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

A. David Redish,
Department of Neuroscience, 6-145 Jackson Hall, 321 Church St. SE, Minneapolis MN 55455, A. David Redish: redish@umn.edu, phone: 612-626-3738, fax: 612-626-5009

Anna Blumenthal,
address at time of project: Drew University, work done at the University of Minnesota

Adam Steiner, and
Graduate Program in Neuroscience, University of Minnesota

Kelsey Seeland
Department of Neuroscience, University of Minnesota

Anna Blumenthal: ABlumenthal@NKI.RFMH.ORG; Adam Steiner: stein704@umn.edu; Kelsey Seeland: seela005@umn.edu

Abstract

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 *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.

Keywords

NMDA-receptor; vicarious trial and error; decision-making; CPP; DCS; VTE

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 *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 (Tolman, 1938; Brown, 1992; Hu and Amsel, 1995; Johnson and Redish, 2007; 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, Johnson, Schmitzer-Torbert, & Redish, 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 and Amsel, 1995) and is related to activity levels in hippocampus as measured by cytochrome-oxidase staining (Hu et al., 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 and Redish, 2002; van der Meer, Johnson, Schmitzer-Torbert, & Redish, 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, Johnson, Schmitzer-Torbert, & Redish, 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 and Redish, 2002; van der Meer, Johnson, Schmitzer-Torbert, & Redish, 2010).

NMDA-receptors are involved in many aspects of learning and memory, particularly the induction of long-term potentiation (LTP) in many brain structures (Morris et al., 1986; Butcher et al., 1990; 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; Ohno et al., 1992; Morris, 2003; Nakazawa et al., 2003). 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 et al., 1990; Kentros et al., 1998; Ekstrom et al., 2001). In contrast, other studies have also shown that D-cycloserine (DCS, a partial NMDA-receptor agonist, Hood et al., 1989), facilitates learning in both humans and rodents, particularly in extinction and during reversals (Monahan et al., 1989; Ledgerwood et al., 2003; Ressler et al., 2004; Golden and Houpt, 2007; Kalisch et al., 2009). The disruption of behavioral learning and the place field learning-related changes 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 hour 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 minutes 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 15g 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 three 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 Figure 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 45mg pellets onto the track, so that a rat received four pellets during a correct lap. Each rat ran one 40minute session each day. Sessions began with the rat being placed at the base of the first T stem (Maze-Start, MS, see Figure 1). Rats ran the maze as a continuous loop and were not removed from the maze during the 40 minute session. All four T components were interchangable, 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 on encountering 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 (Dember and Richman, 1989; Ainge et al., 2007), which makes it comparable in memory requirements to that of simply running left or simply running right.

For the first six 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-Carboxypiperazin-4-yl)propyl-1-phosphonic acid, Tocris Cookson Inc., Ellsville, 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 minutes), 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 six 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 minutes) (Gupta et al., 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 minute 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 (Neuralynx, 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 Figure 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, Johnson, Schmitzer-Torbert, & Redish, 2010). Measuring the ratio between time spent at 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 et al., 2007, see also Johnson and Redish, 2007, and van der Meer, Johnson, Schmitzer-Torbert, & Redish, 2010).

In order to control for potential motoric effects of CPP, speed was measured using the Janabi-Sharifi et al. (2000) algorithm (Masimore et al., 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 $\langle x, y \rangle$ coordinates through each pass through the choice point (CP, Figure 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 Figure 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 Figure 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 Figure 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 Figure 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 accomodate it.

As in our other Multiple-T experiments (Schmitzer-Torbert and Redish, 2002; van der Meer, Johnson, Schmitzer-Torbert, & Redish, 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 reappeared after the switch in reward contingency. As can be seen in Figure 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. [1]

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 vs. CPP: $F = 9.0$, $p = 0.0028$, CPP vs. DCS: $F = 13$, $p = 0.0004$, CPP vs. saline: $F = 11$, $p = 0.0011$, *After*: wash vs. CPP: $F = 8.4$, $p = 0.0038$, CPP vs. DCS: $F = 8.2$, $p = 0.0044$, CPP vs. saline: $F = 10$, $p = 0.0017$). Rats under the influence of *DCS* showed a trend[2] towards increased levels of pausing relative to the *wash* condition before ($F = 3.0$, $p =$

---

[1]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. These supplemental figures show that most of the VTE effect is occurring because of pausing 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 (Figure 5).
[2]A Bonferronni correction for six tests gives a corrected $p < 0.05$ significance at $p < 0.0083$.

0.082) and a significant effect after the contingency switch ($F = 7.3$, $p = 0.0070$). No significant effects of saline against wash were seen after correcting for multiple tests[2] (*Before*: wash vs. saline: $F = 1.81$, $p = 0.1791$, *After*: wash vs. saline: $F = 4.5$, $p = 0.0342$). These results are summarized in Supplementary Table S2.

To determine whether the excess pausing seen at the choice point is actually due to an increase in vicarious trial and error (VTE) behaviors, we visually categorized all passes through the choice point as either containing a VTE event or not. The people doing the categorization were blind to condition when doing the categorization. (See Methods section for details.) As can be seen in Figure 8, the likelihood of seeing VTE behaviors began high and decreased with lap run at the start of each session. After the switch in reward-delivery contingency, VTE behaviors again reappeared. These effects (both at the start of the session and in response to the change in reward-delivery contingency) were more robust in the wash, DCS, and saline conditions than in the CPP conditions.

As before, 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. As before, we only examined the first 30 laps of each condition. 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: $F = 3$, $p = 0.0108$, effect of lap: $F = 5.89$, $p < 0.0001$, effect of drug condition: $F = 6.64$, $p = 0.0002$; *After*: ANOVA[df=5,29,3], effect of rat: $F = 2.3$, $p = 0.044$, effect of lap: $F = 1.39$, $p = 0.083$, effect of drug condition: $F = 4.1$, $p = 0.0068$). Rats under the influence of *CPP* showed significantly reduced levels of VTE from rats in the three other conditions both before and after the switch. (*Before*: wash vs. CPP: $F = 13$, $p = 0.0002$, CPP vs. DCS: $F = 12.5$, $p = 0.0005$, CPP vs. saline: $F = 19$, $p < 0.0001$, *After*: wash vs. CPP: $F = 8.4$, $p = 0.0038$, CPP vs. DCS: $F = 8.2$, $p = 0.0044$, CPP vs. saline: $F = 10$, $p = 0.0017$). However, although a significant effect was found in rats under the influence of *DCS* relative to the *wash* condition after the contingency change ($F = 7.3$, $p = 0.0070$), no other significant effects were seen. These results are summarized in Supplementary Table S3.

4 Discussion

These 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 in the session or immediately following the switch sequence, 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 (McDonald and White, 1994; Day and Schallert, 1996, see Redish, 1999, for review). Our data suggest that the NMDA-blockade by CPP led to perseveration in the task, particularly in the inability to change behaviors in response to the switch in reward contingency. Our data suggest that these perseveration behaviors may be partially due to an inability to recognize that the situation has changed — VTE behaviors that occured in response to situation changes (e.g. in our *wash* condition) also vanished in the *CPP* condition.

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; Kentros et al., 1998; Ekstrom et al., 2001), and the similar effects of systemic CPP and hippocampal disruption on maze-based behaviors (Ohno et al., 1992; Morris, 2003; Nakazawa et al., 2003). 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 et al., 2010; Del Arco et al., 2010). Systemic DCS has been much more often interpreted in terms of a role in prefrontal cortex and the effect on extinction (McCallum et al., 2010; Langton and Richardson, 2010), which can be seen as an ability to recognize a change in reward-contingency and react to it (Capaldi and Lynch, 1968; Bouton, 2004; Redish et al., 2007; Gershman et al., 2010). The relation between reward-contingency changes, hippocampus, prefrontal cortex, and other structures is complex (Hirsh et al., 1978; Quirk et al., 2006; Fuhs and Touretzky, 2007; Redish et al., 2007).

Our task provides an interesting view towards 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 pausing behaviors to these recognition effects. Although the increases in pausing 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 begining 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 signficantly diminished under the influence of CPP, and the excess pausing 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Abbreviations

VTE	vicarious trial and error
CPP	3-(2-Carboxypiperazin-4-yl)propyl-1-phosphonic acid
DCS	D-cycloserine

Acknowledgments

References

Ainge JA, van der Meer MAA, Langston RF, Wood ER. Exploring the role of context-dependent hippocampal activity in spatial alternation behavior. Hippocampus. 2007; 17(10):988–1002. [PubMed: 17554771]

Austin KB, Fortin WF, Shapiro ML. Place fields are altered by NMDA antagonist MK-801 during spatial learning. Society for Neuroscience Abstracts. 1990; 16:263.

Bouton ME. Context and behavioral processes in extinction. Learn. Mem. 2004; 11(5):485–494. [PubMed: 15466298]

Brown MF. Does a cognitive map guide chioces in the radial-arm maze? Journal of Experimental Psychology. 1992; 18(1):56–66. [PubMed: 1578200]

Butcher SP, Davis S, Morris RG. A dose-related impairment of spatial learning by the NMDA receptor antagonist, 2-amino-5-phosphonovalerate AP5. European Neuropsychopharmacology. 1990; 1(1): 15–20. [PubMed: 1983778]

Cain DP. Testing the NMDA, long-term potentiation, and cholinergic hypothesis of spatial learning. Neuroscience and Biobehavioral Reviews. 1998; 22(2):181–193. [PubMed: 9579310]

Capaldi EJ, Lynch AD. Magnitude of partial reward and resistance to extinction: Effect of N-R transitions. Journal of Comparative and Physiological Psychology. 1968; 65(1):179–181. [PubMed: 5648456]

Carli M, Calcagno E, Mainini E, Arnt J, Invernizzi RW. Sertindole restores attentional performance and suppresses glutamate release induced by the NMDA receptor antagonist CPP. Psychopharmacology. 2010 Epub ahead of print.

Davies J, Evans RH, Herrling PL, A W Jones HJO, Pook P, Watkins JC. CPP, a new potent and selective NMDA antagonist. depression of central neuron responses, affinity for [^{3}H]-AP5 binding sites on brain membranes and anticonvulsant activity. Brain Research. 1986; 382(1):169–173. [PubMed: 2876749]

Day LB, Schallert T. Anticholinergic effects on acquisition of place learning in the Morris water task: Spatial mapping deficit or inability to inhibit non-place strategies. Behavioral Neuroscience. 1996; 110(5):998–1005. [PubMed: 8919002]

Del Arco A, Ronzoni G, Mora F. Prefrontal stimulation of $GABA_A$ receptors counteracts the corticolimbic hyperactivity produced by NMDA antagonists in the prefrontal cortex of the rat. Psychopharmacology. 2010 Epub ahead of print.

Dember, WN.; Richman, CL., editors. Spontaneous Alternation Behavior. New York: Springer; 1989.

Ekstrom AD, Meltzer J, McNaughton BL, Barnes CA. NMDA receptor antagonism blocks experience-dependent expansion of hippocampal "place fields". Neuron. 2001; 31:631–638. [PubMed: 11545721]

Fuhs MC, Touretzky DS. Context learning in the rodent hippocampus. Neural Computation. 2007; 19(12):3172–3215.

Gershman SJ, Blei D, Niv Y. Context, learning and extinction. Psychological Review. 2010; 117(1): 197–209. [PubMed: 20063968]

Golden GJ, Houpt TA. Nmda receptor in conditioned flavor-taste preference learning: blockade by mk-801 and enhancement by d-cycloserine. Pharmacology, Biochemistry, and Behavior. 2007; 86(3):587–596.

Gupta AS, van der Meer MAA, Touretzky DS, Redish AD. Hippocampal replay is not a simple function of experience. Neuron. 2010; 65(5):695–705. [PubMed: 20223204]

Hirsh R, Leber B, Gillman K. Fornix fibers and motivational states as controllers of behavior: A study stimulated by the contextual retrieval theory. Behavioral Biology. 1978; 22:463–478. [PubMed: 697681]

Hood WF, Compton RP, Monahan JB. D-cycloserine: a ligand for the n-methyl-d-aspartate coupled glycine receptor has partial agonist characteristics. Neuroscience Letters. 1989; 98(1):91–95. [PubMed: 2540460]

Hu D, Amsel A. A simple test of the vicarious trial-and-error hypothesis of hippocampal function. PNAS. 1995; 92:5506–5509. [PubMed: 7777539]

Hu D, Xu X, Gonzalez-Lima F. Vicarious trial-and-error behavior and hippocampal cytochrome oxidase activity during Y-maze discrimination learning in the rat. International Journal of Neuroscience. 2006; 116(3):265–280. [PubMed: 16484053]

Janabi-Sharifi F, Hayward V, Chen C-SJ. Discrete-time adaptive windowing for velocity estimation. IEEE Transactions on Control Systems Technology. 2000; 8(6):1003–1009.

Johnson A, Redish AD. Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. Journal of Neuroscience. 2007; 27(45):12176–12189. [PubMed: 17989284]

Johnson A, van der Meer MAA, Redish AD. Integrating hippocampus and striatum in decision-making. Current Opinion in Neurobiology. 2007; 17(6):692–697. [PubMed: 18313289]

Kalisch R, Holt B, Petrovic P, De Martino B, Klöppel S, Büchel C, Dolan RJ. The nmda agonist d-cycloserine facilitates fear memory consolidation in humans. Cerebral Cortex. 2009; 19(1):187–196. [PubMed: 18477687]

Kentros CEH, Hawkins RD, Kandel ER, Shapiro M, Muller RV. Abolition of long-term stability of new hippocampal place cell maps by NMDA receptor blockade. Science. 1998; 280:2121–2126. [PubMed: 9641919]

Langton JM, Richardson R. The effect of d-cycloserine on immediate vs. delayed extinction of learned fear. Learning and Memory. 2010; 17(11):547–551. [PubMed: 20971937]

Ledgerwood L, Richardson R, Cranney J. Effects of d-cycloserine on extinction of conditioned freezing. Behavioral Neuroscience. 2003; 117(2):341–349. [PubMed: 12708530]

Masimore B, Schmitzer-Torbert NC, Kakalios J, Redish AD. Transient striatal [gamma] local field potentials signal movement initiation in rats. NeuroReport. 2005; 16(18):2021–2024. [PubMed: 16317346]

McCallum J, Kim JH, Richardson R. Impaired extinction retention in adolescent rats: effects of D-cycloserine. Neuropsychopharmacology. 2010; 35(10):2134–2142. [PubMed: 20592716]

McDonald R, White N. Parallel information processing by hippocampal and dorsal striatal memory systems in the Morris water maze. Behavioral and Neural Biology. 1994; 61(3):260–270. [PubMed: 8067981]

Monahan JB, Handelmann GE, Hood WF, Cordi AA. D-cycloserine, a positive modulator of the n-methyl-d-aspartate receptor, enhances performance of learning tasks in rats. Pharmacology, Biochemistry, and Behavior. 1989; 34(3):649–653.

Morris RG, Anderson E, Lynch GS, Baudry M. Selective impairment of learning and blockade of long-term potentiation by an N-methyl-D-aspartate receptor antagonist, AP5. Nature. 1986; 319(6056):774–776. [PubMed: 2869411]

Morris RGM. Long-term potentiation and memory. Philosophical Transactions of the Royal Society, London B. 2003; 358(1432):643–647.

Muenzinger KF. Vicarious trial and error at a point of choice. I. a general survey of its relation to learning efficiency. J. Genet Psychol. 1938; 53:75–86.

Nakazawa K, Sun LD, Quirk MC, Rondi-Reig L, Wilson MA, Tonegawa S. Hippocampal CA3 NMDA receptors are crucial for memory acquisition of one-time experience. Neuron. 2003; 38(2):305–315. [PubMed: 12718863]

Ohno M, Yamamoto T, Watanabe S. Effects of intrahippocampal injections of N-methyl-D-asparatate receptor antagonists and scopolomine on working and reference memory assessed in rats on a three-panel runway task. Journal of Pharmacological and Experimental Therapeutics. 1992; 263(3):943–951.

Quirk GJ, Garcia R, González-Lima F. Prefrontal mechanisms in extinction of conditioned fear. Biological Psychiatry. 2006; 60(4):337–343. [PubMed: 16712801]

Redish, AD. Beyond the Cognitive Map: From Place Cells to Episodic Memory. Cambridge MA: MIT Press; 1999.

Redish AD, Jensen S, Johnson A, Kurth-Nelson Z. Reconciling reinforcement learning models with behavioral extinction and renewal: Implications for addiction, relapse, and problem gambling. Psychological Review. 2007; 114(3):784–805. [PubMed: 17638506]

Ressler KJ, Rothbaum BO, Tannenbaum L, Anderson P, Graap K, Zimand E, Hodges L, Davis M. Cognitive enhancers as adjuncts to psychotherapy: use of d-cycloserine in phobic individuals to facilitate extinction of fear. Archives of General Psychiatry. 2004; 61(11):1136–1144. [PubMed: 15520361]

Schmitzer-Torbert NC, Redish AD. Development of path stereotypy in a single day in rats on a multiple-T maze. Archives Italiennes de Biologie. 2002; 140:295–301. [PubMed: 12228982]

Schmitzer-Torbert NC, Redish AD. Neuronal activity in the rodent dorsal striatum in sequential navigation: Separation of spatial and reward responses on the multiple-T task. Journal of Neurophysiology. 2004; 91(5):2259–2272. [PubMed: 14736863]

Steele RJ, Morris RGM. Delay-dependent impariment of a matching-to-place task with chronic and intrahippocampal infusion of the NMDA-antagonist D-AP5. Hippocampus. 1999; 9:118–136. [PubMed: 10226773]

Tang YP, Shimizu E, Dube GR, Rampon C, Kerchner GA, Zhuo M, Liu G, Tsien JZ. Genetic enhancement of learning and memory in mice. Nature. 1999; 401(6748):63–69. [PubMed: 10485705]

Tolman EC. The determiners of behavior at a choice point. Psychological Review. 1938; 45(1):1–41.

Tolman EC. Prediction of vicarious trial and error by means of the schematic sowbug. Psychological Review. 1939; 46:318–336.

Tolman EC. Cognitive maps in rats and men. Psychological Review. 1948; 55:189–208. [PubMed: 18870876]

van der Meer MAA, Redish AD. Covert expectation-of-reward in rat ventral striatum at decision points. Frontiers in Integrative Neuroscience. 2009; 3(1):1–15. [PubMed: 19225578]

van der Meer MAA, Johnson A, Schmitzer-Torbert NC, Redish AD. Triple Dissociation of Information Processing in Dorsal Striatum, Ventral Striatum, and Hippocampus on a Learned Spatial Decision Task. Neuron. 2010; 67(1):25–32. [PubMed: 20624589]
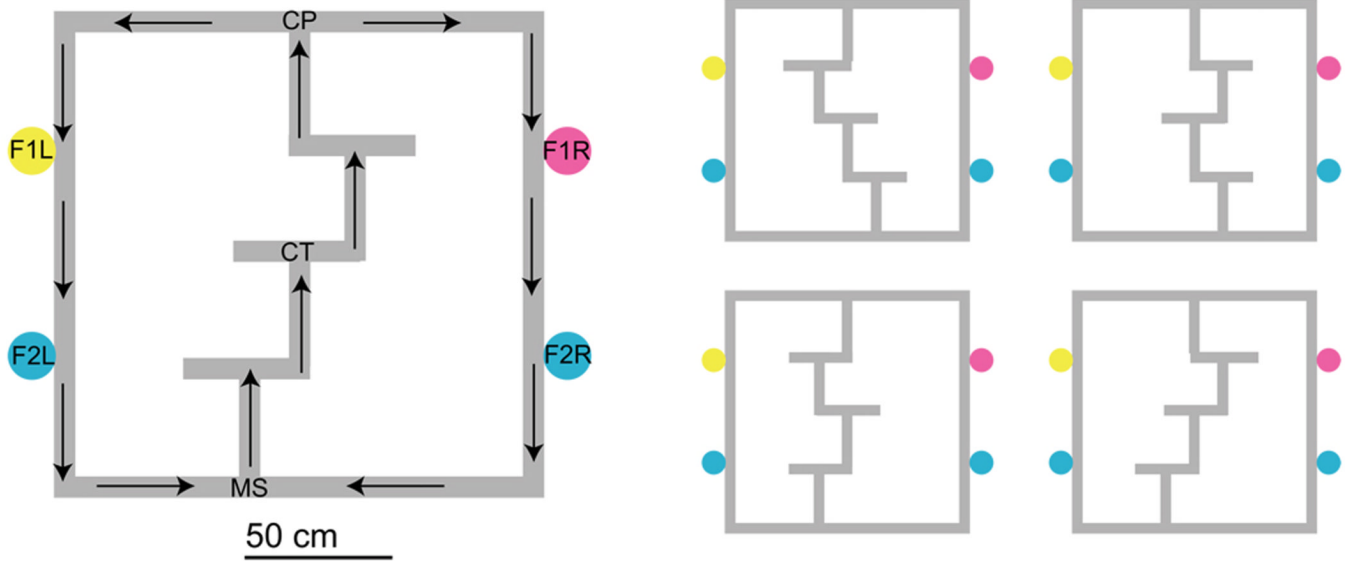
Figure 1.
The Multiple-T maze. Rats run through four choice points along a *navigation sequence* in order to receive food at two sites on either the left or right return rails. On each lap, only one return rail was rewarded. MS: Maze-start. CT: control-point. CP: choice-point. F1L: Feeder-1-left, providing banana-flavored food pellets. F1R: Feeder-1-right, providing fruit-flavored food pellets. F2L, F2R: Feeder-2-left and -right, providing unflavored (white) food pellets. Rats ran a continuous loop and were not removed from the task during the 40-minute experimental session See text for reward-contingency details. During training, all 8 potential maze configurations were used, however only the 4 maze configurations on the right were used during experimental sessions to ensure that the control point remained at a constant position in the room during experimental sessions.
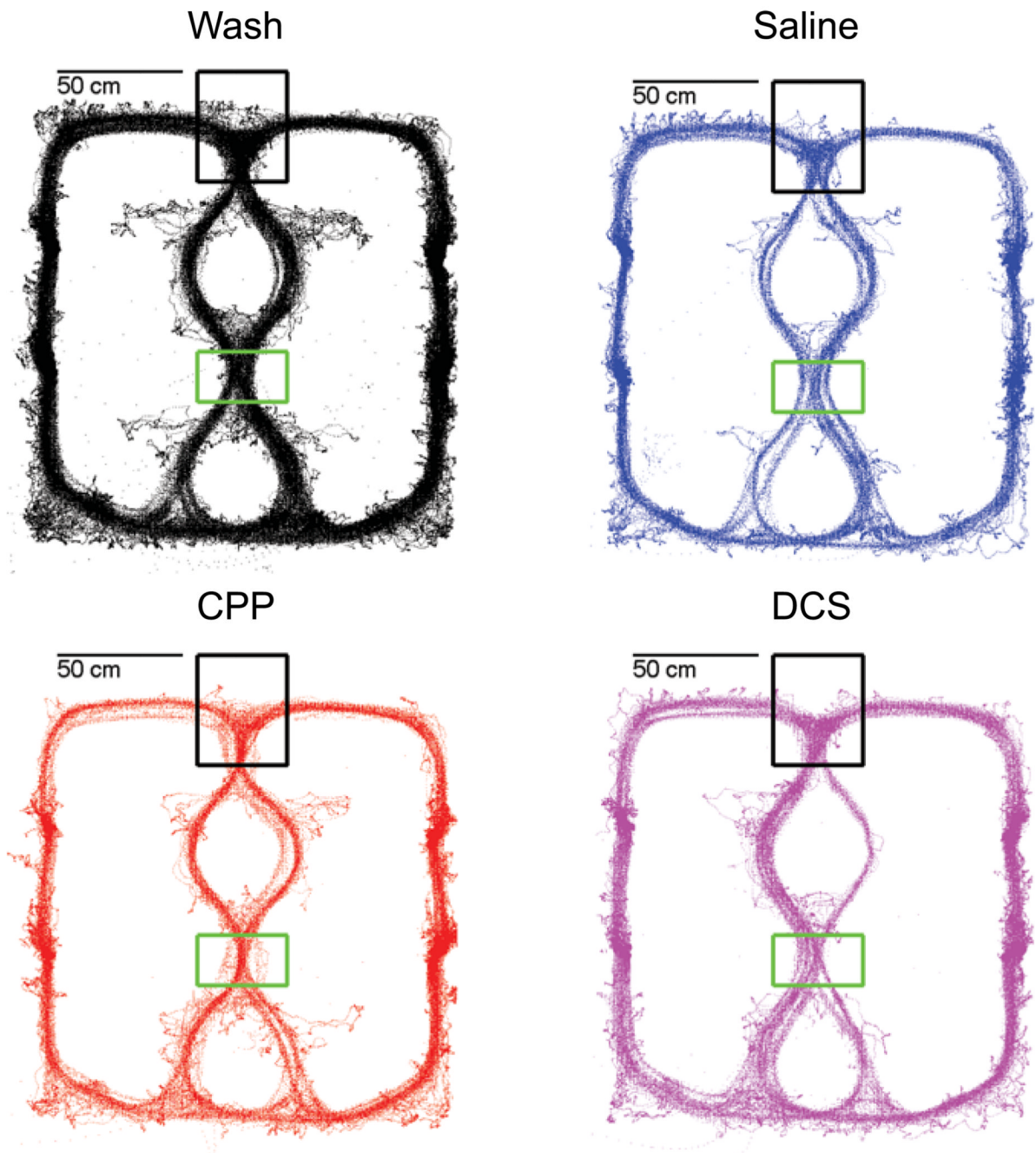
Figure 2.
Tracking on the Multiple-T maze. All tracked positions of all rats from all analyzed sessions are plotted on top of each other. Black dots are from *wash* sessions. Blue dots are from *saline* sessions. Red dots are from *CPP* sessions. Magenta dots are from *DCS* sessions. The choice-point region-of-interest (CP, at T4) is shown in black. The control-point region-of-interest (CT, at T2) is shown in green.
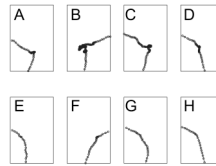
Figure 3.
Vicarious trial and error behavior at the choice point. Each panel shows one example of a pass through the choice point. Rats entered at the bottom and exited to either the left or the right. (A–D) Examples of passes categorized as VTE. (E–H) Examples of passes categorized as not-VTE. See text for details.
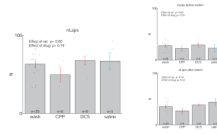
Figure 4.
Number of laps run on the Multiple-T maze. A lap was counted as a pass through the choice-point, separated by a full return along one of the return rails (rewarded or not). No significant effect of drug was seen.
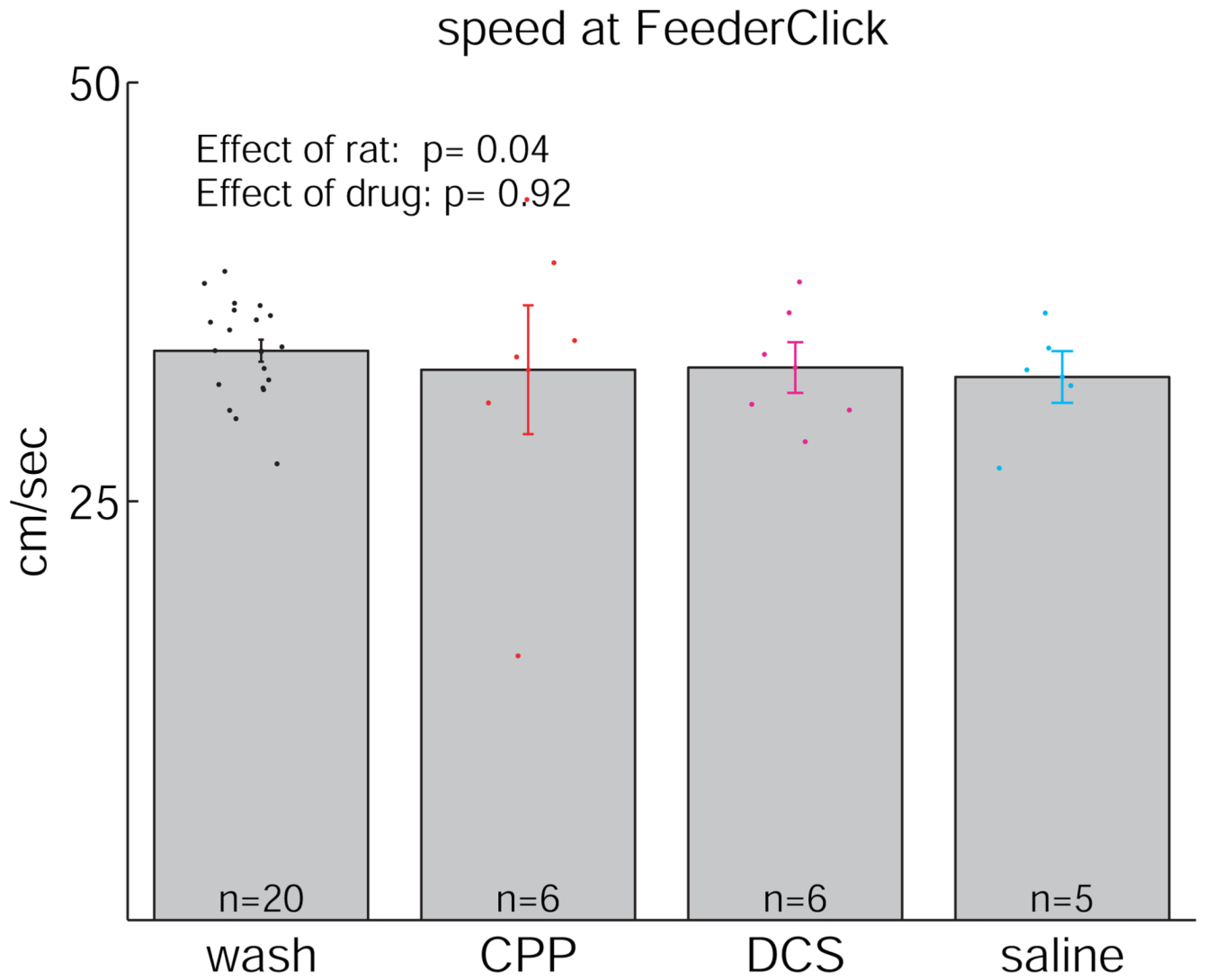
Figure 5.
Speed 0.25 s after triggering of the first feeder. Speed was measured using the Janabi-Sharifi et al. (2000) algorithm (see Masimore et al., 2005). No significant effect of drug was seen.
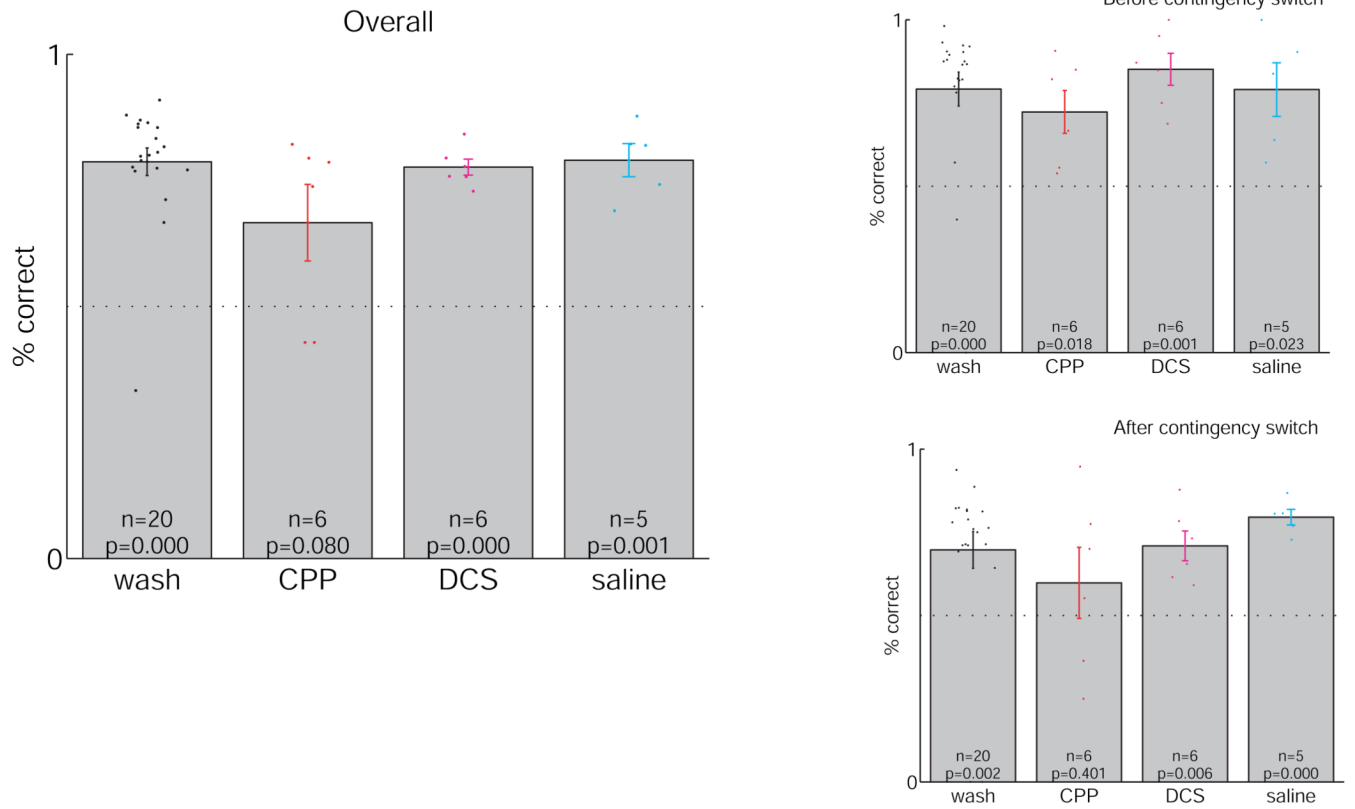
Figure 6.
Success rates on the Multiple-T task. *p*-values at the bottom of each bar indicate significance test relative to chance (0.5). Success rates siginificantly different from chance are indicated in red. Note that after the reward contingency switch, rats receiving CPP were not significantly different from chance.
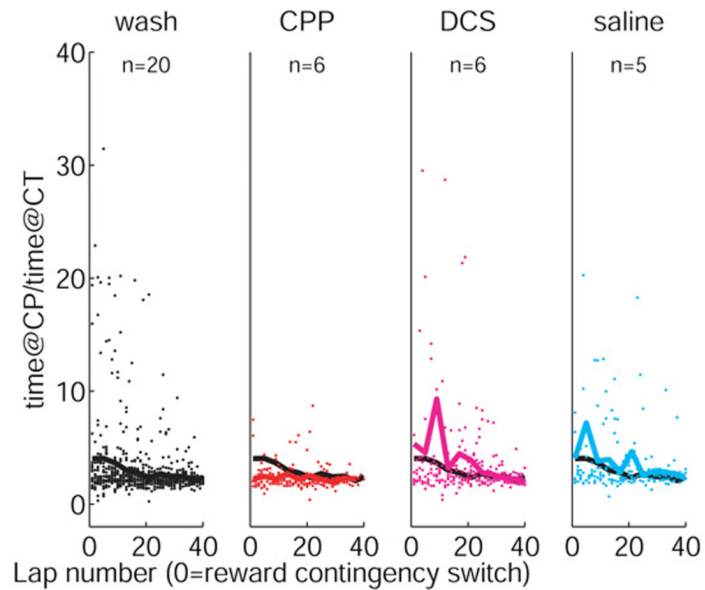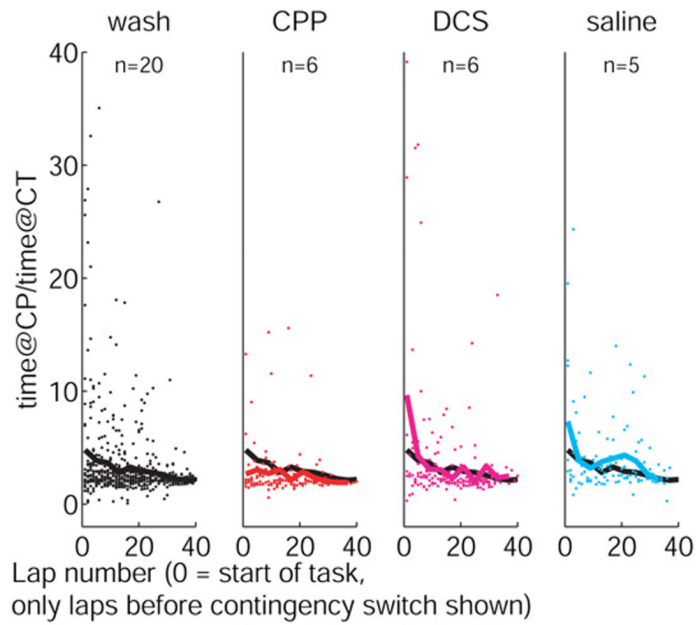
Figure 7.
Pausing time as a function of lap and drug condition. (Before contingency switch: Note the initial high ratio decreasing over lap for the *wash*, *DCS*, and *saline* conditions. But the smaller ratio in the *CPP* condition. After contingency switch: Note the re-initiation of VTE behaviors after the contingency switch, decreasing over lap for the *wash*, *DCS*, and *saline* conditions. Note the clear lack of VTE effects 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.
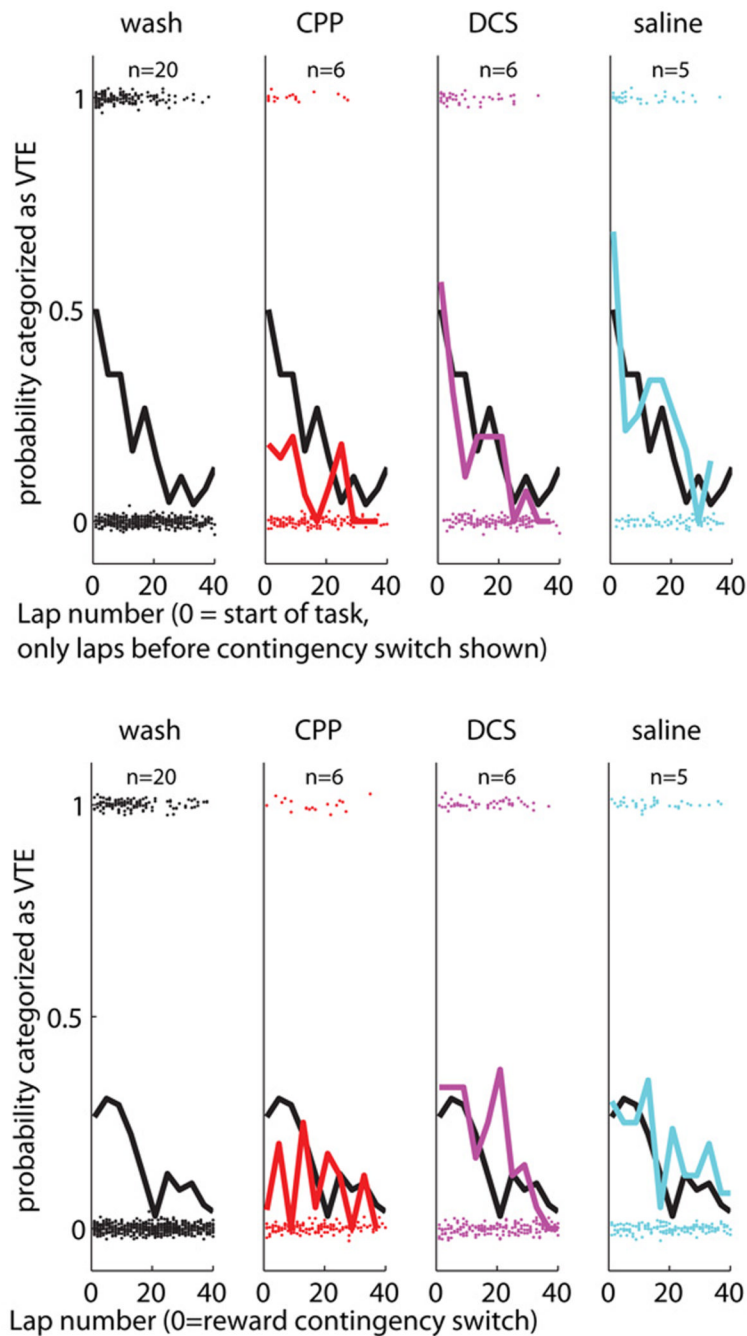
Figure 8.
Probability of pass being categorized as containing a vicarious trial and error (VTE) event as a function of lap and drug condition. (Before contingency switch: Note the initial high ratio decreasing over lap for the *wash*, *DCS*, and *saline* conditions and the smaller ratio in the *CPP* condition. After contingency switch: Note the re-initiation of VTE behaviors after the contingency switch, decreasing over lap for the *wash*, *DCS*, and *saline* conditions. Note the clear lack of VTE effects in the *CPP* condition after the switch in reward contingency. The black line repeated on each condition is the mean probability for the corresponding *wash* condition for comparison.