



Incrementing non-matching- but not matching-to-sample is rapidly learned in an automated version of the odor span task

Thomas J. Wagner¹ · Katherine Bruce¹ · Mark Galizio¹

Received: 3 October 2021 / Revised: 20 January 2022 / Accepted: 16 February 2022 / Published online: 25 February 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

The odor span task (OST) is frequently used to assess memory capacity in rodents. Odor stimuli are presented in a large arena and choices of session-novel odors produce food reward. The procedure can be described as an incrementing non-matching-to-sample contingency because on each trial one new stimulus is presented along with one or more previously presented (non-reinforced) comparison odors. An automated version of this task has recently been developed in which odors are presented with an olfactometer in an operant chamber using a successive conditional discrimination procedure. The present study compared the acquisition of matching- vs. non-matching-to-sample versions of the task with six rats tested under each procedure. All six rats trained on the non-matching variation showed rapid acquisition of the discrimination with high rates of responding to odor stimuli when they were session-novel and low rates of responding to subsequent presentations of those odors. However, only three of the six rats trained on the matching variation met acquisition criteria, and two of the three that did acquire the task required extensive training to do so. These results support findings from the OST that rats can differentiate between stimuli that are session-novel and those previously encountered, but also that a matching contingency is more difficult to learn than a non-matching arrangement. These findings parallel differences observed between acquisition of simple matching- and non-matching-to-sample tasks, but accounts such as novelty preference or the oddity preference effect may not be sufficient to explain the present results.

Keywords Incrementing non-matching-to-sample · Incrementing matching-to-sample · Odor span task · Oddity preference effect · Successive conditional discrimination · Rats

Introduction

Research using the odor span task (OST) has demonstrated that rats readily learn to discriminate odors that they have already encountered during a session from those that are session-novel (e.g., April et al. 2013; Dudchenko et al. 2000; MacQueen et al. 2011). The OST is generally conducted in a large arena in which cups containing scented sand (e.g., Dudchenko et al. 2000), wood chips (e.g., Rushforth et al. 2011) or covered with a scented lid (e.g., MacQueen et al. 2011) are placed. At the beginning of the session, the rat is presented with a cup scented with a particular odor, for instance, cinnamon. When the rat digs in the substrate/removes the scented lid, it obtains a food reward. In the next

trial, two cups with two different odors are presented, with one containing the cinnamon scent and one containing a new scent, e.g., clove. Only selection of the session-novel odor (clove) is reinforced, and selection of the familiar odor (cinnamon) is under extinction. This procedure continues with one new scent added each trial. As the session progresses, the memory load becomes greater, and accuracy tends to decline (Dudchenko et al. 2000; Galizio et al. 2013); thus, the OST is often portrayed as a measure of working memory capacity.

However, evidence is accumulating that raises questions about a memory capacity interpretation of OST performances. Both rats and dogs can remember 70–100 different odors in a single OST session (April et al. 2013; Bratch et al. 2016; Krichbaum et al. 2020). Such findings suggest that animals' memory capacity is quite high or, alternatively, that other processes such as familiarity detection may be involved (April et al. 2013; Krichbaum et al. 2020). Furthermore, variables not directly related to memory capacity

✉ Mark Galizio
galizio@uncw.edu

¹ Department of Psychology, University of North Carolina Wilmington, Wilmington, NC, USA

have been shown to influence OST performances. For example, increasing the number of distractor stimuli available on each trial can affect accuracy on the task (April et al. 2013; Galizio et al. 2019). In addition, Krichbaum et al. (2020) noted decreases in accuracy as a function of the number of trials separating the presentation of a given odor in dogs. This finding is consistent with a working memory account, but emphasizes the potential importance of the retention interval/interference rather than capacity per se. A similar finding has been observed in rats, as Galizio et al. (2020) showed that accuracy decreased when the retention interval was lengthened using an automated version of the OST.

The basis for the remarkable memory observed in the OST is not well understood. One account would emphasize the importance of the use of olfactory stimuli. Indeed, several studies have demonstrated that rats show more rapid concept learning and better remembering with olfactory stimuli relative to stimuli presented in other modalities (e.g., Nigrosh et al. 1975; Slotnick et al. 1991; Thomas and Noble 1988). Thus, rats would be expected to do well on any task requiring odor discriminations. An alternative account might stem from the well-known tendency for rats to spend more time exploring novel objects relative to more familiar ones, or novelty preference (Ennaceur 2010). This tendency might lead to a strong likelihood of selecting the novel odor in the arena and a bias against errors (selecting a previously presented stimulus) very early in OST training, producing rapid acquisition and accurate responding across extended delays regardless of stimulus modality. Of course, in most OST studies the same odors are used in every session so none of the odors are truly novel. However, numerous studies show that object familiarity is transient and a tendency to approach the less familiar object or odor could produce the rapid acquisition seen in the OST (Ennaceur 2010; Spinetta et al. 2008).

Perhaps related to the preference for novel objects is the frequently noted phenomenon that non-matching-to-sample (NMTS) is generally more rapidly acquired than matching-to-sample (MTS), a finding that is sometimes termed the oddity preference effect (Wright and Delius 2005). Rats do show oddity preference (Marks et al. 2018), but evidence that NMTS is acquired more rapidly than MTS has been mixed in rats (Bruce et al. 2018; Lazarowski et al. 2019). Furthermore, although the OST does resemble more traditional NMTS procedures, the oddity preference effect does not seem a likely interpretation of rodent memory performances in this task.

The OST may be described as an incrementing NMTS procedure with each stimulus serving both as a sample and a reinforced comparison stimulus when it is first presented and then as a non-reinforced comparison stimulus when presented again later in the session. In the traditional NMTS arrangement, the sample, as well as matching and

non-matching comparison stimuli, are all presented in an array and selection of the odd (non-matching) stimulus is reinforced. In the OST, however, each of the odors displayed in the arena is different and the selection of the non-matching odor is based only on remembering which odors have been presented earlier in the session. Thus, in the OST there is not an “odd” stimulus in the sense usually used to define the oddity preference effect. Further, in a recent study using an automated version of the OST, a go/no-go procedure was used in which only one odor was presented on each trial and rats learned the task rapidly and showed accurate performances with up to 15 odors even when it was not possible for the rat to directly compare odors during a trial (Galizio et al. 2020). Although direct comparison of odors was not possible in Galizio et al., novelty preference could still provide an account of the rapid learning that was observed.

Thus, the present study sought to replicate the Galizio et al. (2020) study with the addition of an MTS version of the task to assess whether novelty preference was the basis for previous findings with the OST. Up to this point, no data have been published on a matching variant of the OST. The present procedure used the same methods as the non-matching version with the only difference being that responses were reinforced for stimuli that had previously been presented during each session, thus making it an incrementing MTS task. Thus, if rats performed similarly on this matching variant compared to the nonmatching procedure, it would underscore the importance of the olfactory modality rather than a novelty preference bias.

Method

Subjects

Twelve male Sprague–Dawley rats served as subjects, ranging from 72 to 254 days old at the beginning of testing (see Table 1). The rats were housed individually on a reversed 12/12-h light/dark cycle and provided water ad libitum. Food was restricted so that subjects maintained approximately 85% of their free-feeding weight. Animal care and treatment was consistent with the NIH Guide for the Care and Use of Laboratory Animals and the experimental protocol was approved by the UNCW Animal Care and Use Committee.

Apparatus and stimuli

Med Associates operant chambers (30.5 cm × 24 cm × 21 cm) were housed in a well-ventilated room with white noise (70 dB) presented throughout the testing sessions. The chambers were connected

Table 1 Ages, Experience, and Odor Sets of Subjects

Rat	Group	Age at beginning of experiment (Days)	Prior experimental experience ^a	Odor Sets ^b
W6	NMIT	254	Yes	A, B, D
Y14	NMIT	219	Yes	A, C, D
Y12	NMIT	225	Yes	A, C, D
Y11	NMIT	218	Yes	A, B, D
Z1	NMIT	70	No	A, C, D
Z10	NMIT	93	No	A, B, D
W5	MIT	221	Yes	A, C, D
Y13	MIT	220	Yes	A, C, D
Y9	MIT	218	Yes	A, C, D
Y10	MIT	218	Yes	A, B, D
Z2	MIT	99	No	A, C, D
Z7	MIT	70	No	A, C, D

Average ages of rats for NMIT and MIT groups were 179 and 174 days, respectively

^aRats were trained in a fixed ratio lever-pressing task for an undergraduate learning and behavior class

^bOdorants in each set can be seen in Table 2

to three separate five-port olfactometers (Med Associates ENV-275-5) which delivered the olfactory stimuli (via Teflon tubing and air delivery) through a port located in the center of the front chamber wall. The port included a photo beam to detect nose pokes, as well as an opening through which scented air could enter. Air was pumped into glass jars containing liquid odorants (Linear AC0102, 2.84 pound per square inch with an airflow of 0.177 cubic feet per min) and odor entry into the chamber was controlled by computer-operated solenoids. A vacuum pump (Linear VP0125, - 9.84 Hg vacuum and air displacement of 0.247 cubic feet/min) removed air from the chamber through a second opening in the port. In addition, each port contained a stimulus light that accompanied the presentation of odor stimuli. A house light at the top of the chamber, above the ports, was also illuminated during trials. Sucrose pellets (45 mg Bio Serv) served as reinforcement and were delivered via a hopper in the rear of the chamber opposite the nose poke port. A light above the hopper was illuminated during reinforcement delivery.

Odorants

Glass jars connected to the olfactometer contained solutions of 6.7 ml commercially available odorants mixed with 100 ml distilled water (except geraniol—1 ml/100 ml). The odorants used, along with their purchase sources, are listed in Table 2. Each rat was assigned to 13 of the 20 possible scents which were then used for that rat throughout the experiment (see Table 1 for scent assignments).

Table 2 Odorants used in the NMIT and MIT procedures

Scent	Manufacturer
Almond ^b	Great American Spice Co
Apple ^c	Great American Spice Co
Apricot ^a	Great American Spice Co
Blueberry ^d	Great American Spice Co
Brandy ^b	Great American Spice Co
Bubble Gum ^a	Great American Spice Co
Caramel ^c	Nature's Garden
Cinnamon ^a	Great American Spice Co
Clove ^d	Great American Spice Co
Coconut ^c	Great American Spice Co
Fresh Brewed Coffee ^a	Nature's Garden
Fresh Cut Grass ^c	Nature's Garden
Geraniol ^d	Sigma Aldrich
Honey ^d	Great American Spice Co
Peach ^d	Great American Spice Co
Pistachio ^b	Great American Spice Co
Root Beer ^a	Nature's Garden
Sandalwood ^c	Nature's Garden
Strawberry ^b	Great American Spice Co
Toasted Marshmallow ^d	Nature's Garden
Vanilla Butternut ^b	Great American Spice Co

Superscript denotes scent set of odors, referenced in Table 1

Procedure

Initial training

Testing in the olfactometers occurred from approximately 8:00 a.m. to 5:00 p.m. Monday through Friday. Magazine training was the first step with response-independent delivery of sugar pellets every 60 s. Once subjects were reliably consuming the sucrose pellets during a 30-min session, center nose poke responses were shaped by delivering a pellet for each response. When responding developed, odorants were pumped into the center port while reinforcement continued to be delivered for each center nose port response. When regular responding was observed, a fixed-interval (FI) 2-s schedule was arranged and was gradually thinned to a FI 5-s schedule which was in place throughout the remainder of the experiment.

Incremental matching and non-matching to sample training

Once responding was regular on the FI 5-s schedule, matching and non-matching training began. Subjects were randomly assigned to either a non-matching incrementing task (NMIT) or a matching incrementing task (MIT) with the constraint that an equal number of older experienced and

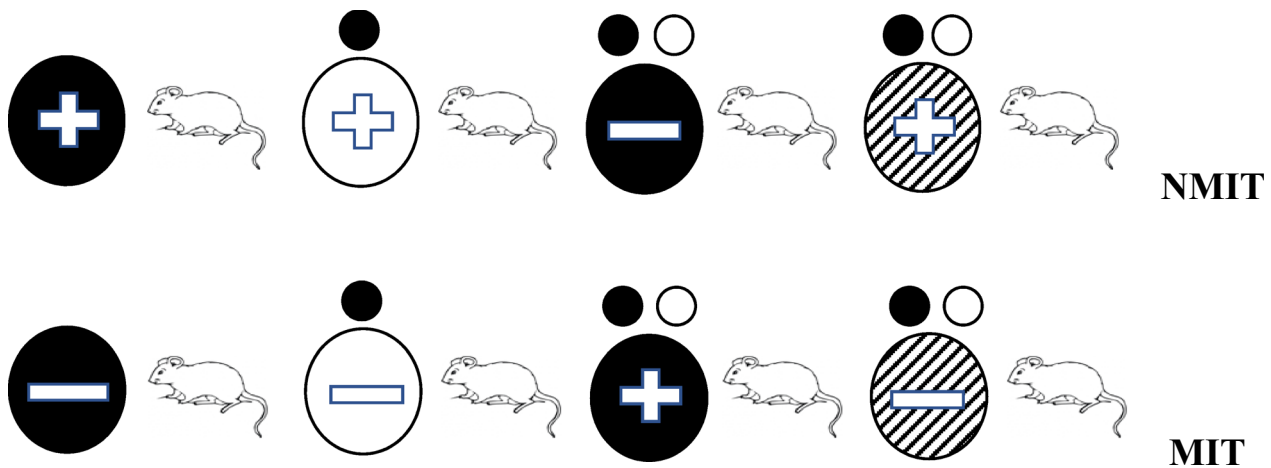


Fig. 1 Visual depiction comparing trial arrangements in the NMIT and MIT procedures. Black, white, and striped circles represent different odors. The large circle represents the stimulus presented on the

current trial, with a plus sign indicating the reinforcement of nose-poke behavior. Small circles represent stimuli encountered on previous trials that must be remembered to permit correct responding

younger naïve rats were represented in each group (see Table 1).

NMIT acquisition All trials were initiated by illumination of the house light and center stimulus port light. An observing nose poke response was required to initiate the delivery of an odor stimulus. During the first presentation of any given odor, responses were reinforced on an FI 5-s schedule, as that odor was session-novel at that point. This type of trial was referred to as a positive trial (S+). Nose pokes were recorded for the first 5 s of odor presentation, and the first response after 5 s turned off the house and center port lights, dispensed a pellet, and illuminated the hopper for 5 s, allowing the rat to retrieve the reinforcer. A 25-s intertrial interval was programmed after all trials to remove odors from the chamber between trials. Once an odor stimulus was presented, it subsequently served as a non-reinforced stimulus when presented again during the session. On a negative trial (S−), a randomly selected, session-familiar stimulus was presented after an observing nose poke and remained on for 5 s. No reinforcement was provided on S− trials, but nose pokes were still recorded. For analysis purposes, only responses made during the initial 5 s of S+ trials were compared to responses during S− presentations. The top panel of Fig. 1 gives a visual depiction of an example stimulus arrangement for four trials with three different odors: Odor 1 (black), Odor 2 (white), and Odor 3 (striped). In Trial 1, “black” is presented as a session-novel odor, so reinforcement is provided for responding according to the FI-5 s schedule. On Trial 2, “white” is presented, which is also a novel odor to that session. Therefore, responding after a 5-s interval results in reinforcement. “Black,” no longer session-novel, is presented again in Trial 3. This is an S− trial with no reinforcement for responding. Finally, responding to

“striped” on Trial 4 provides reinforcement (S+), as it has not been presented yet during the session.

The order of stimulus presentation during and across each 26-trial session (13 S+ and 13 S− trials) was random with the following constraints: no more than four consecutive trials could be S+ or S− and the same stimulus could not serve as an S− for more than two consecutive S− trials. Further, a given stimulus could not serve as an S− more than three times total during a session.

MIT acquisition The same procedures described for NMIT acquisition were followed except that responding to session-familiar stimuli was reinforced and to session-novel stimuli was not. For example, a scent that was presented on Trial 1 was always an S−, but once it was encountered again, it served as an S+ on all subsequent trials because it now matched a previously presented sample odor. The bottom panel of Fig. 1 provides a visual depiction of how stimuli might be arranged in the first four trials of the MIT procedure and offers a comparison to the NMIT procedure. In Trial 1, “black” is novel, and therefore, responses do not provide reinforcement. Likewise, the first presentation of “white” in Trial 2 provides no opportunity for reinforcement. “Black” is presented again on Trial 3, and because it now matches a previously presented odor, it is an S+ trial. Last, Trial 4 is a session-novel presentation of “striped,” and no reinforcement is provided.

Design and statistical analyses In addition to analysis of response rates, performance was assessed using a discrimination index (DI): responses on S+ trials divided by the total responses on both S+ and S− trials during a given session. All subjects were tested for a minimum of 20 sessions and a mean DI of 0.70 across 10 consecutive sessions was designated as the criterion for mastery of the task. If a

subject's 10-day DI averaged less than 0.60 after 75 sessions of training, or if a subject failed to reach the 0.70 mastery criterion after 100 sessions, the experiment was terminated for that subject.

A *t*-test analysis was conducted for acquisition rates of the two groups using the number of sessions until either the mastery criterion was reached or the animal failed to reach criterion (75 or 100 sessions were entered into the analysis as appropriate). Due to a laboratory closure in the wake of Hurricane Florence, training was interrupted for nearly a month for three of the six MIT group rats after approximately 40 training sessions. To both account for this interference and to show how behavior differentially changed over time between the two groups, a two-way mixed analysis of variance (ANOVA) was conducted on the DI of the first 20 training sessions (before the hurricane interruption occurred). This ANOVA was a 2×5 mixed design, with group (NMIT, MIT) as the between-subjects factor and blocks of four sessions (1–4, 5–8, 9–12, 13–16, 17–20) as the within-subjects factor. Further, in order to better characterize how training influenced responding, separate within-subject ANOVAs were conducted for each group to analyze changes in response rates during acquisition to S+ and S– stimuli. The design was a 2×5 within-subjects ANOVA with trial type (S+, S–) and blocks of four sessions (1–4, 5–8, 9–12, 13–16, 17–20) as factors.

Results

Figure 2 shows DI for each subject in the two groups across sessions. All six members of the NMIT group (top panels) successfully and quickly acquired the discrimination, with 14 sessions as the minimum (Z1) and 30 as the maximum (Z10) for time to reach criterion DI. These subjects tended to show rapid and steep increases in DI, as most improved from at or slightly below-chance performance in the first few sessions to above 0.70 in between 9 and 16 sessions. The changes in response rates which provided the basis for the rapid improvement varied across rats (see Fig. 3, top panels). For example, Rats Z10, Y11 and Z1 showed little change in responding to session-novel S+ odors across sessions, but response rates dropped markedly to S– odors as training progressed. Rats Y14, Y11 and Y12 both showed increased S+ responding and decreased S– responding across sessions. Response rates were clearly differentiated for all rats within 10 to 15 sessions.

In contrast, rats trained in the MIT condition showed slower acquisition (Figs. 2 and 3, lower panels). Animals in this group had a difficult time achieving the mastery criterion. Only three of the six animals (W5, Y9, Z7) ever achieved a 10-day DI average of 0.70, and even in these successful cases, Y9 and W5 required 73 and 99 sessions,

respectively, to meet criterion DI (see Fig. 2). In contrast, Rat Z7 showed a rapid acquisition that more resembled that of the NMIT rats. He first reached a DI of 0.70 on his ninth session and met the DI criterion in 18 sessions. However, the remaining three animals failed to ever meet the DI criterion. Rat Z2 showed signs of acquisition after 20 sessions but performance declined and was at near chance levels when testing was discontinued after 75 sessions. Rats Y10 and Y13 also developed above chance accuracy after extensive training but their DIs were quite variable and the experiment was terminated for them after 75 sessions without their meeting criterion. Rats in the MIT condition generally showed DIs that were at or slightly above chance (0.5) levels at the beginning of training. For the three rats that did successfully acquire the discrimination (W5, Y9, Z7), increasingly high rates of responding to the repeated S+ odors seemed to be the main basis for improvement, with S– response rates remaining relatively unchanged across sessions (see Fig. 3).

Figure 4 presents the mean number of sessions to reach the mastery criterion for the two groups (or sessions until the experiment was terminated for subjects that did not meet criterion). Figure 4 also provides a breakdown of performances for the MIT animals that were and were not interrupted by Hurricane Florence. Regardless of the interruption, it is clear that the NMIT group mastered the task in fewer sessions than the MIT group. An independent samples *t*-test of number of sessions to reach the criterion revealed a significant difference $t(11) = 4.48$, $p < .01$, $d = 2.58$. Subjects that experienced no interruption in training appeared to require slightly fewer sessions, but this was largely due to the rapid acquisition of one rat (Z7). In this regard it is worth noting that one animal (Z2) not exposed to the interruption during training failed to acquire the task, and another that did experience it (Y9) successfully acquired the task.

In order to clarify whether the group differences were impacted by the interruption in training, a second analysis was performed on performances early in training before any interruptions occurred. Figure 5 shows mean DI for the first 20 sessions of training in blocks of four sessions. During the early few sessions (Blocks 1 and 2), both MIT and NMIT subjects hovered around chance performance, though MIT animals slightly outperformed their NMIT counterparts during the first block of four sessions. By Block 4 (Sessions 13–16), the NMIT animals began to show higher DIs and this discrepancy grew larger as training continued. A 2×5 mixed ANOVA analysis of the DI in 4-session blocks produced results complementary to the prior *t*-test analysis and statistically confirmed the trends seen in Fig. 5. There were significant main effects of both Session Block, $F(4, 40) = 42.05$, $p < .0001$, and Task Type, $F(1, 10) = 1721.40$, $p < .0001$, indicating that both groups improved their DIs over time, but the NMIT animals outperformed MIT animals overall. Further,

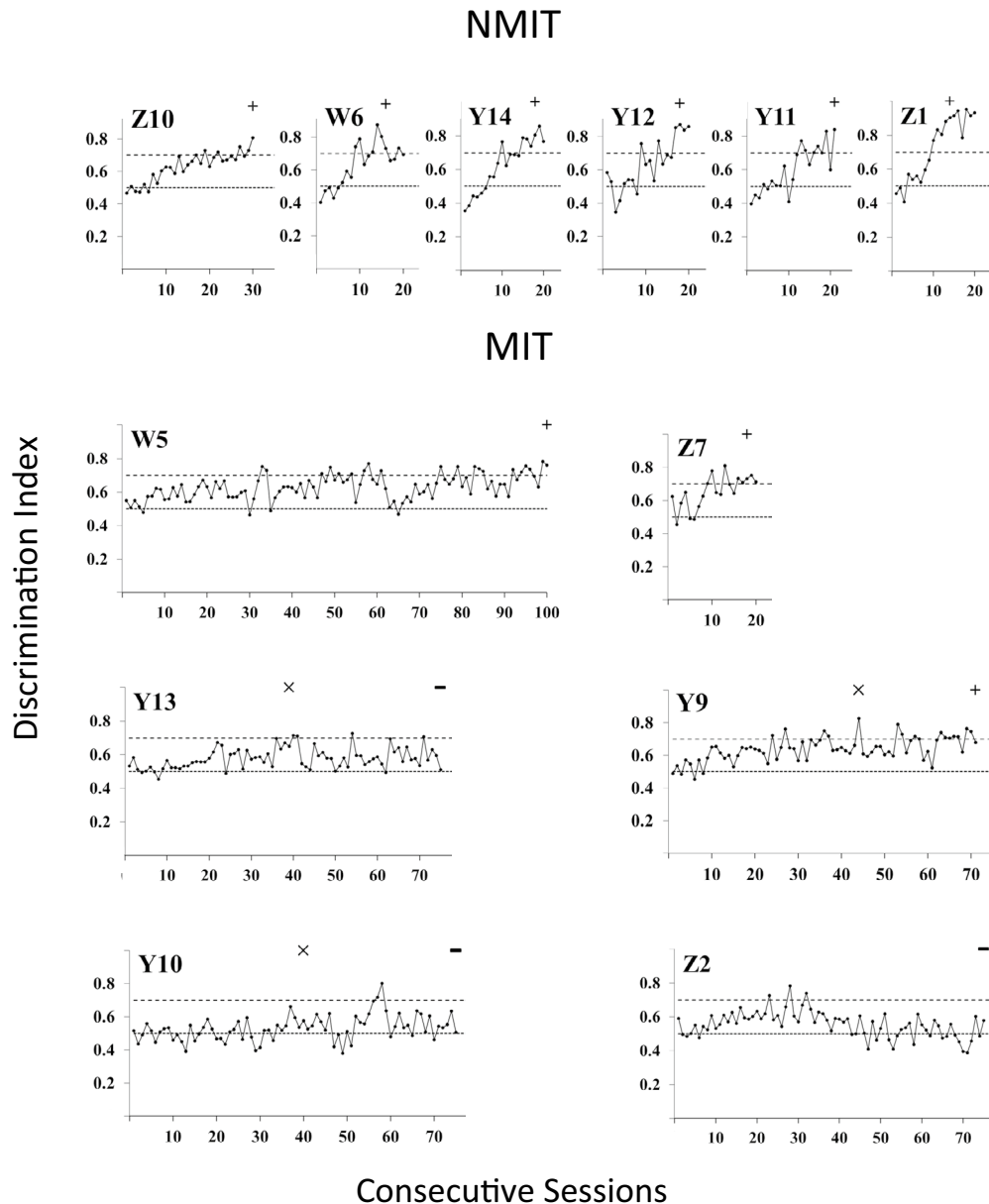


Fig. 2 Discrimination Index (DI) across Sessions for Individual Rats. Horizontal axis indicates the number of sessions until the animal either met the set criterion or was dropped from the study. The “X” icon represents the 25-day interruption of testing due to Hurricane Florence. Plus signs indicate the completion of mastery criterion,

while a minus sign denotes an animal failing to meet criterion. The dotted line at .5 indicates chance performance, or responding to S+ and S− equally, and the dotted line at .7 represents the mastery criterion threshold

those effects were qualified by a significant interaction, $F(4, 40) = 13.84$, $p < .0001$. Tests of simple main effects (t -tests at each block) showed that MIT animals started the acquisition phase performing significantly better than the NMIT group in Block 1 ($p < .001$), but by Blocks 4 and 5, the NMIT group consistently out-performed the MIT group (both $p < .02$).

Figure 6 shows mean response rates on positive and negative trials across the first 20 sessions in blocks of four

sessions which provides additional insight into the differences between MIT and NMIT performances. In general, it is clear that response rates were consistently higher for the MIT group than the NMIT group regardless of type of trial and across session blocks. In the second half of acquisition, an interesting divergence began to appear as follows: responses rates to positive trials increased for the MIT, but not the NMIT, group and rates to negative trials decreased for the NMIT group, but much less so for the MIT group.

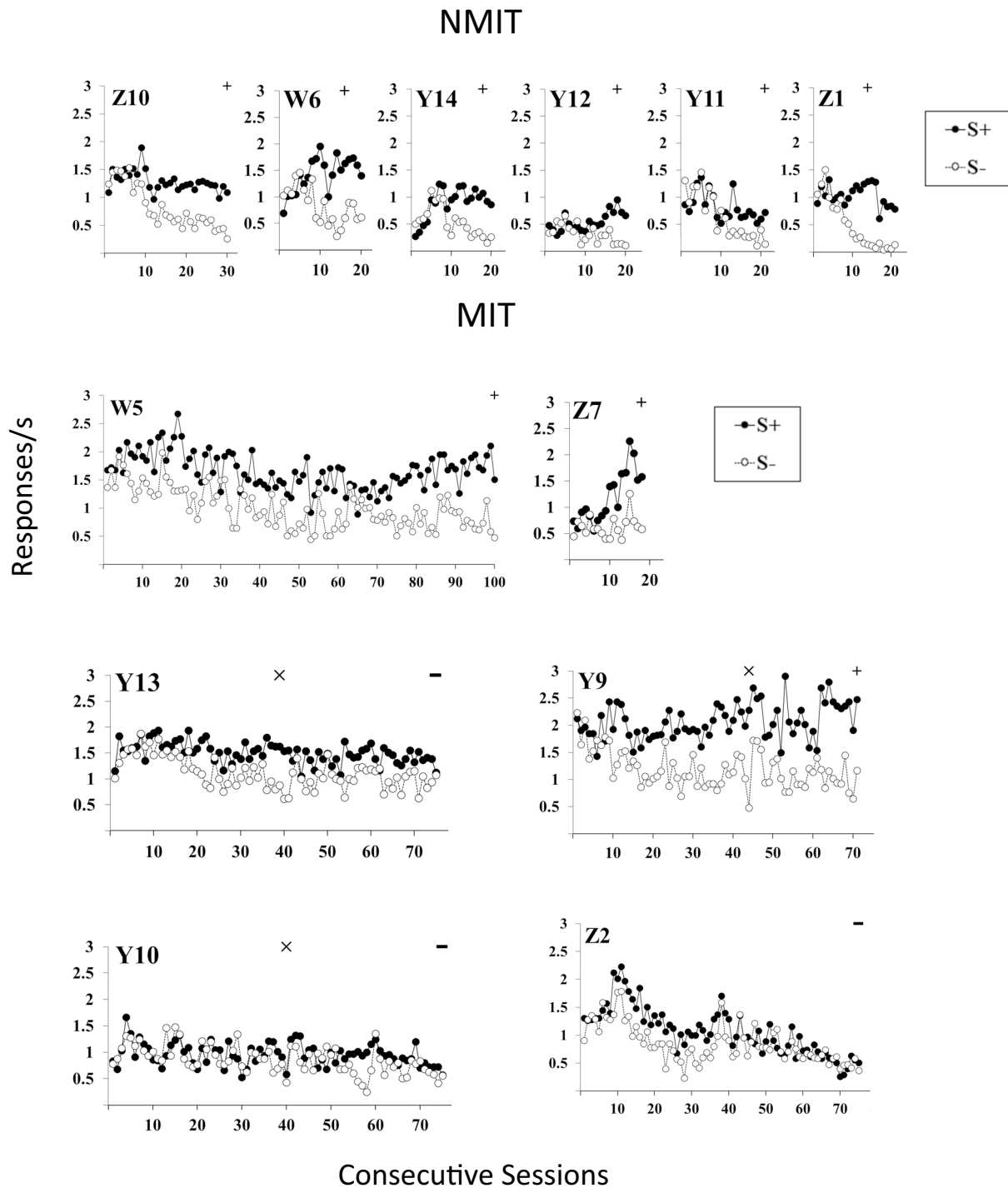


Fig. 3 Responses/s across Sessions for Individual Rats. The x-axis represents session number, while the y-axis indicates the number of responses per s to both S+ (filled circles and solid lines) and S- (open circles with dotted lines). Plus and minus signs indicate mas-

tery criterion being met or not, respectively. “X” markings indicate the beginning of a testing break in between sessions due to the hurricane

By the final block shown in Fig. 6, the highest rates were clearly evident for positive trials for the MIT group and the lowest rates for the NMIT group on negative trials. Rates

for negative trials for the MIT group were almost identical to those for positive trials in the NMIT group.

These observations were confirmed via two, 2×5 within subjects ANOVAs. For the MIT group, the main effect of

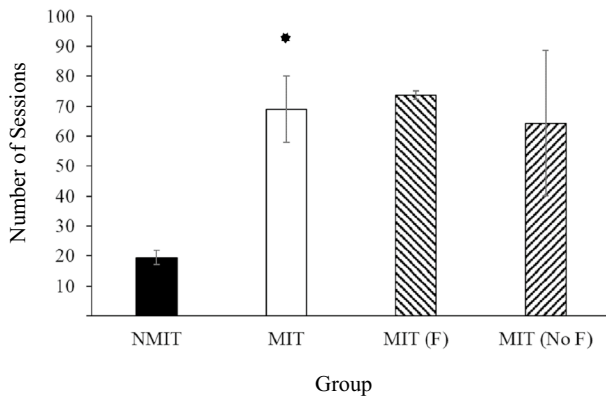


Fig. 4 Mean number of sessions to reach criterion. The left two bars show the means for the NMIT and MIT groups, respectively. The third and fourth bars show the means of the three animals that were (F) and were not (no F) interrupted by the impact of Hurricane Florence, respectively. Error bars indicate SEM and the asterisk indicates a significant difference between NMIT and MIT groups at the $p < .01$ level

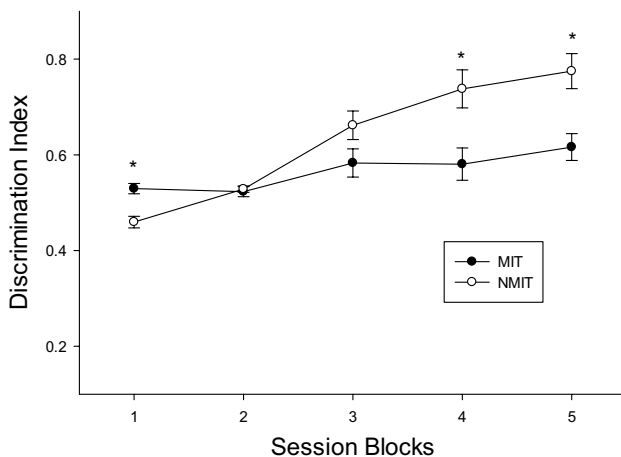


Fig. 5 Mean DI through the first 20 sessions separated into four-session blocks. Error bars indicate SEM. Asterisks indicate significant difference between groups, $p < .05$

trial type [$F(1, 5) = 13.81, p < .02$] and the interaction of trial type and session blocks [$F(4, 20) = 6.85, p < .001$] were significant. MIT rats did respond more overall on positive trials indicating some evidence for acquisition. While the interaction effect was significant, tests of simple main effects showed that there was little change in response rates across blocks, with the exception of a dip in responding on negative trials in Block 5 ($p < .03$).

For the NMIT group, both main effects of trial type [$F(1, 5) = 50.75, p < .001$] and Session Block [$F(4, 20) = 4.67, p < .008$] were significant, showing that, in general, NMIT rats responded more on positive than negative trial types

and their overall response rates generally decreased across the 20 sessions. However, these effects were qualified by the significant interaction, $F(4, 20) = 30.69, p < .0001$, showing that response rates on positive trials were consistent across session blocks, but response rates on negative trials dropped significantly in the second half of acquisition. This was confirmed by a simple main effects one-way ANOVA, $F(4, 20) = 26.04, p < .004$, followed by pairwise comparisons showing the significant drop in responding in Blocks 3, 4, and 5 compared to Blocks 1 and 2 ($p < .05$).

Because overall response rates were generally higher for all trial types for subjects in the MIT group, a signal detection analysis was developed to provide an estimate of discriminability separate from response bias (White 2013). In order to calculate d' , two or more responses had to occur on a trial to count as a hit or false alarm. This was because the first response on any given trial could be considered an olfactory observing response (although rats likely could detect odor presentations without entering the nose port, at least one nose poke did occur on most trials). Thus, an S+ trial on which the rat responded two or more times was considered a hit, an S− trial on which two or more responses occurred was counted as a false alarm, a trial on which 1 or 0 responses occurred was considered a miss on S+ trials and a correct omission on S− trials. Figure 7 shows the results of the signal detection analysis for the first 20 trials plotted in the same fashion as in Fig. 5—as a function of blocks of four trials. The overall pattern of results was quite similar to the DI results. As in that analysis, there were significant main effects for Task Type, $F(1, 5) = 12.66, p = .005$ and Blocks, $F(4, 20) = 42.96, p < .0001$, but the key finding was a Block \times Trial type interaction, $F(4, 20) = 10.80, p < .001$. As with the DI analysis (Fig. 5), the interaction was based on more rapid acquisition in the NMIT group, but some differences were apparent in Fig. 7. For example, the group differences for d' in Block 1 were smaller and not significant, and d' was significantly higher for the NMIT group as early as Block 2 ($p = .008$), indicating more rapid discrimination formation. Consistent with the DI analysis, the NMIT group had higher d' scores during the final two blocks ($ps < .001$ and $.004$, respectively).

In sum, the patterns of learning were different in the two groups with NMIT subjects showing rapid acquisition based largely on learning to reduce responding on S− trials (when stimuli were repeated) with little change on S+, and the slower acquisition in the MIT group based on learning to increase responding to S+ (repeated) trials with no real change in response rates on S− (novel odor) trials. Reconsidering the response rate data for the MIT subjects (see Fig. 3), it should be noted that Rat Z7 showed a rapid enough increase in S+ responding to meet criterion quickly. The other two MIT rats that did meet criterion (W5, Y9) eventually did show some reduction

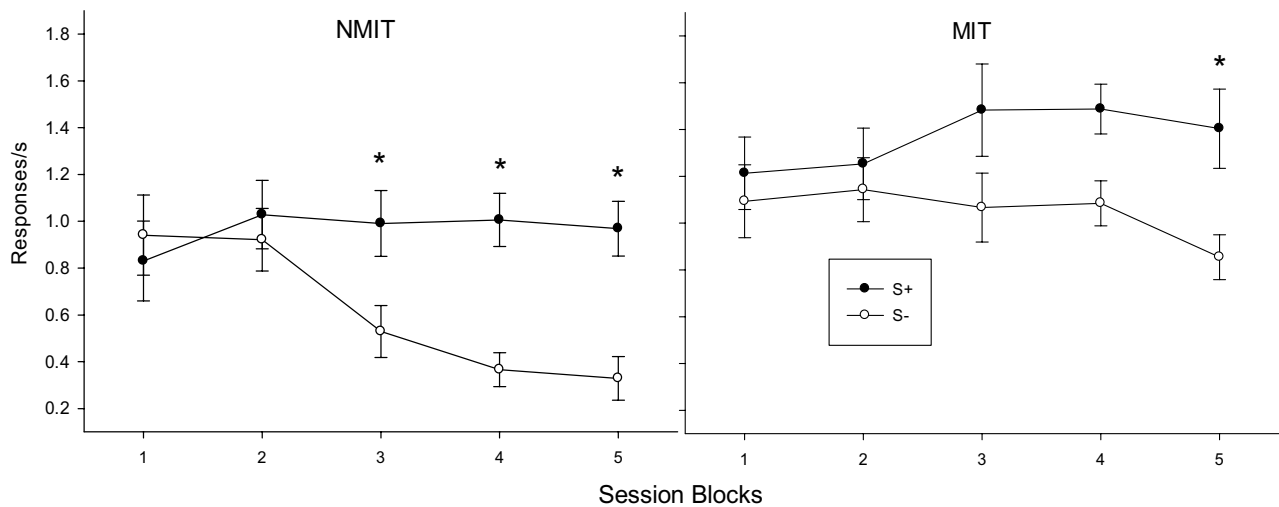


Fig. 6 Responses/s on positive and negative trials through the first 20 sessions separated into four-session blocks. Error bars indicate SEM and asterisks indicate blocks on which t responses rates on negative trials were significantly lower than ($p < .05$)

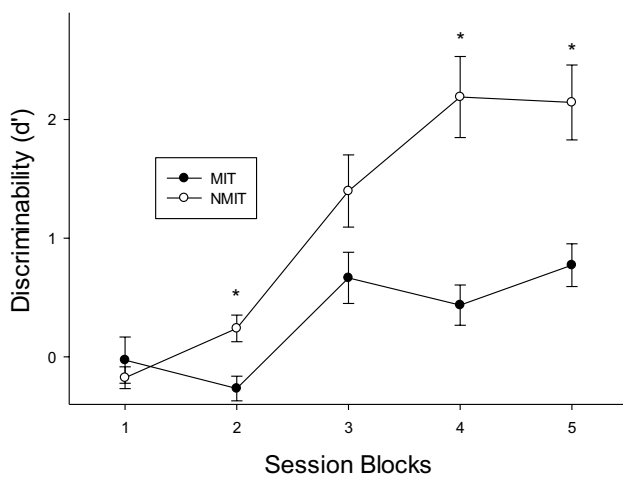


Fig. 7 Mean d' through the first 20 sessions separated into four-session blocks. Error bars indicate SEM. Asterisks indicate significant difference between groups, $p < .05$

in responding on S– trials. In contrast, all NMIT rats (Fig. 3) showed a decrease in S– responding that was evident within 8–12 sessions and allowed them to meet criterion quite rapidly. Finally, a signal detection analysis yielded essentially similar results with a measure not biased by overall response rate differences.

Discussion

Previous research has shown that rats readily acquire an incrementing non-matching to sample task with odor stimuli in an arena-based setting (OST—Dudchenko et al. 2000) as well as in an automated operant apparatus (NMIT—Galizio et al. 2020). A key finding from the present study was that rats can also learn a matching variation of the incrementing task (MIT), although acquisition was more difficult than in the non-matching condition. Every member of the NMIT group met the mastery criterion rather quickly (generally within 20 sessions) while only three of the six MIT subjects met criterion even after extensive training. Further, two of the three rats in the MIT condition that did meet criterion did so at a much slower pace than any member of the NMIT group. It may be worth noting that the one MIT rat that did show rapid acquisition (Z7) was relatively young when training began, but it must be added that the other young rat in that group (Z20) failed to meet criterion (see Table 1). The rapid acquisition observed in the NMIT group was consistent with previous results using the automated NMIT (Galizio et al. 2020) and with findings using the arena-based OST (e.g., Dudchenko et al. 2000). Using the OST and NMIT procedures, rats quickly learn to discriminate between odors they have encountered previously during the session and session-novel odors. Despite the acquisition difficulties in the MIT group, successful discrimination between session-novel and familiar odors was also demonstrated under matching conditions in three of the six animals. However, the substantial differences across the two procedures are puzzling.

The generally slower and inconsistent acquisition in the MIT group may be taken as support for the hypothesis that rapid acquisition of the NMIT (and the related OST) is

based, at least in part, on a preference for novelty. It suggests that the innate preference for less familiar stimuli in rodents (Antunes and Biala 2012; de Lima et al. 2005; Ennaceur 2010), including in the olfactory domain (Spinetta et al. 2008), was pervasive enough to interfere with the acquisition of the matching variation of the current task. From this perspective, the rapid acquisition in the NMIT condition can be understood as deriving from the tendency to approach session-novel odors which would lead to reinforcement and the relative ease of inhibiting approach to familiar odors. In contrast, the failures or slower acquisition in the MIT condition may have stemmed from difficulty in learning to inhibit responding to the preferred session-novel stimuli even though responding was not reinforced in their presence. Consistent with this interpretation, response rates to session-novel negative stimuli remained relatively high throughout the experiment in the MIT group. Indeed, the three rats that met criterion in the MIT group did so by developing even higher response rates to the repeated positive stimuli despite continuing to respond to the session-novel negative stimuli at rates similar to, or only moderately lower than, those produced early in training.

Although an account of the differences in acquisition in terms of novelty preference is plausible, some aspects of the data remain puzzling from this perspective. For example, rats in both the MIT and NMIT groups showed DIs that were near 0.5 (chance levels) for the first several sessions of training. If there was a bias for approaching novel and avoiding familiar stimuli, the NMIT group should have had above chance DIs, and the MIT group below chance DIs, early in training. However, the MIT group had slightly (but significantly) higher DIs on the first few sessions and between group differences only emerged after eight or more sessions of training (see Fig. 5). These findings would seem to contradict a novelty-preference explanation of the group difference; indeed, they suggest a slight initial bias toward responding to recently presented stimuli. Thus, it seems likely that other differences between the MIT and NMIT procedures played a role in the differential acquisition.

An alternative hypothesis might emphasize the non-reinforcement of responding during the initial exposure to each stimulus in the MIT. Although positive and negative trials were quasi-randomly distributed across both procedures, of necessity in the MIT each odor had to be initially presented as an S-. Non-reinforcement of responding to the initial presentation of each odor may have weakened tendencies to respond to subsequent presentations of that odor and thus delayed acquisition. Indeed, this feature of not reinforcing responses to the sample in a more traditional MTS procedure was shown to impair acquisition by Wright and Delius (2005). Further, Wright and Delius also found that reinforcing sample responses impaired NMTS acquisition, a procedural feature shared in our NMIT conditions. Thus, based

on the Wright and Delius account, the sample reinforcement conditions should have biased against acquisition in both the NMIT and MIT conditions and so cannot account for the more successful acquisition of the NMIT.

The Wright and Delius (2005) analysis is based on the premise that reinforcement of responses made to a sample stimulus increases the likelihood of responding to that stimulus (response strength) and non-reinforcement of sample stimulus responding reduces response strength. In fact, the rapid acquisition of the NMIT and OST procedures seems problematic for theories that emphasize the response-strengthening properties of reinforcement. However, as a number of recent theorists have noted, reinforced responses are only likely to reoccur when their reoccurrence is correlated with additional reinforcement (Cowie and Davison 2020, 2021; Shahan 2017). In the NMIT and OST, responses are reinforced on the initial presentation of each stimulus; traditionally viewed, this experience would be expected to strengthen tendencies to respond when that odor appeared again. This account would predict slow learning of these tasks. Thus, that rats rapidly learn to inhibit responding in the presence of stimuli previously associated with reinforcement in the OST and NMIT is more consistent with theories that emphasize the role of learning about the structure of environmental contingencies and then responding in accordance with likely future outcomes (Cowie and Davison 2020). However, the reasons why accurate MIT performances were slow to develop or failed to develop in some animals remains unclear from any theoretical perspective. Further research will be needed to determine the basis of differential acquisition between the MIT and the NMIT procedures. For example, it would be of interest to determine whether the effect can be replicated in other species and with different sensory modalities.

Finally, the substantial differences found here between performances in the NMIT and MIT may have implications for interpretation of research using the OST. The effects of drugs and other neurobiological manipulations on OST performance are generally interpreted in terms of memory capacity (e.g., Galizio et al. 2013; MacQueen et al. 2016; Scott et al. 2020), but all of these studies used an incremental non-matching-to-sample procedure. If OST performance is determined in important ways by novelty preference, then it is possible that the neurobiological manipulations are influencing that bias rather than memory per se. Similarly, De Falco et al. (2019) found unique patterns of neural activity in the medial prefrontal cortex in the OST when rats encountered a novel (rewarded) odor versus a familiar (unrewarded) odor. Thus, it would be of considerable interest to replicate these published studies using a matching version of the OST similar to the MIT.

However, it should be noted that the arena-based OST is not readily converted from a non-matching to a matching

task. In the traditional OST, it is easy to guarantee that the rat has observed the novel odor on each trial as the trial continues until that odor is selected and reinforced. It is not clear this can be arranged in a matching-to-sample version in the arena because on a given trial with, say, two comparison stimuli, the rat could select the previously presented (correct) odor without ever encountering the novel (incorrect) odor. On a subsequent trial that odor would still be novel to the rat and there would be no basis for correct selection. The automated MIT used here solved that difficulty by presenting only one odor on each trial and the observing response requirement assured that the rat was at least exposed to the odor. Despite its potential utility, the difficulty in training accurate MIT performance is a limitation of this procedure and it would be valuable to develop alternative training procedures to make acquisition of the MIT faster and more reliable.

Acknowledgements The authors thank Angela Bennett, Ashley Campbell, Genevieve Guidone, Chandler Hicks, Calista Holt, Jonathan Shaw and Shandy Smith for assistance in data collection. Funding was provided by grants from the UNC Wilmington CSURF program.

Funding Not applicable.

Availability of data and material (data transparency) The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Code availability Not applicable.

Declarations

Conflict of interest The authors declare no conflicts of interest or competing interests.

Ethical approval This research complied with the current laws of the United States of America and was reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) at the University of North Carolina Wilmington.

Consent for publication The authors have all approved this version of the manuscript and give consent for publication.

References

- Antunes M, Biala G (2012) The novel object recognition memory: neurobiology, test procedure, and its modifications. *Cogn Process* 13:93–110
- April LB, Bruce K, Galizio M (2013) The magic number 70 (plus or minus 20): variables determining performance in the rodent odor span task. *Learn Motiv* 44(3):143–158
- Bratch A, Kann S, Cain JA, Wu JE, Rivera-Reyes N, Dalecki S, Crystal JD (2016) Working memory systems in the rat. *Curr Biol* 26(3):351–355
- Bruce K, Dyer K, Mathews M, Nealley C, Phasukkan T, Prichard A, Galizio M (2018) Successive odor matching- and non-matching-to-sample in rats: a reversal design. *Behavioral Processes* 155:26–32
- Cowie S, Davison M (2020) Generalizing from the past, choosing the future. *Perspect Behav Sci* 43(2):245–258
- Cowie S, Davison M (2021) Pigeons prefer to invest early for future reinforcers. *J Exp Anal Behav* 115(3):650–666
- De Falco E, An L, Sun N, Roebuck AJ, Greba Q, Lapish CC, Howland JG (2019) The rat medial prefrontal cortex exhibits flexible neural activity states during the performance of an odor span task. *Eneuro* 6(2):1–16
- de Lima MNM, Laranja DC, Bromberg E, Roesler R, Schröder N (2005) Pre- or post-training administration of the NMDA receptor blocker MK-801 impairs object recognition memory in rats. *Behav Brain Res* 156(1):139–143
- Dudchenko PA, Wood ER, 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. *J Neurosci* 20:2964–2977
- Ennaceur A (2010) One-trial object recognition in rats and mice: methodological and theoretical issues. *Behav Brain Res* 215:244–254
- Galizio M, Deal M, Hawkey A, April B (2013) Working memory in the odor span task: effects of chlordiazepoxide, dizocilpine (MK801), morphine, and scopolamine. *Psychopharmacology* 225:397–406
- Galizio M, Deal M, Mathews M, Panoz-Brown D, Prichard A, Bruce KE (2019) Effects of NMDA antagonist dizocilpine (MK-801) are modulated by the number of distractor stimuli in the rodent odor span task of working memory. *Neurobiol Learn Mem* 161:51–56
- Galizio M, Mason MG, Bruce K (2020) Successive incrementing non-matching-to-samples in rats: an automated version of the odor span task. *J Exp Anal Behav* 114(2):248–265
- Krichbaum S, Rogers B, Cox E, Waggoner LP, Katz JS (2020) Odor span task in dogs (*Canis familiaris*). *Anim Cogn* 23(3):571–580
- Lazarowski L, Goodman A, Galizio M, Bruce K (2019) Effects of set size on identity and oddity abstract-concept learning in rats. *Anim Cogn* 22:733–742
- MacQueen DA, Bullard L, Galizio M (2011) Effects of dizocilpine (MK801) on olfactory span in rats. *Neurobiol Learn Mem* 95:57–63
- MacQueen DA, Dalrymple SR, Drobles DJ, Diamond DM (2016) Influence of pharmacological manipulations of NMDA and cholinergic receptors on working versus reference memory in a dual component odor span task. *Learn Mem* 23(6):270–277
- Marks WN, Parker ME, Zabder NK, Greba Q, Snutch TP, Howland JG (2018) T-type calcium channels in the orbitofrontal cortex mediate sensory integration as measured using a spontaneous oddity task in rats. *Learn Mem* 25(7):317–324
- Nigrosh BJ, Slotnick BM, Nevin JA (1975) Olfactory discrimination, reversal learning, and stimulus control in rats. *J Comp Physiol Psychol* 89(4):285–294
- Rushforth SL, Steckler T, Shoaib M (2011) Nicotine improves working memory span capacity in rats following sub-chronic ketamine exposure. *Neuropsychopharmacology* 36(13):2774–2781
- Scott GA, Liu MC, Tahir NB, Zabder NK, Song Y, Greba Q, Howland JG (2020) Roles of the medial prefrontal cortex, mediodorsal thalamus, and their combined circuit for performance of the odor span task in rats: analysis of memory capacity and foraging behavior. *Learn Mem* 27(2):67–77
- Shahan TA (2017) Moving beyond reinforcement and response strength. *Behav Anal* 40(1):107–121
- Slotnick BM, Kufera A, Silberberg AM (1991) Olfactory learning and odor memory in the rat. *Physiol Behav* 50(3):555–561
- Spinetta MJ, Woodlee MT, Feinberg LM, Stroud C, Schallert K, Cormack LK, Schallert T (2008) Alcohol-induced retrograde memory impairment in rats: prevention by caffeine. *Psychopharmacology* 201(3):361–371

- Thomas RK, Noble LM (1988) Visual and olfactory oddity learning in rats: what evidence is necessary to show conceptual behavior? *Anim Learn Behav* 16(2):157–163
- White KG (2013) Remembering and forgetting. In: Madden GJ (ed) *APA handbook of behavior analysis, vol. 1: methods and principles*. American Psychological Association, pp 411–437
- Wright AA, Delius JD (2005) Learning processes in matching and oddity: the oddity preference effect and sample reinforcement. *J Exp Psychol Anim Behav Process* 31(4):425–432

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.