Learning & Behavior, in press

Validation of a rodent model of episodic memory replay

Cassandra L. Sheridan[1], Danielle Panoz-Brown[1], Richard M. Shiffrin[1], and Jonathon D. Crystal[1]

[1]Department of Psychological & Brain Sciences, Indiana University, 1101 E 10TH ST, Bloomington IN 47405.

Correspondence:

Jonathon D. Crystal

jcrystal@indiana.edu

Abstract

Vivid episodic memories in humans have been described as the replay of the flow of past events in sequential order. Recently, Panoz-Brown et al. (2018) developed an olfactory memory task in which rats were presented with a list of trial-unique odors in an encoding context; next, in a distinctive memory assessment context, the rats were rewarded for choosing the second to last item from the list while avoiding other items from the list. In a different memory assessment context, the fourth to last item was rewarded. According to the episodic memory replay hypothesis, the rat remembers the list items and searches these items to find the item at the targeted locations in the list. However, events presented sequentially differ in memory trace strength, allowing a rat to use the relative familiarity of the memory traces, instead of episodic memory replay, to solve the task. Here, we directly manipulated memory trace strength by manipulating the odor intensity of target odors in both the list presentation and memory assessment. The rats relied on episodic memory replay to solve the memory assessment in conditions in which reliance on memory trace strength is ruled out. We conclude that rats are able to replay episodic memories.

Keywords: replay; episodic memory; familiarity; memory trace strength; olfaction; rat

An essential aspect of human cognition is the ability to recall a series of events in sequential order using episodic memory. For example, when describing a car accident, the observer may rely on episodic memory to "replay" what events occurred in sequential order (e.g., the red car was in the intersection when the black vehicle missed the red light and hit the red car). Traditionally, episodic memory was thought to be uniquely human. However, over twenty years of research has provided compelling evidence that nonhuman animals have several elements of episodic memory (Clayton & Dickinson, 1998; Crystal, 2021). Episodic memory consists of representations of unique past personal events (Tulving, 1972, 1983), which is different from semantic memory, which consists of factual knowledge about the world (Tulving, 1972). Importantly, episodic memory is distinct from judgments of familiarity, which are based on different neural substrates (Carlesimo et al., 2007; Eldridge et al., 2000; Henson et al., 1999; Hofer et al., 2007; Schmitter-Edgecombe & Anderson, 2007). Familiarity is the vague judgment that an item is known, whereas episodic memory consists of remembering an event and the contextual details surrounding the event (Eichenbaum, 2007; Eichenbaum et al., 2012; Eichenbaum et al., 2007; Yonelinas, 2001; Yonelinas & Levy, 2002).

In a classic experiment, Clayton and Dickinson (Clayton & Dickinson, 1998) demonstrated the first evidence for memory of the what, where, and when of an event in nonhumans, an ability that they termed episodic-like memory. On some trials, scrub jays cached peanuts followed by worms, whereas on other trials they cached worms followed by peanuts. Opportunities to retrieve the cached food occurred after a short or long delay. Subsets of birds experienced decayed or fresh worms after the long delay. The jays preferred the cache sites with worms over the peanut sites when the worms were fresh but preferred the peanut sites when the worms were decayed. The jays remembered what, where, and when they had cached a particular

food type based on trial-unique caching experiences. In addition to scrub jays, episodic-like memory has subsequently been demonstrated using a variety of approaches in numerous species (Clayton et al., 2020; Davies et al., 2022; Hoffman et al., 2009; Jozet-Alves et al., 2013; Pahl et al., 2007).

The central hypothesis of an animal model of episodic memory proposes that, at the moment of the memory assessment, the animal remembers back in time and retrieves a memory of an earlier event (Crystal, 2013; Crystal, 2016a, 2016b; Crystal, 2021, 2024, in press). However, an alternative hypothesis proposes that an animal can complete the memory assessment without remembering back to the specific earlier event by relying on judgments of memory trace strength (i.e., familiarity) (Clayton et al., 2000; Roberts et al., 2008). Clayton and colleagues (Clayton et al., 2000) pointed out that many tests of memory in animals (e.g., matching to sample) may be solved by use of memory trace strength. An animal can use judgements of memory trace strength to complete a memory assessment because the presentation of a stimulus gives rise to a memory trace and memory trace strength changes as a function of time. Since the age of memories can be determined from a comparison of memory trace strengths, an animal may complete a memory assessment by following a relatively simple rule. For example, a new-old recognition task may be completed by choosing the presented item that currently has the highest memory trace strength. Notably, an animal that uses judgments of memory trace strength does not need to retrieve an episodic memory of an earlier event. To provide compelling evidence of episodic memory in animals, the use of non-episodic memory solutions must be ruled out (Clayton et al., 2000; Crystal, 2021; Roberts et al., 2008).

Various approaches have been used to assess episodic memory of a single event in rats (Babb & Crystal, 2005, 2006a, 2006b; Crystal, 2013; Crystal & Alford, 2014; Crystal et al.,

2013; Crystal & Smith, 2014; Panoz-Brown et al., 2018; Panoz-Brown et al., 2016; Sheridan et al., 2024; Zhou & Crystal, 2009, 2011; Zhou et al., 2012). For example, evidence has shown that what-where-when memory in rats is based on episodic memory (Zhou & Crystal, 2009, 2011). In addition, a number of studies suggest that rats remember the source of encoded information (Crystal & Alford, 2014; Crystal & Smith, 2014; Smith, Dalecki, et al., 2017; Smith, Slivicki, et al., 2017; Smith et al., 2016); source memory is memory for the source or origin of episodic memories (Janowsky et al., 1989; Johnson et al., 1993; Mitchell & Johnson, 2009). Furthermore, it has been documented that rats can use episodic memory to answer an unexpected question after incidental encoding (Sheridan et al., in press; Sheridan et al., 2024; Zhou et al., 2012). However, episodic memory in humans is not limited to a single event or episode. Therefore, several studies have explored the possibility that rats have episodic memories of multiple events. Crystal and Smith (2014) used a source-memory approach with a memory load of two events to document that rats use episodic memory to remember details about two distinct events and that rats remember episodic memories as bound representations. The binding hypothesis proposes that the source memory for an event is stored with the remaining features of the episodic event in an integrated representation (Crystal, 2021). Thus, binding functions to help differentiate between potentially confusable events in episodic memory (i.e., events that share some, but not all, features). The above findings set the stage to ask if rats can use episodic memory to remember the details about many different events. Because episodic memory includes the memory of the items and the contexts in which they were presented (Eichenbaum, 2007), Panoz-Brown et al (2016) developed an approach to test the hypothesis that rats remember many items in context using episodic memory. Panoz-Brown and colleagues' findings that rats are able to remember many unique events using episodic memory set the stage to test the hypothesis that

rats remember the *order* of multiple episodic memories, which would enable a rat to replay its episodic memories.

Vivid episodic memories in humans have been referred to as the replay of the flow of past events in sequential order and mental time travel (Crystal, 2024, in press; Dede et al., 2016; Eichenbaum, 2000; Eichenbaum et al., 2007; Kurth-Nelson et al., 2016; Staresina et al., 2013; Tulving, 2002). Panoz-Brown et al. (2018) proposed that rats represent multiple items in episodic memory and can engage in episodic memory replay, which is a process by which the rat searches its representations in episodic memory to find items at specific points in a sequence. As shown in Figure 1, rats were presented with a list of odors in a distinct encoding context (see top of schematic). The list of odors varied in length from 5 to 12 items (represented by colored circles in Figure 1, with different colors representing different odors at the top of the schematic), which were randomly selected on each trial. Importantly, the rat could not predict when the list would end. This feature of the approach is essential because the correct choice in the memory assessments could not be identified until after the list ended. When a list ended, the rat was placed in one of two distinct arenas (i.e., referred to as a memory assessment context). In each memory assessment context, the rat was given a choice between two items from the list. In one context, the second to last item (represented by a dark blue circle in Figure 1) was the correct choice (i.e., referred to as the second last memory assessment). In the other context, the fourth to last item was the correct choice (i.e., referred to as the fourth last memory assessment). In each memory assessment, the other item (i.e., the incorrect choice) was randomly selected from a different position in the list (represented by a light pink circle in Figure 1). After the rat learned which arena is used in each memory assessment, a rat that can replay a sequence of episodic memories would select the correct item in the second and fourth last memory assessments at

levels above chance (chance is 50%). Each trial used a unique set of odors randomly sampled (without replacement) from a large inventory of odors.

The results from Panoz-Brown et al. (2018) showed that rats chose the second and fourth to last items with an accuracy above chance (i.e., about 80%). As noted above, the episodic memory hypothesis proposes that rats represent multiple items in episodic memory and can engage in episodic memory replay, a process by which the rat searches its representations in episodic memory to find specific information. However, there is a non-episodic memory solution to the memory assessments described above. Notably, when an item is presented, it gives rise to a memory trace whose level of activation (i.e., strength) changes as a function of time; the level of memory trace strength at the moment of the memory assessment is represented by the backward exponential curve in Figure 1. Therefore, at the time of the memory assessment (see right side of schematic), items located in different ordinal positions within the list have memory trace strengths typical for these specific positions (shown by the orange dots on the exponential curve in Figure 1). Notably, the second to last item has a typical memory trace strength (represented by the large black arrow in Figure 1), and the animals may have learned to select the choice that matched the typically rewarded memory-trace strength associated with each context. Thus, the rats could be merely judging the relative strength (i.e., familiarity) of memory traces to successfully solve the memory assessments by avoiding foils that have memory trace strengths that depart from the target level highlighted by the black arrow. Importantly, this familiarity-based solution would not require episodic memory nor the replay of episodic memories.

As noted above, the presentation of an item gives rise to a memory trace that changes as a function of time. There are two sources of change in memory trace strength that occur over time. One source is similar to temporal decay; over time the context in the experiment gradually

changes relative to the context of encoding, resulting in a change in memory trace strength. The other source is inter-item interference, meaning that the encoding of new items interferes with the retrieval of old items. Thus, storing new experiences, especially highly similar experiences, in memory, makes successful memory retrieval difficult (Anderson, 2003).

To dissociate the strength of memory traces and item sequence, Panoz-Brown et al (2018) conducted an experiment in which they doubled the amount of time between the presentation of list items, which impacted memory trace strength without influencing the order of items (see Figure 2). Notably, the incorrect choice in each memory assessment was specifically chosen to match the delay of a typical second or fourth to last item. For example, as shown in Figure 2B, during the second last memory assessment (see right side of schematic), the incorrect choice was the last item presented in the list because it had occurred in the list at the delay that was typical of a second to last item (shown by the arrow in Figure 2), whereas the second to last item (i.e., the correct choice) had an atypical delay. In the fourth last memory assessment, the incorrect choice was the second to last item because that item had a delay that was typical of a fourth to last item, whereas the fourth to last item had an atypical delay. Therefore, if memory trace strength changes gradually as a function of time, then the incorrect item would be an attractive choice (it matches the typical memory trace strength of a second last item, represented by the arrow in Figure 2). By contrast, if a rat replayed the order of items in episodic memory, the correct item would be an attractive choice. In summary, an animal relying on matching memory trace strengths would choose the incorrect item, producing accuracy below chance, whereas an animal that relies on episodic memory replay would choose the correct item, producing above-chance accuracy. Accuracy was above chance (i.e., about 80%), which supports the hypothesis that rats can replay episodic memories and ruled out the hypothesis that rats relied

on memory trace strength that gradually declined as a function of time (Panoz-Brown et al., 2018).

As noted above, both the passage of time and the presentation of new to-be-remembered items may contribute to changes in memory trace strength. If the passage of time is the major contributor to memory strength, then the study conducted by Panoz-Brown et al. (2018) is an adequate test to rule out the hypothesis that rats use judgments of relative familiarity to complete the memory assessments. If inter-item interference is the major contributor to memory trace strength, then further experimentation is needed to validate the replay of episodic memory because doubling the delay between items did not impact inter-item interference.

In the present study, we tested the hypothesis that rats rely on memory trace strength that declined as a function of inter-item interference. We directly manipulated memory trace strength by manipulating the odor intensity of target odors in both the list presentation and memory assessments. We conducted a pilot study to verify that rats are able to discriminate an individual odor when presented at low and high potencies. Accuracy in discriminating odor intensity was 0.97 ($\pm$ 0.02).

Our approach was the same as that used by Panoz-Brown et al (2018). Rats were presented with a list of 5-12 odors in a distinct memory assessment context (Figure 3A, see the top of the schematic). After list presentation, rats were moved to a memory assessment context where they were given a choice between two odors from the previously presented list (see right side of schematic). Here, we focus on memory trace strength mediated by inter-item interference. The presentation of an item gives rise to a memory trace (represented by the orange dots in Figure 3A) that declines as a function of inter-item interference (represented by the backward exponential curve in Figure 3A). Thus, at the moment of the memory assessment (right side of

the schematic), the last item presented has a high memory trace strength because no subsequent item is presented, so there is relatively little inter-item interference to degrade the memory trace. By contrast, the first item presented has a lower memory trace strength because it is followed by many subsequent items. Importantly, since the second to last item always has one item presented after it, each second to last item has a typical memory trace strength, which corresponds to the arrow in Figure 3A. Similarly, the fourth to last item is always followed by three subsequent items, producing its own typical memory trace strength. Therefore, a rat may choose an item in the memory assessment that matches the typical memory trace strength of a second to last item (represented by the arrow in Figure 3A) or fourth to last item without replaying the sequence of items in episodic memory.

To dissociate memory trace strength and ordinal information, we directly manipulated memory trace strength by increasing the potency of the target odor (i.e., second last or fourth last item), thus creating memory traces that have atypical strengths (Figure 3B*,* high potency odor is represented by smoke above the colored circle). Now, as shown on the right side of the schematic in Figure 3B*,* the memory trace strength for the second to last item (high potency item, represented by the black square above the curve in Figure 3B) does not match the typical memory trace strength of a second to last item, which is represented by the arrow in Figure 3. We chose to manipulate the potency (i.e., intensity) of target odors because differences in stimulus intensity impact memory accuracy (Berliner & Durlach, 1973; Killeen & Grondin, 2022). It has been proposed that memory traces of highly intense stimuli take longer to decay than weak stimuli, which suggests that intense stimuli enhance activation of memory (Berliner & Durlach, 1973; Killeen & Grondin, 2022). Previous studies have also shown that retention

increased as a function of opportunities to encode (Nelson & Wasserman, 1978; Roberts, 1972; Roberts & Grant, 1978).

The high potency test does not offer an odor with a memory trace strength that matches the level typical for the memory assessment context. Therefore, a rat that matches memory trace strength will perform at a level below its baseline accuracy. By contrast, a change to the potency does not impact the ordinal representation of items. Therefore, a rat that replays episodic memory will perform at least as high as its baseline accuracy. If increasing the potency of an item enhances representations in episodic memory, then accuracy in the high potency test may exceed baseline accuracy.

Methods

Subjects. Twelve male Sprague-Dawley rats were obtained from Envigo (Indianapolis, IN; 73 days old and weighed 264 g at the beginning of the experiment). Rats were housed individually and maintained on a 12:12 light/dark cycle, with light onset at 7:30 a.m. and offset at 7:30 p.m. Water was available ad libitum, except during behavioral sessions. Rats received 45-mg chocolate pellets and chow pellets (F0229 and F0164, respectively; Bio-Serv, Frenchtown, NJ) in behavioral sessions. Daily rations were 15 g including pellets consumed during sessions and 5102-Rat-Diet (PMI Nutritional International, St. Louis, MO). All procedures followed the Guide for Care and Use of Laboratory Animals and were approved by the Bloomington Institutional Animal Care and Use Committee at Indiana University. Nearby construction projects produced substantial noise and vibrations, which substantially disrupted behavioral testing. A subset of 5 rats received special monitoring to determine if signs of distress were observed during list presentation. If signs of distress were observed during list presentation, then the testing session was immediately stopped. Signs of distress during list presentation included

any of the following: freezing/parking in corners of arena, backing away from scented lids, biting, or excessive porphyrin production. All rats in this subset acquired the task and were advanced to high potency testing. A subset of 7 remaining rats were also disrupted by construction, but they did not receive the special monitoring described above and did not advance to high potency testing.

Arenas. Three open-field arenas constructed from acrylic plexiglass were used as the distinctive contexts for encoding and memory assessments. All arenas consisted of "food holes" that were used for the placement of scented lids. Each food hole was circular (5 cm diameter, 2.5 cm depth), which allowed a cup to be firmly snapped into place so that the cup lay flush with the floor and could be covered with a lid placed loosely on top. The arena used for list encoding was square (61-cm length, 61-cm width, and 30-cm height), black in color, and had 12 equidistant food holes, arranged along the perimeter of the walls. Two open-field arenas that differed in size, shape, and color were used for the second to last and fourth to last memory assessments. The assignment of these two arenas to second to last and fourth to last contexts was counterbalanced across rats. One arena was white, circular, with a 94-cm diameter floor, and contained a 30-cm high wall. The arena consisted of 18 food holes arranged in two concentric circles (6 food holes were located in the inner ring, and 12 food holes were positioned in the outer ring). The second arena was circular, with a 46-cm diameter floor, and a transparent 30-cm high wall. The pattern of the floor consisted of concentric circles, with the inner, middle, and outer circles colored black, white, and black, respectively. The arena had 8 equidistant food holes positioned along the walls. Arenas were cleaned with 2% chlorhexidine solution after each animal completed its daily session.

False-Bottom Cups. To ensure that the rats could not use the odor of the chocolate reward to select the correct lid, false-bottom cups were used in all memory assessments. Each memory assessment contained two false-bottom cups: one for the correct item and another for the incorrect item. The metal grates (44.5 mm diameter, 179 holes, < 1 mm thick) were placed in the plastic cups at an angle above the pellet(s). The correct choice false-bottom cup contained two chocolate pellets, separated by the thin metal grate (i.e., one chocolate pellet was placed underneath the metal grate, and the other pellet was placed above the grate). The angle at which the metal grates were placed in the plastic cup resulted in the two chocolate pellets being approximately at the same level (i.e., height). The incorrect choice cup contained two chocolate pellets placed underneath the metal grate, with no pellets placed above. The metal grates were fabricated to securely fit into the plastic cups, ensuring the rat could not remove the grate. At the end of each session, false-bottom cups were cleaned with 2% chlorhexidine solution. New false-bottom cups were prepared each day.

Stimuli. Odors were presented using opaque lids that were odorized by storing them in sealed plastic containers. Plastic containers for *low potency odors* were filled with 90 ml of an oil odorant or approximately 150 ml of dry spice powder odorant, and lids were odorized for at least two weeks before being presented to the rats, which was the procedure used in Panoz-Brown et al (2018). The odors used for the high potency probes were all oil odorants. Plastic containers for *high potency odors* were filled with 900 ml of an oil odorant, and lids were odorized for at least two weeks before being presented to the rats. A pilot study with the same odors used here verified that these two levels of potency are discriminable to rats. A metal grate with small holes was used in each container to separate the lids and the odorants in order to prevent direct contact. Odorants were refreshed every month in order to maintain scent potency and consistency.

Odorants included: (i) 74 low potency odors: allspice, amaretto oil, anise, apple, apricot, asparagus, banana, bay, black walnut, blackberry, blueberry oil, brandy oil, bubble gum, butterscotch oil, caraway seed, carob powder, celery seed, champagne, cheddar, cherry oil, chicory root, cilantro, cinnamon, clove, coconut, coffee oil, coriander, cotton candy, cumin, dill weed, fenugreek seed, garlic powder, hazelnut, hickory smoke, honey oil, horseradish, Indian curry, Irish cream oil, lavender, lemon zest, maple, marjoram, menthol-eucalyptus, Mexican oregano, mustard seed, nutmeg, onion powder, orange oil, peach oil, pecan oil, pineapple oil, pistachio, pumpkin, raspberry, root beer oil, rosemary leaf, sage leaf, sassafras, sesame oil, spearmint, spinach powder, strawberry oil, sumac, summer savory, sweet basil, tarragon, thyme, tomato, turmeric, vanilla, Mexican, wasabi, watermelon oil, white willow bark, Worcestershire. (ii) 19 high potency odors: apricot, black walnut, blackberry, bubble gum, caramel, champagne, coconut, cotton candy, English toffee, grape, hot chili, marshmallow, peach, piña colada, pistachio, praline, raspberry, saltwater taffy, tropical punch. All stimuli used as odors were purchased from The Great American Spice Company (Rockford, MI).

General Methods

One session was conducted each day, approximately 5 days per week. During each session, the rat was removed from its home cage and placed in a holding cage. Holding cages were the same as cages used in vivarium housing, except bedding, food, and water were not present. The behavioral testing room contained three open-field arenas and a platform for the holding cage. All arenas and platforms were elevated approximately 78-cm above the floor. Each platform used to hold the arenas as well as the holding cage was set on top of two anti-vibration pads (total height is 2 cm). These were utilized to minimize disruption from nearby construction. During sessions, the default position for the experimenter was located in the middle of the room,

approximately 1-m from the center of the arenas. For each trial, the experimenter removed the rat from the holding cage and placed the rat in the designated arena positioned with its head pointed away from the experimenter. Next, the experimenter returned to the default position where they remained with hands at their sides until the rat displaced the lid of the designated odor. Then, the experimenter removed the rat from the arena and returned it to the holding cage. When odors were presented as stimuli more than once during a session (i.e., presented in a list and also in the memory assessments), new lids were used to prevent the rat from relying on scent marking. The following variables were randomized each trial: the number and identity of the odors used, the location of the odors, and the order of second-last and fourth-last memory assessments.

Pre-training. Before odors were introduced, rats were trained to search the arenas and displace unscented lids from food holes to obtain food rewards, as described in Panoz-Brown et al (2018).

Phase 1: Initial training. Since learning the rule in our approach may be difficult for the rats, we used a training paradigm to optimize learning, as described in Panoz-Brown et al (2018). Accordingly, we implemented training strategies that provide immediate feedback, allowed the rat to continue each trial until they displaced the lid to the correct item, and used a large reward for an initial correct choice. Because timely feedback enhances learning, we utilized a strategy that provided rats with immediate feedback to facilitate the acquisition of the second and fourth to last rules. To ensure that the rats could not track the odor of the chocolate reward, false-bottom cups were utilized during memory assessments. Each memory assessment contained two false-bottom cups: one for the correct item and another for the incorrect item. The incorrect item false-bottom cups contained two chocolate pellets, both placed below a metal grate. The false-

bottom cup for the correct item contained two chocolate pellets, separated by a metal grate (i.e., one chocolate was underneath the metal grate, and the other pellet was above the grate). Having one pellet above the grate provided immediate access to the reward after lid displacement, thereby minimizing the delay between the response and the delivery of the food reward. To further optimize learning, we provided the rat with feedback in every trial by allowing the rat to continue each trial if the initial response was incorrect. To this end, if the rat initially displaced the incorrect lid, the trial continued until it displaced the lid of the correct item (the second choice was not included in calculations of accuracy). Because our approach baited the cups of correct items and allowed the rat to continue in each trial following an initial incorrect choice, it is possible for the rat to choose lids randomly and receive many food rewards. Therefore, to incentivize learning while implementing the features described above, we provided the rat with a large reward following a correct initial choice. Specifically, a correct initial response was rewarded with five additional chocolate food pellets, whereas a correct second response was not rewarded with any additional pellets. Thus, immediately following the rat's initial correct lid displacement, the experimenter delivered the additional food pellets to the cup; the experimenter did not initiate delivery of the additional food reward until after the rat's initial correct response occurred.

During Phase 1, sessions consisted of approximately 7 trial-unique lists and corresponding memory assessments. In this phase, each list consisted of 5-8 trial-unique odors. Odors were presented to the rat in the encoding context, one at a time; the number of items in the list was randomly selected for each trial. Each presentation of a list item consisted of a single cup and odorized lid placed at a randomly determined location and baited with a single chow pellet. The rat was removed from the holding cage and placed in the encoding context (i.e., arena)

facing away from the experimenter. A response was defined as a vertical or horizontal displacement of the lid from the cup. The rat remained in the arena until it displaced the lid and consumed the food reward. Immediately following the lid displacement, the rat was removed from the arena and returned to the holding cage. This procedure continued until all items in the list were presented.

Immediately following each list presentation, the rat was moved to a distinctive memory assessment context. Memory assessments consisted of a choice between two odors previously presented in the list. In one context, one odor was from the second last ordinal position within the list and was rewarded (i.e., correct choice), while the other odor was from a different ordinal position in the list and was not rewarded (i.e., incorrect choice). In one context, a correct choice was defined as the first lid displacement for the second last item, whereas an incorrect response was defined as the first lid displacement for the odor from the different ordinal position. The same memory-assessment procedure was followed in the other memory assessment context, except that the rat was rewarded for selecting the fourth last item in the list. Prior to the start of the experiment, assignment of the second to last and fourth to last memory assessment contexts were counterbalanced across rats. After the presentation of a list, the rats received a memory assessment in the second-last context and in the fourth-last context; the order of these memory assessments was randomly selected for each list. Approximately 10 sessions were conducted in Phase 1.

Phase 2: Baseline training. The procedure used in Phase 2 was the same as in Phase 1, except that the list ranged from 5-12 odors. Testing continued for each rat until performance in the memory assessments met the following criteria: Mean accuracy observed in the second to last

and fourth to last memory assessment was at least 75% for at least six consecutive sessions. Approximately 153 sessions were conducted in Phase 2.

Phase 3: High-potency testing. In this phase, we investigated whether rats relied on non-episodic memory trace strength cues or episodic memory replay when selecting items during the memory assessment. To this end, we directly manipulated memory trace strength by increasing the potency of an item in the list. This generated an atypical memory trace strength for the target item while preserving inter-item interference. Consequently, in both the list presentation and memory assessments, some trials contained a high potency odor in the second last or fourth last position whereas, some trials included a high potency odor in a different randomly determined ordinal position in the list. Thus, in some memory assessments, the high potency odor was the correct choice, while in other memory assessments the high potency odor was the incorrect choice. Consequently, the high potency odors did not provide a cue for selecting the correct item during the memory assessments.

Each session consisted of blocks of up to 5 trials. In one trial, the second last item presented in the list was a high potency odor and was the correct choice in the second last memory assessment. In another trial, the fourth last item presented was a high potency odor and was the correct choice in the fourth last memory assessment. In two separate trials, the high potency odor was in a different randomly determined position in the list and was the incorrect choice in the memory assessment. There was also a single training trial, which had no high potency items in the list or memory assessment. For each session, the order of the trials was randomized, and incomplete blocks continued into the next day. For example, if a rat only completed 3 out of the 5 trials described above, then on the next testing session the rat would

begin by completing the 2 trials missed on the earlier day. Testing continued for each rat until 16 blocks of trials were completed for each trial type.

Results

Episodic memory replay and memory trace strength hypotheses make different predictions when a high-potency item appears in the memory assessment as the correct choice (S+) after the presentation in a list. The mean proportion correct on such trials appears in Figure 4 (labeled High potency). Baseline training data (no high-potency items) also appear in Figure 4. Data in Figure 4 are averaged across second-last and fourth-last trials. Notably, the presence of a high-potency item did not significantly impact the accuracy in selecting the correct item relative to the accuracy observed in baseline training ($t(4)=0.21$, $p=0.84$). The memory trace strength hypothesis predicts (contrary to the data in Figure 4) that accuracy would be lower in high-potency trials relative to baseline training. By contrast, the episodic memory replay hypothesis predicts that high-potency and baseline performance would be equal (i.e., the null hypothesis). Because the absence of evidence does not provide evidence for the absence of an effect, we used a Bayesian statical test to "prove the null hypothesis" (Gallistel, 2009; Rouder et al., 2009). The JZS Bayes factor is >3; that is, episodic memory replay (i.e., the null hypothesis) is more than three times more probable than matching memory trace strength (i.e., the alternative hypothesis) given the data. A Bayes factor of this size is referred to as *substantial* evidence that the null hypothesis is correct (Gallistel, 2009).

Table 1 provides training data in which a high-potency item did not appear in the list or in the memory assessment. These data are included for second-last and fourth-last memory assessments for both baseline training and concurrent training. Baseline training occurred before high-potency odors were introduced. Concurrent training trials were intermixed with high-

potency S+ and high-potency S- trials. Our main objective for intermixing training trials with high-potency tests was to minimize potential disruption from the novel high-potency stimuli. Overall, performance on concurrent training trials was similarly high as in baseline training ($t(4)=-.02$, $p=.51$).

Table 2 provides data from high-potency memory assessments in which the high-potency item was the correct choice (S+) and the incorrect choice (S-) separately for second-last and fourth-last memory assessments. Our main objective for including S- high-potency probe (S-HP) trials was to ensure that the presence of a high-potency item in S+ trials was not a cue to select the high-potency item in the memory assessment. Overall, accuracy was similarly high in S+ and S- trials ($t(4)=.03$, $p=.55$).

Discussion

The episodic memory hypothesis proposes that rats represent multiple items and the order in which they occur in episodic memory and can engage in episodic memory replay, a process by which the rat searches its representations in episodic memory to find specific information. Here, we tested the hypothesis that rats rely on memory trace strength that declines as a function of inter-item interference. To this end, we directly manipulated memory trace strength by increasing the intensity of target odors in both the list presentation and memory assessment. In this approach, we created memory traces that have atypical strengths for the second last or fourth last positions in the list. Therefore, the memory trace strength for the second last or fourth last items does not match their typical levels. Because the high potency test does not provide an odor with a memory trace strength that matches the level that is typical for the memory assessment, a rat that relies on matching memory trace strength would perform at a level below its baseline training accuracy during high potency testing. However, a change in the potency of the odors does not

impact the ordinal representation of items; thus, a rat that relied on episodic memory replay would perform at a level at least as high as its baseline training accuracy. We found that the presence of a high potency item produced accuracy in memory assessments that was equal to the accuracy obtained during baseline training (Figure 4). A Bayesian statistical analysis provided *substantial* evidence that the null hypothesis is correct. Overall, the results of this study provide compelling evidence of episodic memory replay in rats and rule out the hypothesis that rats relied on memory trace strength that declines as a function of inter-item interference.

The rationale for our manipulation is that increasing odor potency increases memory strength. Previous work shows that differences in stimulus intensity impact memory accuracy (Berliner & Durlach, 1973; Killeen & Grondin, 2022). It has been proposed that memory traces of highly intense stimuli take longer to decay than weak stimuli, which suggests that intense stimuli enhance activation of memory (Berliner & Durlach, 1973; Killeen & Grondin, 2022). Although the literature supports our rationale, an additional test of this approach could assess accuracy for weak and more potent to-be-remembered odors.

The use of false-bottom cups in the memory assessments eliminated the possibility that food odors contributed to accuracy. Equivalent performance during high potency testing and baseline training (low-potency odors) was not due to the rats being unable to discriminate high potency from low potency odors. Rats' ability to distinguish between high- and low-potency odors was examined in a pilot study prior to the start of this experiment. The pilot study, with the same odors used here, asked rats to distinguish between a low-potency odor versus the same odor with a high potency. For example, some rats were trained to discriminate between high-potency grape oil and low-potency grape oil by being rewarded for choosing high-potency grape oil while choosing low-potency grape oil was never rewarded. Similarly, other rats were rewarded for

choosing the low-potency oil while the high-potency oil was never rewarded. The average accuracy at the end of training in the pilot study was 0.97 ± 0.02.

Overall, the main threat to the episodic memory replay hypothesis is a non-episodic memory solution (Clayton et al., 2000; Crystal, 2021, 2024, in press; Roberts et al., 2008), which we refer to as memory trace strength (i.e., familiarity). If the major influencer of memory trace strength is the passage of time (i.e., memory trace strength declines as a function of elapsed time), then the second experiment conducted by Panoz-Brown et al (2018), in which the amount of time between the list items and the memory assessment was doubled, ruled out the use of familiarity. However, it is more likely that inter-item interference between subsequent items may produce the different levels of memory trace strength for each ordinal position in the list (see Figure 3A). Thus, we developed a test in which inter-item interference would be similar to the trained profile, but we directly manipulated memory trace strength by increasing the intensity of a target item. Now, having ruled out rats' use of memory trace strength that declines as a function of time and memory trace strength that declines as a function of inter-item interference, we can make a strong claim regarding the original episodic memory hypothesis. The episodic memory hypothesis proposes that the rats represent all items in the list and the order in which they occur in a representational space within episodic memory and that rats are capable of replaying episodic memory to search that representational space to find the second last or fourth last item presented. Therefore, the results obtained here combined with the results found in Panoz-Brown et al (2018) provide compelling evidence that rats are capable of episodic memory replay.

In the current experiment, we directly manipulated memory trace strength and found that accuracy did not decline when an atypical strength item was presented in a list and subsequent

memory assessment. Because the length of the list varied randomly from trial to trial, the rat could not predict the end of the list until it was placed in the memory assessment context. It is possible that training causes more storage strength of items toward the end of the list. However, the rat does not know the list length. Therefore, any enhancement toward the end of the list would have to be due to retrospective "rehearsal" at list end prior to test; according to this proposal, the rat searches backwards from the end of the list to find a target item. To accomplish this enhancement, the rat would have to do something like episodic replay. It is unlikely that list memory performance is based on short term memory or working memory because we have shown that high accuracy is maintained after at least a 60-minute delay and after potential interference memory from several other odors (Panoz-Brown et al., 2018). Converging evidence for the above conclusion comes from studies in which we inhibited neurons in the hippocampus which selectively and reversibly reduced performance on the list memory task, while sparing other aspects of memory (new-old recognition memory and a +/- discrimination rule) (Panoz-Brown et al., 2018).

Recently, we have documented that rats are able to replay incidentally encoded episodic memories during an unexpected assessment of memory (Sheridan et al., 2024). In one task, rats were explicitly trained to report the third from last item in a trial-unique list of odors. All of the rats successfully learned the replay task. In a second task, rats foraged in the absence of odors in an eight-arm radial maze. In a critical test, scented lids covered food while the rats foraged in the radial maze on a single occasion. Finally, memory of the third last odor was assessed, also on a single occasion. The rats correctly answered the unexpected question after incidental encoding odors while foraging. A control condition ruled out the use of stimulus generalization. The data suggest that rats are able to encode multiple pieces of putatively unimportant information and

subsequently replay a stream of episodic memories when confronted with an unexpected problem (Sheridan et al., 2024). In a further test, we showed that rats can also answer an unexpected question after incidental encoding of entirely novel odors (Sheridan et al., in press).

Animal models have documented that place cells in the hippocampus of rodents undergo the sequential reactivation in neuronal coding for where events occurred (sharp wave ripple complexes consisting of transient, brief, high-frequency network oscillations), a phenomenon known as *hippocampal replay* (Carr et al., 2011; Davidson et al., 2009; Diba & Buzsáki, 2007; Foster & Wilson, 2006; MacDonald et al., 2011; O'Keefe & Dostrovsky, 1971; Pfeiffer & Foster, 2013). Although hippocampal replay may be implicated in memory (Fernández-Ruiz et al., 2019; Gillespie et al., 2021; Jadhav et al., 2012; Liu et al., 2023; Zielinski et al., 2020), neurophysiological evidence that it involves searching memory to guide future behavior is lacking (Gillespie et al., 2021; Gupta et al., 2010; Redish, 2020). Other work suggests that time cells in the hippocampus and prefrontal cortex fire sequentially (MacDonald et al., 2013; MacDonald et al., 2011; Pastalkova et al., 2008; Tiganj et al., 2016), which may be used to track the sequences of events (Howard et al., 2014). Therefore, we developed an approach that gave rats opportunities to report, via their behavior, episodic memories about a stream of events in sequential order.

Our data provide evidence that rats are able to remember multiple trial-unique events using episodic memory and can engage in episodic memory replay which is a process by which the rat searches its representations in episodic memory, to find items at specific points within the sequence. Our findings provide supporting evidence for the view that rats can be used to model fundamental features of human memory and that the ability to replay multiple episodic memories is old in the evolutionary timescale.

Acknowledgements This article is in honor of the contributions of Nicky Clayton to the study of comparative cognition. Support by National Science Foundation grant NSF/BCS-1946039 to J.D.C. and NSF/DBI-2050311. C.L.S. was supported by the Harlan Scholars program. We thank the following individuals for assistance: Cami Albers, Sarah Flannery, Stephen Lang, Ryleigh Looper, Risha Mehta, Kennedy Quintana, Priyanka Thomas, and Clara Hall.

Declarations

Funding Support by National Science Foundation grant NSF/BCS-1946039 to J.D.C.

and NSF/DBI-2050311.

Competing interests The authors declare no competing interests.

Ethics approval Approval by the IACUC at Indiana University (Protocols 18-024 and 21-016).

Availability of data Data for each rat are provided in supplemental tables (Tables S1, S2, and S3).

Code availability This paper does not report original code.

Open practices statement The data are available in the supplemental tables, and the study was not preregistered.

References

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language*, *49*(4), 415-445. https://doi.org/https://doi.org/10.1016/j.jml.2003.08.006

Babb, S. J., & Crystal, J. D. (2005). Discrimination of what, when, and where: Implications for episodic-like memory in rats. *Learning & Motivation*, *36*, 177-189. https://doi.org/https://doi.org/10.1016/j.lmot.2005.02.009

Babb, S. J., & Crystal, J. D. (2006a). Discrimination of what, when, and where is not based on time of day. *Learning & Behavior*, *34*, 124-130. https://doi.org/10.3758/bf03193188

Babb, S. J., & Crystal, J. D. (2006b). Episodic-like memory in the rat. *Current Biology*, *16*, 1317-1321. https://doi.org/https://doi.org/10.1016/j.cub.2006.05.025

Berliner, J. E., & Durlach, N. I. (1973). Intensity perception. IV. Resolution in roving-level discrimination. *The Journal of the Acoustical Society of America*, *53*(5), 1270-1287. https://doi.org/10.1121/1.1913465

Billard, P., Schnell, A. K., Clayton, N. S., & Jozet-Alves, C. (2020). Cuttlefish show flexible and future-dependent foraging cognition. *Biology Letters*, 16(2), 20190743. https://doi.org/doi:10.1098/rsbl.2019.0743. https://doi.org/http://doi.org/10.1098/rsbl.2019.0743

Carlesimo, G., Serra, L., Fadda, L., Cherubini, A., Bozzali, M., & Caltagirone, C. (2007). Bilateral damage to the mammillo-thalamic tract impairs recollection but not familiarity in the recognition process: a single case investigation. *Neuropsychologia*, *45*(11), 2467-2479.

Carr, M. F., Jadhav, S. P., & Frank, L. M. (2011). Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. *Nature neuroscience*, *14*(2), 147-153. https://doi.org/10.1038/nn.2732

Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*, *395*(6699), 272-274.

Clayton, N. S., Griffiths, D. P., & Dickinson, A. (2000). Declarative and episodic-like memory in animals: Personal musings of a Scrub Jay. In C. Heyes & L. Huber (Eds.), *Evolution of cognition* (pp. 273-288). The MIT Press.

Crystal, J. D. (2013). Remembering the past and planning for the future in rats. *Behavioural Processes*, *93*, 39-49. https://doi.org/10.1016/j.beproc.2012.11.014

Crystal, J. D. (2016a). Animal models of source memory. *Journal of the Experimental Analysis of Behavior*, *105*(1), 56-67. https://doi.org/https://doi.org/10.1002/jeab.173

Crystal, J. D. (2016b). Comparative Cognition: Action Imitation Using Episodic Memory. *Current Biology*, *26*(23), R1226-R1228. https://doi.org/https://doi.org/10.1016/j.cub.2016.10.010

Crystal, J. D. (2021). Evaluating evidence from animal models of episodic memory. *Journal of Experimental Psychology: Animal Learning and Cognition*, *47*(3), 337. https://doi.org/10.1037/xan0000294

Crystal, J. D. (2024). Temporal foundations of episodic memory. *Learning & Behavior*, *52*, 35-50. https://doi.org/10.3758/s13420-023-00608-x

Crystal, J. D. (in press). Mental time travel in the rat. *Philisophical Transactions of the Royal Society B*. https://doi.org/10.1098/rstb.2023.0404

Crystal, J. D., & Alford, W. T. (2014). Validation of a rodent model of source memory. *Biology Letters*, *10*(3), 20140064. https://doi.org/10.1098/rsbl.2014.0064

Crystal, J. D., Alford, W. T., Zhou, W., & Hohmann, A. G. (2013). Source memory in the rat. *Current Biology*, *23*(5), 387-391. https://doi.org/http://dx.doi.org/10.1016/j.cub.2013.01.023

Crystal, J. D., & Smith, A. E. (2014). Binding of episodic memories in the rat. *Current Biology*, *24*(24), 2957-2961. https://doi.org/10.1016/j.cub.2014.10.074

Dalecki, S. J., Panoz-Brown, D. E., & Crystal, J. D. (2017). A test of the reward-contrast hypothesis. Behavioural Processes, 145, 15-17. https://doi.org/https://doi.org/10.1016/j.beproc.2017.09.018

Davidson, T. J., Kloosterman, F., & Wilson, M. A. (2009). Hippocampal replay of extended experience. *Neuron*, *63*(4), 497-507.

Davies, J. R., Garcia-Pelegrin, E., Baciadonna, L., Pilenga, C., Favaro, L., & Clayton, N. S. (2022). Episodic-like memory in common bottlenose dolphins. *Current Biology*, *32*(15), 3436-3442.e3432. https://doi.org/https://doi.org/10.1016/j.cub.2022.06.032

Dede, A. J. O., Frascino, J. C., Wixted, J. T., & Squire, L. R. (2016). Learning and remembering real-world events after medial temporal lobe damage. *Proceedings of the National Academy of Sciences*, *113*(47), 13480-13485. https://doi.org/10.1073/pnas.1617025113

Diba, K., & Buzsáki, G. (2007). Forward and reverse hippocampal place-cell sequences during ripples. *Nature neuroscience*, *10*(10), 1241-1242.

Eichenbaum, H. (2000). A cortical–hippocampal system for declarative memory. *Nature reviews neuroscience*, *1*(1), 41-50.

Eichenbaum, H. (2007). Comparative cognition, hippocampal function, and recollection. *Comparative Cognition & Behavior Reviews*, *2*.

Eichenbaum, H., Sauvage, M., Fortin, N., Komorowski, R., & Lipton, P. (2012). Towards a functional organization of episodic memory in the medial temporal lobe. *Neuroscience & Biobehavioral Reviews*, *36*(7), 1597-1608.

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The Medial Temporal Lobe and Recognition Memory. *Annual Review of Neuroscience*, *30*(1), 123-152. https://doi.org/10.1146/annurev.neuro.30.051606.094328

Eldridge, L. L., Knowlton, B. J., Furmanski, C. S., Bookheimer, S. Y., & Engel, S. A. (2000). Remembering episodes: a selective role for the hippocampus during retrieval. *Nature neuroscience*, *3*(11), 1149-1152.

Fernández-Ruiz, A., Oliva, A., Fermino de Oliveira, E., Rocha-Almeida, F., Tingley, D., & Buzsáki, G. (2019). Long-duration hippocampal sharp wave ripples improve memory. *Science*, *364*(6445), 1082-1086. https://doi.org/10.1126/science.aax0758

Foster, D. J., & Wilson, M. A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*, *440*(7084), 680-683.

Gallistel, C. R. (2009). The importance of proving the null [Article]. *Psychological Review*, *116*(2), 439-453. https://doi.org/10.1037/a0015251

Gillespie, A. K., Maya, D. A. A., Denovellis, E. L., Liu, D. F., Kastner, D. B., Coulter, M. E., Roumis, D. K., Eden, U. T., & Frank, L. M. (2021). Hippocampal replay reflects specific past experiences rather than a plan for subsequent choice. *Neuron*, *109*(19), 3149-3163. e3146.

Gupta, A. S., Van Der Meer, M. A., Touretzky, D. S., & Redish, A. D. (2010). Hippocampal replay is not a simple function of experience. *Neuron*, *65*(5), 695-705.

Henson, R. N., Rugg, M., Shallice, T., Josephs, O., & Dolan, R. J. (1999). Recollection and familiarity in recognition memory: an event-related functional magnetic resonance imaging study. *Journal of neuroscience*, *19*(10), 3962-3972.

Hofer, A., Siedentopf, C. M., Ischebeck, A., Rettenbacher, M. A., Verius, M., Golaszewski, S. M., Felber, S., & Fleischhacker, W. W. (2007). Neural substrates for episodic encoding and recognition of unfamiliar faces. *Brain and cognition*, *63*(2), 174-181.

Hoffman, M. L., Beran, M. J., & Washburn, D. A. (2009). Memory for "what", "where", and "when" information in rhesus monkeys (Macaca mulatta). *Journal of Experimental Psychology: Animal Behavior Processes*, *35*(2), 143-152. https://doi.org/https://doi.org/10.1037/a0013295

Howard, M. W., MacDonald, C. J., Tiganj, Z., Shankar, K. H., Du, Q., Hasselmo, M. E., & Eichenbaum, H. (2014). A unified mathematical framework for coding time, space, and sequences in the hippocampal region. *Journal of neuroscience*, *34*(13), 4692-4707.

Jadhav, S. P., Kemere, C., German, P. W., & Frank, L. M. (2012). Awake Hippocampal Sharp-Wave Ripples Support Spatial Memory. *Science*, *336*(6087), 1454-1458. https://doi.org/10.1126/science.1217230

Janowsky, J. S., Shimamura, A. P., & Squire, L. R. (1989). Source memory impairment in patients with frontal lobe lesions. *Neuropsychologia*, *27*(8), 1043-1056. https://doi.org/http://dx.doi.org/10.1016/0028-3932(89)90184-X

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*(1), 3-28. https://doi.org/10.1037/0033-2909.114.1.3

Jozet-Alves, C., Bertin, M., & Clayton, N. S. (2013). Evidence of episodic-like memory in cuttlefish. *Current Biology*, *23*(23), R1033-R1035. https://doi.org/http://dx.doi.org/10.1016/j.cub.2013.10.021

Killeen, P. R., & Grondin, S. (2022). A trace theory of time perception. *Psychological Review*, *129*(4), 603-639. https://doi.org/10.1037/rev0000308

Kurth-Nelson, Z., Economides, M., Dolan, Raymond J., & Dayan, P. (2016). Fast Sequences of Non-spatial State Representations in Humans. *Neuron*, *91*(1), 194-204. https://doi.org/https://doi.org/10.1016/j.neuron.2016.05.028

Liu, C., Todorova, R., Tang, W., Oliva, A., & Fernandez-Ruiz, A. (2023). Associative and predictive hippocampal codes support memory-guided behaviors. *Science*, *382*(6668), eadi8237. https://doi.org/doi:10.1126/science.adi8237

MacDonald, C. J., Carrow, S., Place, R., & Eichenbaum, H. (2013). Distinct hippocampal time cell sequences represent odor memories in immobilized rats. *Journal of neuroscience*, *33*(36), 14607-14616.

MacDonald, C. J., Lepage, K. Q., Eden, U. T., & Eichenbaum, H. (2011). Hippocampal “time cells” bridge the gap in memory for discontiguous events. *Neuron*, *71*(4), 737-749.

Mitchell, K. J., & Johnson, M. K. (2009). Source monitoring 15 years later: What have we learned from fMRI about the neural mechanisms of source memory? *Psychological Bulletin*, *135*(4), 638-677. https://doi.org/10.1037/a0015849

Nelson, K. R., & Wasserman, E. A. (1978). Temporal factors influencing the pigeon's successive matching-to-sample performance: sample duration, intertrial interval, and retention interval. *Journal of the Experimental Analysis of Behavior*, *30*(2), 153-162. https://doi.org/https://doi.org/10.1901/jeab.1978.30-153

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map: preliminary evidence from unit activity in the freely-moving rat. *Brain research*.

Pahl, M., Zhu, H., Pix, W., Tautz, J., & Zhang, S. (2007). Circadian timed episodic-like memory - a bee knows what to do when, and also where. *Journal Of Experimental Biology*, *210*(Pt 20), 3559-3567.

Panoz-Brown, D., Iyer, V., Carey, L. M., Sluka, C. M., Rajic, G., Kestenman, J., Gentry, M., Brotheridge, S., Somekh, I., Corbin, H. E., Tucker, K. G., Almeida, B., Hex, S. B., Garcia, K. D., Hohmann, A. G., & Crystal, J. D. (2018). Replay of Episodic Memories in the Rat. *Current Biology*, *28*(10), 1628-1634.e1627. https://doi.org/https://doi.org/10.1016/j.cub.2018.04.006

Panoz-Brown, D. E., Corbin, H. E., Dalecki, S. J., Gentry, M., Brotheridge, S., Sluka, C. M., Wu, J.-E., & Crystal, J. D. (2016). Rats remember items in context using episodic memory. *Current Biology*, *26*(20), 2821-2826. https://doi.org/http://dx.doi.org/10.1016/j.cub.2016.08.023

Pastalkova, E., Itskov, V., Amarasingham, A., & Buzsaki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. *Science*, *321*(5894), 1322-1327.

Pfeiffer, B. E., & Foster, D. J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*, *497*(7447), 74-79. https://doi.org/https://doi.org/10.1038/nature12112

Redish, A. D. (2020). Beyond replay: Introduction to the special issue on hippocampal replay. *Hippocampus*, *30*(1), 3-5. https://doi.org/10.1002/hipo.23184

Roberts, W. A. (1972). Short-term memory in the pigeon: Effects of repetition and spacing. *Journal of Experimental Psychology*, *94*(1), 74-83. https://doi.org/10.1037/h0032796

Roberts, W. A., Feeney, M. C., MacPherson, K., Petter, M., McMillan, N., & Musolino, E. (2008). Episodic-Like Memory in Rats: Is It Based on When or How Long Ago? *Science*, *320*(5872), 113-115. https://doi.org/doi:10.1126/science.1152709

Roberts, W. A., & Grant, D. S. (1978). An analysis of light-induced retroactive inhibition in pigeon short-term memory. *Journal of Experimental Psychology: Animal Behavior Processes*, *4*(3), 219-236. https://doi.org/10.1037/0097-7403.4.3.219

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225-237. https://doi.org/10.3758/PBR.16.2.225

Schmitter-Edgecombe, M., & Anderson, J. W. (2007). Feeling of knowing in episodic memory following moderate to severe closed-head injury. *Neuropsychology*, *21*(2), 224.

Sheridan, C. L., Bonner, L., & Crystal, J. D. (in press). Replay of incidentally encoded novel odors in the rat. *Animal Cognition*. https://doi.org/https://doi.org/10.1007/s10071-024-01880-8

Sheridan, C. L., Lang, S., Knappenberger, M., Albers, C., Loper, R., Tillett, B., Sanchez, J., Wilcox, A., Harrison, T., Panoz-Brown, D., & Crystal, J. D. (2024). Replay of incidentally encoded episodic memories in the rat. *Current Biology*, *34*(3), 641-647.e645. https://doi.org/https://doi.org/10.1016/j.cub.2023.12.043

Smith, A. E., Dalecki, S. J., & Crystal, J. D. (2017). A test of the reward-value hypothesis. *Animal Cognition*, *20*(2), 215-220. https://doi.org/10.1007/s10071-016-1040-z

Smith, A. E., Slivicki, R. A., Hohmann, A. G., & Crystal, J. D. (2017). The chemotherapeutic agent paclitaxel selectively impairs learning while sparing source memory and spatial memory. *Behavioural Brain Research*, *320*, 48-57. https://doi.org/http://dx.doi.org/10.1016/j.bbr.2016.11.042

Smith, A. E., Xu, Z., Lai, Y. Y., Kulkarni, P. M., Thakur, G. A., Hohmann, A. G., & Crystal, J. D. (2016). Source memory in rats is impaired by an NMDA receptor antagonist but not by PSD95-nNOS protein–protein interaction inhibitors. *Behavioural Brain Research*, *305*, 23-29. https://doi.org/http://dx.doi.org/10.1016/j.bbr.2016.02.021

Staresina, B. P., Alink, A., Kriegeskorte, N., & Henson, R. N. (2013). Awake reactivation predicts memory in humans. *Proceedings of the National Academy of Sciences*, *110*(52), 21159-21164. https://doi.org/10.1073/pnas.1311989110

Tiganj, Z., Jung, M. W., Kim, J., & Howard, M. W. (2016). Sequential Firing Codes for Time in Rodent Medial Prefrontal Cortex. *Cerebral Cortex*, *27*(12), 5663-5671. https://doi.org/10.1093/cercor/bhw336 %J Cerebral Cortex

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of Memory* (pp. 381-403). Academic Press.

Tulving, E. (1983). *Elements of episodic memory*. Oxford University Press.

Tulving, E. (2002). Episodic memory: from mind to brain. *Annu Rev Psychol*, *53*, 1-25. https://doi.org/10.1146/annurev.psych.53.100901.135114

Yonelinas, A. P. (2001). Components of episodic memory: the contribution of recollection and familiarity. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *356*(1413), 1363-1374.

Yonelinas, A. P., & Levy, B. J. (2002). Dissociating familiarity from recollection in human recognition memory: Different rates of forgetting over short retention intervals. *Psychonomic Bulletin & Review*, *9*(3), 575-582. https://doi.org/10.3758/BF03196315

Zhou, W., & Crystal, J. D. (2009). Evidence for remembering when events occurred in a rodent model of episodic memory. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(23), 9525-9529. https://doi.org/10.1073/pnas.0904360106

Zhou, W., & Crystal, J. D. (2011). Validation of a rodent model of episodic memory. *Animal Cognition*, *14*(3), 325-340. https://doi.org/10.1007/s10071-010-0367-0

Zhou, W., Hohmann, A. G., & Crystal, J. D. (2012). Rats answer an unexpected question after incidental encoding. *Current Biology*, *22*(12), 1149-1153. https://doi.org/10.1016/j.cub.2012.04.040

Zielinski, M. C., Tang, W., & Jadhav, S. P. (2020). The role of replay and theta sequences in mediating hippocampal-prefrontal interactions for memory and cognition. *Hippocampus*, *30*(1), 60-72.

Baseline Training	
2nd Last	**4th Last**
0.907 ± 0.035	0.788 ± 0.011
Concurrent Training	
2nd Last	**4th Last**
0.867 ± 0.052	0.785 ± 0.056

Table 1. Training data.

The table displays the mean proportion correct ± 1 SEM for the 2nd last and 4th last memory assessments during training (all items were low potency in the encoding context and memory assessment). Baseline training data were collected during the last six sessions, before high-potency testing. Concurrent training data were intermixed with high-potency trials. Data for each rat are provided in Supplemental Information Table S2.

S+ HP	
2nd Last	**4th Last**
0.843 ± 0.031	0.864 ± 0.041
S- HP	
2nd Last	**4th Last**
0.853 ± 0.037	0.797 ± 0.050

Table 2. High-potency data.

This table displays the mean proportion correct ± 1 SEM for the 2nd last and 4th last memory assessments. All items in the list were low potency, except for a single item that was high potency. The high-potency item from the list sometimes appeared in the memory assessment as the correct choice (S+), and on other occasions, it appeared as the incorrect choice (S-). Data for each rat are provided in Supplemental Information Table S3.

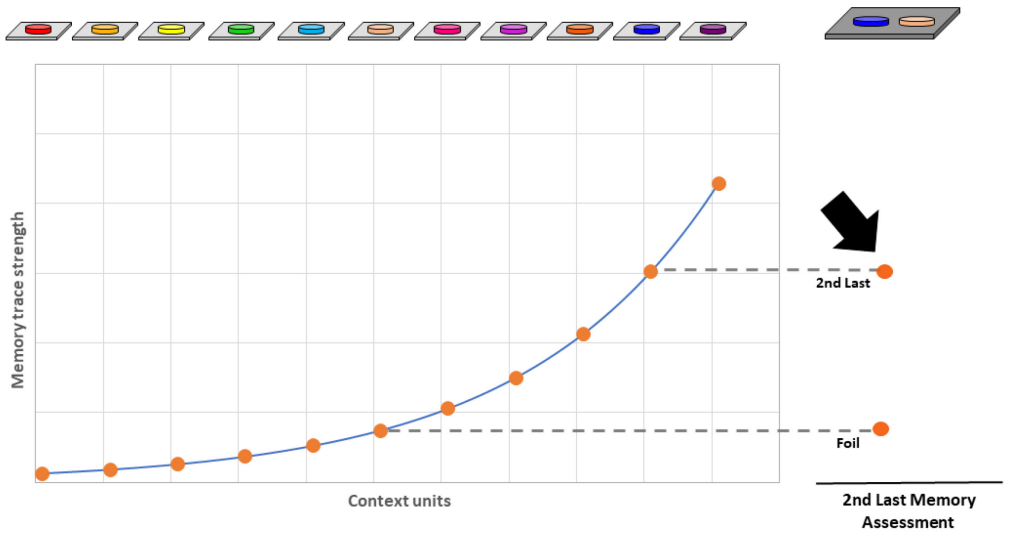

Figure 1. Panoz-Brown et al (2018) Experiment 1. Rats were presented with a list of 5-12 odors (depicted at the top of the schematic, different colors represent different odors) in a distinct encoding context (depicted by a small rectangle). After the list presentation, rats were moved to a memory assessment context (depicted by the large rectangle) where they were given a choice between two odors from the previously presented list. One odor was the second to last odor (shown in blue) and the other odor came from a different position within the list (shown in light pink). If the rats selected the second to last odor, they were rewarded. At the moment of the memory assessment (right side of schematic), each item has a typical memory trace strength (shown by the orange dots). If the presentation of an item gives rise to a memory trace that degrades as a function of time, the orange dots take the form of a backward exponential curve. Note that a second to last item has a typical memory trace strength which is indicated by the large black arrow.

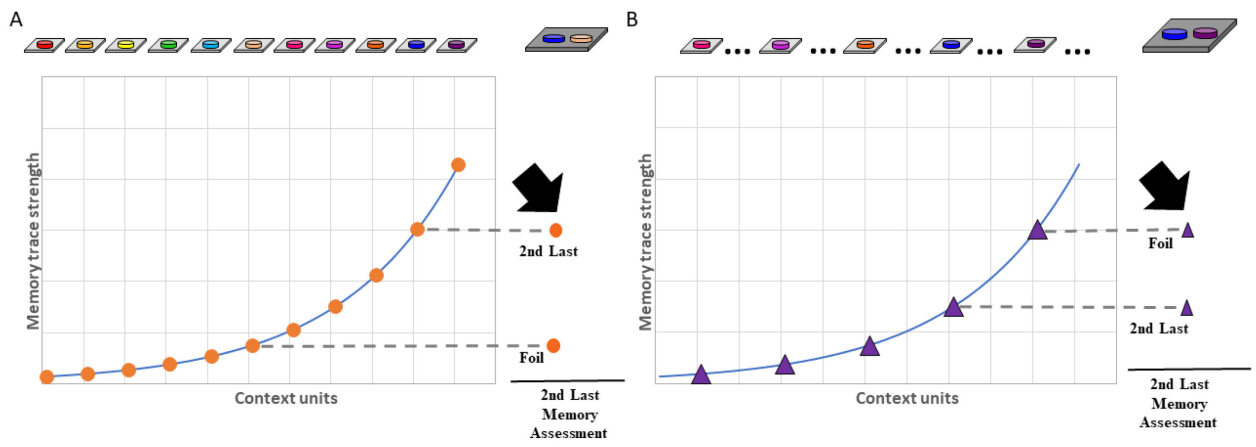

Figure 2. Panoz-Brown et al (2018) Experiment 1 vs. Experiment 2. A: Left panel is the same as Figure 1 (included for comparison) and depicts Experiment 1 from Panoz-Brown et al. **B:** This panel depicts the manipulation used in Experiment 2 in Panoz-Brown et al. The rat is presented with a list of odors in a distinct encoding context, but the time between the presentation of items and the memory assessment was doubled. Thus, at the moment of the memory assessment, the memory trace strength for the second last item does not match the large black arrow which corresponds to the typical memory trace strength of a second last item. Instead, the incorrect choice matches the large black arrow creating a dissociation between memory trace strength and the order of items. An animal that matches memory trace strength would be inclined to choose the incorrect item, producing accuracy below chance, whereas an animal that replays its episodic memories would choose the correct item, producing accuracy above chance.

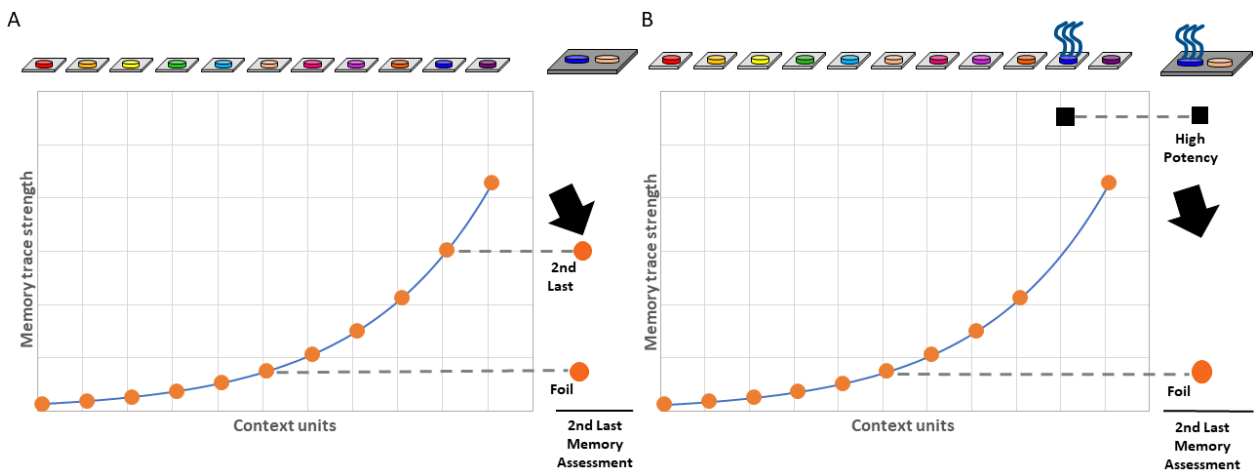

Figure 3. Training vs. high potency conditions. A. The left panel is the same as Figure 1 (included for comparison) depicting a list training trial. **B.** The right panel depicts the high potency manipulation. The rat is presented with a list of odors in a distinct encoding context (top of the schematic), but the second last odor has a higher potency (depicted as blue smoke) than normal. This created a memory trace that has an atypical strength for a second last odor (shown by the black square). Thus, at the moment of the memory assessment (right side of schematic), the memory trace strength for the second last item does not match the large black arrow which corresponds to the typical memory trace strength of a second last item.

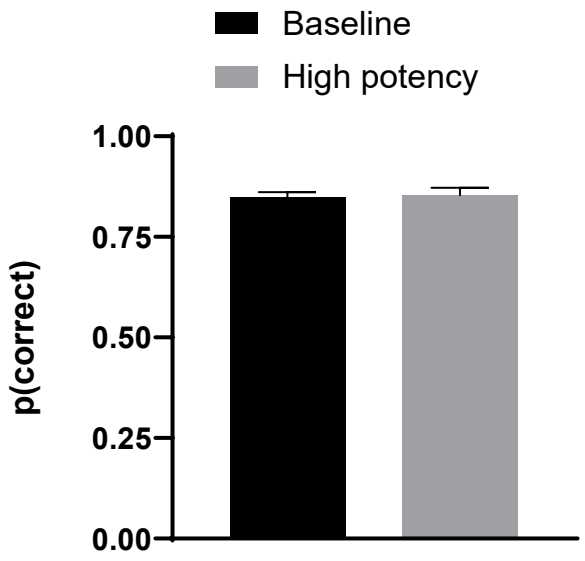

Figure 4. Baseline Training vs. High-Potency Testing. This figure displays the mean proportion of correct responses averaged across the second- and fourth-last memory assessments during baseline training and high-potency testing. The baseline training data are from the last six sessions before high-potency testing. The high-potency data are from trials in which the high-potency item was the correct choice in the memory assessment. Error bars represent 1 SEM. Data for each rat are provided in Supplemental Information Table S1.

Figure Captions

Figure 1. Panoz-Brown et al (2018) Experiment 1. Rats were presented with a list of 5-12 odors (depicted at the top of the schematic, different colors represent different odors) in a distinct encoding context (depicted by a small rectangle). After the list presentation, rats were moved to a memory assessment context (depicted by the large rectangle) where they were given a choice between two odors from the previously presented list. One odor was the second to last odor (shown in blue) and the other odor came from a different position within the list (shown in light pink). If the rats selected the second to last odor, they were rewarded. At the moment of the memory assessment (right side of schematic), each item has a typical memory trace strength (shown by the orange dots). If the presentation of an item gives rise to a memory trace that degrades as a function of time, the orange dots take the form of a backward exponential curve. Note that a second to last item has a typical memory trace strength which is indicated by the large black arrow.

Figure 2. Panoz-Brown et al (2018) Experiment 1 vs. Experiment 2. A: Left panel is the same as Figure 1 (included for comparison) and depicts Experiment 1 from Panoz-Brown et al. **B:** This panel depicts the manipulation used in Experiment 2 in Panoz-Brown et al. The rat is presented with a list of odors in a distinct encoding context, but the time between the presentation of items and the memory assessment was doubled. Thus, at the moment of the memory assessment, the memory trace strength for the second last item does not match the large black arrow which corresponds to the typical memory trace strength of a second last item. Instead, the incorrect

choice matches the large black arrow creating a dissociation between memory trace strength and the order of items. An animal that matches memory trace strength would be inclined to choose the incorrect item, producing accuracy below chance, whereas an animal that replays its episodic memories would choose the correct item, producing accuracy above chance.

Figure 3. Training vs. high potency conditions. A. The left panel is the same as Figure 1 (included for comparison) depicting a list training trial. **B.** The right panel depicts the high potency manipulation. The rat is presented with a list of odors in a distinct encoding context (top of the schematic), but the second last odor has a higher potency (depicted as blue smoke) than normal. This created a memory trace that has an atypical strength for a second last odor (shown by the black square). Thus, at the moment of the memory assessment (right side of schematic), the memory trace strength for the second last item does not match the large black arrow which corresponds to the typical memory trace strength of a second last item.

Figure 4. Baseline Training vs. High-Potency Testing. This figure displays the mean proportion of correct responses averaged across the second- and fourth-last memory assessments during baseline training and high-potency testing. The baseline training data are from the last six sessions before high-potency testing. The high-potency data are from trials in which the high-potency item was the correct choice in the memory assessment. Error bars represent 1 SEM. Data for each rat are provided in Supplemental Information Table S1.

Supplemental Information

Table S1. Performance of individual rats in baseline and high potency conditions, corresponding to data presented in Figure 4.

Subject	Baseline	High potency
NU05	0.833	0.833
NU06	0.885	0.806
NU07	0.851	0.833
NU11	0.865	0.906
NU12	0.804	0.889
Mean ± SEM	**0.848 ± 0.014**	**0.853 ± 0.019**

Table S2. Performance of individual rats in baseline and concurrent training conditions, corresponding to data presented in Table 1.

Subject	Baseline training		Concurrent training	
	2nd last	4th last	2nd last	4th last
NU05	0.889	0.778	0.778	0.944
NU06	1	0.769	0.944	0.611
NU07	0.938	0.765	0.889	0.833
NU11	0.923	0.808	1	0.813
NU12	0.786	0.821	0.722	0.722
Mean ± SEM	**0.907 ± 0.035**	**0.788 ± 0.011**	**0.867 ± 0.052**	**0.785 ± 0.056**

Table S3. Performance of individual rats in S+ and S- high potency conditions, corresponding to data presented in Table 2.

Subject	S+ high potency		S- high potency	
	2nd last	4th last	2nd last	4th last
NU05	0.889	0.778	0.944	0.889
NU06	0.833	0.778	0.889	0.833
NU07	0.778	0.889	0.833	0.611
NU11	0.938	0.875	0.875	0.875
NU12	0.778	1	0.722	0.778
Mean ± SEM	**0.843 ± 0.031**	**0.864 ± 0.041**	**0.853 ± 0.037**	**0.797 ± 0.050**