

IDENTITY MATCHING-TO-SAMPLE WITH OLFACTORY STIMULI IN RATS

TRACY PEÑA, RAYMOND C. PITTS, AND MARK GALIZIO

DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF NORTH CAROLINA AT WILMINGTON

Identity matching-to-sample has been difficult to demonstrate in rats, but most studies have used visual stimuli. There is evidence that rats can acquire complex forms of olfactory stimulus control, and the present study explored the possibility that identity matching might be facilitated in rats if olfactory stimuli were used. Four rats were trained on an identity match-to-sample procedure with odors mixed in cups of sand as stimuli. Digging in the sample cup produced two comparison cups, and digging in the comparison cup that contained the same scent as the sample was reinforced. When criterion accuracy levels were reached, novel stimuli were added to the baseline training regimen. All 4 rats reached terminal performance of above 90% correct matching with more than 20 different baseline stimuli and matched novel stimulus combinations with above-chance accuracy; 3 of the 4 rats matched novel stimuli at levels significantly above chance. Accurate matching performance was demonstrated both with 2- and 3-comparison procedures. These results suggest that generalized matching-to-sample can be observed in rats when olfactory stimuli are used and, furthermore, that multiple-exemplar training may be important for its emergence.

Key words: Matching-to-sample, concept learning, discrimination, olfactory learning, stimulus control, digging, rat

Identity matching-to-sample (MTS) is a widely used procedure in studies of concept learning. MTS procedures generally involve the presentation of a sample stimulus along with two or more comparison stimuli. In identity matching, a response to the comparison stimulus that is physically identical to the sample is reinforced (S+) and a response to any non-identical comparison is not (S-). These conditional discriminations are readily acquired by pigeons and monkeys, but such performance alone is not sufficient to confirm the learning of an "identity concept" because various alternative sources of stimulus control are possible (e.g., control by specific stimulus configurations, see Carter & Werner, 1978; Cumming & Berryman, 1965). Tests with novel stimuli to determine whether matching performance transfers, or generalizes, are com-

monly used to provide evidence of control by the identity relation between sample and comparison and such transfer is often called generalized matching (Dube, McIlvane, & Green, 1992).

Evidence supporting generalized matching has been obtained in a number of non-human species including old- and new-world primates (Barros, Galvao, & McIlvane, 2002; Katz, Wright, & Bachevalier, 2002; Mishkin & Delacour, 1975; Oden, Thompson, & Premack, 1988), sea lions (Kastak & Schusterman, 1994; Pack, Herman, & Roitblat, 1991), dolphins (Herman & Gordon, 1974) and pigeons (Cook, Katz, & Cavoto, 1997; Wright, 1997; Wright, Cook, Rivera, Sands, & Delius, 1988; Zentall, Edwards, Moore, & Hogan, 1981), but such effects have been more difficult to demonstrate in rats. For example, Iversen (1993, 1997) found that extensive training was required for rats to develop accurate conditional responding with visual stimuli, and found no evidence of control by the identity relation.

In contrast, studies using go/no-go variations of the MTS procedure with olfactory stimuli in rats have had success in demonstrating stimulus control by identity or difference relations (Lu, Slotnick, & Silberberg, 1993; Otto & Eichenbaum, 1992). For example, Lu et al. used an olfactometer to present rats with

Tracy Peña, Raymond C. Pitts, and Mark Galizio, Department of Psychology, University of North Carolina at Wilmington.

This research was supported in part by Grant DA12879 (M.G.) and is based on a Master's thesis by T.P. under the supervision of M.G. The authors thank Erica Blackwell for helpful comments on the manuscript and Patrick McKinney and David Peña for assistance in various phases of the project.

Corresponding author: Mark Galizio, Department of Psychology, University of North Carolina at Wilmington, Wilmington, NC 28403 (e-mail: galizio@uncwil.edu).

doi: 10.1901/jeab.2006.111-04

successive pairs of odors. Licking a drinking tube was reinforced with water delivery if the odors were identical, but not when they were different. Although fairly extensive training was required to reach high levels of accuracy on the initial set of stimuli, accurate performance developed very rapidly with subsequent sets of stimuli. Two of the three rats tested showed 90% correct or higher levels of accuracy on the first session with a novel stimulus set. However, because these performances were averaged across the session (180–200 trials), rapid learning of specific stimulus configurations, rather than generalized control by identity relations, might be invoked to account for the results of the Lu *et al.* study.

The rapid learning and high level of accuracy noted by Lu *et al.* (1993) suggest that methods involving olfactory stimuli may prove fruitful in the study of stimulus control in rats. The present study was an attempt to develop a simple olfactory MTS procedure. We adapted a procedure developed by Eichenbaum and his colleagues (Bunsey & Eichenbaum, 1996; Dudchenko, Wood, & Eichenbaum, 2000; Dusek & Eichenbaum, 1998; Wood, Dudchenko, & Eichenbaum, 1999; see also Mihalick, Langlois, Krienke, & Dube, 2000) in which rats were trained to dig in cups of scented sand to obtain buried food reinforcement. In the present study, digging in a sample cup led to the presentation of comparison cups, which were scented with the same odor as the sample (S+) or a different one (S-). Novel stimuli were added to the baseline when performance criteria were met to assess generalized matching, and accuracy on trials with novel stimuli and novel combinations of stimuli formed a critical part of the analyses.

METHOD

Subjects

Subjects were 4 male Holtzman (Sprague-Dawley) albino rats approximately 6 months old at the beginning of the experiment. Water was available *ad lib*, but access to food was restricted to a 1-hr period each day that began approximately 1 hr after testing was completed. Rats were housed individually and were maintained under a 12/12 hr reversed light/dark cycle.

Apparatus

An operant chamber (28 cm long \times 26 cm wide \times 30 cm high) was modified for use in the study. The front wall of the chamber was constructed of clear Plexiglas that allowed the experimenter to observe the rat. A 5-cm section was removed from the bottom of the front wall of the chamber so that a removable plastic tray could be inserted. The tray that was used with the two-comparison MTS procedure was constructed of clear plastic and had three 5-cm holes drilled into the top. The holes held clear 2-oz plastic cups. The cups were arranged in an inverted triangle (from the rat's view) such that when the tray was inserted approximately 12 cm into the apparatus, a single cup (sample) was accessible to the rat. When the tray was completely inserted, the two comparison cups (8 cm apart) also were accessible (see Figure 1). A second tray used for the three-comparison procedure had four 5-cm holes with a sample hole placed in the same position as with the two-comparison tray and three comparison holes arranged in a row with the center hole directly behind the sample and left and right comparison holes 3 cm on either side.

Stimuli

Olfactory stimuli were generated by mixing household spices (Great American Spice Co.) or liquid extracts (Durkee) with sterilized play sand. See Table 1 for a list of the odorants used, and the sequence with which the stimuli appeared for each subject. Spices were mixed at a ratio of 1 g of spice per 100 g of sand. This ratio of spice to sand was chosen because pilot research with several spices (celery, cinnamon, coffee, garlic, ginger, mustard, onion, paprika, and sage) suggested that these spice levels were sufficient to mask the scent of the sucrose pellet. In these pilot studies several rats, including J6, J11 and J16 of the present study, were exposed to two cups containing the same spice with only one cup baited. No evidence of above-chance pellet detection was obtained at the 1 g spice/100 g sand concentrations. For liquid odorants, the ratio consisted of two drops of liquid per stimulus cup (these odorants were added later in the experiment as the availability of novel spices became limited). Stimulus cups were filled to approximately 1 cm below the rim with sand. The

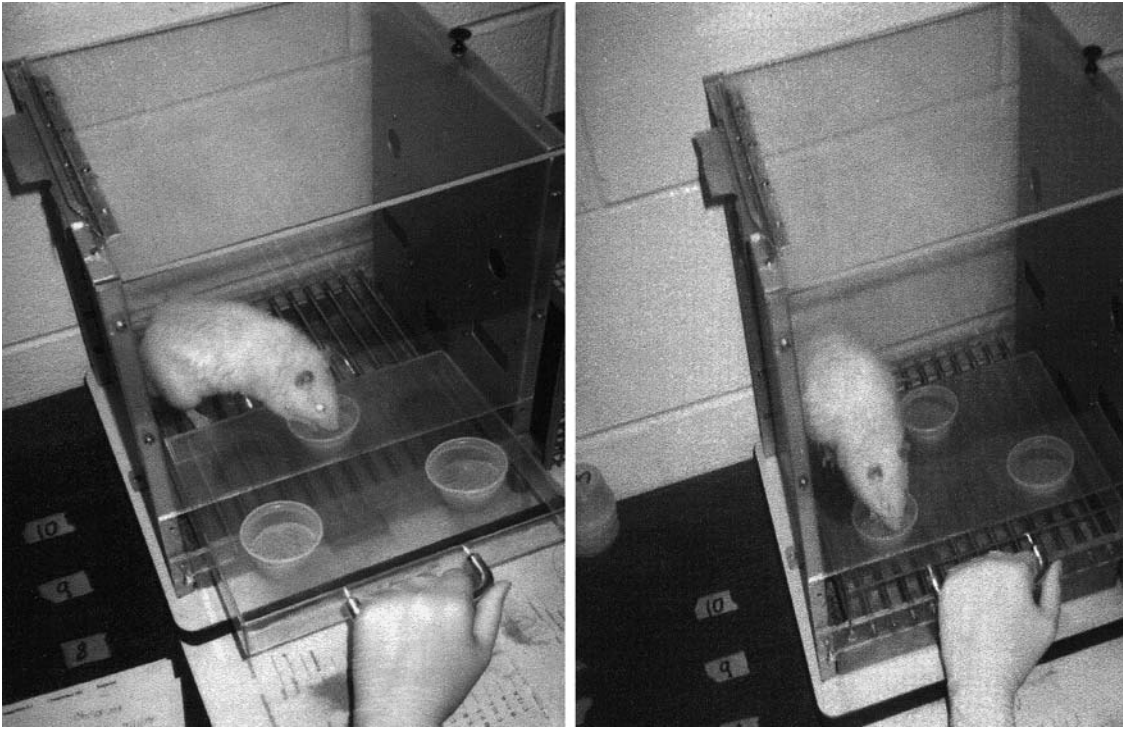


Fig. 1. Photographs of the two-comparison apparatus are shown with the rat at the sample cup in the left panel and at one of the comparison cups in the right panel.

experimenter wore latex gloves during preparation of the stimuli and tweezers were used to place sucrose pellets 1 cm deep in stimulus cups; the tweezers also were inserted into unbaited cups prior to each trial. In a few instances, subjects persistently failed to dig in sample cups when scented with a particular odorant. When this occurred, use of that odorant was discontinued for that animal (see Table 1).

Procedure

Testing usually was conducted 5 days per week (M-F) in the presence of a 70-dB continuous white noise.

Pretraining. Initial sessions were designed to shape digging in scented sand. First, pairs of stimulus cups containing only 45-mg sucrose pellets were presented until rats readily consumed them. In the next stage of pretraining, pairs of cups were filled with sand and a pellet was placed on top of the sand. When these pellets were reliably consumed, additional trials were conducted with the pellets embedded progressively deeper in the sand until rats

successfully retrieved 10–20 consecutive pellets buried approximately 1 cm deep.

Match-to-sample procedure. A trial began with the insertion of the stimulus tray approximately 12 cm into the chamber such that only the sample stimulus cup, which was baited with a pellet, was accessible to the rat. A digging response was recorded when the rat's paws or nose made contact with the sand in a cup in such a way that sand was displaced. Immediately after the rat dug in the sample cup, the tray was fully inserted into the chamber (approximately 20 cm), allowing access to the comparison stimulus cups (note that the sample cup remained available as well and full tray insertion did not interfere with consumption of the sample food pellet when one was available). One comparison was always the same scent as the sample, and this cup always contained a sucrose pellet (S+). The other comparison (S−) was a different scent and was not baited (except on specially designated test trials—see below). A response on a given trial was recorded as correct if the subject dug in the cup containing S+ first and recorded as

Table 1
Sequence of Odors for Each Subject.*

# Odors	Rat J6	Rat J10	Rat J11	Rat J16
2	(1) Mustard, Cinnamon	(1) Garlic, Celery	(1) Garlic, Celery	(1) Mustard, Cinnamon
3	(36) Sage	(34) Cinnamon	(33) Cinnamon (47) Dropped Cinnamon	(47) Sage
3			(47) Mustard	
5	(41) Celery, Onion	(36) Paprika, Coffee	(54) Paprika, Coffee	(52) Celery, Onion
7	(47) Garlic, Paprika	(43) Sage, Onion	(60) Sage, Onion	(57) Garlic, Paprika
6			(61) Dropped Sage	
7			(63) Turmeric	
9	(50) Coffee, Ginger	(45) Mustard, Ginger	(67) Clove, Coriander	(66) Coffee, Ginger
11	(57) Marjoram, Thyme	(51) Marjoram, Thyme	(73) Cumin, Nutmeg	(68) Marjoram, Thyme
10		(60) Dropped Ginger	(79) Dropped Nutmeg	
11		(67) Cumin	(82) Vanilla	
13	(59) Turmeric, Nutmeg	(69) Turmeric, Nutmeg	(97) Pineapple, Strawberry	(91) Turmeric, Nutmeg
15	(61) Orange, Cumin	(74) Clove, Coriander	(101) Orange, Walnut	(93) Orange, Cumin
17	(63) Vanilla, Clove	(77) Vanilla, Oregano	(106) Coconut, Almond	(102) Vanilla, Clove
19	(67) Pineapple, Coriander	(79) Pineapple, Orange	(111) Caraway, Rosemary	(109) Pineapple, Coriander
21	(69) Strawberry, Lemon	(81) Strawberry, Lemon	(113) Lemon, Cherry	(121) Strawberry, Lemon
23	(71) Oregano, Almond	(83) Walnut, Almond	(116) Anise, Brandy	
25	(73) Cherry, Root Beer	(94) Cherry, Root Beer	(118) Allspice, Root Beer	
27	(75) Walnut, Brandy	(96) Coconut, Brandy	(126) Rum, Savory	
29	(77) Rosemary, Caraway		(128) Oregano, Chocolate	
31			(139) Maple, Peppermint	
33			(142) Dill, Bay	
35			(145) Marjoram, Thyme	
34			(162) Dropped Clove	
33			(164) Dropped Rosemary	
35			(178) Sumac, Fennel	
37			(187) Butter, Fenugreek	

* Number in parenthesis indicates the session in which stimuli were introduced or removed from the Baseline.

incorrect if the subject dug in the cup containing S- first. A correction procedure was used such that, when the subject made an incorrect response, access to both comparisons continued until the subject dug in S+. The trial was terminated when the subject dug in S+ and consumed the pellet (or when 5 s had elapsed after digging in S+ terminated). Trials also were terminated in cases in which the rat failed to dig in S+ within 3 min. When a trial terminated, the experimenter removed the tray from the chamber, recorded the rat's response, replaced the stimulus cups with pre-mixed cups designed for the next trial and, after an inter-trial interval of approximately 15 s, inserted the tray to begin the next trial.

On each trial, one stimulus was presented in the sample position of the tray and in one of the comparison positions and a different stimulus was presented in the other comparison position (see Figure 1). During the initial MTS training with two stimuli, each session consisted of 24 trials, and each stimulus served

as the sample 12 times per session. Within each session, a stimulus could appear as the sample on no more than two consecutive trials; each stimulus appeared as a comparison in each position an equal number of trials (until the number of stimuli prevented this—see below).

Although pilot procedures noted above indicated that rats could not discriminate cups on the basis of olfactory cues provided by the presence of the pellet, these tests were conducted only with the first nine odorants that were used in the present study. Thus, when additional, untested stimuli were added to the procedure it was important to determine whether accurate performance might be under the control of the scent of the pellet. Beginning with the transition to 11 baseline stimuli and throughout the rest of the experiment, two randomly selected trials in each session were arranged with both comparison cups baited. These "Double-Baited" trials were analyzed separately.

Reinforcement reduction phase. During the initial phase of MTS training each subject was exposed to only two stimuli. Subjects J16 and J6 began the MTS procedure with mustard and cinnamon and subjects J11 and J10 started with celery and garlic. A pellet was always placed in the sample cup and it was consumed prior to digging in the comparison cups on most trials. Once a rat achieved a criterion of 75% or higher accuracy for two consecutive sessions, the probability that the sample cup contained a pellet on a given trial was reduced to .75. Subsequently, a more stringent criterion of 90% correct for two consecutive sessions was implemented and the probability of reinforcement for sample observing responses (hereafter referred to as "sample reinforcement probability") was further reduced to .5 and maintained at this level throughout the remainder of the experiment (rat J16 was briefly exposed to a sample reinforcement probability of .25 and then 0, but when this manipulation disrupted his performance, the .5 sample reinforcement probability was reinstated).

Novel probe phase 1. In this phase, tests for generalized identity matching were introduced by arranging sessions in which one or more novel odors were presented on specific trials. In these sessions, novel odors were presented as samples after one or two trials with the familiar stimuli (baseline trials) had been completed. In the first of these sessions only a single novel stimulus was introduced. The novel stimulus was presented as a sample and correct comparison on the first trial on which it appeared, along with one of the baseline stimuli as the incorrect comparison. The novel stimulus then became part of the baseline stimulus set and was used throughout the rest of the experiment. Thus, after the first novel stimulus introduction, sessions included three stimuli until the rat met a criterion of two consecutive sessions at 90% correct or higher for all trials. From this point on, two novel stimuli generally were introduced each time criterion was met. On the first trial that included the novel odors, one novel odor served as the sample and the correct comparison and the other novel stimulus served as the incorrect comparison. At least one trial with baseline stimuli then separated the next presentation involving the novel odors. On the next trial involving the novel stimuli, the

Table 2
Session Configurations in the Novel Probe Phases.

Number of stimuli	Sample presentations	Comparison presentations	Trials per session
3	8	16	24
5	5	10	25
7	3	5 or 6	21
9	3	6	27
11	2	4	22
13-24	1 or 2	2 or 3	24
>24	0 or 1	0-3	24

stimulus that had previously appeared as the incorrect comparison was now presented as the sample and the other served as the incorrect comparison. Subsequent trials during the session included combinations of the baseline and novel stimuli as samples and comparisons. Once a stimulus was introduced it continued to be used in all later sessions except where noted (see Table 1).

For analysis purposes, response accuracy for trials involving a novel stimulus as sample (Novel Probe trials) were considered separately from trials involving stimuli that had been encountered previously, but had never been presented together in the particular sample-comparison combination (Novel Combination trials), and from trials involving sample-comparison combinations that *had* been encountered at least once before during the experiment (trials). Note that a trial was considered a novel combination only the first time two odors appeared in a particular sample-comparison configuration—regardless of the particular spatial configuration of the comparisons (i.e., which stimulus was in the left or right position). Thus, sessions in which novel stimuli were added contained three trial types: Novel Probe trials (a stimulus served as a sample for the first time), Novel Combination trials (stimuli that had never been presented together in a particular sample-comparison combination), and baseline trials (stimulus combinations that had been encountered previously).

Rules for composition of sessions varied somewhat across this phase (see Table 2). Initially, sessions were programmed so that each stimulus appeared an equal number of times as a sample and as a comparison in random order with the same constraints specified in the Reinforcement Reduction

Phase. For example, when the baseline consisted of three stimuli, each stimulus served as a sample on 8 trials and as a comparison on 16 trials (8 as S+ and 8 as S- paired with various sample stimuli) in each 24-trial session. When the baseline was increased to five stimuli, each stimulus served as a sample on 5 trials and as a comparison on 10 trials in each 25-trial session. As the baseline increased from three to nine stimuli, trial types used in the initial sessions were repeated, such that some possible combinations of sample and comparison stimuli were not presented. When the total number of odors reached 11, there were 22 trials presented with each odor appearing twice as samples and four times as comparisons. At this stage the sessions were organized such that each trial of every session was a Novel Combination until all possible combinations had been used at least once. Subsequently, all possible stimulus combinations were pooled and combinations were chosen randomly without replacement until each had been presented once, and then the process was repeated.

Once the number of baseline stimuli exceeded 11, all subsequent sessions were composed of 24 trials. At this point each stimulus appeared at least once and no more than twice per session as a sample. When the number of baseline stimuli exceeded 24, those odors that did not appear within a particular session were programmed to appear in the next, and no stimulus served as a sample more than once per session. Digging in the sample was reinforced with a probability of .5, except that samples were always baited on Novel Probe trials. The correct comparison appeared an equal number of times in the left and right positions in each session.

Novel probe phase 2. Novel Probe samples were always baited in Phase 1, so it was possible that accurate performance on novel probes could derive from a "win-stay" response pattern. Also, only the correct comparison cup was baited on Novel Probe trials, so accurate performance could have been due to olfactory detection of the pellet with some of the odorants that had not been tested in the pilot study. The Novel Probe Phase 2 was an effort to examine these two possibilities. Usually, subjects were advanced to the Phase 2 condition after they had met a criterion for generalized matching by responding correctly

on five out of the most recent six Novel Probe trials.

Phase 2 was different from Phase 1 in two ways. First, reinforcement was never provided for digging in the sample cup on Novel Probe trials. Sample reinforcement probability remained at .5 for baseline trials during this phase. Second, sucrose pellets were placed in both comparison cups on all Novel Probe trials during this phase, thus ensuring that correct responding could not be based on the scent of a pellet. In sessions that did not include novel probes, two randomly selected baseline trials had comparison cups that were both baited as before. Otherwise, this phase of the experiment progressed as the previous one, with novel stimuli added when a 90% correct criterion was met for two consecutive sessions. Phase 2 was continued until correct responding on five out of six consecutive Novel Probe trials was observed.

Novel probe phase 3. Only three of the subjects were tested in this phase as baseline performance deteriorated for rat J16 and he was removed from the study. This phase of the experiment differed from the previous phases in that Novel Probe sessions involved two different trial configurations. In one type (Novel-Familiar), the sample/correct comparison odor was novel but the incorrect comparison was one of the previously encountered baseline stimuli. The other type (Familiar-Novel) involved a familiar sample/correct comparison and a novel incorrect comparison (Novel Probes in Phases 1 and 2 usually involved novel sample and comparison stimuli). The initial session with each novel stimulus in this phase included one trial with each probe type. These probe types tested whether the animals were responding on the basis of rejecting a novel stimulus or on the basis of rejecting a familiar comparison. As in the previous phase, samples were never baited on Novel Probe trials and both correct and incorrect comparisons contained pellets. Otherwise, the criterion for introduction of novel odors and trial and session composition was the same as in the previous phase. Two subjects (J6 and J10) met the criterion for this phase (five out of six consecutive correct on Novel Probe trials of either type). Rat J6 was advanced to the next phase. After rat J10 met criterion, however, he failed to dig in the sample stimuli across several sessions, and was

Table 3

Number of Sessions each for Pellet Detection and MTS Procedures.

Phase	Subject			
	J16	J6	J11	J10
Pellet Detection (Pilot)	12	17	10	*
Reinforcement Reduction	46	35	32	33
Novel Probe, Phase 1	46	27	68	43
Novel Probe, Phase 2	68	8	25	6
Novel/Familiar	*	6	22	17
3 Comparisons	*	25	39	*

* Subject did not participate in phase.

removed from the study. Rat J11 was advanced to the Three Comparison Phase before meeting the Phase 3 criterion, because of the importance of observing more than one rat under these conditions.

Three comparison phase. This phase of the experiment tested rats' MTS performance when there were three comparisons available. This was of particular importance because in the two-comparison tests, it could be posited that the stimulus with the stronger overall scent was controlling responding. Because the incorrect comparison involved only a single stimulus cup, the two cups of the sample and correct comparison may have produced a detectably more intense stimulus. In order to assess this possibility, half of the trials in this phase were arranged such that the two incorrect comparisons were the same scent (Two-Odor Control). Thus, each scent within a given trial should have been equally intense. The other trials were arranged such that two different stimuli served as the incorrect comparisons.

For the two subjects (J6 and J11) that advanced to this phase, two novel odors were introduced each time criterion was met (two sessions at 90% or higher for rat J6 and two sessions at 88% or higher for rat J11). The sample was not baited, but all three comparisons were baited on Novel Probe Trials in this phase. Each probe session consisted of two probes that were either novel-familiar or familiar-novel (just as in the previous phase of the experiment). The experiment was terminated during this phase because of the advanced age of the two subjects (nearly two years). Table 3 depicts the overall number of

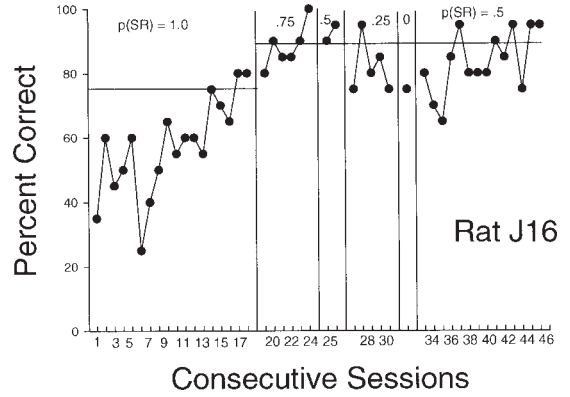


Fig. 2. Percent correct for each session during the initial exposure to the match-to-sample procedure with two stimuli for rat J16. Panel labels indicate the sample reinforcement probability. Horizontal line indicates criterion level performance.

sessions each subject was tested in each phase of the study.

Interrater reliability and blind sessions. Throughout the experiment, each rat was exposed to two or more sessions (J16, 5; J6, 7; J10, 2; J11, 3) in which a second investigator was present. The second investigator independently rated whether the animal dug in the right or left comparison (or center during the three-comparison phase) and was blind with respect to which comparison was correct. There also were several sessions (rat J16, 4; rat J6, 7; rat J11, 5) in which rats were tested by different experimenters under blind conditions. Separate analyses were conducted for these sessions.

RESULTS

Reinforcement reduction phase. Figures 2 and 3 show the results from the initial MTS training which involved only two stimuli. Each panel shows the percent correct matches across consecutive sessions. For the first 12–14 sessions, accuracy for all four rats fluctuated about chance levels (50%), but improved to the criterion level of two sessions at 75% or higher after 15 to 24 sessions. At this point, reinforcement probability for digging in the sample was reduced to .75. Rats J16 (Figure 2) and J11 (Figure 3) reached the more stringent criterion (two sessions at 90% or above) in six and four sessions, respectively, whereas subjects J6 and J10 (Figure 3) required somewhat

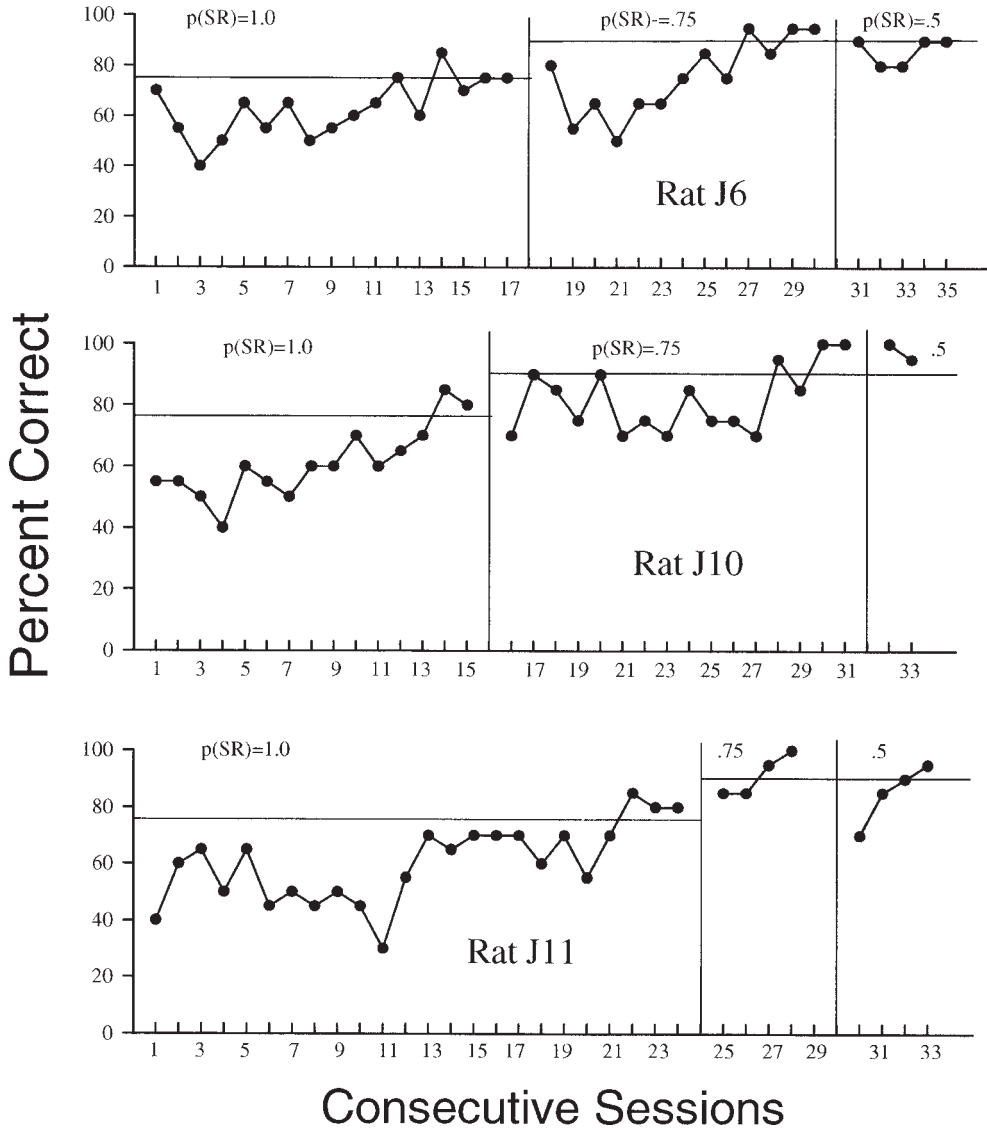


Fig. 3. Percent correct for each session during the initial exposure to the match-to-sample procedure with two stimuli for rats J6, J10 and J11. Other characteristics are as described in Figure 2.

more training (13 and 16 sessions, respectively).

Rat J16 was the first subject to meet criterion when the sample reinforcement probability was .5 (in two sessions). Sample reinforcement probability then was reduced to .25 for five sessions, and then to 0 for one session. Accuracy decreased under these conditions, and in order to maintain digging in the sample, the sample reinforcement probability was set at .5 for J16 and the other subjects throughout the remainder of the study. The

other three subjects met criterion under these conditions within two to five sessions (see Figure 3). By the end of this phase then, all four rats were showing very accurate matching with two stimuli.

Novel probe phase 1. Figures 4 and 5 show individual rats' performances for the Novel Probe Phase 1 of the study as the number of stimuli was increased. Within each graph, successive panels show overall percent correct after one or two novel stimuli were added to the baseline. Fractions inserted in each panel

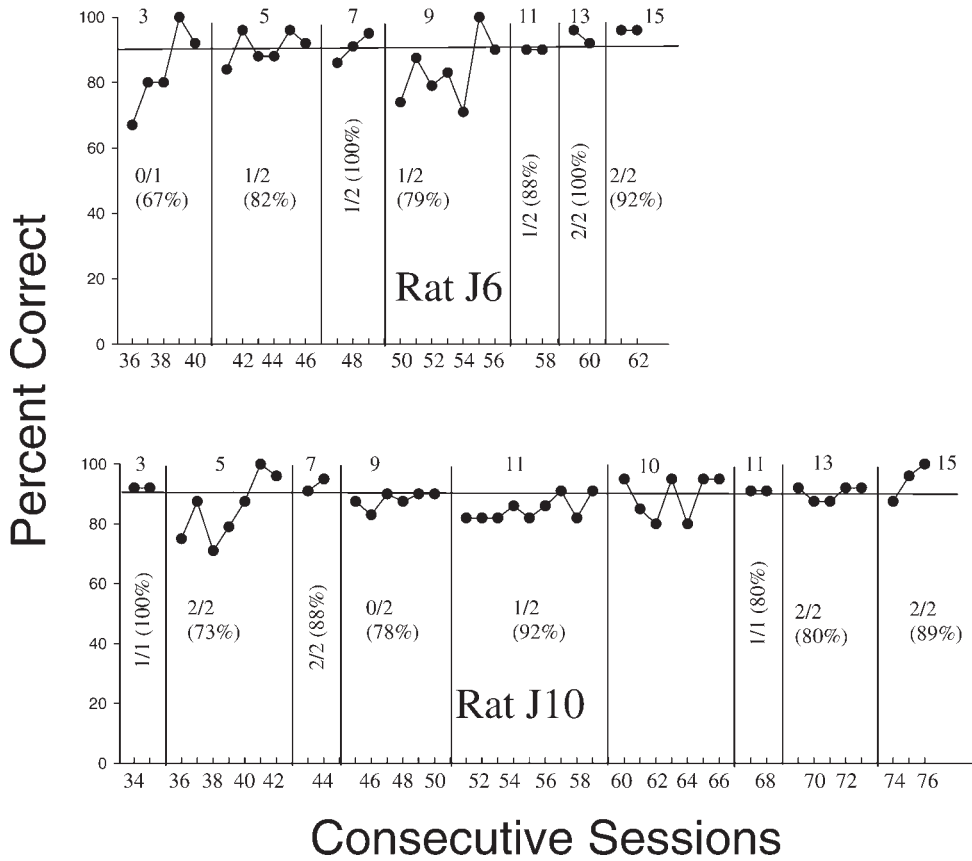


Fig. 4. Performance of rats J6 and J10 for the Novel Probe Phase 1 of the study as the number of stimuli was increased. Panel labels indicate number of stimuli. Within each graph, successive panels show overall percent correct after one or two novel stimuli were added to the baseline. Fractions inserted in each panel indicate number of correct responses over trials for Novel Probe trials (the first trial on which a novel stimulus appeared as a sample). The numbers in parentheses indicate percent correct for all sessions within the panel on trials involving Novel Combinations (trials that involved stimuli that had not previously been presented in the particular sample-comparison combination). Horizontal lines indicate the criterion (90% correct) level of performance.

indicate number of correct responses over trials for Novel Probe trials—the first trial on which a novel stimulus appeared as a sample; and percentages in each panel indicate accuracy on Novel Combination trials—trials involving stimuli not previously presented in that particular sample-comparison combination. For example, the leftmost panel of Figure 4 for rat J6 indicates that there was some disruption in accuracy when a third stimulus (sage) was added to the baseline (first three sessions in Figure 4 versus last two sessions in Figure 3 for rat J6). Performance recovered quickly and after five sessions with three stimuli in the baseline sequence, rat J6 met the criterion of two consecutive sessions with 90% overall accuracy. Accuracy when two more

novel stimuli (celery and onion) were added is indicated in the next panel of the graph. Rat J6 made one correct response on the two Novel Probe trials, but showed evidence of rapid learning with the new stimuli by achieving 82% accuracy on Novel Combination trials and meeting the 90% overall criterion in six sessions. When the number of stimuli was increased to seven, rat J6 matched correctly on one of the first two Novel Probe trials, performed correctly on all Novel Combination trials and met criterion in only three sessions. As the program advanced from 11 to 15 stimuli for rat J6, he met the criteria for generalized matching by performing correctly on the five out of six Novel Probe trials while maintaining 90% or better correct on baseline and per-

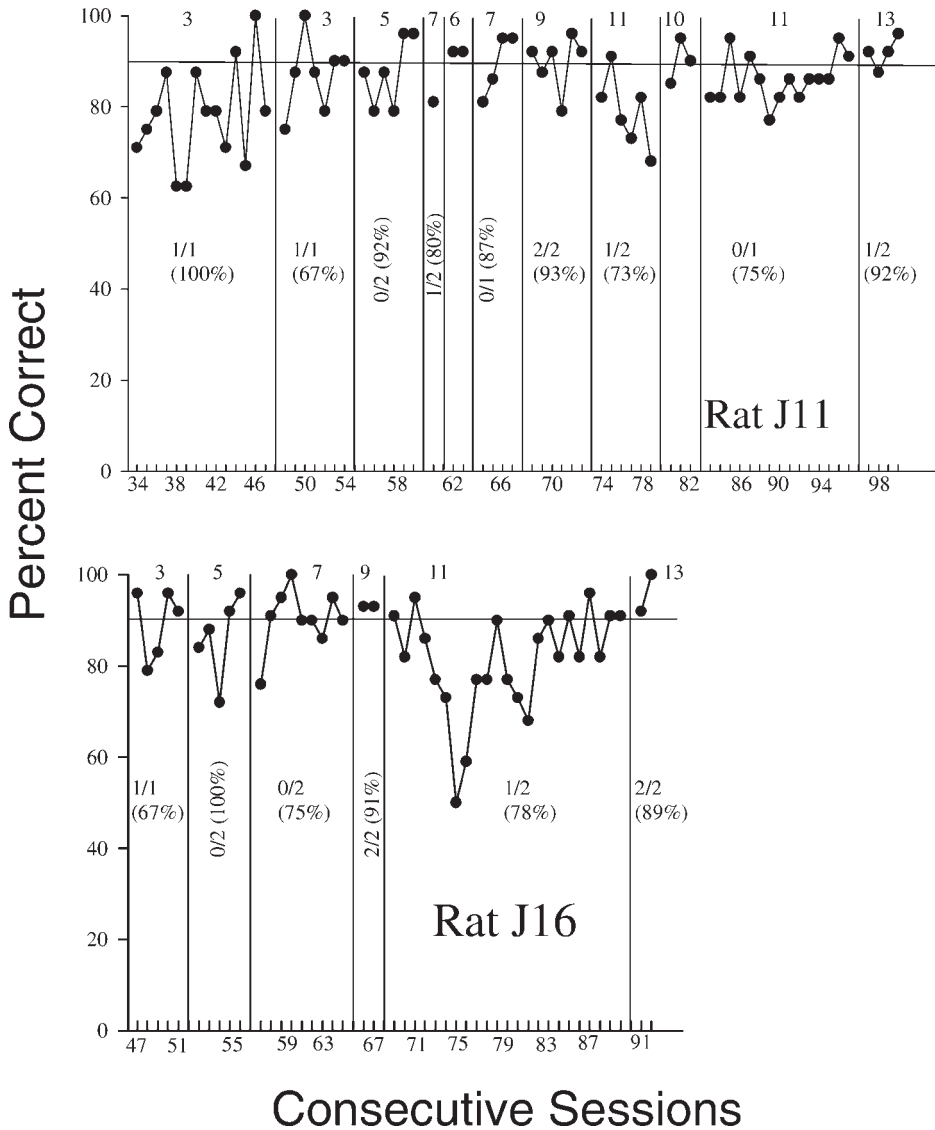


Fig. 5. Performance of rats J11 and J16 for Novel Probe Phase 1. The format is the same as in Figure 4.

forming at 93% on Novel Combination trials over six consecutive sessions (last three panels of Figure 4 for rat J6).

Rat J10 (Figure 4) showed high levels of accuracy through most of this phase. As the number of stimuli increased from three to seven, this rat performed correctly on the first five Novel Probe trials and at 87% on Novel Combination trials. Rat J10 then responded incorrectly on both Novel Probe trials (mustard and ginger) when the number of stimuli increased to nine. On trials with ginger as the sample, rat J10 consistently showed long

latencies or failed to make the observing response at all and, thus, ginger was removed from the protocol and replaced with cumin. Following this change, rat J10 performed correctly on five consecutive Novel Probe trials, showed high levels of overall accuracy, and was advanced to the next phase.

Figure 5 shows that rat J11 required somewhat more training than rats J6 and J10 to develop criterion level performance. In part this may have been due to the same sorts of difficulties noted with rat J10, as rat J11 often failed to approach certain sample stimuli. After

several sessions with persistent failures to dig in the sample cup with particular stimuli, three (cinnamon, sage, and nutmeg) were removed from this subject's protocol and replaced with other novel stimuli during this phase (see Table 1). High levels of accuracy eventually were achieved for both baseline and Novel Combination trials as rat J11 was advanced to a 13-stimulus baseline, but no evidence of accurate responding on Novel Probe trials was seen during this phase (correct matching occurred on only four of the final eight Novel Probe trials). Despite failing to meet the criteria for generalized matching, rat J11 was advanced to the next phase.

Finally, rat J16 (Figure 5) met the criterion fairly rapidly with each addition of novel stimuli up through nine (rat J16 actually met criterion on his third session with seven stimuli, but was inadvertently exposed to six additional sessions before being moved to the nine-stimulus condition). His performance declined temporarily with 11 stimuli. Criterion was met, however, after 23 sessions, and subsequent performance with 13 stimuli was nearly perfect. Rat J16 met the overall criterion for generalized matching by performing correctly on five out of six consecutive Novel Probe trials as the number of stimuli was increased from 9 to 13.

At this stage of the experiment all four of the animals were showing high levels of accuracy on baseline trials and trials involving novel combinations of 13 to 15 different stimuli. In addition, three of the four rats had met criterion for generalized matching by responding correctly on five of their last six Novel Probe trials. However, two features of the Novel Probe procedures in this phase raised questions. First, although the sample reinforcement probability was .5 on baseline and Novel Combination trials, the sample cup was always baited on Novel Probe trials during Phase 1. Thus, it could be argued that on Novel Probe trials the animals simply were responding to the scent that had most recently been associated with food. To address this concern, in the Novel Probe Phase 2 condition, food pellets were *never* available in the sample cup on Novel Probe trials. Second, to ensure that responding was not controlled by the scent of the sucrose pellet, in Phase 2 *both* comparison cups always were baited on Novel Probe trials.

Novel probe phase 2. The results of the Novel Probe Phase 2 conditions are shown in Figures 6 and 7. Rat J6 moved through this phase in just two more than the minimum possible number of sessions, and rat J10 required only the minimum of six sessions to meet criteria for generalized matching (Figure 6). Both of these rats performed correctly on all six of the Novel Probe trials encountered during this phase and showed high levels of accuracy on Novel Combination and baseline trials as the number of stimuli was increased from 17 to 21. Rat J11 displayed high levels of baseline accuracy, but required more sessions than rats J6 and J10 to meet the criterion for generalized matching (Figure 7). As the number of baseline stimuli increased from 15 to 25, rat J11 performed correctly on 9 of 12 Novel Probe trials and, ultimately, on five of his last six Novel Probes to meet criterion. Recall that rat J11 did not meet the generalized matching criterion in Novel Probe Phase 1, so perhaps the additional exposure to Novel Probe trials in Phase 2 was critical in his case. Finally, as Figure 7 shows, baseline accuracy for rat J16 was not as high as for the other three rats. Although both baseline and probe trial performances were maintained at above-chance levels of accuracy, performance was persistently below the 90% criterion levels. Although J16 did meet the generalized matching criterion with five of six correct on his final Novel Probe trials, his baseline performance with 21 stimuli declined. Thus, after 40 sessions without meeting the 90% criterion for 2 consecutive sessions, J16 was removed from the study.

Novel probe phase 3. In the Novel Probe trials of this phase, different stimulus configurations were studied in which either a novel stimulus appeared as the sample with a familiar incorrect comparison stimulus or one of the familiar baseline stimuli served as the sample with a novel stimulus as the incorrect comparison. Figures 8 and 9 show performance in Phase 3 for the three rats tested. Rat J6 met the overall criterion by performing correctly on all six Novel Probe trials in the minimum number of sessions and maintained nearly 100% accuracy on baseline trials which now included up to 27 different stimuli. Rat J10 performed correctly on the first two Novel Probe trials in this phase with up to 23 stimuli in the baseline, but inadvertently was taken back to the 21-

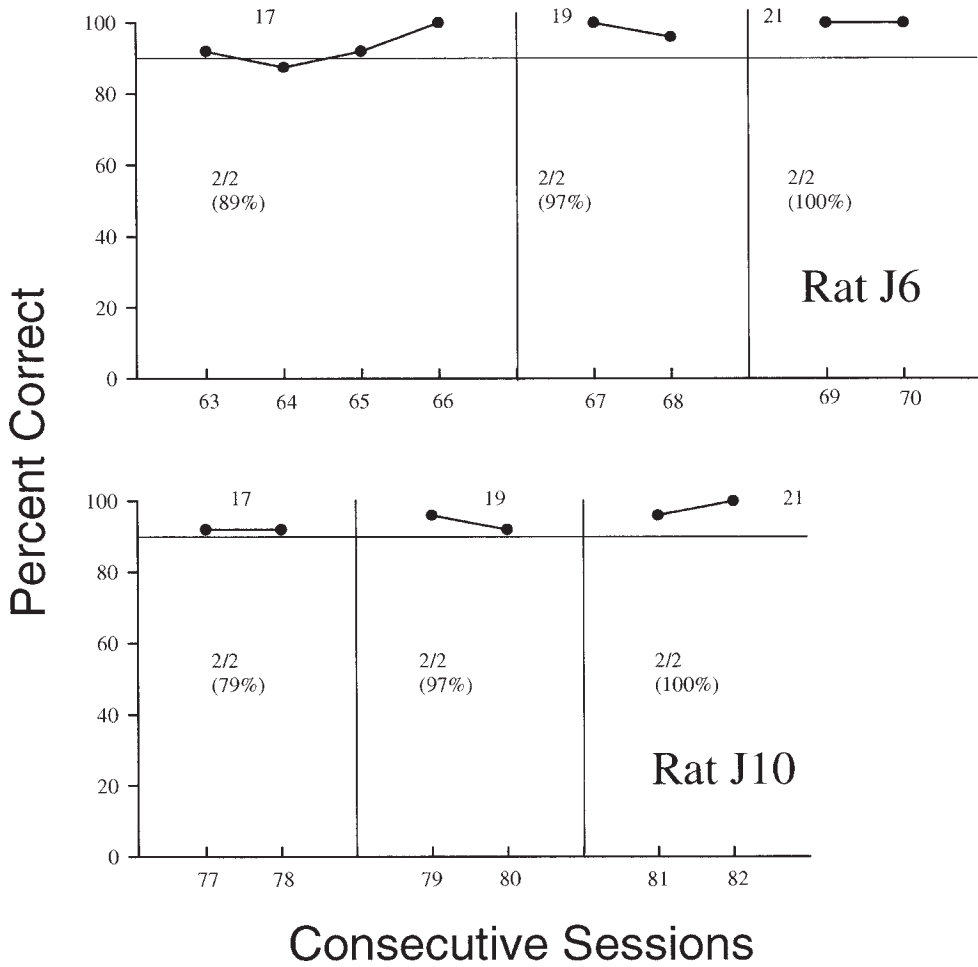


Fig. 6. Performance of rats J6 and J10 for Novel Probe Phase 2 of the study. The format is the same as in Figure 4.

stimulus baseline and required three sessions to meet criterion again. Subsequently, rat J10 performed correctly on four additional consecutive Novel Probe trials and responded at or near 90% on Novel Combination and baseline trials with 27 stimuli to meet the generalized matching criteria. After four sessions with 27 baseline stimuli, however, J10 became ill and was removed from the study. Rat J11 performed with high levels of accuracy on baseline and Novel Combination trials with up to 35 stimuli during this phase. However, accuracy on Novel Probe trials was only at chance levels. In order to study more than one rat under the Three-Comparison conditions, rat J11 was moved to that phase despite not yet meeting the generalized matching criteria for Phase 3.

Summary of novel probe phases. Figure 10 summarizes the results of Novel Probe Phases 1–3 and shows percent correct on Novel Probe trials for the last three criterion test sessions (leftmost bars), and performance of each animal on Novel Combination trials throughout the phase (center bars). Percent correct for baseline trials on which both comparison cups contained sucrose pellets (Double-Baited trials) also are depicted (rightmost bars). In order to provide statistical confirmation of the accuracy of performance with Novel Probes, a one-tailed binomial test was conducted for all Novel Probe trials beginning with the criterion-level performance of Phase 1 and extending through the end of Phase 2. Three of the four animals showed levels of accuracy that were significantly different from chance (J16,

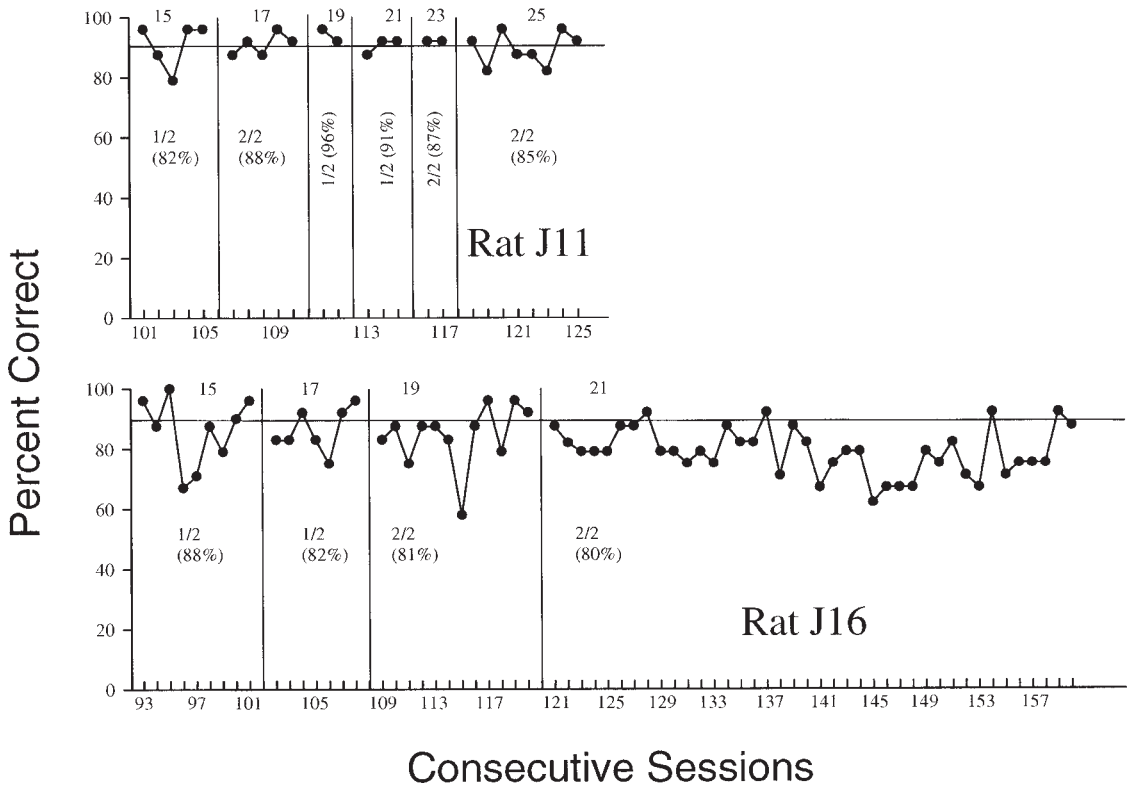


Fig. 7. Performance of rats J11 and J12 for Novel Probe Phase 2 of the study. The format is the same as in Figure 4.

11 out of 14, $p < .05$; J6, 11 out of 12, $p < .05$; J10, 11 out of 11, $p < .05$). Additional analyses conducted for J6 and J10 that included the novel probes introduced during Phase 3 also

were statistically significant (17 out of 18, and 19 out of 19, $p < .05$ respectively). Although rat J11 met the criterion of 5 out of 6 correct on Novel Probe trials in Phase 2, his overall

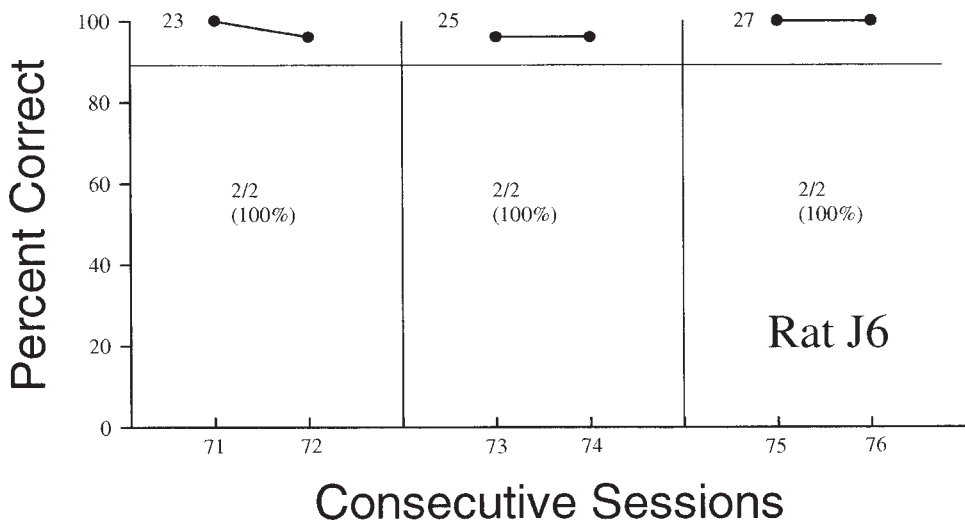


Fig. 8. Performance of rat J6 for Novel Probe Phase 3 of the study. The format is the same as in Figure 4.

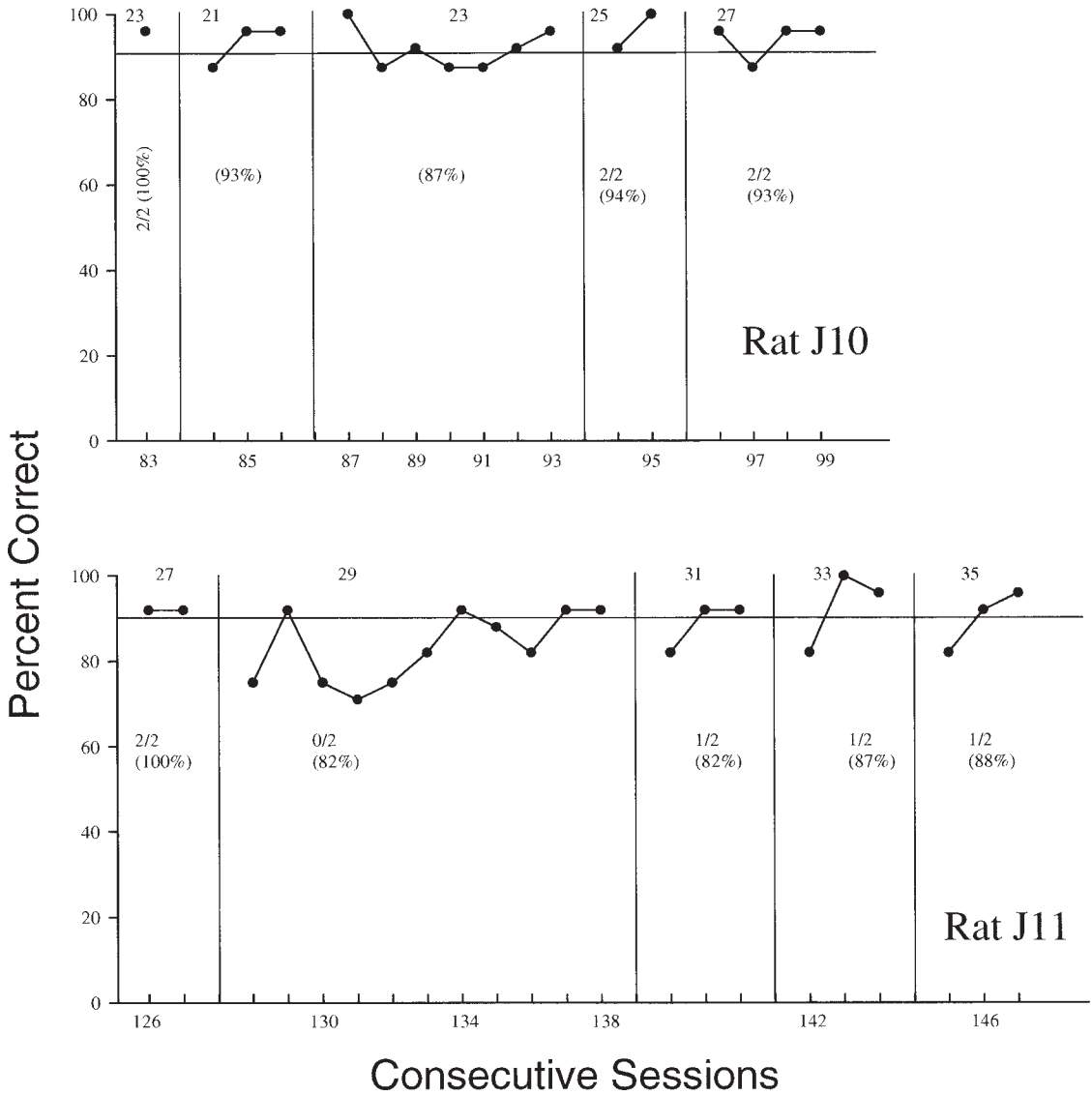


Fig. 9. Performance of rats J10 and J11 for Novel Probe Phase 3 of the study. The format is the same as in Figure 4.

performance on novel probes was not significantly different from chance (Phases 1 and 2: 9 out of 12, $p > .05$; all phases: 14 out of 22, $p > .05$). Performance on Novel Combination trials was consistently high, averaging above 80% correct and significantly ($p < .05$) above chance levels for all phases and all four rats throughout the study. Finally, performance on Double-Baited trials was usually comparable to other baseline and probe trials and was significantly ($p < .05$) above chance in all cases except for the performance of rat J11 in Phase 3.

Three-comparison phase. The Three-Comparison Phase of the experiment was designed to answer two questions: would subjects be able to maintain baseline accuracy when more than two comparisons were available, and were they solving the MTS problems by choosing the odor that was most intense? Thus, in the Three-Comparison Phase some trials (Two-Odor Control) arranged for the two incorrect comparison cups to be identical so that both stimulus scents would be represented in two cups. Figure 11 shows the overall percent correct for rats J6 and J11 on a session-by-

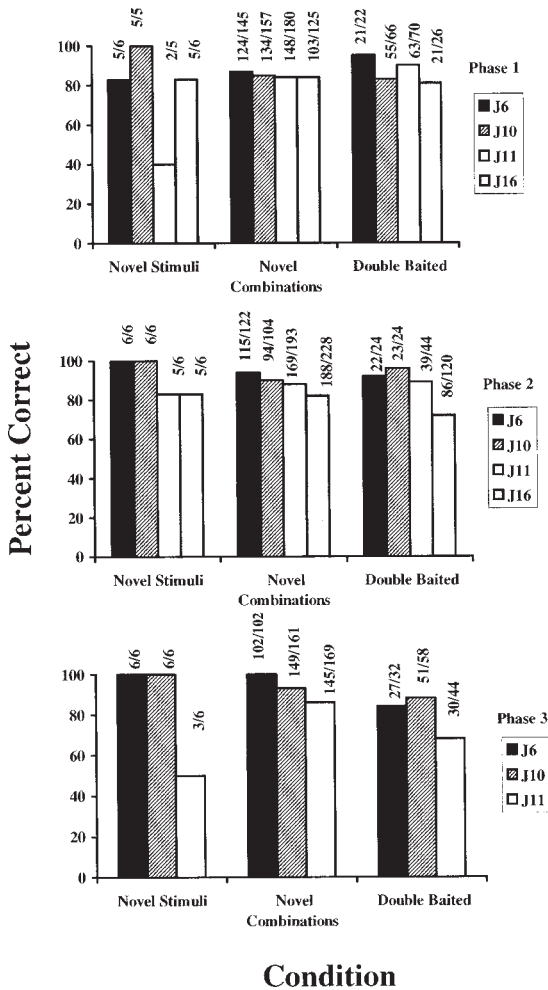


Fig. 10. Summary of the Novel Probe Phases (Phase 1, top panel; Phase 2, middle panel; Phase 3, bottom panel) for each subject. Percent correct on Novel Probe trials for the last three criterion test sessions (leftmost bars), and performance of each animal on Novel Combination trials throughout the phase (center bars). Percent correct for baseline trials on which both comparison cups contained sucrose pellets (Double-Baited trials) are depicted in the rightmost bars.

session basis, as well as the percent correct for the Two-Odor Control trials. Accuracy declined somewhat for both rats when three comparisons were available, although performance remained well above chance. Thus, accurate matching did not depend on the two-comparison stimulus configuration. Importantly, performance with three different comparison stimuli were comparable to those obtained with the Two-Odor Control trials in

both cases, which suggests that odor intensity was not controlling responding.

Interrater reliability and blind sessions. Because the experimenter potentially was visible to the rat during this study, several sessions were conducted with the experimenter blind with respect to the correct comparison position for three of the rats. Subject J16 was tested under such blind conditions for a total of four sessions and maintained an average performance of 81% correct (compared to an 83% average across Phases 1–3 of the experiment). Subject J6 received a total of seven blind testing sessions averaging 91% correct (versus 91% across the experiment), and subject J11 received a total of five blind testing sessions averaging 74% correct (compared to 88% across the experiment). Thus, performance during the blind sessions was well above chance accuracy and, with the possible exception of J11, very close to each animal's average performance across the entire experiment, suggesting that the experimenter was not inadvertently cuing the subjects. In addition, interrater agreement was determined for all subjects. Table 4 shows the interrater agreement for sessions in which there was a blind rater present, as well as the animals' percent correct on those sessions. High interrater reliability was demonstrated on nearly all sessions with an overall mean of 98% and a range of 88–100%.

DISCUSSION

Initial acquisition of the MTS conditional discrimination with just two stimuli took place fairly rapidly, with rats reaching the criterion of 90% accuracy in 24 to 31 sessions (576–720 trials). As novel stimuli were added to the baseline, high levels of matching, generally between 80%–95% correct, were maintained throughout Phases 1 and 2. Such accurate performance as the number of stimuli in the baseline was increased to more than 20 for each of the four rats is consistent with an interpretation of control by the identity relation between sample and comparison, but also could be interpreted as rapid learning of specific stimulus pairs or configurations. Thus, the analyses of performance on trials involving novel stimuli and stimulus combinations are critical. Throughout Phases 1 and 2, accuracy was consistently above 80% (Figure 10) for all

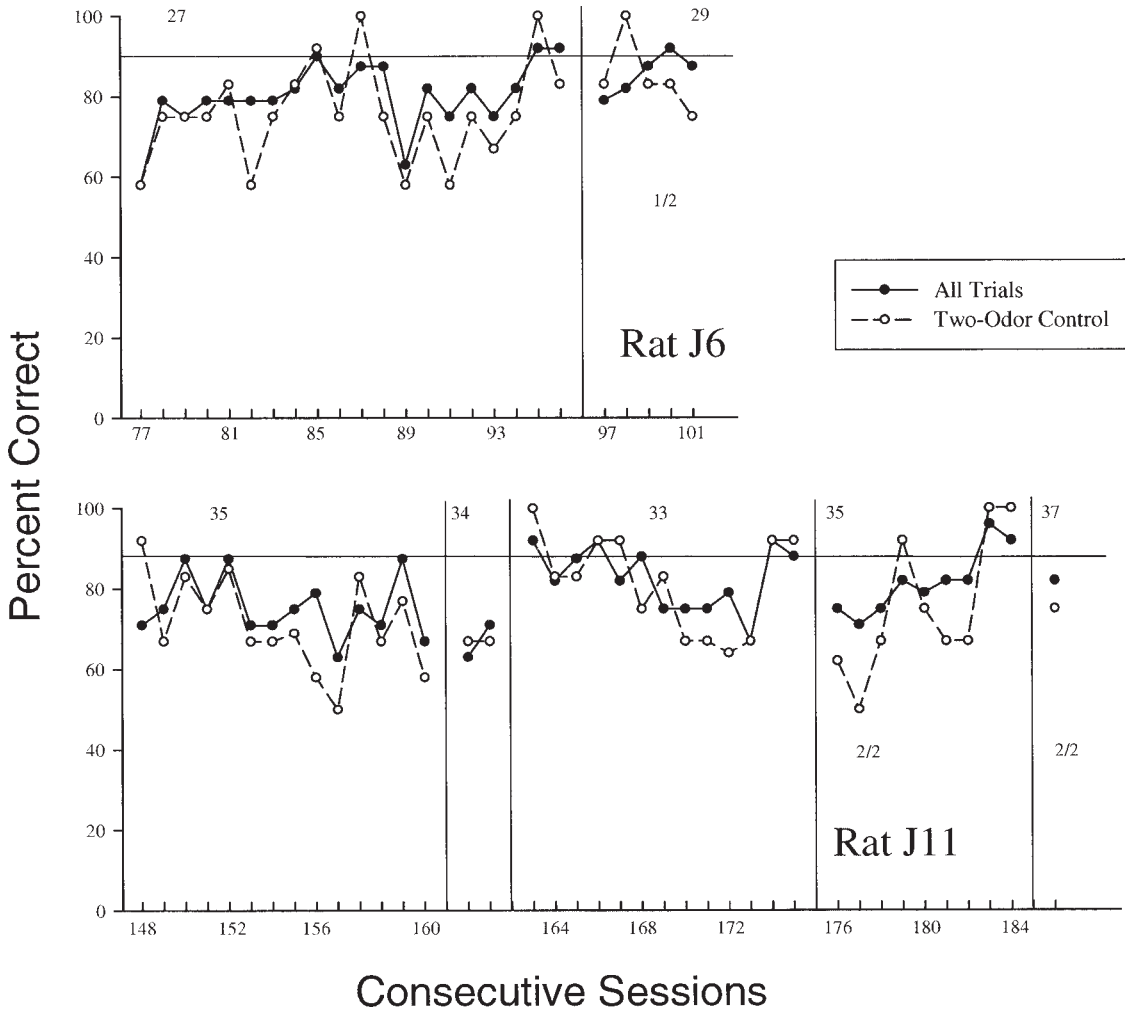


Fig. 11. Results of the three comparison conditions are shown for the 2 rats tested. Percent correct is plotted separately for all trials of the session (black circles) and for the two-odor control trials (white circles). Horizontal lines indicate criterion (90%) levels of performance.

four rats on Novel Combination trials, suggesting that rapid learning of specific stimulus configurations cannot account for the observed matching. Importantly, the above-chance accuracy on Novel Probe trials across Phases 1 and 2 met the criterion for generalized matching for three of the four animals (J6, J10 and J16). Phase 3 performance on Novel Probe trials, that introduced novel stimuli with a familiar baseline stimulus as a sample or comparison, also was highly accurate for two of the three rats tested (J6 and J10). In summary, high levels of accuracy were demonstrated on trials when both stimuli were novel (J6, J10 and J16), when only the sample or only the comparison was novel (J6

and J10), and when familiar stimuli appeared together for the first time (all four rats). These outcomes of Phases 1–3 support an interpretation that control by the identity relation between sample and comparison stimuli was the basis for the performance.

Additional tests were used to evaluate the likelihood of various alternative accounts of the observed performance. For example, although pilot work had determined that the spice concentrations used were sufficiently high to mask the scent of the food pellet for many of the odorants, by the end of Phase 1 many untested novel concentrations were in use. From Phase 2 on then, all comparison cups were baited on Novel Probe trials, and

Table 4
Interrater agreement.

<u>J16 Interrater reliability</u>		
Number of stimuli	Inter-rater agreement	S performance
15	100%	100%
21	100%	79%
21	100%	79%
21	100%	88%
21	100%	79%
<u>J6 Interrater reliability</u>		
Number of stimuli	Interrater agreement	S performance
17	100%	88%
19	100%	100%
21	100%	88%
27	88%	88%
27	100%	88%
27	96%	92%
29	100%	92%
<u>J10 Interrater reliability</u>		
Number of stimuli	Interrater agreement	S performance
2	90%	70%
27	100%	90%
<u>J11 Interrater reliability</u>		
Number of stimuli	Interrater agreement	S performance
2	100%	90%
25	96%	92%
35	96%	75%

the maintenance of high levels of accuracy further rules out the possibility that the outcomes were influenced by olfactory detection of the sucrose pellet location.

Food reinforcement often followed digging in the sample cup in the present study and, because direct reinforcement of sample observing responses is unusual in match-to-sample research, the importance of this aspect of the procedure as a determinant of stimulus control should be considered in the present study. This procedure was used to ensure exposure to the sample stimulus. Note, however, that reduction of sample reinforcement probability to .5 early in the experiment demonstrated that direct reinforcement of digging in the sample on each trial was not necessary for accurate matching. However, although the sample was baited on all Novel Probe trials during Phase 1, successful performance on Novel Probe trials in Phase 2 (when

the sample was *never* baited) and beyond showed that the results were not due to the development of a pattern of selecting the comparison that had most recently been associated with food. One could make the case that digging in the sample produced the conditioned reinforcement of the presentation of the comparisons whether or not it also produced food, and thus, that it is still possible to account for the behavior in terms of selection of the comparison that was most recently reinforced. However, the reinforcement of sample observing or orienting behavior by presentation of the comparison stimuli is a common, virtually unavoidable, feature of match-to-sample procedures in general. The importance of sample responding in the development of matching has been demonstrated (e.g., Wright, 1997) and further exploration of the importance of direct reinforcement of sample observing in match-to-sample procedures would be of interest.

Another possible source of stimulus control in the present study was odor intensity. Because the sample cup remained in the chamber when the comparisons were presented, two cups of the sample spice and only one of the comparison spices were present, and this raised the possibility that odor intensity, rather than identity, might have come to control responding. To determine whether successful matching could be accounted for by responding to the strongest odor, the experiment was extended to include a three-comparison task for two of the animals. In this phase, half of the trials contained three odors (sample, matching odor, and two different alternative odors) and the remaining half contained just two odors (sample, matching odor, and two of the same alternative odors). This two-odor condition was an effort to equate the strength of the sample and incorrect comparison odors. Although J6 and J11 both showed slight declines in accuracy with three comparisons, their performance was still well above chance (with averages of 81% and 78%, respectively). When the two-odor control performance was analyzed separately, it was evident that performance was not controlled by digging in the odor most strongly represented because accuracy on these trials nearly matched those of their session-by-session performance. Results from this phase also showed that high levels of

matching transferred to the novel stimulus combinations and spatial comparison configurations associated with three comparisons. Finally, high inter-rater reliability and accuracy under blind testing conditions suggest that experimenter bias was not responsible for the high success rates. In sum, the various control conditions consistently supported an account of the present results in terms of generalized identity matching.

Consistent with MTS studies with pigeons and monkeys, there was evidence here that training with multiple exemplars enhanced control by the identity relation. As the experiment progressed and more stimuli were added to the baseline, all four subjects required fewer trials to meet the generalized matching criteria (at least until very late in the experiment). For example, in Phase 1, the animals required an average of 7.5 test sessions to meet criterion. However, in Phase 2 an average of just four test sessions was needed. This decrease was particularly interesting because the MTS sessions were becoming progressively more complex: as more stimuli were added to each animal's baseline, there were more trials per session involving novel stimulus combinations. However, it should be noted that as the number of baseline stimuli increased, it may have become more likely that novel sample odors were sufficiently similar to previously learned odors to produce stimulus generalization. It would be interesting to determine whether matching acquired with the current techniques would transfer to an orthogonal stimulus domain (e.g., visual) and the case for control by identity relations would certainly be strengthened by such a demonstration (see Zentall & Hogan, 1974, 1976).

Although the successful MTS performance of the rats in the present study may have been related to the use of multiple exemplars, other features of the procedure likely were important. For example, the use of olfactory stimuli may have been particularly crucial. The present results replicate and extend previous findings using olfactory stimuli (Lu *et al.*, 1993). The present study extends the results of the Lu *et al.* study in demonstrating above-chance performance with novel stimuli and stimulus combinations and high levels of accuracy with a matching procedure, rather than with a go/no-go task. In contrast, rats' performance on MTS with visual stimuli typically has not been very successful

(Iversen, 1993, 1997; Rothblat & Hayes, 1987), although there have been reports of generalized matching with visual stimuli (Nakagawa, 2000; Prusky, Douglas, Nelson, Shabanpoor, & Sutherland, 2004). The extent to which the olfactory modality, multiple exemplar training, and other procedural features contributed to the successful outcomes in the present study remains to be determined. In any case, it is clear that conditional discriminations with olfactory stimuli are learned rapidly and, in the present study, this made it possible to establish a repertoire involving many different stimuli, a factor which may be critical for the development of generalized matching-to-sample.

Finally, a few limitations of the present study should be noted. Performance for two of the rats (J10 and J16) declined in the later stages of the study. In both of these cases, the advanced age of the subjects (20 months and 16 months respectively) may have adversely affected their performance. For example, the decline for Rat J10 tended to involve a failure to dig in the sample, rather than a failure of matching; this may have reflected a loss in reinforcer efficacy. Also, the present procedure required extensive experimenter intervention and relied on experimenter judgment regarding response definition. An automated version of this MTS procedure using an olfactometer would eliminate these limitations and would be an important extension. Note, however, that the interrater reliability and blind testing data collected in the present study support the validity of the present outcomes. Thus, despite the limitations listed above, this methodology appears to provide an effective means of studying complex stimulus control in rats. Indeed, the present findings of identity matching suggest the possibility that this methodology might permit the analysis of other emergent behaviors in rats.

REFERENCES

- Barros, R., Galvão, O., & McIlvane, W. J. (2002). Generalized identity matching-to-sample in *cebus apella*. *The Psychological Record*, 52, 441–460.
- Bunsey, M., & Eichenbaum, H. (1996). Conservation of hippocampal memory function in rats and humans. *Nature*, 379, 255–257.
- Carter, D. E., & Werner, T. J. (1978). Complex learning and information processing by pigeons: A critical analysis. *Journal of the Experimental Analysis of Behavior*, 29, 565–601.

- Cook, R. G., Katz, J. S., & Cavoto, B. R. (1997). Pigeon same-different concept learning with multiple stimulus classes. *Journal of Experimental Psychology: Animal Behavior Processes*, *23*, 417-433.
- Cumming, W. W., & Berryman, R. (1965). The complex discriminated operant: Studies of matching-to-sample and related problems. In D. I. Mostofsky (Ed.), *Stimulus generalization* (pp. 284-330). Stanford: Stanford University Press.
- Dube, W. V., McIlvane, W. J., & Green, G. (1992). An analysis of generalized matching-to-sample test procedures. *The Psychological Record*, *42*, 17-28.
- Dudchenko, P. A., Wood, E. R., & Eichenbaum, H. (2000). Neurotoxic hippocampal lesions have no effect on odor span and little effect on odor recognition memory, but produce significant impairments on spatial span, recognition and alternation. *The Journal of Neuroscience*, *20*, 2964-2977.
- Dusek, J. A., & Eichenbaum, H. (1998). The hippocampus and transverse patterning guided by olfactory cues. *Behavioral Neuroscience*, *112*, 762-771.
- Herman, L. M., & Gordon, J. A. (1974). Auditory delayed matching in the bottlenose dolphin. *Journal of the Experimental Analysis of Behavior*, *21*, 19-26.
- Iversen, I. (1993). Acquisition of matching-to-sample performances in rats using visual stimuli on nose keys. *Journal of the Experimental Analysis of Behavior*, *59*, 471-482.
- Iversen, I. (1997). Matching-to-sample performance in rats: A case of mistaken identity? *Journal of the Experimental Analysis of Behavior*, *68*, 27-45.
- Kastak, D., & Schusterman, R. (1994). Transfer of visual identity matching-to-sample in two California sea lions (*Zalophus californianus*). *Animal Learning & Behavior*, *22*, 427-435.
- Katz, J. S., Wright, A. A., & Bachevalier, J. (2002). Mechanisms of same/different abstract concept learning by rhesus monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*, *28*, 358-368.
- Lu, X. M., Slotnik, B. M., & Silberberg, A. M. (1993). Odor matching and odor memory in the rat. *Physiology and Behavior*, *53*, 795-804.
- Mihalick, S. M., Langlois, J. C., Krienke, J. D., & Dube, W. V. (2000). An olfactory discrimination procedure for mice. *Journal of the Experimental Analysis of Behavior*, *73*, 305-318.
- Mishkin, M., & Delacour, J. (1975). An analysis of short-term visual memory in the monkey. *Journal of Experimental Psychology: Animal Behavior Processes*, *1*, 326-334.
- Nakagawa, E. (2000). Transfer of learning between matching (or non-matching)-to-sample and same-different discriminations in rats. *The Psychological Record*, *50*, 771-805.
- Oden, D. L., Thompson, R. K. R., & Premack, D. (1988). Spontaneous transfer of matching by infant chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 140-145.
- Otto, T., & Eichenbaum, H. (1992). Complementary roles of the orbital prefrontal cortex and the perirhinal-entorhinal cortices in an odor-guided delayed non-matching-to-sample task. *Behavioral Neuroscience*, *106*, 762-775.
- Pack, A. A., Herman, L. M., & Roitblat, H. L. (1991). Generalization of visual matching and delayed matching by a California sea lion (*Zalophus californianus*). *Animal Learning & Behavior*, *19*, 37-48.
- Prusky, G. T., Douglas, R. M., Nelson, L., Shabanpoor, A., & Sutherland, R. J. (2004). Visual memory task for rats reveals an essential role for hippocampus and perirhinal cortex. *Proceeding of the National Academy of Sciences*, *101*, 5064-5068.
- Rothblat, L. A., & Hayes, L. L. (1987). Short-term object recognition memory in the rat: Nonmatching with trial-unique junk stimuli. *Behavioral Neuroscience*, *101*, 587-590.
- Wood, E. R., Dudchenko, P. A., & Eichenbaum, H. (1999). The global record of memory in hippocampal neuronal activity. *Nature*, *397*, 613-616.
- Wright, A. A. (1997). Concept learning and learning strategies. *Psychological Science*, *8*, 119-123.
- Wright, A. A., Cook, R. G., Rivera, J. J., Sands, S. F., & Delius, J. D. (1988). Concept learning by pigeons: Matching-to-sample with trial-unique video picture stimuli. *Animal Learning & Behavior*, *16*, 436-444.
- Zentall, T. R., Edwards, C. A., Moore, B. S., & Hogan, D. E. (1981). Identity: The basis for both matching and oddity learning in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *7*, 70-86.
- Zentall, T. R., & Hogan, D. E. (1974). Abstract concept learning in the pigeon. *Journal of Experimental Psychology*, *102*, 393-398.
- Zentall, T. R., & Hogan, D. E. (1976, 30 January). Pigeons can learn identity or difference, or both. *Science*, *191*, 408-409.

Received: September 29, 2004
 Final acceptance: October 18, 2005