Effects of pharmacological manipulations of NMDA-receptors on deliberation in the Multiple-T task

Anna Blumenthal\textsuperscript{a,1}, Adam Steiner\textsuperscript{b}, Kelsey Seeland\textsuperscript{c}, A. David Redish\textsuperscript{d,*}

\textsuperscript{a}Drew University, United States
\textsuperscript{b}Graduate Program in Neuroscience, University of Minnesota, United States
\textsuperscript{c}Department of Neuroscience, University of Minnesota, United States
\textsuperscript{d}Department of Neuroscience, 6-145 Jackson Hall, 321 Church St. SE, Minneapolis, MN 55455, United States

\textbf{A R T I C L E  I N F O}

Article info

Received 8 October 2010
Revised 4 January 2011
Accepted 25 January 2011
Available online 4 February 2011

\textbf{A B S T R A C T}

Both humans and non-human animals have the ability to navigate and make decisions within complex environments. This ability is largely dependent upon learning and memory processes, many of which are known to depend on NMDA-sensitive receptors. When humans come to difficult decisions they often pause to deliberate over their choices. Similarly, rats pause at difficult choice points. This behavior, known as \textit{vicarious trial and error} (VTE), is hippocampally dependent and entails neurophysiological representations of expectations of future outcomes in hippocampus and downstream structures. In order to determine the dependence of VTE behaviors on NMDA-sensitive receptors, we tested rats on a Multiple-T choice task with a reward-delivery reversal known to elicit VTE. Rats under the influence of NMDA-receptor antagonists (CPP) showed a significant reduction in VTE, particularly at the reward reversal, implying a role for NMDA-sensitive receptors in the generation of vicarious trial and error behaviors.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

Under certain conditions, particularly during learning and after changes in reward contingencies, rats pause at difficult decision-points and serially turn back and forth towards the available options (Muenzinger, 1938; Tolman, 1938, 1939, 1948). This behavior was called \textit{vicarious trial and error} (VTE) and was hypothesized to entail the serial consideration of possible paths, that is, deliberation. This behavior was originally described by Tolman (1938) as a conflict-driven behavior at a choice point, and can be seen during odor or visual discrimination tasks, as well as on the radial-arm maze and on T-maze tasks (Brown, 1992; Hu & Amsel, 1995; Johnson & Redish, 2007; Tolman, 1938; van der Meer, Johnson, Schmitzer-Torbert, & Redish, 2010). VTE generally occurs early in learning and decreases with time and experience (Tolman, 1938; van der Meer et al., 2010). Increased levels of early VTE behavior have been correlated with better performance and more efficient learning (Muenzinger, 1938; Tolman, 1939).

VTE is abolished with hippocampal lesions (Hu & Amsel, 1995) and is related to activity levels in hippocampus as measured by cytochrome-oxidase staining (Hu, Xu, & Gonzalez-Lima, 2006). Recently, Johnson and Redish (2007) discovered that decoded hippocampal representations transiently swept forward down possible choices during VTE events, and van der Meer and Redish (2009) discovered that ventral striatal reward representations transiently reactivated during VTE events, confirming Tolman’s hypotheses that VTE reflects a serial representation of possibilities (in hippocampus) and the development of an expectation of reward contingencies (in ventral striatum). The Johnson and Redish (2007) and van der Meer and Redish (2009) results were found in rats running on a Multiple-T task, in which animals ran through a sequence of low-cost choices until they reached a high-cost choice where they turned left or right to receive food reward (Schmitzer-Torbert & Redish, 2002; van der Meer et al., 2010). VTE behaviors were observed to primarily occur at the final (high-cost) choice point, and could be quantitatively measured by comparing the time spent at the final choice point relative to one of the earlier (control) points (van der Meer et al., 2010). Interestingly, animals showed VTE behavior during early laps on the task, and the behavior decreased or vanished altogether with repeated laps of unchanging sequence within the day. During later laps within the session, animals ran straight through the choice point without pausing (Schmitzer-Torbert & Redish, 2002; van der Meer et al., 2010).
NMDA-receptors are involved in many aspects of learning and memory, particularly the induction of long-term potentiation (LTP) in many brain structures (Butcher, Davis, & Morris, 1990; Morris, Anderson, Lynch, & Baudry, 1986; Morris, 2003). Pharmacological blockage of these receptors impairs behavioral performance on a number of tasks, particularly spatial, hippocampally-dependent tasks involving novel and flexible representations of the environment (Butcher et al., 1990; Morris, 2003; Nakazawa et al., 2003; Ohno, Yamamoto, & Watanabe, 1992). On the other hand, genetically mutated mice with an over expression of NMDA-receptors showed superior abilities in a wide variety of learning and memory tasks, including the Morris Water Maze (Tang et al., 1999).

Studies have shown that 3-(2-carboxypiperazin-4-yl)propyl-1-phosphonic acid (CPP, an NMDA-receptor antagonist, Davies et al., 1986) specifically affects hippocampally-dependent behavioral abilities (Ohno et al., 1992), as well as stability and learning-related changes in hippocampal place cells (Austin, Fortin, & Shapiro, 1990; Ekstrom, Meltzer, McNaughton, & Barnes, 2001; Kentros et al., 1998). In contrast, other studies have also shown that d-cycloserine (DCS, a partial NMDA-receptor agonist, Hood et al., 1986) specifically affects hippocampally-dependent behavioral learning seen in rats under the influence of NMDA-receptor antagonists and the enhancement of behavioral learning seen in rats under the influence of NMDA-receptor agonists led us to hypothesize that pharmacological manipulation of NMDA-receptors may affect VTE behavior and overall performance on the Multiple-T task. Specifically, under the hypothesis that NMDA-receptors facilitate hippocampally-dependent learning and thus likely facilitate deliberative decision-making, we predicted that NMDA-receptor antagonists (CPP) would impair both flexible decision-making and decrease VTE behaviors, and that NMDA-receptor agonists (DCS) would increase VTE behaviors.

2. Methods

2.1. Subjects

A within-subjects design was used. Male Fisher–Brown-Norway hybrid rats (n = 6, aged 4–6 months at the start of behavioral training) were housed in single cages and maintained on a 12 h light-dark cycle with lights-on at 8:00 am. Throughout the entire experiment, animals were maintained at more than 80% of their free-feeding body weight. All procedures were in accordance with the National Institutes of Health guidelines for animal care and were approved by the Institutional Animal Care and Use Committee at the University of Minnesota.

2.2. Behavioral training

Animals were handled approximately 10–15 min each day for 2 weeks prior to behavioral training. On the sixth day of handling, standard food was removed from cages and replaced with reward pellets in order for rats to familiarize and adjust to the pellets (5TUL Research Diets, New Brunswick, NJ) used in behavioral training and testing. Each pellet weighed 45 mg, and animals were given 15 g worth of pellets placed in a small bowl in their home cage. Pellets were white, yellow (“banana” flavored), and pink (“fruit” flavored). A different flavor was offered to each animal each day. During the final 3 days of handling, animals were handled with a “backpack” (a Velcro strap with an LED light) secured around their body to get them used to wearing the backpack. The LED-light on the backpack allowed tracking of the animal on the behavioral task. Animals had complete freedom of motion even while wearing the backpack.

Animals went through 18 days of behavioral training on an elevated multiple T-task prior to testing. The Multiple-T maze was similar to that described by Schmitzer-Torbert and Redish (2002, 2004) aside from changes in specific measurements. See Fig. 1. The maze consisted of three movable T’s made out of plywood boards covered in carpet, the stem of each was 40.5 cm long, and the two choice arms were both 28 cm long. The three T’s connected two 166.5 cm long rails which were connected at either end to 177.5 cm long return rails. Each return rail was equipped with two automated feeders (Med-Associates, St. Albans, VT) spaced 45 cm apart from each other. Each feeder released two 45 mg pellets onto the track, so that a rat received four pellets during a correct lap. Each rat ran one 40 min session each day. Sessions began with the rat being placed at the base of the first T stem (Maze-Start, MS, see Fig. 1). Rats ran the maze as a continuous loop and were not removed from the maze during the 40 min session. All four T components were interchangeable, as were the components making up the top and bottom rails, and the components making up the side return rails. Which component formed which T or return
rail was randomized each day to control for odor cues across sessions.

The first three choice points (T1, T2, T3) were considered “low-cost” choice points because rats were allowed to turn around the dead-end. However, the final choice point T4 was considered a “high-cost” choice point CP because rats had to complete the unrewarded lap before getting another pass through the navigation sequence. Reward direction was programmed at the start of the session to be either always on the left side of the maze or always on the right side of the maze for the first 6 days of training, and either left, right, or alternating during the final 12 days of training. During the alternation contingency, rats were rewarded for alternately making left and right choices at T4. The alternation contingency was programmed so that the rat was rewarded for making the opposite choice from that made on the previous lap, so a rat running consistently leftward on an alternation contingency would only be rewarded on the first lap and would not be rewarded until he made a rightward choice at T4. Reward contingency varied randomly across training sessions, but was kept constant within each training session.

Including the alternation condition ensures that there are three reward contingencies available. This means that when animals are faced with a change in reward contingency (see below), they cannot simply switch to the other behavior. Instead, a change in reward contingency requires a re-determination of the new contingency. Because the alternation condition did not include a delay, it is unlikely that the alternation condition itself is hippocampally dependent (Ainge, van der Meer, Langston, & Wood, 2007; Dember & Richman, 1989), which makes it comparable in memory requirements to that of simply running left or simply running right.

For the first 6 days animals were trained on the maze with wooden blocks placed on the rail at the starting point and final high-cost choice point (T4, CP) to block the animal from going to the non-rewarded side for that day, so that the animal received a reward with every completed lap. Wooden blocks were also placed at T2 on the non-rewarded side to serve as a control point (CT). On the seventh day, no blocks were placed and animals could run freely left or right from the choice point with no explicit external cues. This forced them to learn reward direction anew each day through trial and error. Alternation conditions were included in the pseudo-random cycle during training after the blocks were removed.

### 2.3. Drug administration

DCS (D-cycloserine, Sigma-Aldrich, St. Louis, Missouri) or CPP (±3-(2-carboxyxyperazin-4-yl)propyl-1-phosphonic acid, Tocris Cookson Inc., Ellicsille, Missouri) was dissolved in sterile 0.9% saline. Sterile 0.9% saline vehicle was used as a control. All drugs were prepared and then frozen at −20°C in small (approximately single doses) centrifuge tubes. Shortly before injection, the centrifuge tube was taken out of the freezer, warmed in a gloved hand until completely thawed and at room temperature (approximately 5 min), and then vortexed for 30 s. The proper dose of the drug condition for the session (DCS, (10 mg/kg); CPP, (5 mg/kg); saline, 0.2 mL) was drawn into a sterile syringe and injected IP. The person running the animal was not present at the time of injection and was blind to the injection condition. After injection, animals were returned to their home cage for 1 h before being run on the task. Drug doses and post-injection waiting times for CPP and DCS were chosen based on the results of other behavioral studies: (Kentros et al., 1998, CPP), and (Ledgerwood et al., 2003, DCS), respectively.

### 2.4. Behavioral testing

After completing the 18-day training sequence, rats received one saline dose session to control for the novelty of receiving an IP injection and to get them used to running after an IP injection. This first saline dose session was not analyzed. The behavioral testing regimen consisted of a 6 day testing paradigm of three drug days separated by “wash” days in which no injection was given. “Wash” days served as “no-injection controls”, interspersed behaviorally between drug days to control for across-day learning effects. Each animal received a different drug (saline, CPP, or DCS) on one of three drug days, resulting in all animals receiving each drug treatment once. The experimenter running the animal on the task was blind to drug condition. Multiple-T sessions were identical to those described in behavioral training with one significant difference: the reward contingency was switched approximately halfway through each session (20 ± 2 min) (Gupta, van der Meer, Tourzetzky, & Redish, 2010). The original and the switched contingencies were different each day for each animal, and were counterbalanced across drug condition, so that each animal got a different reward contingency each drug day and each wash day. See Supplemental Table S1.

### 2.5. Analysis

During all training and testing sessions, animals wore an in-house-made backpack with an LED for tracking. Each 40 min trial was video recorded in order to observe animals’ general behavior during the task with a specific focus on low-cost control point (T2, CT) vs. high-cost choice point (T4, CP) behavior. The experimenter noted the number of times the animal had to be blocked from turning around during each session. The position of the animal was recorded at 60 Hz from a camera in the ceiling using the video tracking available in Cheetah (Neuralyx, Bozeman, Montana). Food-delivery signals were also recorded using the digital-input event signals available in Cheetah. Two sessions (one wash session and one saline session) were not included in the analysis because of bad tracking.

Because all mazes were spatially aligned such that the final choice point (T4, CP) and the control point (T2, CT) remained in a constant position in space, the choice point (CP) and control point (CT) regions-of-interest were defined spatially. See Fig. 2. Entry and exit times into and out of CP and CT were calculated from the tracking data. Errors were defined as leaving the CP in the wrong direction. Vicarious trial and error (VTE) behavior was quantitatively measured as the time spent at the choice point divided by the time spent at the control point (van der Meer et al., 2010). Measuring the ratio between the choice point and time spent at the control point controls for overall changes in slowing and controls for random increases in pausing on the task. That this ratio is being driven by changes in pause-time at the choice point rather than at the control point can be seen in Supplemental Figures S1 and S2. Although head position was not quantitatively available from this data (because position was measured from a backpack), in other studies where head position is available, these increased pausing measures tend to be correlated to actual VTE behaviors (Johnson & Redish, 2007, see also Johnson, van der Meer, & Redish, 2007, & van der Meer et al., 2010).

In order to control for potential motoric effects of CPP, speed was measured using the Janabi-Sharifi, Hayward, and Chen (2000) algorithm (Masimore, Schnitzler-Torbert, Kalkiros, & Redish, 2005). In particular, average speed 0.25 s after the first feeder fired (F1L or F1R) was compared across drug conditions.

In order to directly address the question of whether there were actual differences in vicarious trial and error (VTE) behaviors, we categorized each pass through the choice point as either containing a VTE event or not. The \( (x,y) \) coordinates through each pass through the choice point (CP, Fig. 1) was extracted from the tracking data. 2373 passes were found. Two independent observers were presented with the set of 2373 unlabeled passes in a random order.
This made the raters blind to session, rat, and drug-delivery condition. Because the passes were presented in random order, no within-session or sequential effects would be expected to occur. The independent raters then categorized each pass as either VTE, not-VTE, or uninterpretable (for example because tracking was poor). (See Fig. 3 for examples of passes categorized as VTE and not-VTE.) The two raters had >95% agreement between assignments, implying that these ratings were reliable. The two raters then met to discuss the passes on which they disagreed categorization or which they had declared “uninterpretable”. The final categorization included only 17 uninterpretable passes out of 2373. It is important to note that so far in this new analysis, both raters were completely blind to which passes came from which sessions, which rats, or which conditions. The 17/2373 uninterpretable passes are not considered further in the analyses. The identity markers were then re-attached to the data and the proportion of passes on each lap in each condition that were categorized as VTE or not-VTE were compared. Comparisons were done using ANOVAs as before.

3. Results

All rats ran the task. As can be seen in Fig. 4, rats ran a similar number of laps under all four conditions (ANOVA, significant effect of rat \( p = 0.00 \), but no significant effect of drug condition \( p = 0.18 \)). Similarly, there was no significant difference in the number of laps run before or after the reward-contingency switch.
To control for potential motoric effects of the injection itself and for potential motoric effects of CPP or DCS, we also measured the speed 0.25 s after the triggering of the feeder (which makes a small “click” sound). This is the point where animals are running the fastest, presumably because they know that food is waiting at the feeder site for them. As can be seen in Fig. 5, no significant effect of drug was seen (ANOVA, \( p = 0.92 \)).

Although rats ran a similar number of laps, and reached similar maximum speeds, not all rats made more correct choices than chance. See Fig. 6. Overall, the average percentage of correct laps were significantly above chance (0.5) for the wash, DCS, and saline conditions, both overall, and before and after the switch in reward contingency. However, while the rats running under the influence of CPP were significantly above chance before the contingency switch, they were not above chance after the reward contingency switch, indicating that the rats running under the influence of CPP either did not recognize the switch or were unable to change their behavior to accommodate it.

As in our other Multiple-T experiments (Schmitzer-Torbert & Redish, 2002; van der Meer et al., 2010), VTE behaviors were predominantly seen at the start of the session as animals learned the reward contingency for the day. In addition, VTE also appeared after the switch in reward contingency. As can be seen in Fig. 7, rats in the wash, DCS, and saline conditions all showed significant VTE effects (measured as the time spent at the choice point, CP, divided by the time spent at the control point, CT). However, rats under the influence of CPP showed a lower VTE pausing ratio in the early laps before the contingency switch and almost no VTE pausing after the contingency switch.\(^2\)

To control for potential motoric effects of the injection itself and for potential motoric effects of CPP or DCS, we also measured the speed 0.25 s after the triggering of the feeder (which makes a small “click” sound). This is the point where animals are running the fastest, presumably because they know that food is waiting at the feeder site for them. As can be seen in Fig. 5, no significant effect of drug was seen (ANOVA, \( p = 0.92 \)).

Fig. 4. Number of laps run on the Multiple-T maze. A lap was counted as a pass through the choice point, separated by a full return along one of the return rails (rewarded or not). No significant effect of drug was seen.

Fig. 5. Speed 0.25 s after triggering of the first feeder. Speed was measured using the Janabi-Sharifi et al. (2000) algorithm (see Masimore et al., 2005). No significant effect of drug was seen.

In order to test for the significance of the differences between these conditions, we performed an ANOVA controlling for three factors: rat, lap-number, and drug condition. Because not all rats ran more than 30 laps before the switch and again after the switch, we restricted the analyses to only the first 30 laps before or after the switch. Both before and after the reward-contingency switch, there were overall significant effects of drug condition (Before: ANOVA [df = 5,29,3], effect of rat: \( F = 8.05, p < 0.0001 \), effect of lap: \( F = 2.94, p < 0.0001 \), effect of drug condition: \( F = 4.93, p = 0.0021 \); After: ANOVA [df = 5,29,3], effect of rat: \( F = 5.76, p < 0.0001 \), effect of lap: \( F = 1.05, p = 0.39 \), effect of drug condition: \( F = 5.43, p = 0.001 \)). Rats under the influence of CPP showed significantly reduced levels of pausing at the choice point from rats in the three other conditions both before and after the switch. (Before: wash

\(^2\) See also Supplemental Figures S1–S4, showing the time spent at the choice point (CP), before (S1) and after the reward contingency switch (S2), as well as the time spent at the control point (CT), before (S3) and after the reward contingency switch (S4). These supplemental figures show that most of the VTE effect is occurring because of passing at the choice point (CP). We report VTE as the ratio of the two times because this controls for average speed of the animal, which did differ between animals (Fig. 5).
Fig. 6. Success rates on the Multiple-T task. p-values at the bottom of each bar indicate significance test relative to chance (0.5). Success rates significantly different from chance are indicated in red. Note that after the reward contingency switch, rats receiving CPP were not significantly different from chance. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The results indicate that NMDA-receptor manipulation affected animals’ behavior and performance on the Multiple-T maze. In particular, animals under the influence of CPP showed little to no vicarious trial and error (VTE) behaviors early on or post-switch and made the most errors both pre and post switch. Rats under the influence of CPP did not seem to recognize the switch in reward contingency, being unable to improve beyond chance after the reward contingency switch and showing no VTE after it.

The results suggest that NMDA-receptors may be crucial for both early learning within a day and relearning in response to change in reward-delivery contingency. The parallel change in VTE under the influence of NMDA-receptor modulators suggests that early learning may be related to the VTE process and that this flexible early learning characterized by VTE may also be necessary for successful later performance, since animals that did not show VTE early on or post-switch had the most difficulty learning the reward pattern and making correct decisions. Previous research has shown that NMDA-receptor antagonists, such as CPP, cause impairment in early acquisition learning. For example, Steele and Morris (1999) found that rats under NMDA-receptor blockade (infused intrahippocampally with AP5) were unable to find the hidden platform on the water maze despite recent trials in the same location, suggesting the NMDA-receptor antagonist disrupted the usually
rapid spatiotemporal storage of memory. Evidence shows that this NMDA-receptor-dependency may be concentrated on early learning: animals' with extensive overtraining are still able to perform the water maze despite NMDA-receptor blockage (Cain, 1998).

Hippocampal disruption often leads to perseverative behaviors (Day & Schallert, 1996; McDonald & White, 1994, see Redish, 1999, for review). Our data suggest that this perseverative behaviors may be partially due to an inability to recognize that the situation has changed – VTE behaviors that occurred in response to situation changes (e.g. in our wash condition) also vanished in the CPP condition after the switch in reward contingency. The black line repeated on each condition is the mean VTE for the corresponding wash condition for comparison.

We have interpreted these results in terms of effects on hippocampal function, due to the known effects of NMDA-antagonists, particularly systemic CPP on normal hippocampal function (Austin et al., 1990; Ekstrom et al., 2001; Kentros et al., 1998), and the similar effects of systemic CPP and hippocampal disruption on maze-based behaviors (Morris, 2003; Nakazawa et al., 2003; Ohno et al., 1992). However, NMDA-receptors are ubiquitous throughout the brain and it is possible that other structures may be involved. For example, systemic CPP affects glutamate release in medial prefrontal cortex and direct manipulations of prefrontal cortex can counter the effects of CPP (Carli, Calcagno, Mainini, Arnt, & Invernizzi, 2010; Del Arco, Ronzoni, & Mora, 2010). Systemic DCS has been much more often interpreted in terms of a role in prefrontal cortex and the effect on extinction (Langton & Richardson, 2010; McCallum, Kim, & Richardson, 2010), which can be seen as an ability to recognize a change in reward contingency and react to it (Bouton, 2004; Capaldi & Lynch, 1968; Gershman, Blei, & Niv, 2010; Redish, Jensen, Johnson, & Kurth-Nelson, 2007). The relation between reward-contingency changes, hippocampus, prefrontal cortex, and other structures is complex (Fuhls & Touretzky, 2007;
May 95 (11) 376–384

A. Blumenthal et al. / Neurobiology of Learning and Memory 95 (2011) 376–384

Hirsch, Leber, & Gillman, 1978; Redish et al., 2007; Quirk, Garcia, & González-Lima, 2006).

Our task provides an interesting view toward this question because on the Multiple-T task, the environment is familiar, as are the potential reward contingencies, but animals must still learn which contingency is being rewarded on a particular day. For the probe trials analyzed here, animals had to not only recognize which contingency was being rewarded, but also recognize that a contingency change had taken place and switch their behavior to reflect that new contingency. Our results suggest that animals who were unable to make use of NMDA-receptor dependent flexible learning were able to learn the original contingency as sufficiently as controls, but were impaired in recognizing the contingency change. These findings suggest NMDA-receptors are particularly important for perseveration and recognizing changes in reward contingency.

Our results relate the presence of VTE and related passing behaviors to these recognition effects. Although the increases in passing under the influence of DCS did not reach significance before the contingency switch, there was a significant increase after the contingency switch. However, in our study CPP clearly and significantly reduced VTE in the initial set at the beginning of each day (where the animal has to determine which reward-contingency is in place). In response to the change in reward contingency, VTE behaviors were significantly diminished under the influence of CPP, and the excess passing normally seen at the choice point was nearly absent under the influence of CPP, suggesting a role for NMDA-receptors in the performance of VTE behaviors.

Acknowledgements

This project was run as part of the 2009 NSF-sponsored REU Summer Undergraduate Research program in the Behavioral Sciences and Cognitive Sciences at the University of Minnesota. We thank the other members of the Redish lab and the other participants in the BSC REU program for helpful discussion. Funded by NSF (0648715) and by NIH-R01-MH08318.

Appendix A. Supplementary data


References


May 95 (11) 376–384


