

BIOL 499 SENIOR HONOURS PROJECT

Where will it eat?

Investigating the effect of varying reward sites for
learned foraging in rats

Kris Wu

Submitted on: 4/18/2013

Supervised by:

Dr. Matthijs van der Meer

Second Reader:

Dr. Britt Anderson

Approx. 5,600 words

Acknowledgements

This undergraduate senior honours thesis project would not have been possible without the guidance, supervision, support, and mentorship from Dr. Matthijs van der Meer. I am very grateful for having had the opportunity extended to work in the lab. I am also thankful for the participation of my second reader, Dr. Britt Anderson, who has been an especially helpful teacher and mentor over the last two years of my undergraduate career. I extend a great amount of appreciation to the van der Meer lab, all of whom have been available for discussion and advice along the way. I am especially thankful for the animal handling training I received from Robert Cross and Sushant Malholtra. Lastly, I would like to thank Radhika Shankar, my fellow 499 student and a friend who was there working with me every step of the way.

Abstract

Research in foraging behaviour depicts foraging as a hierarchy of decisions sensitive to possible trade-offs and rewards (e.g. predator presence, quantity of reward, etc.). A recently proposed reward-based decision making model classifies one of these decisions as a process of valuation (i.e. attributing relative values to differing options). Neuroelectrophysiology reveals that the ventral striatum contains anticipatory ramping cells that may play a large role in influencing the valuation process in reward-based decision making. Before neural recording can take place however, there must be behavioural evidence that a task can be learned. The principle aim of this project is, on a novel set-up, to demonstrate that rats are capable of adapting to a foraging strategy based on how reward is varied by quantity and location. To accomplish this, we constructed an elevated square track with four independently variable reward locations. Rats were trained to run counter-clockwise on the track, breaking a photo beam at each site in order to trigger reward release. Behavioural data revealed patterns consistent with learning. A two-way analysis of variance with reward locations and days as independent variables revealed significant difference between unrewarded and rewarded sites, providing evidence for place-preference learning. After rats demonstrated evidence of learning the basic task, reward locations were modified to determine what strategies were employed in learning the task.

Contents

Acknowledgements.....	ii
Abstract.....	iii
List of Figures.....	v
Introduction.....	1
Methods.....	6
Apparatus.....	6
Handling – Training Timeline.....	7
Place-Reward Task.....	7
Further Manipulations.....	8
Results.....	9
Place-Reward Task.....	9
Extinction.....	10
Rotated Reward Map.....	11
Reward Map Reversal.....	12
Reward Map Reversal and Switch.....	13
General Discussion.....	14
Appendix A: Results Figures.....	19
Place-Reward Task.....	19
Extinction.....	21
Rotated Reward Map.....	23
Reward Map Reversal.....	27
Reward Map Reversal and Switch.....	29
References.....	30

List of Figures

Figure 1 Ventral Striatal Ramp cells shows increasing activity while approaching first reward site (F1) (van der Meer, et al., 2011).	3
Figure 2: Square Maze Simple Diagram.....	6
Figure 3: R019 Place Reward % nose poke across days	19
Figure 4: R019 Place Reward median latency across days.....	19
Figure 5: R023 Place Reward % nose poke across days	20
Figure 6: R023 Place Reward median latency across days.....	20
Figure 7: R019 Extinction % nose poke across days	21
Figure 8: R019 Extinction median latency across days	21
Figure 9: R023 Extinction % nose poke across days	22
Figure 10: R023 Extinction median latency across days	22
Figure 11: R019 Rotated Reward Map % nose poke across days	23
Figure 12: R019 Rotated Reward Map median latency across days	23
Figure 13: R019 Day 1 Rotated Reward Map – Within Session	24
Figure 14: R019 Day 3 Rotated Reward Map – Within Session	24
Figure 15: R023 Rotated Reward Map % nose poke across days	25
Figure 16: R023 Rotated Reward Map median latency across days	25
Figure 17: R023 Day 1 Rotated Reward Map – Within Session	26
Figure 18: R023 Day 3 Rotated Reward Map – Within Session	26
Figure 19: R019 Reversed Reward Map % nose pokes across days	27
Figure 20: Reversed Reward Map median latency across days.....	27
Figure 21: R023 Reversed Reward Map % nose poke across days.....	28
Figure 22: R023 Reversed Reward Map median latency across days	28
Figure 23: R019 Reversed Reward Map and Switch – Within Session	29
Figure 24: R023 Reversed Reward Map and Switch – Within Session	29

Introduction

For us, eating options vary immensely. The possibilities range from multitudes of restaurants to preparing a home cooked meal. Even when we are only faced with choosing a restaurant, we must weigh countless decisions including choosing the type of cuisine, how long we are willing to wait, how far we want to travel, and even which specific dishes we want. Given the vast number of factors, it is astounding that we are able to finally decide where and what to eat. While perhaps not as complex as our experience, most other animals are also presented with many foraging options that they must evaluate. Factors including possible predator presence, the size of the food reward, and the amount of effort required must all be considered in order to choose the most optimal foraging location. This project will examine a basic framework for these decisions, design and run a basic behavioral experiment, seek evidence behind what learning strategies are present, and establish some foundations for future neuroscience research needed to identify what neural mechanisms drive the associated decision making behaviour.

While many theories for animal foraging have been studied over the years (for a brief review, see Perry, G., & Pianka, E.R., 1997), re-examining and reviewing them it is out of the scope of this study. Foraging theories emphasize that many factors contribute to the decision of what patches/flowers/areas should be harvested. What underlies each foraging theory or model however, is the fundamental need for the animal to make a decision. In fact, behavioural ecologists view foraging as a hierarchy of decisions, with each decision point being sensitive to the possible trade-offs of each possible outcome (Stephens, D.W., 2008). Whenever an animal is

required to make a decision based on how it evaluates a list of possible alternatives, it employs what Rangel, Camerer, & Montague (2008) call value-based decision making.

In their Neuroeconomics review, Rangel et al. (2008) outline a framework for value-based decision making. This framework is the focus of this proposal. More specifically, we wish to explore a particular process of value-based decision making in rats. Rangel et al. (2008) separates their decision framework's computation into five basic processes. The first involves computing a representation of the problem, identifying internal (e.g. fatigue) states, external (e.g. distance to travel) states, and possible actions (e.g. travel to obtain food). The second process assigns a specific value to each of the possible actions given the internal and external states. The third process compares and selects an action to follow based on the relative values. Once the action has been completed, the fourth process measures the desirability of the outcome actually obtained. Finally, the last process facilitates learning by updating the first three processes based on the fourth's feedback.

To demonstrate how the framework may work, we will look to a basic foraging situation. Assuming a rat is hungry, alert, and knows where several food locations are, the first process will construct a representation for its internal and external states as well as a possible action given the known food sites. The valuation process will assign values to the act of visiting (or not) each foraging site. Assuming there are three foraging sites: one with a large amount of food, one with little food, and the last one with a modest amount of food, it is likely that the valuation process will assign the largest value to the first, and smaller respective values to the second and third. The third process will select the action with the largest assigned value (visiting the first site). Once the visit is complete, the fourth process assigns a value to the outcome, and

the last process will update the other processes based on this feedback. If the visit to the first site began to have decreasing amounts of food, it is likely that over time, the valuation process will be updated to assign a different value to the act of visiting the first site.

Rangel et al. (2008) are transparent in saying that these categories, or processes, are not rigid. Many questions still remain as to whether their descriptions follow actual neural behaviour. Must valuation take place before selecting an action, are they performed in parallel, or, depending on the context of choice, does a separate process for action selection exist at all? They are quick to point out that though they have delineated each process, the processes themselves do not necessarily map onto neural networks in the same way. Nonetheless, they provide a framework with which we can use as a starting point to ask questions about decision making behaviour and their possible neural correlates.

Although the specific neural processes that underlie valuation are not yet well understood, the ventral striatum has been demonstrated as a central node in processing reward or value information in both rats and humans (van der Meer, M.A.A., & Redish, A.D., 2009; van der Meer, M.A.A., & Redish, A.D., 2011; Schönberg, T., Daw, N.D., Joel, D., & O’Doherty, J.P.,

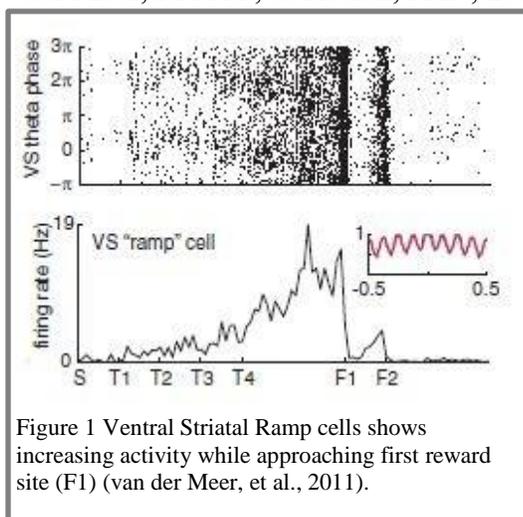


Figure 1 Ventral Striatal Ramp cells shows increasing activity while approaching first reward site (F1) (van der Meer, et al., 2011).

2007). Particularly noteworthy is the finding that in rats (Figure 1), the ventral striatum follows a definite ramping firing pattern as the rat approaches reward sites (van der Meer, M.A.A., & Redish, A.D., 2009; van der Meer, M.A.A., & Redish, A.D., 2011; Malhotra, S., Cross, R.W.A., & van der Meer, M.A.A., 2012). Since this ramping firing pattern is observed as rats

approach specific reward outcomes (and sometimes trigger cues [Malhotra et al., 2012]), it is not therefore, farfetched to assume that these ventral striatal cells play a role in Rangel et al.'s described valuation process. Indeed, in their review, Malhotra et al. (2012) noted that ventral striatal ramp cells can distinguish between predicted reward amounts, outcomes, and that some neurons ramp more for larger rewards and some ramp more for smaller rewards. It is very possible that this neural activity represents, physically, what Rangel et al. (2008) describe as the valuation process.

Before neural recording can take place on a novel task set-up, however, rats must first demonstrate being able to learn the behaviour required to accomplish the desired task. The primary goal of this project is to demonstrate that rats are indeed capable of adapting to a foraging strategy based on the value of the reward. In order to accomplish this, we will employ a square maze with four independently variable reward sites. Rats will be introduced to the maze having each reward site primed with a different but stable amount, or value, of reward. It is expected that given some time, the rats will learn to associate different levels of rewards with their specific sites.

Indeed, previous animal research has clearly demonstrated that rodents of all kinds are able to locate food rewards in various task settings (for a review, see Redish, 1999). In fact, animals have shown the ability to employ both local (i.e. landmarks located within their environment) and distal (i.e. landmarks outside of their immediate vicinity) cues in order to navigate to a reward site, escape route, or other favourable location. After training rats on the basic place-reward association task, we will then employ several manipulations to help seek evidence for what strategies were used in learning the task (i.e. what cues the rat employed).

In order to help determine whether or not the rats have truly learned the task, we will employ two main measures: reward location visits as measured by nose pokes, and running speed as measured by the time required for the rat to run from the track's midpoint to the reward site. It is naturally assumed that as a rat learns to associate a specific sites with greater reward and other sites with smaller or no reward, it will begin to display a bias in sampling, that is to say, it will continue visiting sites with reward and begin skipping unrewarded sites. Indeed, drug-related conditioned place preference studies (see Tzschentke, T.M., 1998 for a review; see also Carboni, E., & Vacca, C., 2003, for a review conditioned place preference study designs) demonstrate that experiences of reward lead to location preference where reward was experienced. In the same way, we expect for rats to show preference for reward sites in which they are actually rewarded.

In much the same way, rats have demonstrated preference for reward site shown in their running speed (Crespi, 1942; see also Miller, N.E. & Miles, W.R., 1935; and Reynolds, W.F. & Pavlik, W.B., 1960). Critically, the magnitude or value of the reward object determines the response strength, in this case, running speed. We would also then expect the time a rat will take to run from the midpoint of a track to the reward site will be shortest when running towards greater reward sites and longest for unrewarded sites.

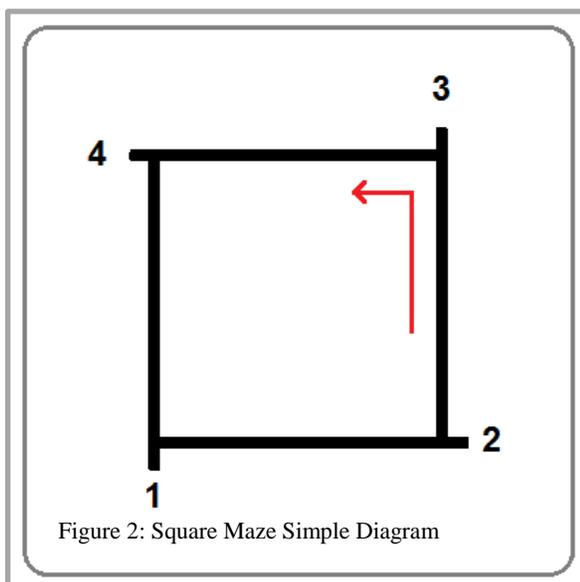
While neural recording is not possible in the eight months provided for this project, the work completed in the project will set the foundation for future research. Future long term goals include uncovering how activity in the ventral striatum influences value-based decision making in situations like the aforementioned foraging example. By recording ventral striatal

activity as rats conduct the place reward tasks, we gain insight into how this area of the brain influences value-based decision making.

Methods

Apparatus

We employed a version of the Take5 task by Schmitzer-Torbert, N.C., & Redish, A.D. (2008). Rats were trained to run clockwise on an elevated square track (Figure 2). The track



was constructed using wooden boards and mounted on a separate metal frame. The track was 100 centimeters long and 10 centimeters wide. Each corner had a 10x10 centimeter square extension where a food pellet receiver was placed. Photo beam sensors were placed directly in front of the pellet receiver in order to detect food site visits, or nose pokes. Food pellet dispensers were placed at

each corner of the track and connected to the receptacles. These dispensers were triggered when the rat disrupted a photo beam placed at the entrance of the extension. Photo beam sensors and pellet feeders were connected to a computer using a National Instrument's Data Acquisition system. A ceiling mounted webcam was connected to the computer to be used for position tracking.

Handling – Training Timeline

Two rats (R019 and R023) were handled and trained on the task over the course of just under two months. All data collected for this project was collected during these training sessions. Rats underwent one week of pre-training handling where they were allowed to explore on the experimenter's lap for ten to twenty minutes per session. This ensured that each rat was comfortable with the experimenter as well as allowing the experimenter to gauge each rat's individual temperament. After, each rat was placed on the apparatus for pre-task training. During this time, each feeder was primed to release food pellets to condition the rats to associate reward with the reward sites. Reward sites were only rewarded if rats proceeded around the track in a counter-clockwise fashion (Figure 2). After one week of pre-task training, the rats learned to run counter-clockwise in order to receive food reward. Once pre-training was complete, they were trained on a basic place-reward task for twelve days in order to investigate the basic aim of this project. Further manipulations including extinction (5 days) and others (rotated reward map: 4 days, reversed reward map: 4 days) were conducted to determine what strategies the rats had used to learn the task.

Place-Reward Task

In order to demonstrate that rats are capable of adapting to a learned foraging strategy, we set up a basic task where different reward values were assigned to different feeding sites, or places, hence a place-reward task. For the first four days of training, site 1 (for site references, see Figure 2) dispensed 0 pellets, sites 2 and 3 dispensed 2 pellets, and site 4 dispensed 4 pellets. From day 5 onwards, site 3 became the 4 pellet reward site and site 4 became a 2 pellet reward site.

In addition, we implemented a delay in the pellet dispensing process to ensure that the rats would learn to associate visiting and poking their noses into a site in order to obtain food reward. In order to trigger pellet release when a delay is put in place, a rat would have to keep the photo beam broken for the amount of time set for the delay. We began with a 500-ms delay on day 8 and increased it to a 1000-ms delay on day 10. We continued with a 1000-ms delay for the remainder of the training period.

Further Manipulations

Once the rats demonstrated that they were capable of learning the basic place-reward task, we employed several manipulations to seek evidence that the rats had actually remembered the task one day to the next, and for the possible learning strategies they employed in learning the task. In order to ensure the rats were not engaging in new learning during each session and actually remembered the task from a prior training day, we employed the use of an extinction phase during the experiment. Extinction also helps to provide evidence that the rats are not merely learning a reward pattern (i.e. being able to predict what reward is next based on a one-back and/or two-back check of what reward was received, e.g. “I am going to receive 0 rewards because my last reward was 2 pellets and the one before that was 4 pellets.”). To accomplish this, we included a 5 minute extinction phase at the beginning of the experiment where rats would not be rewarded any food pellets. Once the extinction phase was over, the regular place reward task continued. This was done for 5 days.

To help determine what strategies the rat was using to associate reward with place (i.e. feeding site), we employed three manipulations. In order to investigate evidence for whether the rats were using local or distal cues to help determine place-reward associations, we rotated the track during the last three days of the extinction period. The reward sites remained the same

in relation to the experiment room. In this way, the local cues have changed but the distal cues remained the same. This was the first manipulation.

To provide evidence for how the rats learned the task we employed a second and third manipulation. The second was a rotated reward map where we rotated the amount of reward distributed per site. The basic task had a [site: reward] pattern that resembled [1:0; 2:2; 3:4; 4:2]. Once the reward map was rotated, the pattern changed to [1:2; 2:0; 3:2; 4:4]. After 5 days of the rotated reward map manipulation, we reversed the reward map for 3 days, giving a [1:2; 2:4; 3:2; 4:0] reward pattern and the third manipulation. It is expected that as On the last day of the reversed reward map manipulation, we included a switch after half the session time elapsed, returning the reward map to the rotated reward map pattern. This was done to briefly observe how rats would update their learning even within a session.

Results

For all associated graphs, please see the figures included in Appendix A.

Place-Reward Task

Qualitative behavioural data indicate that both rats learned the place reward task within the 12 day training period (see Figures 3-6). Graphs clearly depict fewer and decreasing frequency of nose pokes at the unrewarded site, as well as longer travel times while approaching the unrewarded site. Since latency is easily affected by outliers, graphs display the median value for easier viewing. To determine whether there was a statistical difference in reward site visits as measured by nose pokes, a two-way analysis of variance (ANOVA) was conducted using day and feeder/reward site as independent variables. Data for both rats was pooled when a two-way

ANOVA with rat and reward site revealed no significant difference for performances between R019 and R023. There was overall significant variance in the model, $F(47, 95) = 6.23$, $MSE = 0.003$, $p < 0.001$. Most importantly, there was a significant main effect of reward site, $F(3, 95) = 81.01$, $MSE = 0.003$, $p < 0.001$, meaning there was significant variance between nose pokes at specific reward locations. Pairwise post-hoc (Tukey HSD) comparisons revealed only feeder 1, the unrewarded site to be visited significantly less than all other feeders ($M_{diff} = 0.195$, $p < 0.001$, $M_{diff} = 0.203$, $p < 0.001$, $M_{diff} = 0.221$, $p < 0.001$). No significant difference was found between all rewarded feeders.

The same ANOVA was carried out for the measure of latency and it followed the same pattern as with nose pokes. There was significant variance in the overall model, $F(47, 95) = 2.22$, $MSE = 3.56$, $p < 0.01$. As expected, there was a significant main effect of reward site, but not days, $F(3, 95) = 15.22$, $MSE = 3.56$, $p < 0.001$. Pairwise post-hoc comparisons once again showed that latencies while approaching feeder 1 (unrewarded site) were significantly larger than all other feeders ($M_{diff} = 3.15$, $p < 0.001$, $M_{diff} = 3.21$, $p < 0.001$, $M_{diff} = 2.27$, $p = 0.001$). Once again, there was no significant difference between all rewarded feeders.

Extinction

Figures 7-10 display data from the 5 minutes of extinction during the 5 days of extinction. As with the place-reward task, we pooled data for both rats in our statistical calculations. Also, only the 5 minutes of extinction were considered for our calculations.

A two-way ANOVA with reward site and day as independent variables was conducted for nose pokes resulting in significant variation in the overall model, $F(19, 39) = 4.26$, $MSE = 0.002$, $p = 0.001$. Once again, there was a significant main effect of reward site, $F(3, 39)$

= 19.53, $MSE = 0.002$, $p < 0.001$. Post-hoc comparisons once again revealed that feeder 1, the unrewarded site, was visited significantly fewer times than the other rewarded sites ($M_{diff} = 0.127$, $p < 0.001$, $M_{diff} = 0.104$, $p < 0.001$, $M_{diff} = 0.044$, $p < 0.05$).

The same ANOVA was carried out for latency measures. There was no significant variation in the overall model $F(19, 39) = 1.29$, $MSE = 4.58$, $p > 0.1$. There was, however, still a significant main effect of reward site, $F(3, 39) = 5.68$, $MSE = 4.58$, $p < 0.01$. Post-hoc comparisons revealed that latencies for approaching the “unrewarded site” (site 1) was significantly longer than approaching sites 3 and 4 ($M_{diff} = 3.50$, $p < 0.01$, $M_{diff} = 3.32$, $p < 0.05$), but not site 2. During non-extinction (i.e. place-reward task) sites 3 and 4 were 4 pellet and 2 pellet rewards respectively and site 2 was a 2 pellet reward site. Once again, rewarded sites were not significantly different from each other.

Rotated Reward Map

Behavioural data as represented by the graphs (figures 11, 12, 15, and 16) were analyzed by a two-way reward site by day ANOVA. Once again, data was pooled between the two rats. Data patterns were consistent with the basic place-reward task.

For nose pokes per reward site, there was significant variation in the overall model, $F(15, 31) = 15.30$, $MSE = 0.002$, $p < 0.001$. As expected there was a significant main effect of reward site, but not day, $F(3, 31) = 70.52$, $MSE = 0.002$, $p < 0.001$. Post-hoc comparisons revealed that the unrewarded site (this time site 2) was visited significantly fewer times than all rewarded sites ($M_{diff} = 0.215$, $p < 0.001$, $M_{diff} = 0.261$, $p < 0.001$, $M_{diff} = 0.257$, $p < 0.001$). As expected, there was no significant difference between all rewarded sites.

For the latency measure, there was no significant variation in the overall model, but a significant main effect, as expected, $F(15, 31) = 1.69, p > 0.05$, and $F(3, 31) = 4.79, MSE = 0.123, p < 0.05$, respectively. Post-hoc comparisons revealed that the previously unrewarded site (now 2 pellet reward site) only had significantly longer latency times compared to the one of the 2 pellet rewarded sites ($M_{diff} = 0.587, p < 0.05$). The current unrewarded site was not significantly different than any of the rewarded sites.

In addition to the two measures, we compiled a simple moving average across the laps in a single session. Averages were calculated per 15 laps (complete single run around the track). This provided a clearer view of the rat's behaviour as he engaged in learning during a session. It is interesting to observe that during the first day of introducing the rotated reward map, the proportion of nose pokes at the new 0 reward site is high but quickly declines (figures 13, 14, 17, and 18). During Day 3, it is clear that the unrewarded site visits begins and remains low in comparison to the other reward sites. Of particular interest is R023's behaviour for the rotated reward map's 2 pellet reward that was previously an unrewarded site (Figures 17 and 18). During the first day, he behaves as if the old unrewarded site was still dispensing 0 rewards and visits it less and less. By day 3 however, he has updated his behaviour to reflect the basic place-reward task.

Reward Map Reversal

Data patterns (figures 19 – 22) for the reversed reward map resemble that of the rotated task and the place-reward task. For nose pokes per feeder, there was significant variation in the overall model, $F(11, 23) = 135.1, MSE = 0.000, p < 0.001$. As expected, there was a main effect of reward site, $F(3, 23) = 488.8, MSE = 0.000, p < 0.001$. Post-hoc comparisons revealed that the unrewarded site was visited significantly fewer times than all three rewarded sites (M_{diff}

= 0.264, $p < 0.001$, $M_{diff} = 0.259$, $p < 0.001$, $M_{diff} = 0.259$, $p < 0.001$). The unrewarded sites were once again not significantly different from one another.

For latency measures, there was no significant overall variation, $F(3, 23) = 2.44$, $p > 0.05$. There was a significant main effect of reward site, $F(3, 33) = 6.66$, $MSE = 0.064$, $p < 0.01$. Post-hoc comparisons revealed the unrewarded site to be significantly different than the largest reward site ($M_{diff} = 0.652$, $p < 0.05$) but not the other two rewarded sites. The rewarded sites were not significantly different from each other.

Reward Map Reversal and Switch

Out of interest, a reward map reversal “switch” was implemented after half the training session had elapsed. As shown in the moving average within session learning graphs (figures 23 and 24), both rats demonstrate the place-reward bias in that they seldom visit the no reward site. Halfway through the task, they begin to update their behaviour by visiting the new no reward site less and less, and begin to visit the new 4 reward site more often for the remainder of the training session.

General Discussion

The challenges presented in this project have been numerous. When we began in September of 2012, the experiment room was not yet set up, the apparatus did not exist, and I did not possess many of the skills needed to make this experiment possible. Needless to say, this senior honours thesis project has involved learning and developing many skills along the way including but not limited to: basic wood and metal work construction, making electrical wiring, some computer programming, and most importantly animal handling.

This project sought to determine whether rats were able to associate different values of rewards to different feeding sites and thus learn to maximize their reward, or minimize their lost time, by only sampling sites that provided reward. Based on our data patterns, it is quite clear that rats show a tendency to visit unrewarded sites fewer than rewarded ones. It also appears that they are more hesitant in their approach. Their response strength as displayed by their running speeds slow, sometimes dramatically, as they approach an unrewarded feeding site.

While the basic place-reward task data seemingly demonstrates that rats had learned the task, it is possible that they did not maintain place-reward memory for the task if they employed one of two strategies. Had the rats memorized either the reward sequence, or engaged in new learning within the session each day, they could circumvent remembering the place-reward associations between days. Data from the extinction periods where no reward was given provide evidence that this is not the case. While the graphical representations (Figures 7-10, Appendix A) do not show a distinct difference between unrewarded and rewarded sites, our statistical analyses reveal that there was a significant difference between site for both nose pokes

and approach latencies. While the data pattern is not quite as clear for latencies, this is expected as this measure is more easily influenced by outliers. That combined with having a frustrating extinction phase is perhaps why no significant difference was observed between site 1 and 2, and unrewarded and one of the rewarded locations. Nevertheless, the bias the rats displayed towards the rewarded locations during extinction reveal that they must have retained some place-reward association in their memory from previous training days.

Perhaps a better measure that could have been employed would have been to compare the amount of time spent around each reward site, rather than just the site visits. Conditioned place preference studies demonstrate that animals will not only visit a location where reward was administered/is expected, but will also spend more time there (Carboni, E., & Vacca, C., 2003). In addition, it would be interesting to conduct a statistical analysis to determine whether rats stayed around reward sites significantly longer than unrewarded sites, thus showing conditioned place preference.

Further evidence that some sort of task based memory was retained between days was R023's performance for the first rotated reward map session (see figures 17 and 18, Appendix A). During the first session, he showed clear learning within the session that a new reward site was now no longer being rewarded, and began to visit it less. What displays his having remembered the task across days is his performance at the previously unrewarded site. Even though it is now primed to release a reward of 2 pellets, he retains his behaviour in nose poking at the site less and less, almost as if he assumed there was only one change in the reward map, namely that a new site was being unrewarded. This pattern quickly disappeared however, as we can see from his performance on the third day (figure 18, Appendix A). This simple

behaviour seemingly indicates that the rats do not simply remember a “reward pattern” dissociated from their actual place.

In addition to the basic place reward task, both rotated reward map and reversed reward map tasks displayed clear place-reward learning. Not only are rats able to learn a place-reward association, but they are also able to update their association given changes they observe across days, and even within a session. This provides evidence that place-reward associations learned by the rats are not only retained in memory across days of training, but there is also some sort of reward updating process in place to help them adapt as the task changes. Just as response strength as measured by site visits and running speed would change as the rat’s internal reward map were being updated, it would especially interesting to observe how the ventral striatal ramping cells thought to influence value-based reward would be modulated based on the changing tasks.

Although we only conducted one day of the reversed reward map and switch manipulation, it provides some evidence that place-reward learning, while not necessarily at this point all that significant, can occur during a session and not just across days. Both rats displayed having remembered the reversed reward map from the get go (figures 23 and 24, Appendix A). After the halfway point when the reward map switch (or re-reversal) occurred, both rats began to display a clear behaviour change. They began to visit the new no reward site less and less, and began to visit the new 4 reward site more often for the remainder of the training session. While less dramatic in R019 to R023, this could possibly be attributed the inherent differences in their temperament (being frustrated at change more easily could lead to slower learning) and individual intelligence. Nevertheless, it is clear that rats are able, to some extent; update their place-reward associations during a training session.

It is interesting to note that for the most part, no difference in performance was observed between 2 pellet and 4 pellet rewards. In every pairwise comparison conducted, all rewarded sites were not significantly different from each other in performance. This could be due to the fact that the difference in pellet numbers for this task does not create a large enough contrast in value for the rats to care. Given that all three rewarded sites are rewarded on any given lap around the track, it would make sense for the rat to maximize its reward by visiting all the sites. Perhaps if there was a sizeable enough difference between the largest reward site and other reward sites so that the combined value of the lesser rewards do not significantly compare to the largest reward, the rat might learn a strategy to maximize reward and minimize time spent on the track?

With the primary goal of this project established, we went on to try and observe what strategies rats could use to make their place-reward association. More specifically, we sought to obtain evidence required to determine whether rats were using local or distal cues in their learning strategies. As it is difficult (especially for novices) to make perfectly symmetrical woodworked structures, our apparatus was by no means completely exact. Differing imperfections at and around each corner, or reward site, would provide possible cues for the rats to use in their learning. It is also perfectly probable that the rats could have ignored these features for preference of using the different aspects of the room (e.g. relative locations of things like the computer, shelves, and walls to the reward sites on the track) as distal cues.

During extinction, we rotated the entire track on day three to five to observe their behavioural biases. Since no reward is given during the extinction period, it was expected that if rats used local cues in their navigation, their behaviour would follow the movement of the physical feeders. If they used distal cues, then there should be no observed change in behaviour.

While the data does not look clear (figures 7 – 10, Appendix A), the decrease in clear distinctions between reward site performances can be attributed to both frustration and task adaptation as the rat is learning during the 5 minute period that no feeder sites are being rewarded. Nevertheless, statistical analyses revealed that performance matched the basic place-reward task, demonstrating that local cues most likely do not really play a large role in establishing place-reward associations. Rats most likely use distal features in the room to help them navigate during their task.

What is evident from our data is that rats are indeed very adept at learning to associate reward value with place. They are able to adapt to a foraging strategy where they preferentially select from rewarded sites and skip unrewarded sites. They are able to retain the place-reward association in memory across days and also update their associations as the tasks change. Unfortunately, in this study, we were unable to show any differences between our 2 pellet reward and 4 pellet reward sites. Now that it has been made clear that rats are capable of learning a basic place-reward task, it would be interesting to observe what neural processes underlie their behaviour.

Appendix A: Results Figures

Place-Reward Task

R019

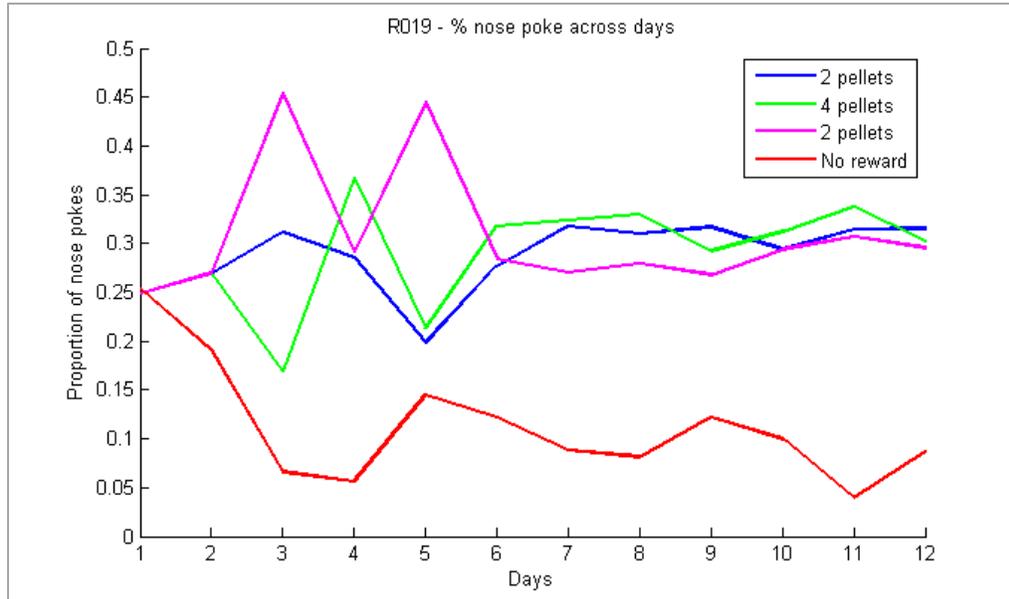


Figure 3: R019 Place Reward % nose poke across days

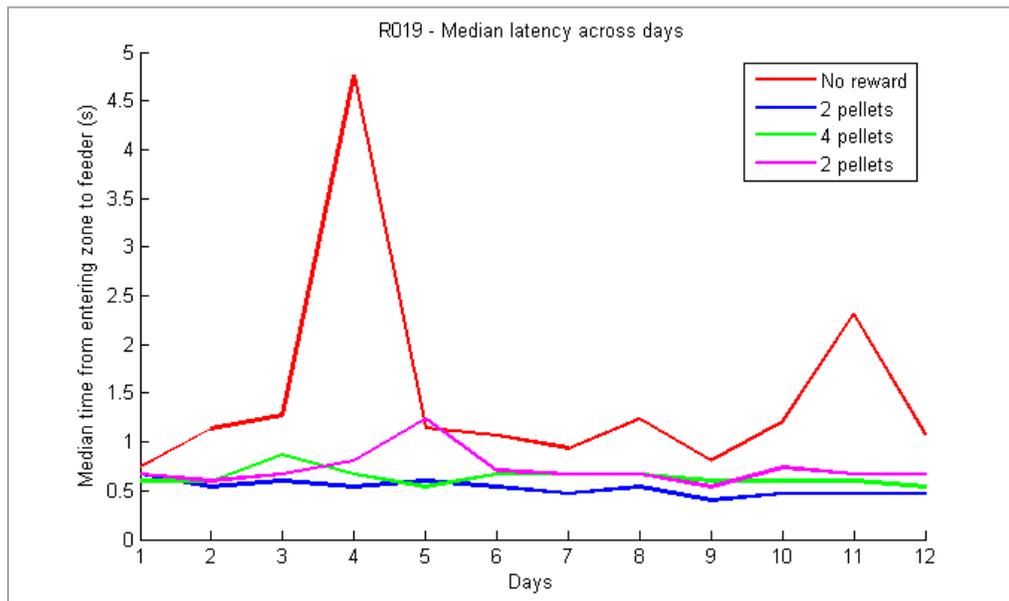


Figure 4: R019 Place Reward median latency across days

R023

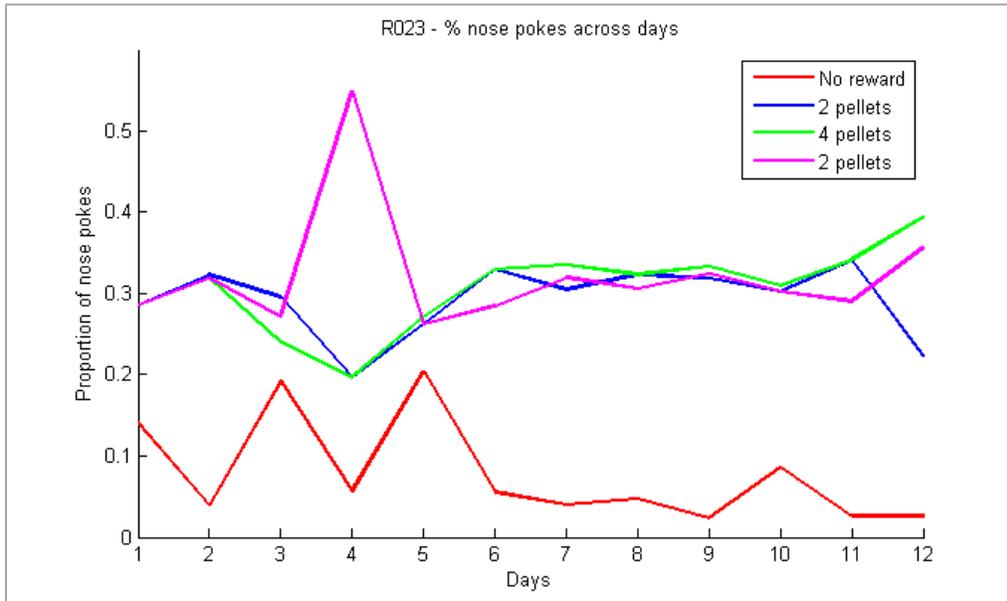


Figure 5: R023 Place Reward % nose poke across days

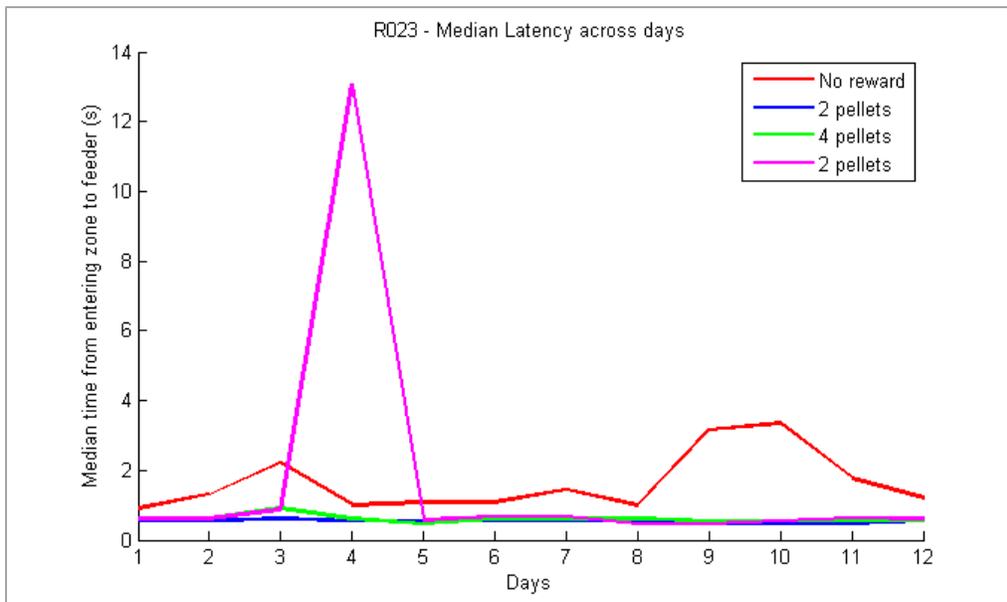


Figure 6: R023 Place Reward median latency across days

Extinction

The following graphs display only the 5 minutes of extinction on each day of training.

R019

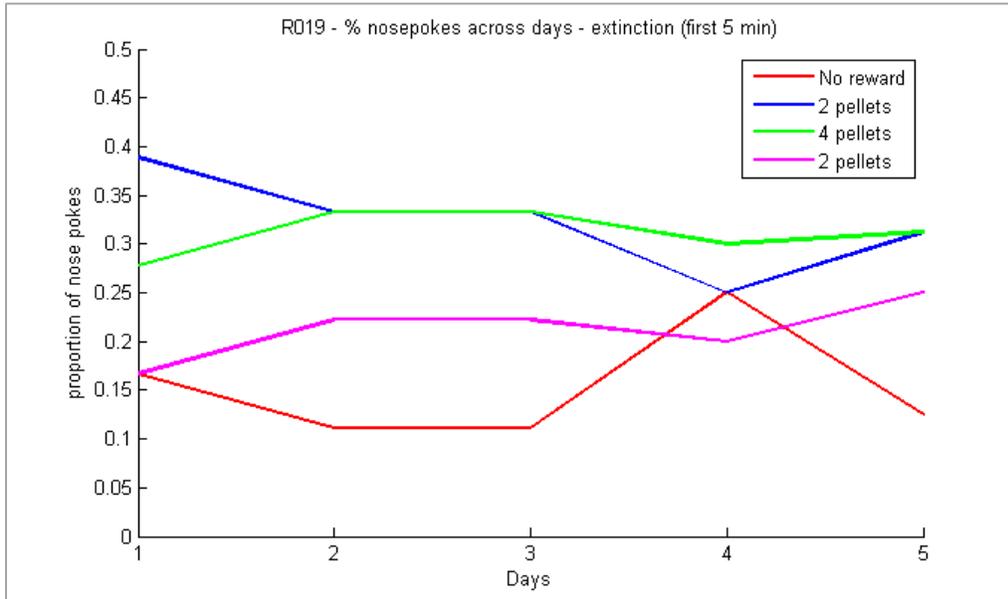


Figure 7: R019 Extinction % nose poke across days

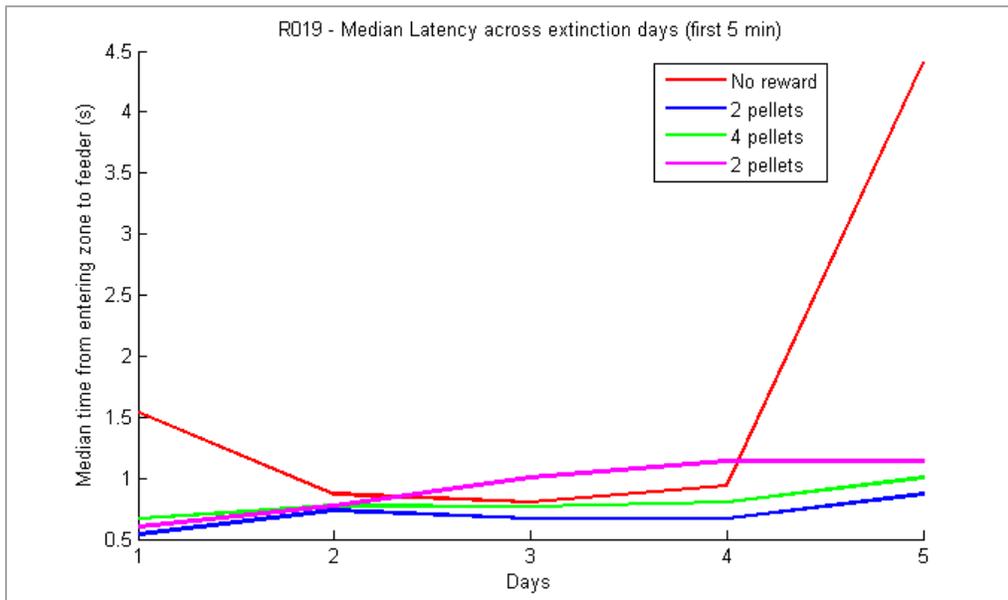


Figure 8: R019 Extinction median latency across days

R023

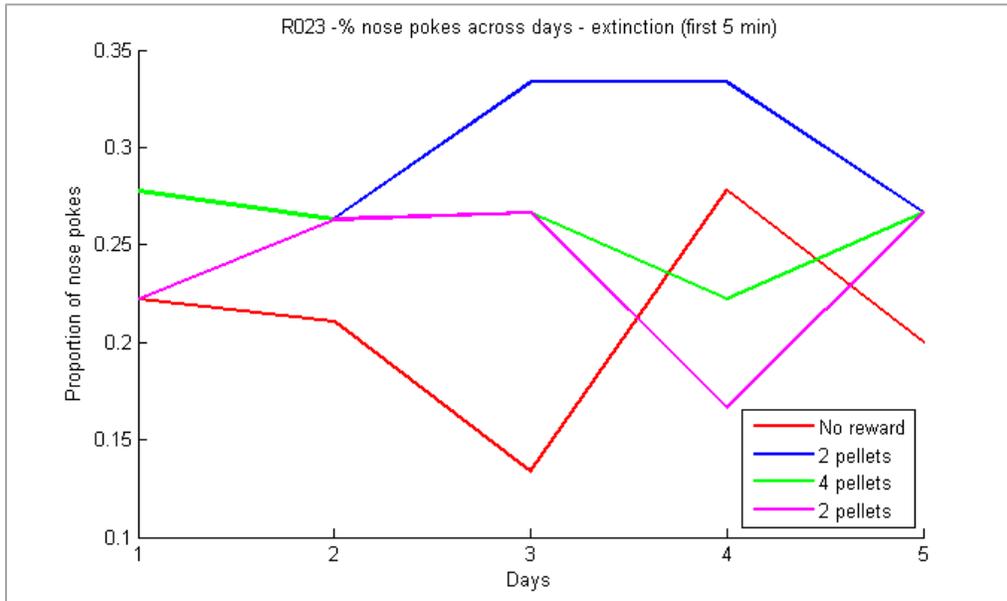


Figure 9: R023 Extinction % nose poke across days

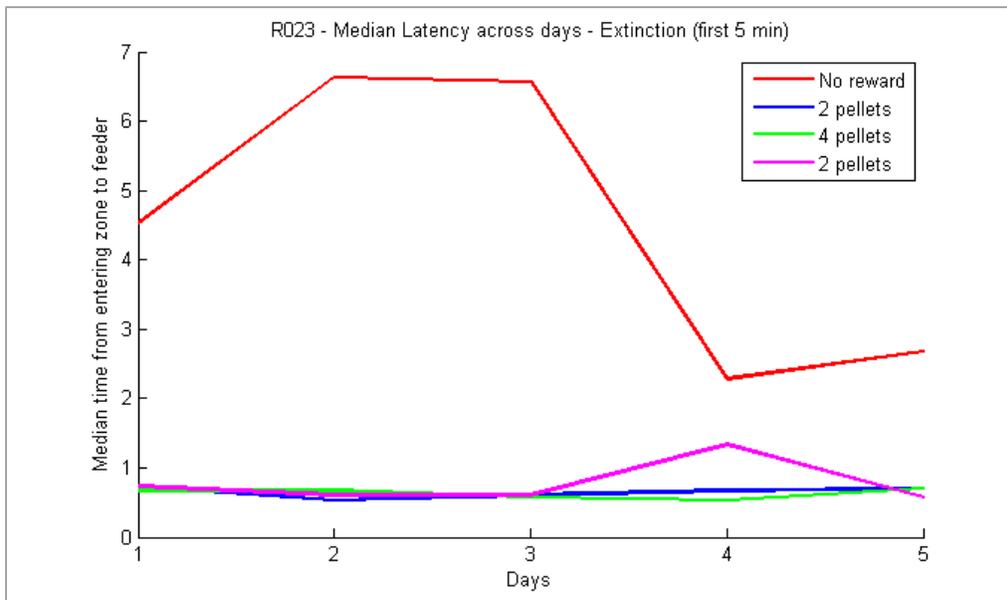


Figure 10: R023 Extinction median latency across days

Rotated Reward Map

R019

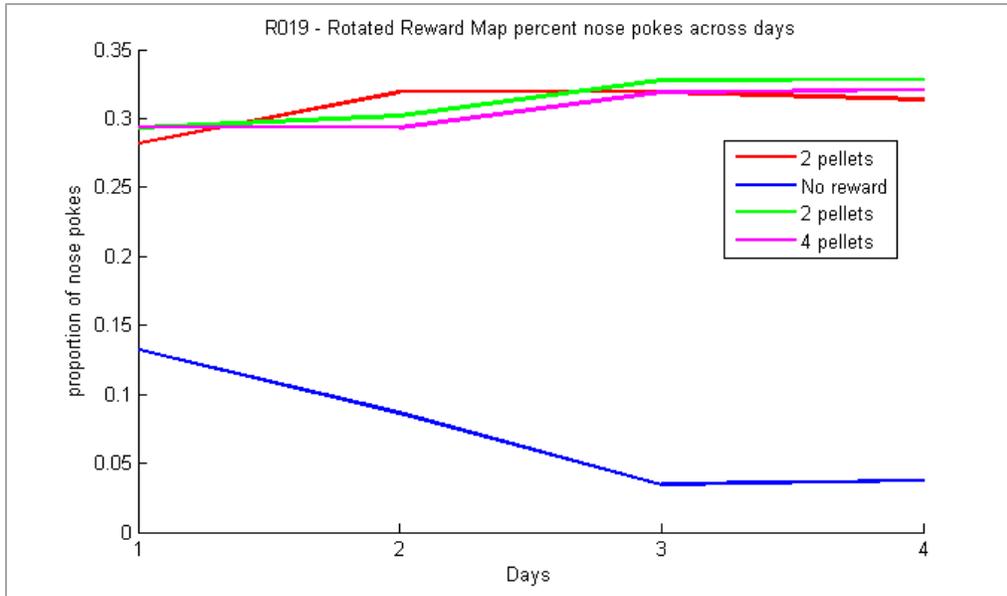


Figure 11: R019 Rotated Reward Map % nose poke across days

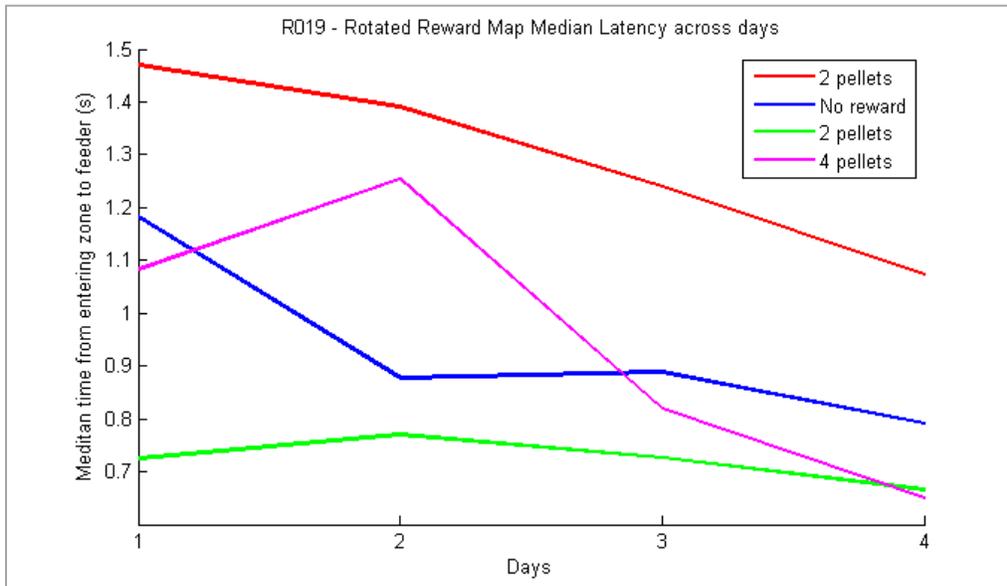


Figure 12: R019 Rotated Reward Map median latency across days

Within Session Learning Plots

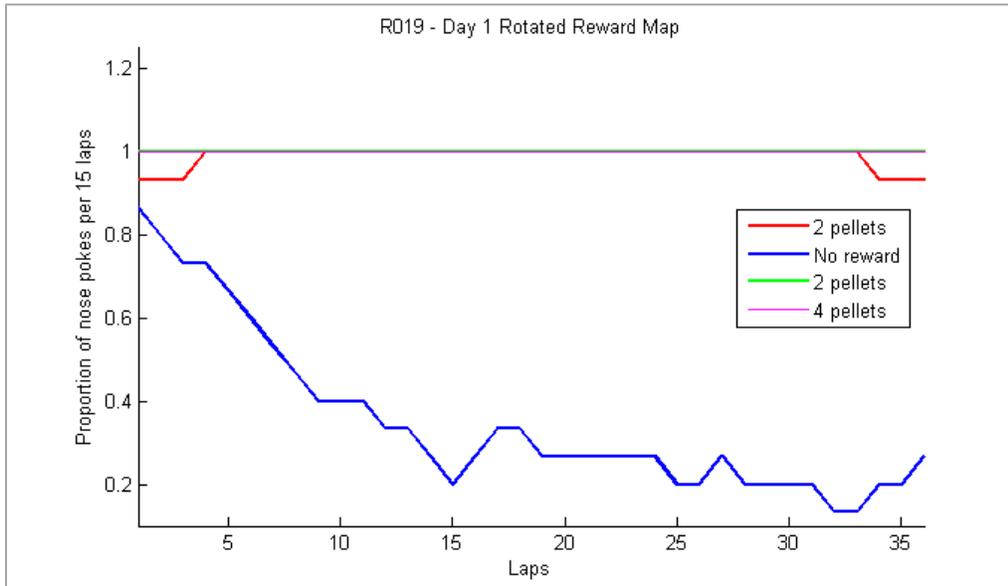


Figure 13: R019 Day 1 Rotated Reward Map – Within Session

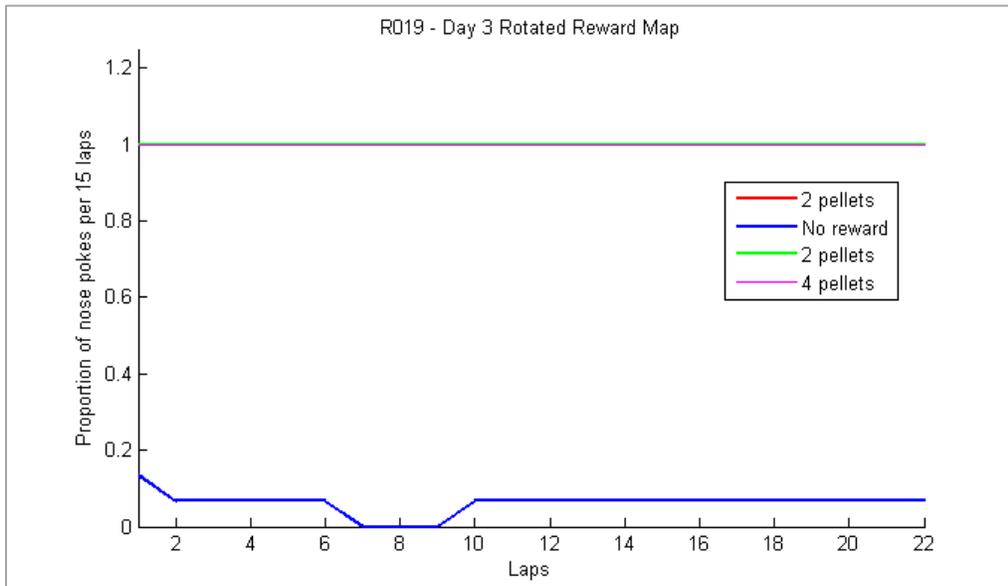


Figure 14: R019 Day 3 Rotated Reward Map – Within Session

R023

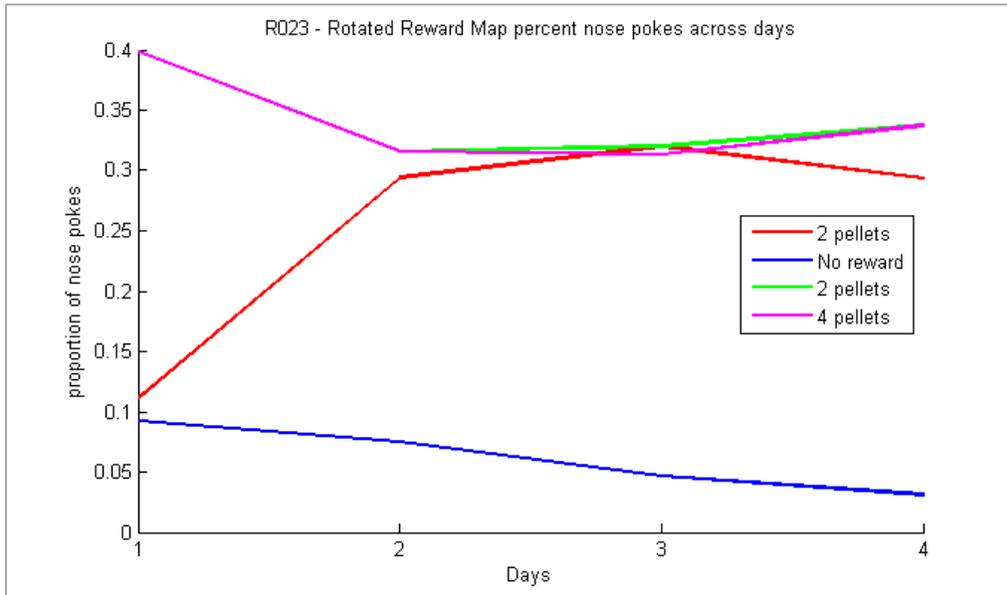


Figure 15: R023 Rotated Reward Map % nose poke across days

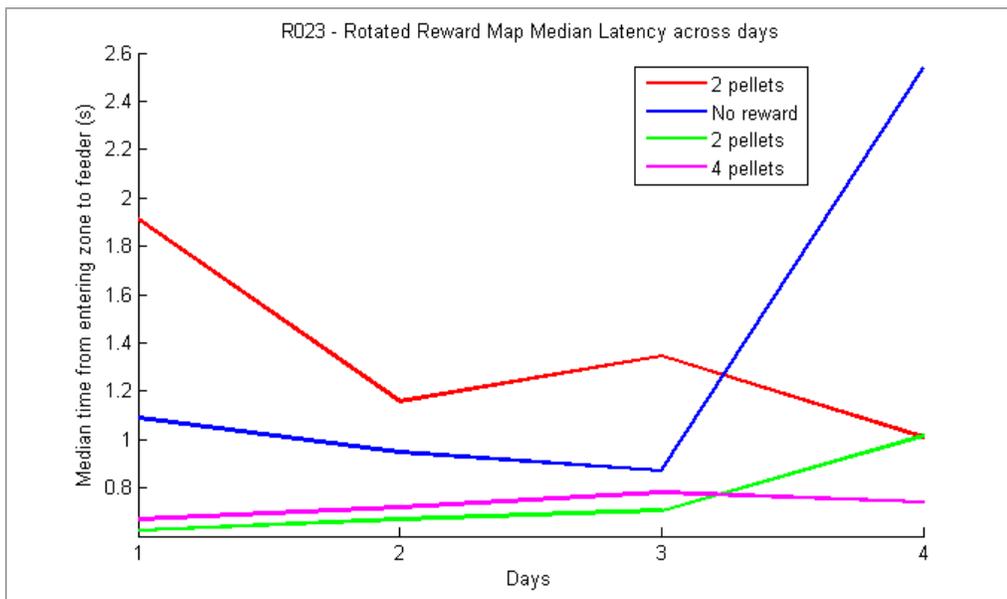


Figure 16: R023 Rotated Reward Map median latency across days

Within Session Learning Plots

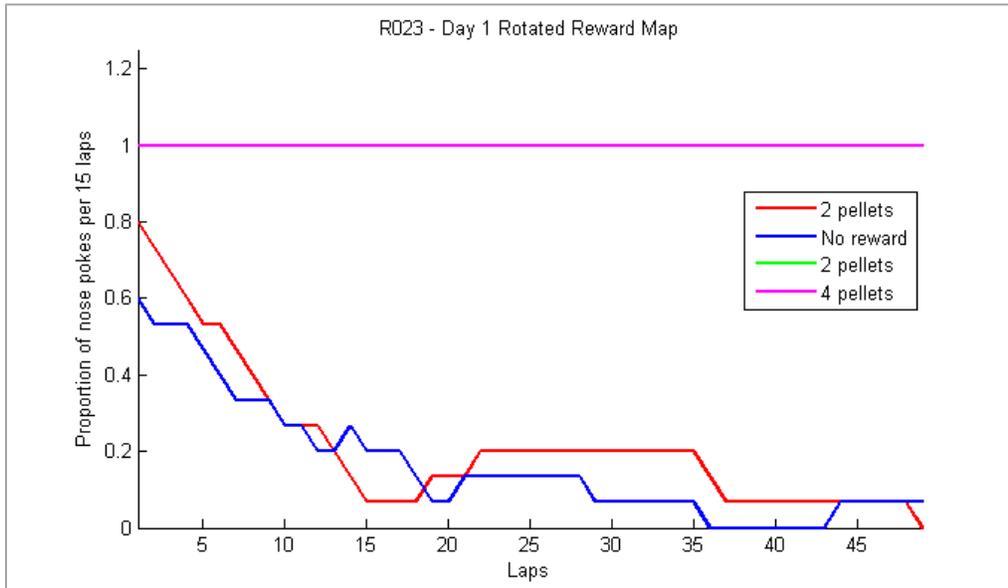


Figure 17: R023 Day 1 Rotated Reward Map – Within Session

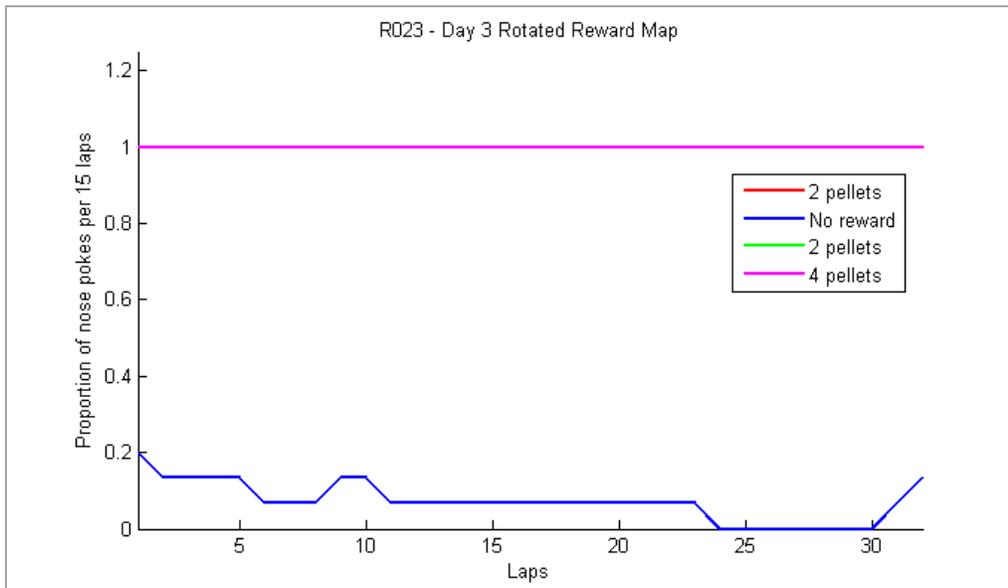


Figure 18: R023 Day 3 Rotated Reward Map – Within Session

Reward Map Reversal

R019

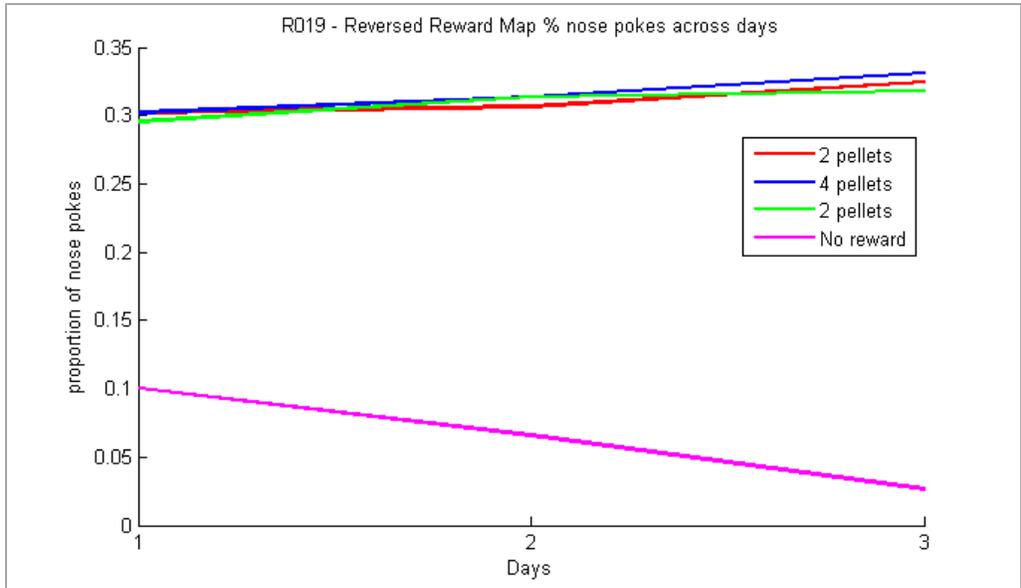


Figure 19: R019 Reversed Reward Map % nose pokes across days

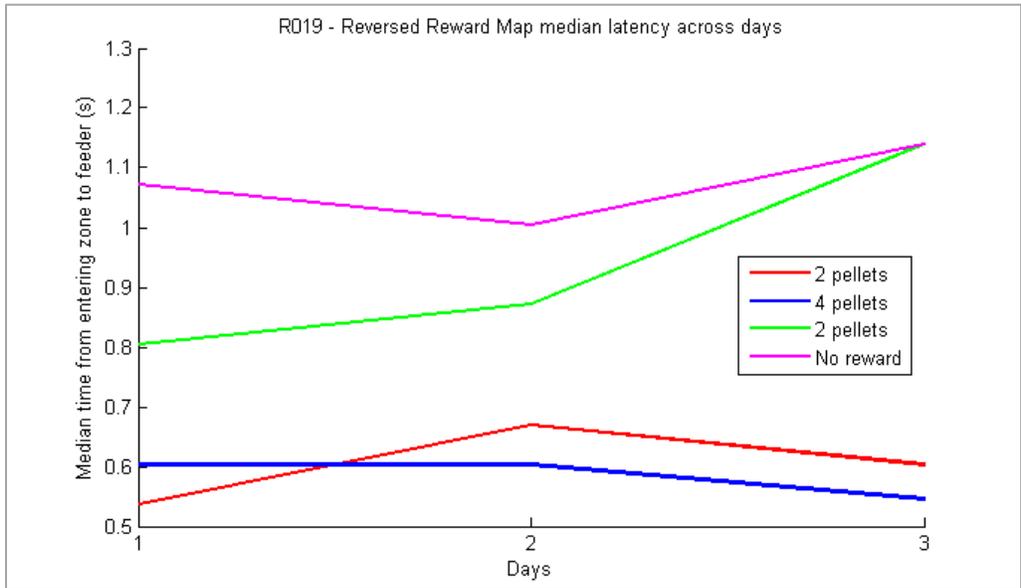


Figure 20: Reversed Reward Map median latency across days

R023

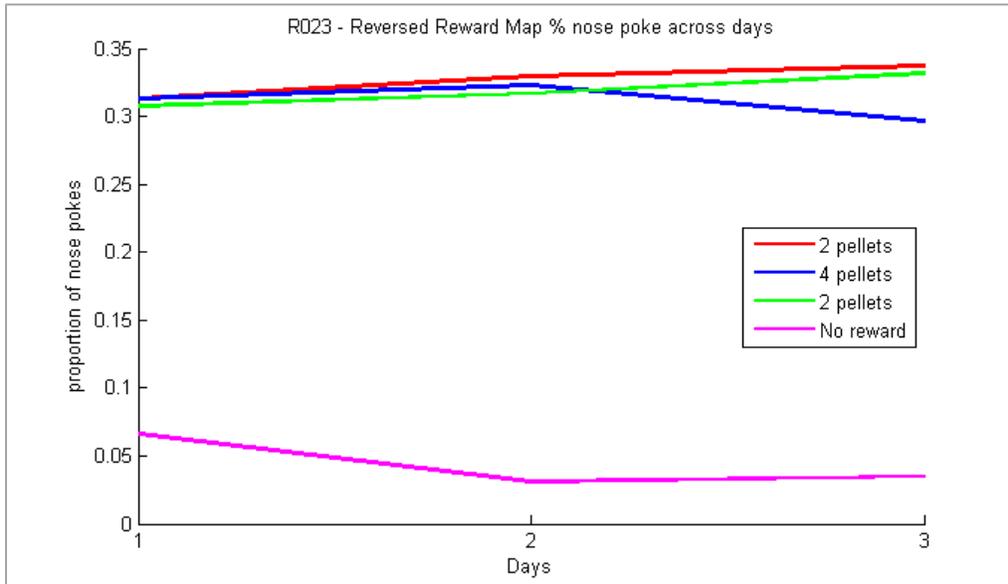


Figure 21: R023 Reversed Reward Map % nose poke across days

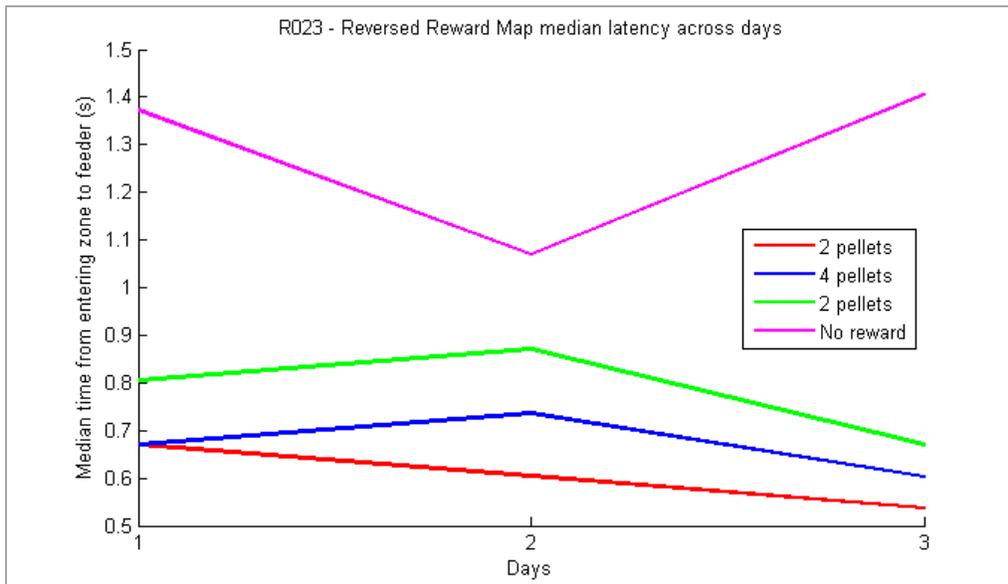


Figure 22: R023 Reversed Reward Map median latency across days

Reward Map Reversal and Switch

R019

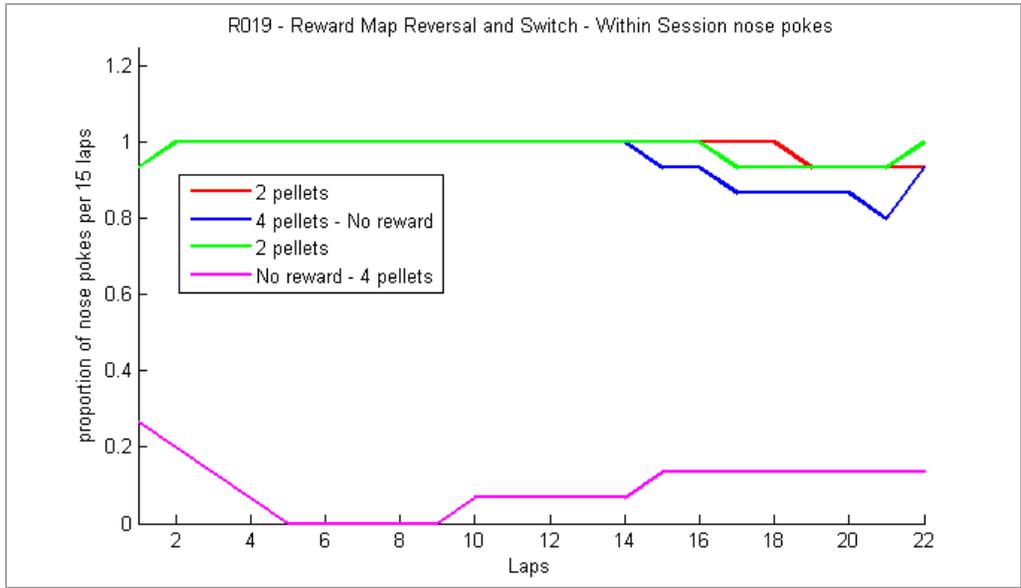


Figure 23: R019 Reversed Reward Map and Switch – Within Session

R023

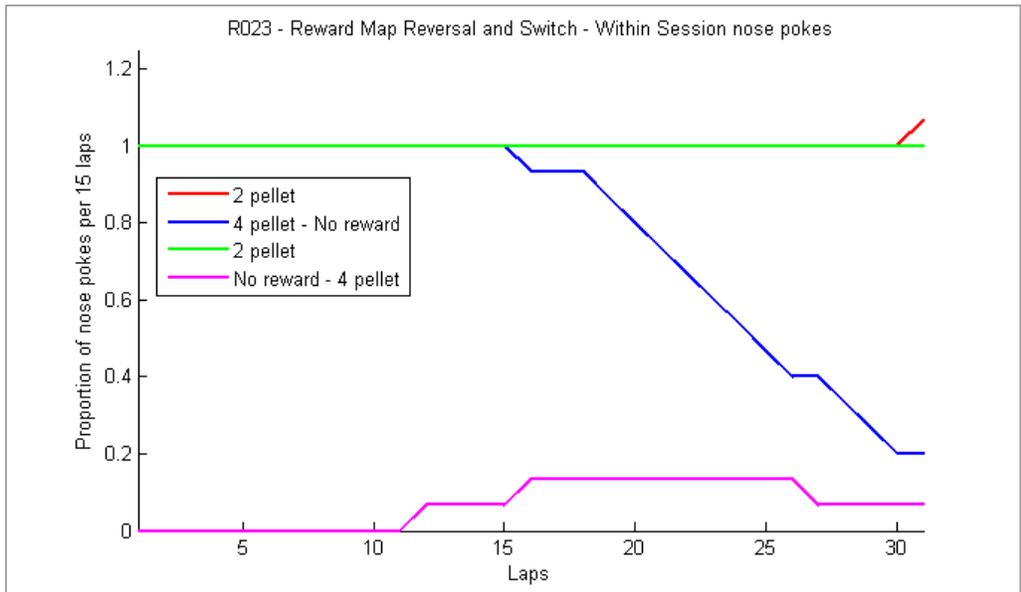


Figure 24: R023 Reversed Reward Map and Switch – Within Session

References

- Carboni, E., & Vacca, C., (2003). Condition Place Preference. *Methods in Molecular Medicine, Vol. 79: Drugs of Abuse: Neurological Reviews and Protocols*: 481 – 498.
- Crespi, L. P., (1942). Quantitative variation of incentive and performance in the white rat. *American Journal of Psychology, Vol. 55*, 467-517.
- Malhotra, S., Cross, R.W.A., & van der Meer, M.A.A., (2012). Theta phase precession beyond the hippocampus. *Reviews in the Neuroscience, Vol 23(1)*: 39 – 65.
- Miller, N.E. & Miles, W.B., (1935). Effect of caffeine on the running speed of hungry, satiated, and frustrated rats. *Journal of Comparative Psychology, Vol 20(3)*: 397 – 412.
- Perry, G., & Pianka, E.R., (1997). Animal foraging: past, present, and future. *Trends in Ecology and Evolution, Vol. 12*: 360 – 364.
- Rangel, A., Camerer, C., & Montague, P.R., (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience, Vol. 9*: 545 – 556.
- Redish, A.D., (1999). *Beyond the Cognitive Map: From place cells to episodic memory*. Cambridge, MA: MIT Press
- Reynolds, W.F. & Pavlik, W.B., (1960). Running speed as a function of deprivation period and reward magnitude. *Journal of Comparative and Physiological Psychology, Vol. 53(3)*: 615 – 618.

- Schmitzer-Torbert, N.C., & Redish, A.D., (2008). Task dependent encoding of space and events by striatal neurons is dependent on neural subtype. *Neuroscience*, Vol. 153: 349 – 360.
- Schönberg, T., Daw, N.D., Joel, D., & O’Doherty, J.P., (2007). Reinforcement Learning Signals in the Human Striatum Distinguish Learners from Nonlearners during Reward-Based Decision Making. *The Journal of Neuroscience*, Vol. 27(47): 12860 – 12867.
- Stephens, D.W., (2008). Decision ecology: Foraging and the ecology of animal decision making. *Cognitive, Affective, & Behavioural Neuroscience*, Vol. 8(4): 475 – 484.
- Tzschentke, T. M. (1998) Measuring reward with the conditioned place preference paradigm: a comprehensive review of drug effects, recent progress and new issues. *Progress in Neurobiology*, Vol. 56, 613–672.
- van der Meer, M.A.A., & Redish, A.D., (2009). Low and high gamma oscillations in rat ventral striatum have distinct relationships to behaviour, reward, and spiking activity on a learned spatial decision task. *Frontiers in Integrative Neuroscience*, Vol. 3(9) doi: 10.3389.
- van der Meer, M.A.A., & Redish, A.D., (2011). Theta Phase Precession in Rat Ventral Striatum Links Place and Reward Information. *The Journal of Neuroscience*, Vol. 31(8): 2843 – 2854.