ; 123(1): 79–89. doi:10.1037/a0013126.

Abstract-Concept Learning Carryover Effects from the Initial Training Set

Tamo Nakamura,

The University of Texas Health Science Center at Houston, Medical School

Anthony A. Wright,

The University of Texas Health Science Center at Houston, Medical School

Jeffrey S. Katz,

Auburn University

Kent D. Bodily, and

Auburn University

Bradley R. Sturz

Auburn University

Abstract

Three groups of pigeons were trained in a *same/different* task with 32, 64, or 1024 color-picture stimuli. They were tested with novel transfer pictures. The training-testing cycle was repeated with training-set doublings. Group 32 learned the *same/different* task as rapidly as a previous 8-item group and transferred better than the 8-item group at the 32-item training set. The 64- and 1024-item groups learned the task only somewhat slower than other groups, but their transfer was better and equivalent to baseline performances. These results show that pigeons trained with small sets (e.g., 8 items) have carryover effects that hamper transfer when the training set is expanded. Without carryover effects (i.e., initial transfer from groups 32 & 64), pigeons show the same degree of transfer as rhesus and capuchin monkeys at these same set sizes. This finding has implications for the general ability of abstract-concept learning across species with different neural architectures.

Author contact: Anthony A. Wright, The University of Texas Health Science Center at Houston, Medical School, Department of Neurobiology and Anatomy, 6431 Fannin Street, Houston, TX 77030 Or PO Box 20708, Houston, TX 77225, anthony.a.wright@uth.tmc.edu, Tel: 713 500 5627, Fax: 713 500 5623. Correspondence may be addressed to any author: Tamo Nakamura, School of Psychology, The University of Newcastle, University Drive, Callaghan NSW 2308 Australia; Anthony A. Wright, Department of Neurobiology and Anatomy, The University of Texas Medical School at Houston, P.O. Box 20708, Houston, TX 77225-0708 USA; Jeffrey S. Katz, Kent D. Bodily, Department of Psychology, Auburn University, 226 Thach Hall, Auburn, AL 36849 USA or Bradley R. Sturz, Department of Psychology, Armstrong Atlantic State University, Savannah, GA 31419 USA. Tamo.Nakamura@newcastle.edu.au, Anthony.A.Wright@uth.tmc.edu, katzjef@auburn.edu, bodilke@auburn.edu, or bradley.sturz@armstrong.edu.

The following manuscript is the final accepted manuscript. It has not been subjected to the final copyediting, fact-checking, and proofreading required for formal publication. It is not the definitive, publisher-authenticated version. The American Psychological Association and its Council of Editors disclaim any responsibility or liabilities for errors or omissions of this manuscript version, any version derived from this manuscript by NIH, or other third parties. The published version is available at <http://www.apa.org/journals/com/>

Keywords

same-different; abstract-concept learning; item-specific learning; relational learning; training set size; pigeons

Among long-standing issues in animal cognition is which (nonhuman) animal species have the cognitive ability to learn abstract concepts (e.g., Premack, 1978). The hallmark of abstract concepts is that they transcend specific features of the training stimuli because they are based upon relationships (i.e., relational learning). In a seminal paper that was instrumental in the modern resurgence of animal cognition, David Premack said "...unlike the primate, the behavior of the pigeon is predominantly controlled by absolute rather than relational factors..." (1978, p. 432). Premack's claim that item-specific learning (i.e., absolute factors) by pigeons has been the accepted explanation for why pigeons, at that time, had failed to learn abstract concepts (e.g., Cumming, Berryman, & Cohen, 1965; Farthing & Opuda, 1974; Santi 1978). Nevertheless, evidence began to emerge that if a large number of training stimuli (e.g., 152 color cartoons) were used, pigeons actually could learn a matching-to-sample (MTS) abstract concept (Wright, Cook, Rivera, Sands, & Delius, 1988).

A somewhat more advantageous procedure than MTS for testing abstract-concept learning is the *same/different* procedure because the matching and non-matching stimuli are presented on separate trials. Details of currently used S/D procedures used to test pigeon abstract-concept learning vary widely including the numbers of stimuli used in training and testing (e.g., Blaisdell & Cook, 2005; Cook, Kelly, & Katz, 2003; Santiago & Wright, 1984; Wasserman, Hugart, & Kirkpatrick-Steger, 1995; Young & Wasserman, 1997). In consideration of the evidence that some species only accurately transfer their performance to novel stimuli following training with a substantial number of training stimuli, we have regularly made the number of training stimuli a parameter in our experiments.

In a series of experiments with rhesus monkeys, capuchin monkeys, and pigeons, we began training with a small set of 8 stimuli and then progressively expanded (doubled) the training set followed by transfer tests with novel stimuli (Katz & Wright, 2006; Katz, Wright, & Bachevalier, 2002; Wright, Rivera, Katz, & Bachevalier, 2003). This regime allowed us to assess whether or not the degree of concept learning would depend on the number of training exemplars and whether or not transfer would eventually reach the level of baseline performance with the training stimuli. We anticipated that large numbers of training stimuli might be necessary to achieve a level of transfer equivalent to baseline performance, and therefore used picture stimuli. We selected pictures so that the training and testing pictures would be distinctive from one another and the transfer stimuli were only tested once, by definition of their being novel. In these experiments when the training set was expanded sufficiently (256 items for pigeons, 128 for monkeys), transfer became equivalent to baseline performance which remained above 80% correct. We termed this full abstract-concept learning because the baseline performance sets an upper limit. This result was the first time that pigeons had shown full concept learning in an S/D task, but pigeons required more training exemplars (i.e., stimulus pairs of *same* or *different*) than rhesus or capuchin monkeys. We characterized this species difference as a quantitative difference in contrast to

the ability to fully learn the S/D abstract concept which we characterized as a qualitative similarity.

The purpose of the experiments of this article was to explore whether or not training pigeons with somewhat larger initial sets of stimuli might produce better transfer than those trained initially with an 8-item set in the prior study. Training with the smaller 8-item set might produce carryover effects that would interfere with generalized relational learning which would in turn be necessary for good performance with the large training sets. To this end, we trained different groups of pigeons with larger initial training sets and compared their transfer to groups trained initially with smaller training sets. Such comparisons, of course, would depend upon pigeons being able to learn the S/D task with large training sets. With these considerations, the first group of pigeons was trained on an intermediate set size of 32 pictures.

Experiment 1

A group of four pigeons was trained on an initial set size of 32 items and then was tested for transfer to novel stimuli. One purpose was to compare this group's learning and transfer to a group that had learned the S/D task with an initial 8-item set (Katz & Wright, 2006). After the 8-item group was tested for transfer, their training set was expanded to 16 items and then to 32 items, followed by a second transfer test. At issue was how acquisition and transfer of the 32-item group would compare to the 8-item group when trained and tested at the same set sizes.

Method

Subjects—The subjects were four experimentally naïve White Carneaux pigeons (*Columba livia*) 5–6 years old from the Palmetto Pigeon Plant in Sumter, South Carolina. The pigeons were maintained at approximately 85% of their free-feeding weights and were individually housed with a 12-hr light-dark cycle. Testing was conducted 5–7 days a week.

Apparatus

Chamber: Pigeons were tested in a custom wooden chamber (35.9-cm wide × 45.7-cm deep × 51.4-cm high). A custom-built wooden grain hopper allowed a video monitor (Eizo T550, Ishikawa, Japan; 800 × 600 pixel resolution) and infrared touch screen (17" Unitouch, Carroll Touch, Round Rock, TX) to be better positioned (i.e., lower) on the stimulus panel. Experimental events were controlled and recorded with custom software written in Visual Basic on a Pentium personal computer. A video card (ATI 3D Rage Pro AGP 2X, Ontario, Canada) controlled the monitor. A computer-controlled relay interface (Model no. PIO-12, Metrabyte, Taunton, MA) operated the food hopper and houselight.

Stimuli: The configuration of the stimulus display, the training stimuli for each set size, the test stimuli and stimulus pairings were identical to those previously used for pigeons (Katz & Wright, 2006), rhesus monkeys (Katz et al., 2002), and capuchin monkeys (Wright et al., 2003). The stimulus display for *same* and *different* trials and the stimuli of the 32-item set used to train these subjects (and the previous pigeons and monkeys) is shown in Figure 1. Stimulus displays consisted of two digitized travel-slide color pictures (each 5.7 × 3.8 cm)

and a white rectangle (2.5×2.4 cm) on a black background. The pictures were vertically aligned with a 1.28-cm gap between them. The top picture was centered 20.63 cm from the left edge and 18.75 cm from the top of the cutout in the front panel. The bottom of the white rectangle was horizontally aligned with the bottom of the lower picture and to the right with a 1.4-cm gap between them.

Procedure

Preliminary training: Preliminary training was very similar to Katz & Wright (2006). The pigeons were first trained to eat from the hopper and then autoshaped to white rectangles separately presented in the two lower positions. Once pigeons consistently responded to the white rectangles (2–6 sessions), S/D training began.

S/D task training: The procedures were identical (except the initial training set) to those used previously with the 8-item group including: A correction procedure (CP) on the initial acquisition, training to a performance criterion of 85% with the 32-item training set, use of the same training stimuli as the 8-item group for the 32-item acquisition, and use of the same transfer stimuli and same transfer stimulus pairs as the 8-item group following the 85% criterion with the 32-item training set.

Training on the S/D task began with presentation of a picture in the upper position. A peck to the picture in the upper position was followed by presentation of a lower picture and the white rectangle, leaving the upper picture in view. If the two pictures were *same* (half the trials), a peck to the lower picture was correct. If the two pictures were *different* (other half of the trials), then a peck to the white rectangle was correct.¹

Correct responses were followed by a 0.5-s, 600.6-Hz tone and reinforcement (3–5 s access to mixed grain). Food reinforcement or an incorrect response was followed by a 15-s intertrial interval (ITI). Beginning with the fifth training session, incorrect choices produced a correction procedure consisting of a 15-s timeout (with the house light turned off), the 15-s ITI (house light on) and a repeat of the trial. The correction procedure was repeated until the subject made the correct response. Performance on correction trials (although recorded) did not figure in any analyses presented in this article. Over the first 15 training sessions, the number of responses to the upper picture was gradually increased to 20 pecks.

Sessions consisted of 100 trials (50 *same*, 50 *different*). The sequence of *same* and *different* trials was pseudorandomly constructed and varied from day to day. The items used to construct the displays were selected with replacement from the 32-item set (see Figure 1). If a session was not completed within 3 hr, then it resumed the next day. Pigeons were trained with the correction procedure until their performance met the criterion of 85% on a single

¹This procedure of training animals to touch/peck the lower picture for a *same* response is learned more rapidly than an alternative procedure with two buttons (for *same* and *different*) on either side of the pictures (Bhatt & Wright, 1992). Although responding to the matching stimulus on *same* trials shares some similarity to MTS, in MTS the identical matching stimulus being present on every trial allows the subject to compare the choices and choose the better of the two. In S/D tasks, by contrast, the correct and incorrect stimuli are not present at the same time; they occur on separate trials. Subjects make each decision in isolation, so to speak. The subject decides whether the test was identical to the sample. If so, the subject makes an (arbitrarily designated) “same” response (response to the lower item in our procedure), or if not, it makes an (arbitrarily designated, but different) “different” response (response to the white rectangle in our procedure).

session and then the correction procedure was removed. (One subject that was slow to acquire had the correction procedure removed after 31 sessions, the mean number of sessions for training with the correction procedure in Katz & Wright, 2006.) After removing the correction procedure, subjects were trained until the 85% performance criterion was again attained.

Transfer testing: Each transfer session consisted of 90 baseline training trials (45 *same*, 45 *different*) plus 5 *same* and 5 *different* transfer trials. Novel stimuli, pictures never before seen, were presented on transfer trials. Examples of pictures for eight transfer trials are shown in Figure 1. The 10 transfer trials were pseudorandomly placed within each session following the seventh trial. Transfer testing was conducted for six sessions. A total of 90 novel stimuli were used in this transfer test. These transfer stimuli were used as training stimuli (as space in the training sets permitted) over the course of training set doublings. Correct performance on transfer trials was reinforced. Reinforcing correct transfer performance was appropriate in this experiment because these stimuli were only tested once and never used again as transfer stimuli (by definition of novelty). Indeed, having enough novel transfer stimuli for all transfer tests and training stimulus sets was our rationale for using these ‘travel slides’ in the first place. Conducting transfer tests in extinction has an inherent problem in that subjects can (and do) learn that the appearance of novel stimuli signals extinction, and the resulting poor performance can reflect this extinction cuing, not the ability or inability to learn the abstract concept.

Set-size expansion: Following the transfer test at the 32-item set, the stimulus set size was doubled to 64 items (the same 64 items used for the 8-item group). The pigeons were re-trained to the criterion of 85% correct, and then tested with a novel set of 90 transfer stimuli, the same transfer stimuli and transfer stimulus pairs as the previous 8-item group following acquisition with the 64-item training set (Katz & Wright, 2006). This training-testing cycle was repeated four more times until the set size reached 1024 items. Because 90 novel pictures were used during each transfer test and then later incorporated into training-set expansions, subjects could see as many as 1114 different pictures (1024 training plus 90 transfer stimuli). The training and testing items were the same as those used in our previous experiments with pigeons (the 8-item group) and monkeys.

Results

Acquisition—Figure 2 shows the acquisition functions for the four pigeons in the 32-item group in the upper panel. All subjects readily learned the S/D task ($M = 2,700$ trials) and their acquisition was no less rapid than the 8-item group ($M = 2,900$ trials) trained otherwise identically, as shown in the lower panel of Figure 2. Mean acquisition functions for the two groups (23 sessions where all subjects participated) were compared by a two-way repeated-measures analysis of variance (ANOVA) yielding a main effect of Session, $F(22, 132) = 42.7$, $p < .00001$, $\eta^2 = .88$, but no effect of Group, $F(1, 6) = 1.1$, $p = .34$, or interaction, $F(22, 132) < 1$.

Transfer—Transfer by the 32-item group is shown on the right in Figure 3. The 32-item group showed 73.8% (95% C.I. = ± 6.08) correct transfer to novel stimuli which was

substantially better (14%) than transfer by the 8-item group ($M = 59.6\%$, 95% C.I. = ± 4.39) following learning with the same 32-item training set, $F(1, 6) = 10.1$, $p < .02$, $\eta^2 = .63$, 95% C.I. for the mean difference = -25.05 to -3.28 . This substantially better transfer for the 32-item group shows that prior training with 8 and 16 items (subsets of the 32-item set) apparently changed the way in which the 32-item set was learned by the 8-item group. Further analyses indicated that transfer performance was constant and stable across the six consecutive testing sessions, $F_s(5, 15) < 1.6$, and did not differ on same and different trials, $F_s(1, 3) < 1$, as shown by separate two-way repeated measures ANOVA for each group at each set size.

Set-Size Expansion and Transfer—Despite the substantial transfer of 73.8% following training with the 32-item set performance, this level of transfer performance was less than this group's 83.2% baseline performance as shown by a one-way repeated measures ANOVA of Trial Type (baseline, transfer), $F(1, 3) = 10.1$, $p = .05$, $\eta^2 = .77$, 95% C.I. for the mean difference = -0.02 to 18.81 . Following this transfer test, the 32-item group's training set was progressively doubled along with transfer testing following each acquisition. These transfer results and baseline performances (during transfer testing) are shown in Figure 4 and compared to the 8-item group trained and tested with the same items. The 32-item groups' transfer performance increased from 32 to 256 items relative to baseline performance, as confirmed by two-way repeated measures ANOVA of Trial Type (baseline, transfer) and Set Size (32, 64, 128, 256, 512, 1024) which yielded a significant interaction, $F(5, 15) = 3.1$, $p = .04$, $\eta^2 = .51$. Post-hoc analyses showed that there was a difference between baseline and transfer performance at the set sizes 32–128, $F(1, 3) = 25.4$, $p < .02$, $\eta^2 = .89$, 95% C.I. for the mean difference = 2.59 to 9.25 , but not at the set sizes 256–1024, $F(1, 3) < 1$, 95% C.I. for the mean difference = -4.25 to 1.78 . Further analyses indicated that transfer performance was constant and stable across the six consecutive testing sessions, $F_s(5, 15) < 1.2$.

Transfer performance did not differ for *same* and *different* trials on set sizes 32, 64, 128, & 256 [$F_s(1, 3) < 5.6$], but there was a similar *same* bias (93.8% vs. 73.3%) on set sizes 512 [$F(1, 3) = 121$, $p < .005$, $\eta^2 = .98$] and 1024 [$F(1, 3) = 24$, $p < .02$, $\eta^2 = .89$], as shown by separate two-way repeated measures ANOVA for each set size. It is important to point out that the bias that occurred at set sizes 512 and 1024 did not occur at the earlier set sizes when the subjects were initially learning the S/D task. Biases could have been a concern if the subjects had learned a default strategy, for example, learn *same* trials and default (to the white rectangle) on *different* trials (cf., Roberts & Grant, 1978 for MTS). If that were the case, then there should have been a strong *different*-trial bias on the novel transfer trials with the earlier training sets. But there was no bias with the earlier training sets. Thus, there is no evidence that these subjects were using a default strategy to make an end-run around the relational strategy of comparing the two pictures presented on each trial.

Baseline performances by the 8-item and 32-item groups were comparable over the set size expansion. Both groups showed a gradual increase in transfer performance as the set size increased. The increasing slope for transfer performance by the 32-item group was less than that for the 8-item group, as shown by the standardized coefficient for linear regression on the set-size transfer function for each group (8-item group: $\beta = .679$, adjusted $r^2 = .44$, $t =$

4.7, $p < .01$; 32-item group: $\beta = .439$, adjusted $r^2 = .16$, $t = 2.3$, $p = .03$) and a one-way ANOVA of Group (8, 32) based on the individual slopes of each pigeon yielded a significant Group effect, $F(1, 6) = 12.75$, $p = .012$, $\eta^2 = .68$, 95% C.I. for the mean difference = 0.004 to 0.03. The smaller increase in transfer performance with set-size expansion for the 32-item group was due to this group's better transfer at 32 items and by the transfer of both groups becoming equivalent to baseline at the same set size of 256 items.

Discussion

The results from this experiment show that the 32-item group has a substantial transfer advantage (14%) over the 8-item group for the 32-item set size. This higher degree of abstract-concept learning by the 32-item group suggests that this group's learning of the relationship between the two items presented on each trial spread to other pairs as the task was learned. Spreading of relational learning throughout the set of training pairs is likely responsible for the comparatively rapid acquisition by the 32-item group. The 32-item group learned 1024 different pairs of items no slower than the 8-item group learned 64 pairs of items. What the 8-item group learned that accounted for their lesser transfer performance is unclear. One possibility is that the 8-item group learned the individual item pairs, that is they learned the task item specifically as opposed to relationally. If so, such an explanation would leave unanswered why the 32-item group did not show full concept learning (transfer equivalent to baseline). Clearly the 32-item group did not learn the individual item pairs or else their acquisition rate would likely have been close to 16 times that of the 8-item group ($1032/64 = 16$). Another possibility for what the 8-item group learned would include domain restricted learning. That is, the range of pictures for which they can accurately apply their relational learning is restricted to the small training set and thereby creates resistance to expanding this limited domain. This latter possibility might gain some support if the learning of the 32-item group also had domain-restricted learning relative to another group trained initially, for example, with a 64-item set.

Experiment 2

Experiment 1 showed that the 8-item group had a carryover effect from the smaller training set that adversely affected subsequent transfer. With an initial transfer advantage for the 32-item group, it is somewhat surprising that this group did not achieve full concept learning with a smaller training set than the 8-item group. Why the 32-item group did not go on to achieve full concept learning more rapidly may have had to do with its own detrimental carryover effect, similar to the one that affected (or afflicted) the 8-item group. The alternative, of course, is that 256 items are necessary for any pigeon to fully learn the S/D abstract concept independent of any carryover effects. To explore these possibilities, another group of experimentally-naïve pigeons was trained with a 64-item set to determine whether this size of training set might produce full concept learning.

Method

Subjects—The subjects were 2 experimentally naïve, 6 year old, pigeons maintained similar to the subjects in Experiment 1.

Apparatus and Procedure—The apparatus and procedures were the same as those used in Experiment 1, except that this group began training with the 64-item training set. This was the same 64-item training set used in all of our previous experiments with pigeons and monkeys. They were trained to the criterion of 85% correct, and tested with a novel set of 90 transfer stimuli, the same transfer stimuli and transfer stimulus pairs as the previous 8-item group following acquisition with the 64-item training set (Katz & Wright, 2006). This training-testing cycle was repeated four more times until the set size reached 1024 items. The training and testing items at each set size were the same as those used in our previous experiments with pigeons (the 8-item group) and monkeys.

Results

Acquisition—Acquisition of the S/D task by the two pigeons in the 64-item group is shown in Figure 5. The two pigeons acquired the task at somewhat different rates. Most notably was Pigeon R9137's slowness to meet the 85% acquisition criterion during the later stages of training. The mean acquisition functions for the three groups are shown in the lower portion of Figure 5. The 64-item group's acquisition was slightly below that for the other two groups from session 13 onward, but these differences were not statistically significant, as shown by a two-way repeated measures ANOVA of Session (1 to 23) and Group (8, 32, 64) with adjusted Greenhouse-Geisser degrees of freedom on percentage correct yielding a main effect of Session, $F(3.3, 23.3) = 38, p < .00005, \eta^2 = .84$, and no effect of Group, $F(2, 7) = 1.1, p = .4$, or interaction, $F(17, 23.3) = 1.1, p = .41$. Although the smaller group size for the 64-item group may have contributed to the lack of differences among the acquisition functions, nevertheless any learning by the 64-item group that was tied to specific training pairs would have been expected to drastically lower this group's learning curve because of the far greater number of pairs to be learned. This lack of a substantial difference between the 64-item group's acquisition function and those of the other groups suggests that this group learned more about the general relationship between the items of each pair than did the other groups.

Transfer—Transfer and baseline performances (during transfer testing) by the 64-item group are shown in Figure 6. In this case, transfer performance ($M = 81.7\%$, 95% C.I. = ± 5.31) was not different from their baseline performance ($M = 83.7\%$, 95% C.I. = ± 3.01), as shown by separate one-way ANOVAs comparing Trial Type (baseline, transfer) at each session for each subject, $F_s(1, 10) < 1$. This result is important because it shows, for the first time, that if pigeons are trained from the beginning with a set of 64 items they learn the S/D abstract concept fully. Further analyses indicated that transfer performance was constant and stable across the six consecutive testing sessions, $F < 8.5$, and did not differ on *same* and *different* trials, $F < 4$, as shown by a two-way repeated measures ANOVA with adjusted Greenhouse-Geisser degrees of freedom.

Set-Size Expansion and Transfer—Following this initial transfer test, the 64-item group's training set was progressively doubled along with transfer testing following each acquisition until a training set of 1024 items was reached. These transfer results and baseline performances are shown in Figure 7 and are superimposed upon those from the 32-item and 8-item groups. Baseline and transfer performances were comparable throughout for the 64-

item group and to those achieved by the other groups following full concept learning at the 256-item set, as shown by a three-way repeated measures ANOVA of Group (8, 32, 64), Trial Type (baseline, transfer), and Set Size (256, 512, 1024) which yielded no effects, $F_s < 2.1$, $p_s > .18$. Further analyses indicated that transfer performance was constant and stable across the six consecutive testing sessions for the 64-item group, $F_s < 2.9$, at each set size. Additionally, transfer performance did not differ for *same* and *different* trials with 128, 256, 512 & 1024 set sizes [$F_s < 9$], but there was a *same* bias (96.7% vs. 71.7%) for set size 1024 [$F = 225$, $p < .05$, $\eta^2 = .99$], as shown by separate two-way repeated measures ANOVA for each set size with adjusted Greenhouse-Geisser degrees of freedom.

Discussion

The 64-item group learned the S/D task at a rate similar to those from the 8- and 32-item groups despite there being four times as many different pairs than for the 32-item group and 64 times as many different pairs than for the 8-item group. This result reinforces the conclusion from Experiment 1 that the task was learned relationally by these subjects trained with larger initial training sets, and that each increment of relational learning aids in the learning of all the other pairs encountered in the future. The result is comparatively rapid acquisition.

Moreover, the 64-item group transferred better than the 32-item group and showed full abstract-concept learning at this initial set size and this level of transfer remained stable across subsequent expansions of the training set. The finding of full concept learning at the 64-item set is to be contrasted with that for the 8- and 32-item groups that did not reach full concept learning until the 256-item set. The finding of full concept learning at the 64-item set suggests that with this initial training set size there were no carryover effects and no restrictions on the visual-picture domain. Should the explanation of a restricted/unrestricted visual-picture domain be correct, then pigeons should be able to learn the S/D task with any other larger training set size—no matter how large. Additionally, their levels of baseline and transfer performances should be equivalent to the 64-item group of pigeons.

Experiment 3

Experiment 2 showed that pigeons could fully learn an S/D abstract concept if the training set contained 64 items. This result is to be contrasted with pigeons trained with 8 and 32 items, where those groups required 256 training items to fully learn the concept. Presumably, any set size greater than 64 items should produce full concept learning. But as the training sets become huge there is the possibility that pigeons might have difficulty learning the task because the items would never become very familiar (cf., Wright et al., 1988). The purpose of Experiment 3 was to determine whether or not pigeons could learn the S/D task if they were trained from the beginning with a set of 1024 items and 1,048,576 different stimulus pairs. We are unaware of any study that successfully trained animals in a two-item S/D task with such a large training set.

Method

Subjects—The subjects were 2 experimentally naïve 1-year-old pigeons maintained like those in the previous experiments.

Apparatus and Procedure—The apparatus, preliminary training, S/D training, and transfer testing were identical to Experiment 2 except that the 1024 training set was used from the beginning with this group. This training set and the transfer stimuli that followed acquisition were the same as the 1024 set of training stimuli and transfer stimuli used with the other groups.

Results & Discussion

Acquisition—Figure 8 shows acquisition of the 1024-item S/D task for two pigeons in the upper panel. Acquisition was very rapid for pigeon I16865. Acquisition was considerably slower for pigeon O13667. Although pigeon O13667 was acquiring the task as well as pigeon I16865 up to session 30, this subject struggled by comparison to reach the 85% correct criterion later in acquisition. Nevertheless, pigeon O13667 did learn the task. The 1024-item group's acquisition was virtually indistinguishable from the 64-item group's function (over the 23 sessions where all subjects participated) shown in the lower panel of Figure 8, and was also comparable to the acquisitions by the other groups, as shown by a two-way repeated measures ANOVA of Session (1 to 23) and Group (8, 32, 64, 1024) with adjusted Greenhouse-Geisser degrees of freedom on percentage correct yielding a main effect of Session, $F(3, 23.6) = 35.4, p < .00001, \eta^2 = .82$, and no effect of Group, $F(3, 8) = 1.1, p = .42$, or interaction, $F(8.9, 23.6) = 1.1, p = .29$. While the smaller group sizes for the 64 and 1024 groups may have contributed to the lack of differences among the mean acquisition functions, it is still remarkable that these groups trained at the larger set sizes did not show substantially slower acquisitions compared to those for the smaller training set groups. With over 4000 different pairs for the 64-item group and over one million pairs for the 1024-item group, neither group could have seen all the pairs (even once) during the 23 sessions (2300 trials) of the acquisitions shown in Figure 8. For example, the 1024-item group actually saw only a mean of 1851 pairs (1150 *different* and 701 *same*) before they met criterion. Learning a general abstract rule of S/D is the only way that such acquisition functions could have been produced by subjects with these large numbers of training stimuli.

Transfer—Transfer and baseline performance (during transfer testing) by the 1024-item group is shown in Figure 9. Not surprisingly, transfer performance ($M = 90\%$, 95% C.I. = ± 6.64) was not different from this group's baseline performance ($M = 84.8\%$, 95% C.I. = ± 4.47) as shown by separate one-way ANOVAs comparing Trial Type (baseline, transfer) at each session for each subject, $F_s(1, 10) < 2.52, p > .14$. Further analyses indicated that transfer performance was constant and stable across the six consecutive testing sessions, $F < 1$, and did not differ on *same* and *different* trials, $F < 1$, as shown by a two-way repeated measures ANOVA with adjusted Greenhouse-Geisser degrees of freedom. The 1024-item group's transfer and baseline performance was also comparable to that of the other groups at the 1024-item set as shown in Figure 10 and by a two-way repeated measures ANOVA of Group (8, 32, 64, 1024) and Trial Type (baseline, transfer), which yielded no effects, $F_s < 3.1, p_s > .11$.

General Discussion

The experiments of this article show that the size of the initial training set greatly affected the level of the initial transfer. The initial levels of transfer by these four groups show a trend that spans virtually the entire range of performance that might be obtained (51.3% to 90% correct). Both the 64-item group as well as the 1024-item group showed full transfer at their initial training sets (i.e., transfer equivalent to baseline). By contrast, the 8-item group showed virtually no transfer (i.e., not statistically different from chance) and the 32-item group was in between, showing partial transfer. Despite a lack of transfer by the 8-item group and partial transfer by the 32-item group, transfer in both groups was equivalent to their training-trial performance when their expanded training set reached 256 items. This eventual full concept learning was maintained as the set-size was further expanded and the absolute value of their transfer was equivalent to transfer from the 64- and 1024-item groups. These findings of stable, high-accuracy transfer and baseline performances attest to the pigeon's ability to learn an abstract S/D concept and maintain it over months of testing and training set-size with expansions.

Previously, we had characterized the pigeon's S/D abstract-concept learning as qualitatively similar to rhesus and capuchin monkeys in that they had achieved full concept learning with transfer equivalent to baseline, but qualitatively different from the monkeys in that the pigeons' transfer grew at a slower rate with set size than the monkeys' transfer. These comparisons are shown in Figure 11. The experiments of the present article further reinforce the finding of qualitative similarity between pigeons and monkeys; pigeons, like monkeys, can fully learn an S/D abstract concept. Even the quantitative difference of a greater number of training exemplars for pigeons to achieve full concept learning may be questionable given the results of the present experiments.

The quantitative difference between pigeons and monkeys might disappear if pigeons' carryover effects were eliminated. In Figure 12 we compare initial transfer by the 8-, 32-, and 64-item groups of pigeons from the present article with transfer performances by the rhesus and capuchin monkeys at the same training set sizes, trained with the same items and tested with the same transfer items. None of the groups transferred with the 8-item set. Transfer by the groups of pigeons without any prior carryover effects were midway between the capuchin and rhesus monkeys at the 32- and 64-item sets as confirmed by separate one-way ANOVAs at set sizes 8, 32, 64: $F_s < 3.9$, $p_s > .09$. These similar pigeon and monkey transfer results are in stark contrast to the pigeon transfer results shown in Figure 12 with carryover effects. Expanding an initial 8-item training set has a severe detrimental effect on later transfer. Of course, there is no evidence about whether monkeys have carryover effects. Nevertheless, our experience is that monkeys would be less likely to have carryover effects, and certainly not of the magnitude seen here with pigeons. In any case, these comparisons show that pigeons do indeed have the cognitive ability to learn an S/D abstract concept quantitatively as well as qualitatively similar to old and new-world monkeys.

What produces the pigeon's carryover effects and resulted in the suggestion that pigeons may need more exemplars of the rule than monkeys to learn the S/D abstract concept is unclear. But there are indications. Historically, the focus has been on item-specific learning

for what subjects learn when they do not learn the abstract concept. The animal cognition literature is rife with elaborate discussion and theory (i.e., speculation) about pigeons (and other animals) learning item-specific association such as the gestalt configurations of the displays or if-then rules for each stimulus combination (e.g., Carter & Werner, 1978; Premack, 1978; Zentall & Hogan, 1974). If pigeons (or monkeys) were learning item-specific associations with the smaller set sizes and changed to relational learning when the number of pairs became too great to be learned, then there should have been an abrupt change in transfer performance. But all of the set-size transfer functions (e.g., Figure 11) show a continuous and gradual increase in transfer without any abrupt change. Moreover, gradual transfer increases occurred for all the individual subjects in the present article as well as our previous tests of pigeons and monkeys.

An extension of the item-specific view of learning is that transfer is due to generalization from these learned item-specific associations (e.g., Mackintosh, 2000). The generalization hypothesis of abstract-concept learning goes hand-in-hand with the hypothesis of item-specific learning. Despite being promulgated for more than 30 years, there had been no tests of the generalization hypothesis (or the item-specific learning hypothesis). We tested the generalization hypothesis by looking at how long it should have taken pigeons and monkeys to learn each newly expanded training set given their prior learning and degree of transfer as the training set was expanded (Wright & Katz, 2007). By progressively doubling the training set it was predicted that trials-to-acquisition should rise steeply according to the generalization hypothesis (and item-specific learning hypothesis) because the number of new pairs to-be-learned were increasing so rapidly and transfer was only gradually increasing. The results, on the other hand, were just the opposite of the generalization predictions. Trials-to-acquisition actually decreased, with the net result that there was as much as a two log unit difference between predicted and obtained data. Indeed, at the largest set sizes the subjects (pigeons and monkeys) reached the 85% criterion well before the majority of stimuli, let alone the different pairs, had even been seen once (also see Bodily, Katz, & Wright, 2008 for a similar finding with the MTS procedure). A similar analysis conducted on the present 32-item group, the group that showed partial concept learning, found that for expanded set sizes there were no significant increases in trials to acquisition (*Ms*: 475, 1175, 675, 325, 550). The only way such a finding could have occurred was if these subjects were learning the S/D task relationally, that is, learning the S/D abstract concept.

If subjects in the S/D task are not learning item-specific associations of the individual training pairs, then they must have been learning the task relationally—even with the small training set of 8 items. This is a strong, if not profound, conclusion. If they are learning relationally with the 8-item training set, then why do they not transfer this relational learning to novel stimuli? Consider the possibility that their lack of novel-stimulus transfer is due to a restricted domain within which they can apply their relational learning. Domain-restricted relational learning would account for the lack of initial transfer and for the gradual growth of transfer as the set size was expanded (as opposed other explanations such as a blend of item-specific learning and relational learning, or generalization from item-specific learning). Additionally, domain-restricted relational learning might help explain the detrimental

carryover effects by pigeons shown in the experiments of the present article. Subjects making relational judgments within a restricted domain (e.g., the 8 pictures of the 8-item group) would need to expand their domain and apply the same relational judgments (i.e., decision strategy) to other new pictures as the set size was expanded. Pigeons appear more reluctant to expand their domain for relational judgments than monkeys as shown by the pigeon's trials-to-acquisition decreasing less rapidly than those for monkeys as the set size was expanded (Wright & Katz, 2007).

The proposal that learning may begin as relational learning but with a restricted domain may apply to human learning as well as animal learning. Even adult humans show evidence of their domain being restricted when they first learn some relational tasks. For example, learning to obtain a specific amount of water using three sizes of jars may not initially transfer to length, area, or weight even when the strategy (i.e., formula) is the same (e.g., Chen & Mo, 2004). Likewise, investigations of children commonly claim that learning is initially tied to stimulus features (e.g., within dimensional size comparisons) which later in development gives way to relational factors (e.g., across dimensional size and brightness comparisons) (e.g., Domus, Hummel, & Sandhofer, 2008; Kotovsky & Gentner, 1996). But when performance is tied to stimulus features it can be unclear whether this performance is an example of item-specific learning or domain-restricted relational learning. Although there is a wide acceptance that failures of novel-stimulus transfer mean that the subjects (humans or animals) are learning the task item specifically, here too in human learning there appears to be a dearth of item-specific learning tests. Because item-specific learning is so widely accepted as the explanation for concept-learning failure, research into alternatives appears to have been prematurely short-circuited by this accepted explanation. In theory, it would seem (at least to us) that the original training items could be manipulated to determine whether or not learning was of the item-specific learning variety or domain-restricted relational learning.

Acknowledgments

Research supported by NIH grants MH-061798, MH-072616, DA10715, and NSF grant IBN-0316113.

References

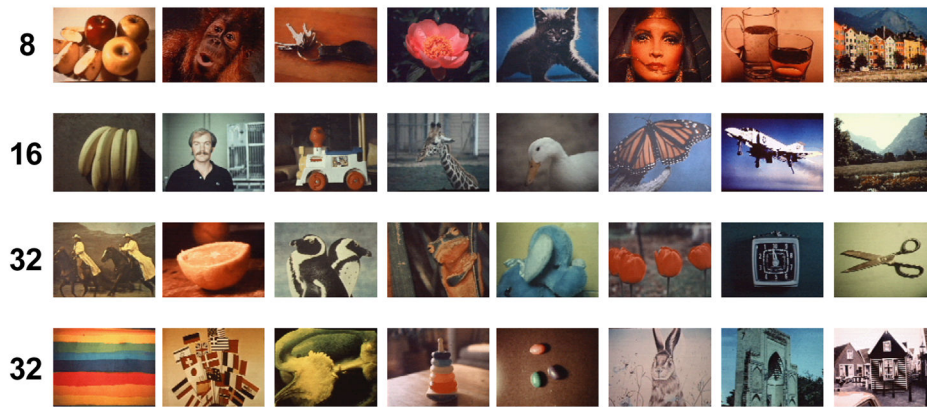
- Blaisdell AP, Cook RG. Two-item *same-different* concept learning in pigeons. *Learning & Behavior*. 2005; 33:67–77. [PubMed: 15971494]
- Bodily KD, Katz JS, Wright AA. Matching-to-sample abstract-concept learning by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*. 2008; 34(1):178–184. [PubMed: 18248124]
- Carter DE, Werner TJ. Complex learning and information processing by pigeons: A critical analysis. *Journal of the Experimental Analysis of Behavior*. 1978; 29:565–601. [PubMed: 16812079]
- Chen Z, Mo L. Schema induction in problem solving: A multidimensional analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2004; 30:583–600.
- Cook RG, Kelly DM, Katz JS. Successive two-item same-different discrimination and concept learning by pigeons. *Behavioral Processes*. 2003; 62:125–144.
- Cumming WW, Berryman R, Cohen LR. Acquisition and transfer of zero-delay matching. *Psychological Reports*. 1965; 17:435–445. [PubMed: 5833264]
- Doumas LAA, Hummel JE, Sandhofer CM. A Theory of the Discovery and Predication of Relational Concepts. *Psychological Review*. 2008; 115:1–43. [PubMed: 18211183]

- Farthing GW, Opuda MJ. Transfer of matching-to-sample in pigeons. *Journal of the Experimental Analysis of Behavior*. 1974; 21:199–213. [PubMed: 16811738]
- Katz JS, Wright AA. Mechanisms of *same/different* abstract-concept learning by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*. 2006; 32:80–86. [PubMed: 16435967]
- Katz JS, Wright AA, Bachevalier J. Mechanisms of *same/different* abstract-concept learning by rhesus monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*. 2002; 28:358–368. [PubMed: 12395493]
- Kotovsky L, Gentner D. Comparison and categorization in the development of relational similarity. *Child Development*. 1996; 67:2797–2822.
- Mackintosh, NJ. Abstraction and discrimination. In: Heyes, C.; Huber, L., editors. *The evolution of cognition*. Cambridge, MA: MIT; 2000. p. 123-141.
- Premack, D. On the abstractness of human concepts: Why it would be difficult to talk to a pigeon. In: Hulse, SH.; Fowler, H.; Honig, WK., editors. *Cognitive processes in animal behavior*. Hillsdale, NJ: Erlbaum; 1978. p. 423-451.
- Roberts WA, Grant DS. Interaction of sample and comparison stimuli in delayed matching to sample with the pigeon. *Journal of Experimental Psychology: Animal Behavior Processes*. 1978; 4:68–82.
- Santi A. The role of physical identity of the sample and correct comparison stimulus in matching-to-sample paradigms. *Journal of the Experimental Analysis of Behavior*. 1978; 29:511–516. [PubMed: 16812074]
- Santiago HC, Wright AA. Pigeon memory: Same/different concept learning, serial probe recognition acquisition, and probe delay effects on the serial-position function. *Journal of Experimental Psychology: Animal Behavior Processes*. 1984; 10:498–512. [PubMed: 6491609]
- Wasserman EA, Hugart JA, Kirkpatrick-Steger K. Pigeons show same–different conceptualization after training with complex visual stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*. 1995; 21:248–252. [PubMed: 7602259]
- Wright AA, Cook RG, Rivera JJ, Sands SF, Delius JD. Concept learning by pigeons: matching-to-sample with trial-unique video picture stimuli. *Animal Learning and Behavior*. 1988; 16:436–444.
- Wright AA, Katz JS. Mechanisms of *same/different* concept learning in primates and avians. *Behavioural Processes*. 2006; 72:234–254. [PubMed: 16621333]
- Wright AA, Rivera JJ, Katz JS, Bachevalier J. Abstract-concept learning and list-memory processing by capuchin and rhesus monkeys. *Journal of Experimental Psychology: Animal Behavior Processes*. 2003; 29:184–198. [PubMed: 12884678]
- Young ME, Wasserman EA. Entropy detection by pigeons: Response to mixed visual displays after same-different discrimination training. *Journal of Experimental Psychology: Animal Behavior Processes*. 1997; 23:157–170. [PubMed: 9095540]
- Zentall TR, Hogan DE. Abstract concept learning in the pigeon. *Journal of Experimental Psychology*. 1974; 102:393–398.

Examples: Training Displays



32-Item Set Training Stimuli



Examples: Transfer Pairs Tested with the 32-Item Set

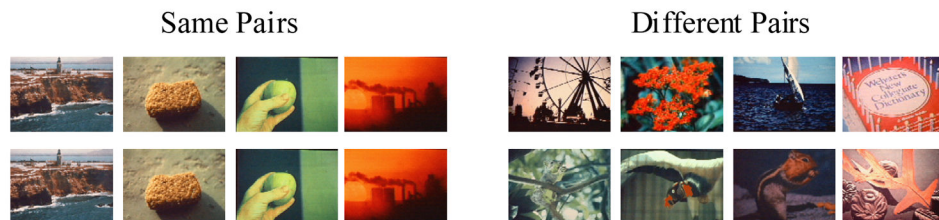


Figure 1. Top: Examples of *same* and *different* trial displays. Middle: The 32 pictures of the 32-item training set. Numbers on left indicate cumulative rows of pictures used for the 8-item, 16-item, and 32-item training sets. Bottom: Four examples of *same* and *different* pairs (from 30 pairs of each type) used to test transfer following acquisition with the 32-item set. The complete set of transfer pairs used in this test, along with those of other tests, and other training stimulus sets can be seen in color in Wright & Katz (2006).

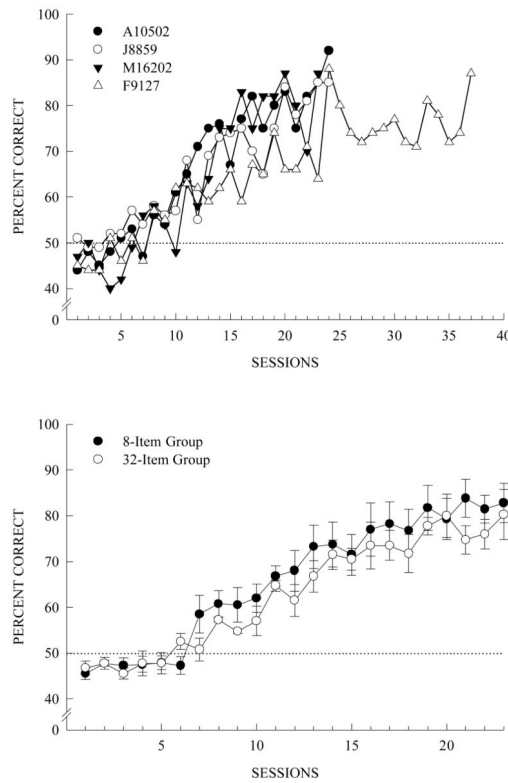


Figure 2. Acquisition of the 32-item S/D task by four pigeons in the top panel. The mean acquisition for the 23 sessions in which the four pigeons in the 32-item group participated compared to the mean acquisition for four pigeons in an 8-item group (Katz & Wright, 2006) in the bottom panel. Dotted line is chance performance. Error bars are standard errors of the mean.

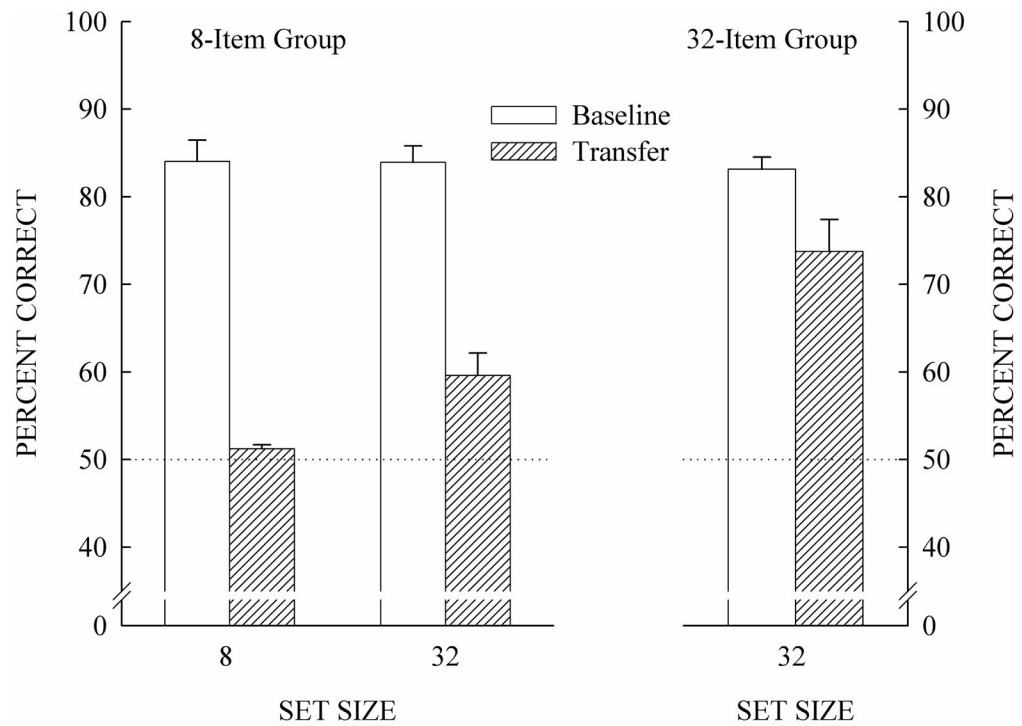


Figure 3. Mean baseline and transfer performance for the 32-item group on the right compared to mean baseline and transfer performance by the 8-item group on the left following learning with the 8-item and 32-item training sets. Dotted line is chance performance. Error bars are standard errors of the mean.

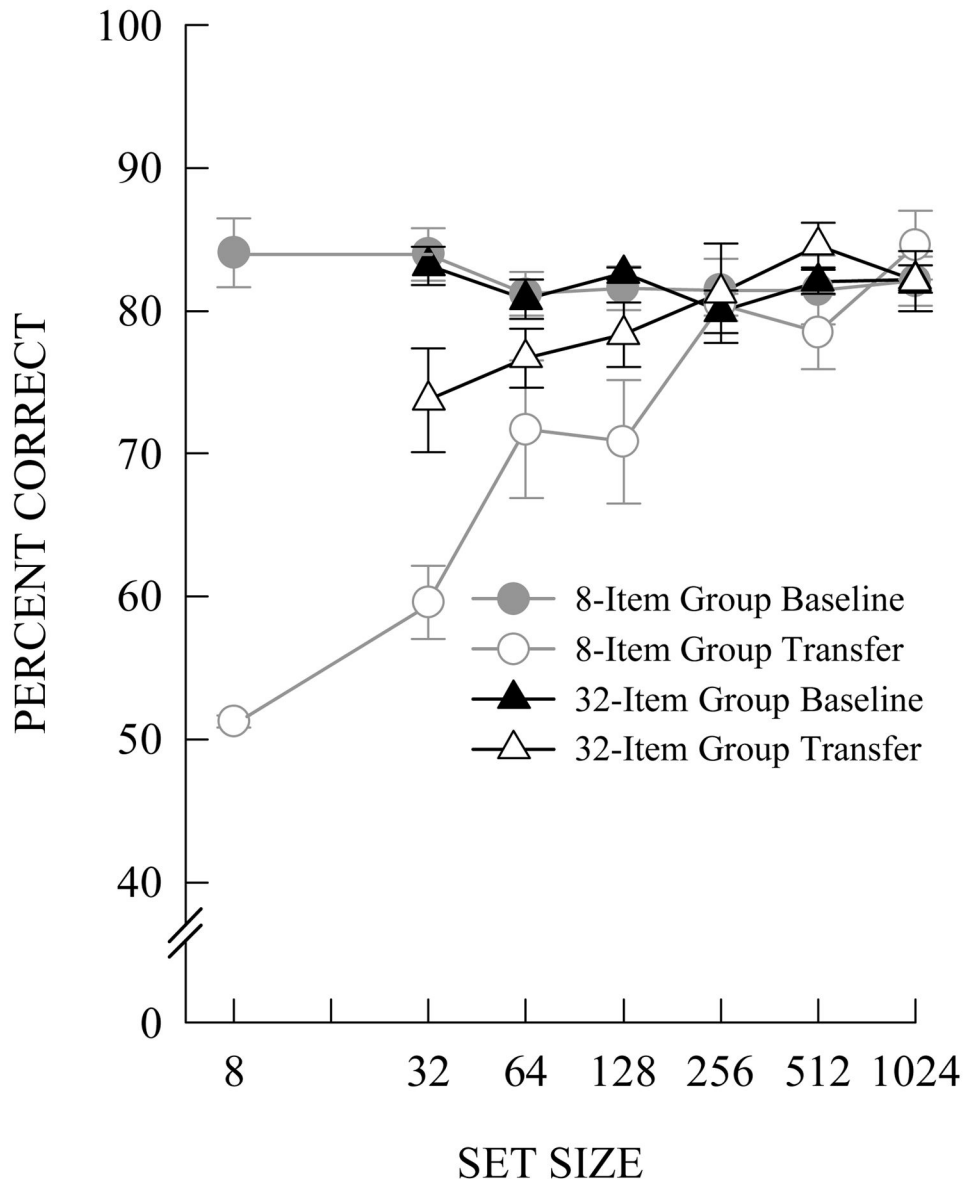


Figure 4. Mean baseline and transfer performance for the 32-item group (black) compared to mean baseline and transfer performance for the 8-item group (gray) as the training sets were progressively expanded to 1024 items. Error bars are standard errors of the mean.

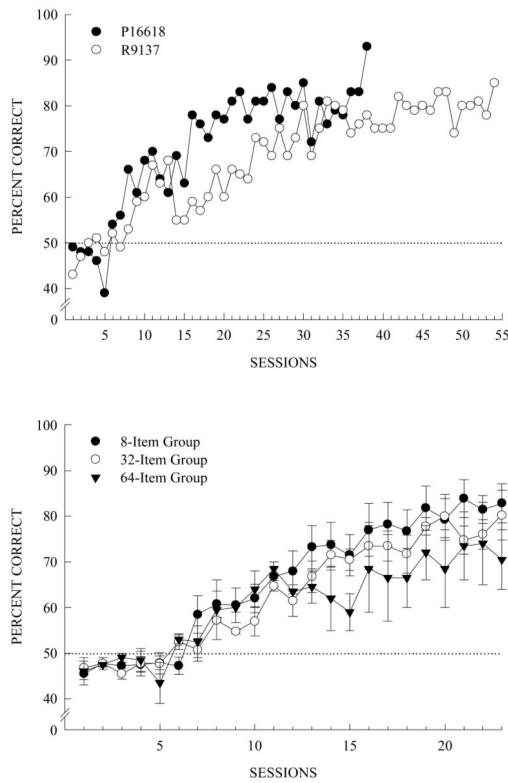


Figure 5. Acquisition of the 64-item S/D task by two pigeons in the top panel. The mean acquisitions for the 64-item group compared to those for the 8-item and 32-item groups in the bottom panel. Dotted line is chance performance. Error bars are standard errors of the mean.

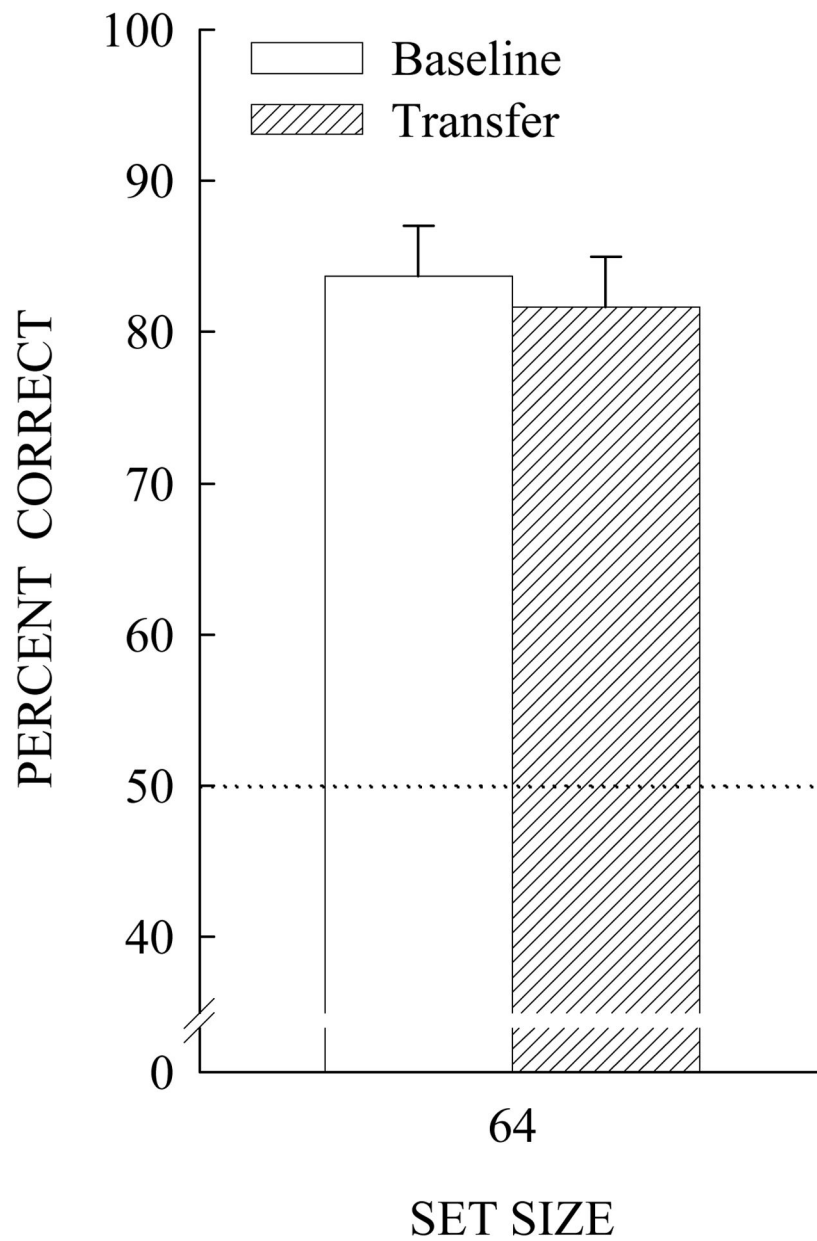


Figure 6. Mean baseline and transfer performance for the 64-item group. Dotted line is chance performance. Error bars are standard errors of the mean.

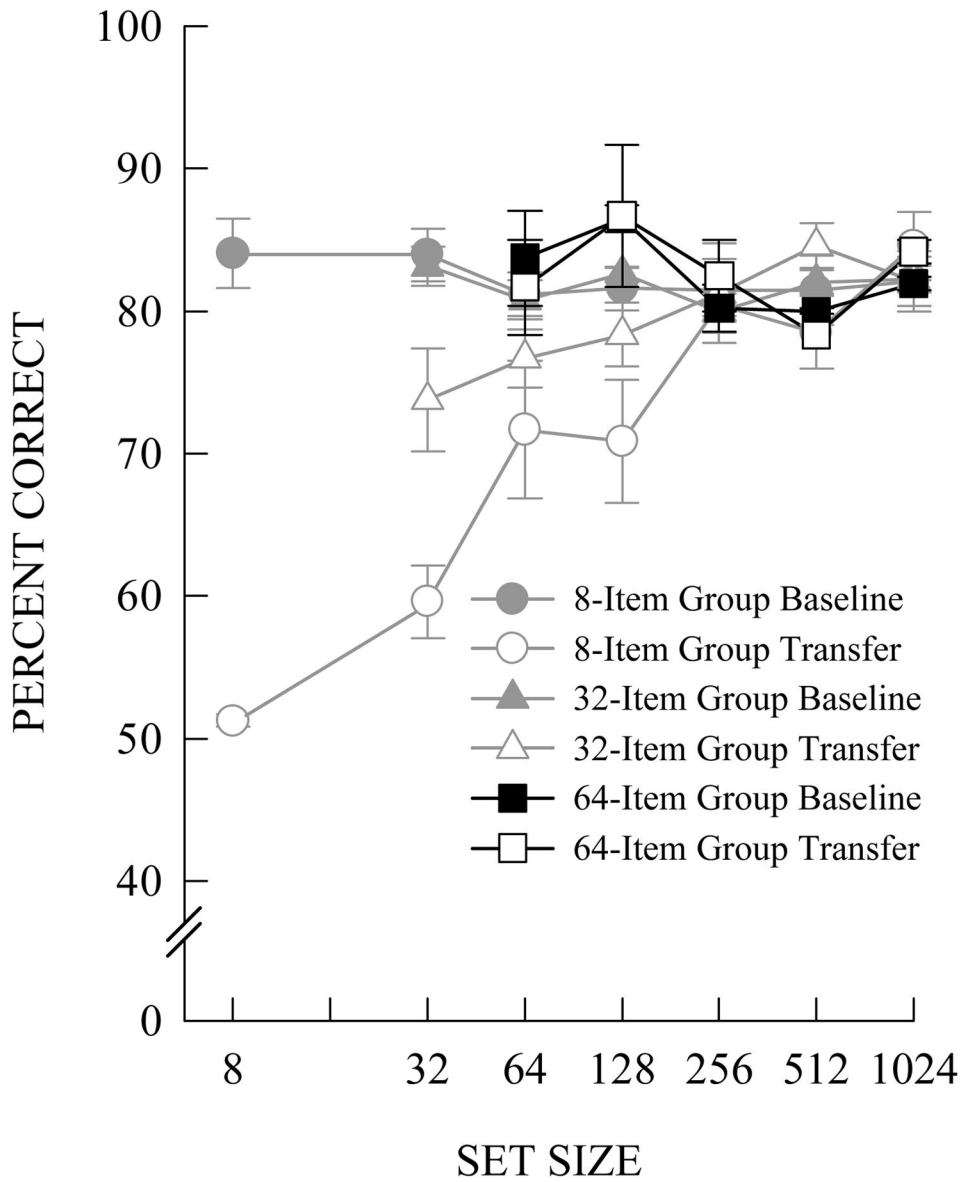


Figure 7. Mean baseline and transfer performance for the 64-item group (black) compared to mean baseline and transfer performance for the 8- and 32-item groups (gray) as the training sets were progressively expanded to 1024 items. Error bars are standard errors of the mean.

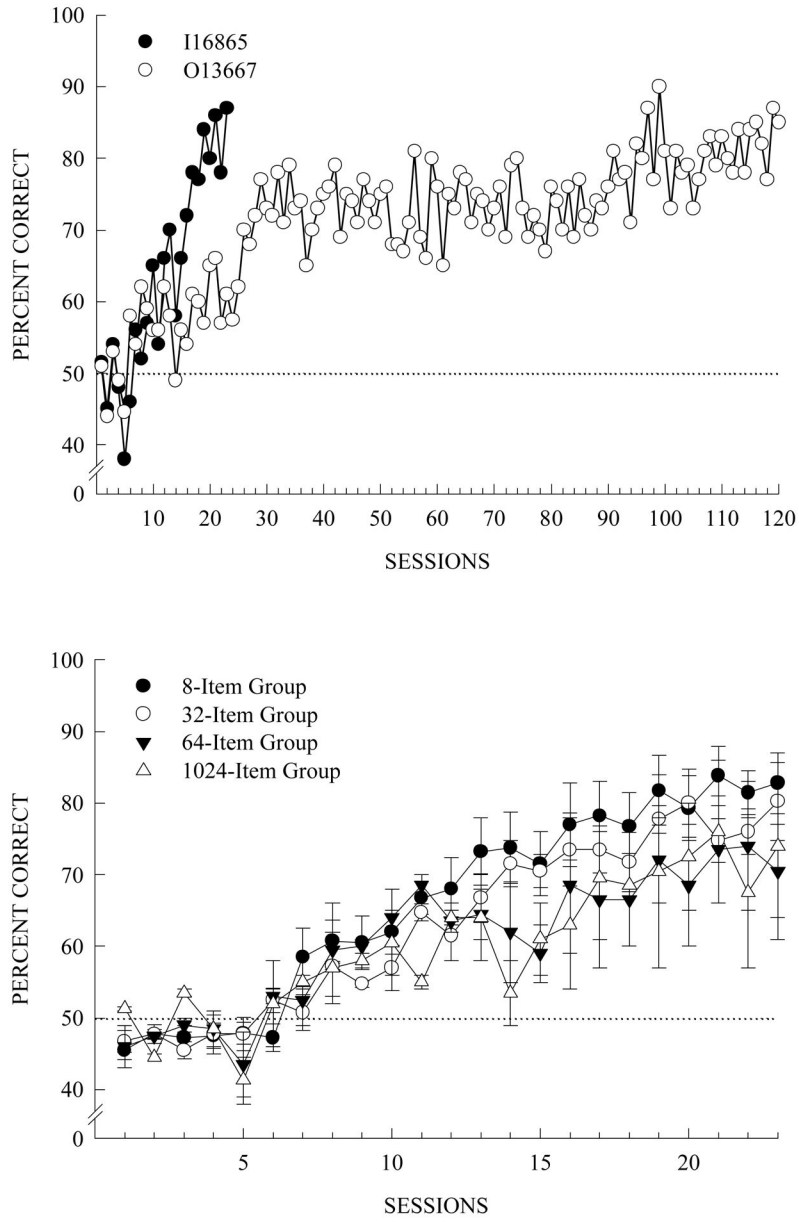


Figure 8. Acquisition of the 1024-item S/D task by two pigeons in the top panel. The mean acquisitions for the 1024-item group compared to those for the 8-item, 32-item, and 64-item groups in the bottom panel. Dotted line is chance performance. Error bars are standard errors of the mean.

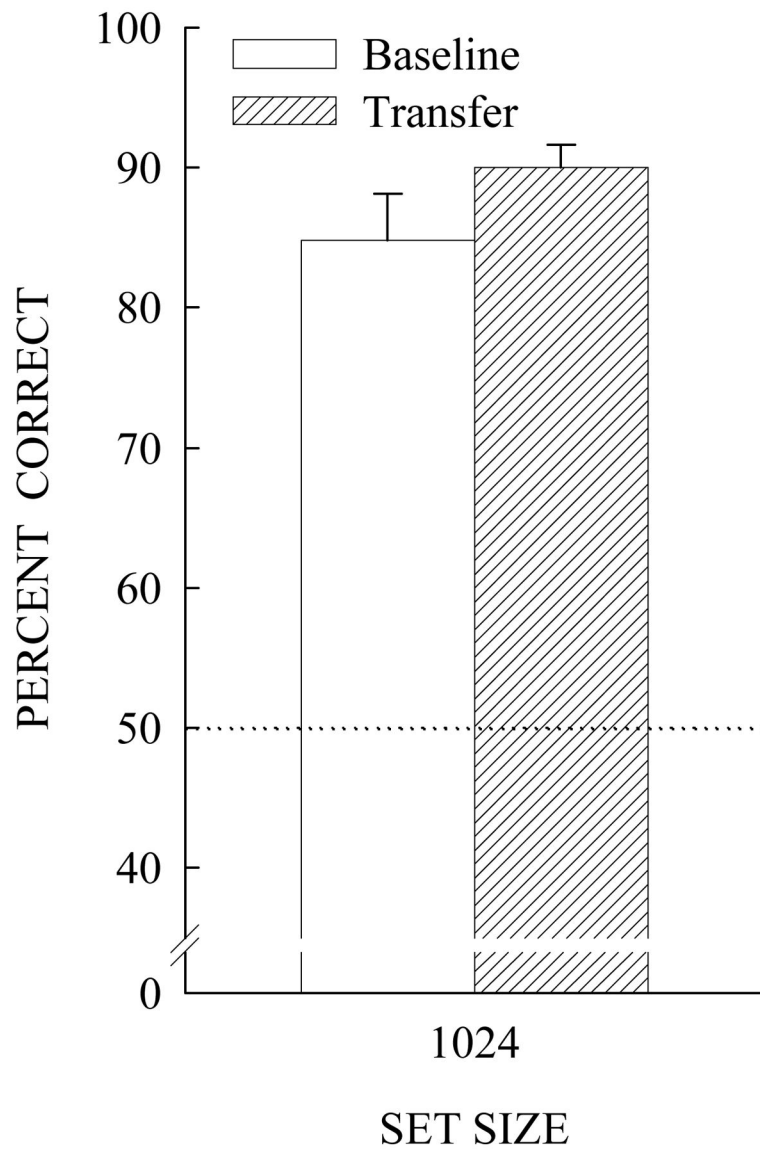


Figure 9. Mean baseline and transfer performance for the 1024-item group. Dotted line is chance performance. Error bars are standard errors of the mean.

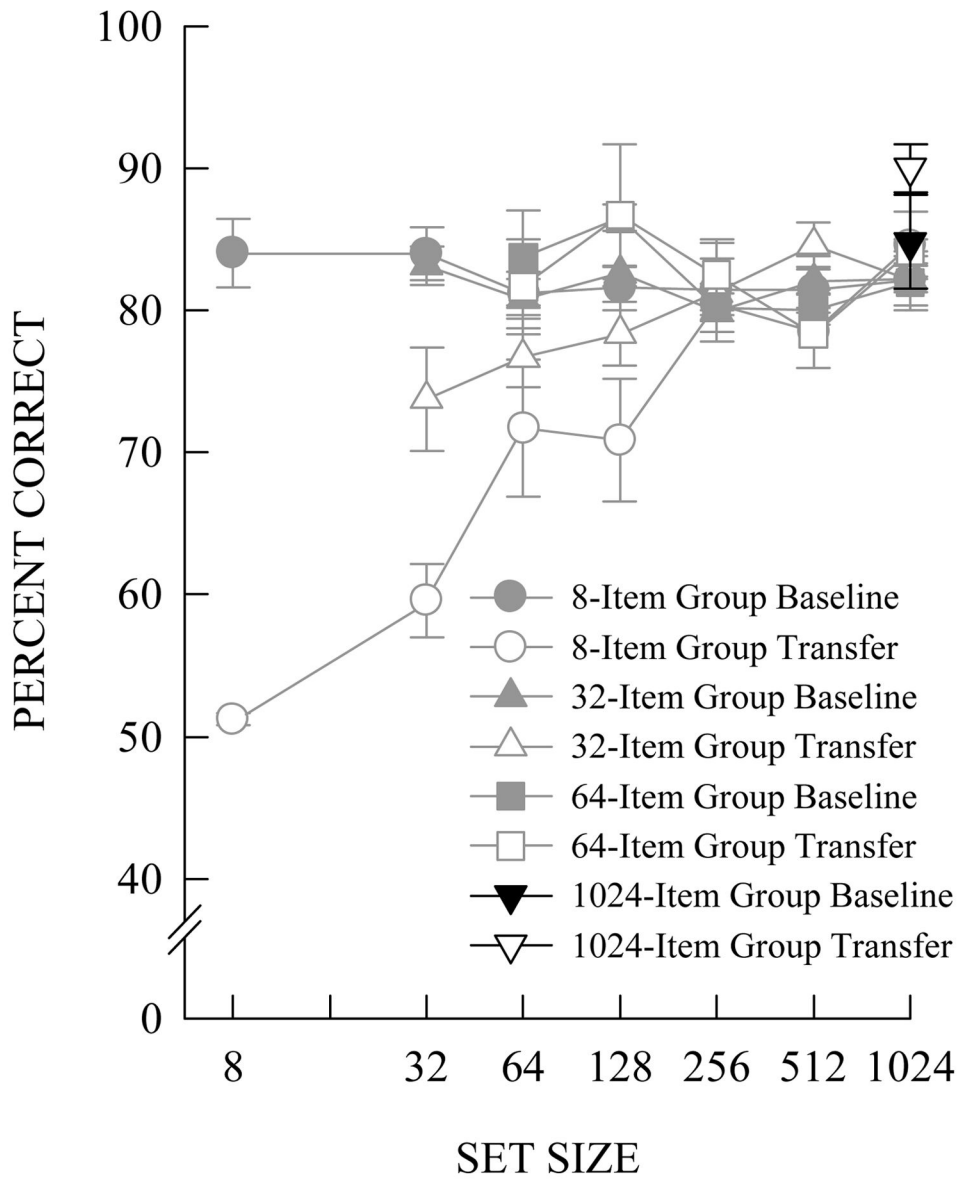


Figure 10. Mean baseline and transfer performance for the 1024-item group (black) compared to mean baseline and transfer performance for the 8-, 32-, and 64-item groups (gray) when their training sets were progressively expanded to 1024 items. Error bars are standard errors of the mean.

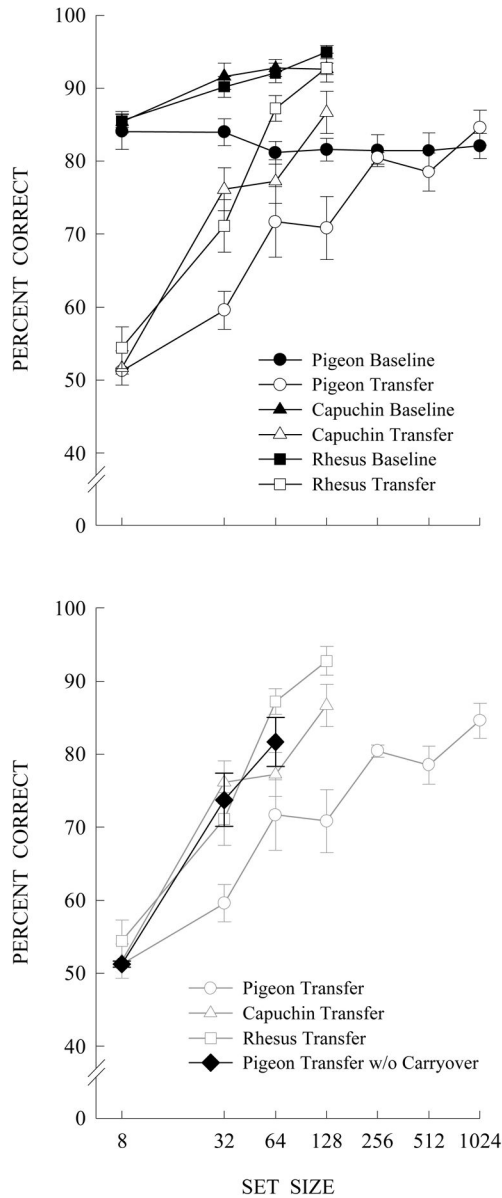


Figure 11. Top: Set-size functions for baseline and transfer performances for pigeons, rhesus monkeys, and capuchin monkeys as their training sets were progressively expanded from 8 items to 128 items (monkeys) or to 1024 items (pigeons). Error bars are standard errors of the mean. Bottom: Transfer by pigeons trained initially with either 8, 32, or 64 items. The resulting set-size function (black diamonds) compared to the set-size functions for groups in the top panel shows a lack of carryover effects that occurred from expanding an initial 8-item set.