Learning What to Expect and When to Expect It Involves Dissociable Neural Systems

Andrew R. Delamater, **Brandon Chen**, **Helen Nasser**, and **Karim Elayouby**

Brooklyn College and Graduate Center – City University of New York

Abstract

Two experiments with Long-Evans rats examined the potential independence of learning about different features of food reward, namely, "what" reward is to be expected and "when" it will occur. This was examined by investigating the effects of selective reward devaluation upon responding in an instrumental peak timing task in Experiment 1 and by exploring the effects of pre-training lesions targeting the basolateral amygdala (BLA) upon the selective reward devaluation effect and interval timing in a Pavlovian peak timing task in Experiment 2. In both tasks, two stimuli, each 60 s long, signaled that qualitatively distinct rewards (different flavored food pellets) could occur after 20 s. Responding on non-rewarded probe trials displayed the characteristic peak timing function with mean responding gradually increasing and peaking at approximately 20 s before more gradually declining thereafter. One of the rewards was then independently paired repeatedly with LiCl injections in order to devalue it whereas the other reward was unpaired with these injections. In a final set of test sessions in which both stimuli were presented without rewards, it was observed that responding was selectively reduced in the presence of the stimulus signaling the devalued reward compared to the stimulus signaling the still valued reward. Moreover, the timing function was mostly unaltered by this devaluation manipulation. Experiment 2 showed that pre-training BLA lesions abolished this selective reward devaluation effect, but it had no impact on peak timing functions shown by the two stimuli. It appears from these data that learning about "what" and "when" features of reward may entail separate underlying neural systems.

1. Introduction

One important problem in the study of associative learning has been specifying the nature of the systems responsible for learning about and encoding different aspects of reward. For instance, Konorski (1967; also Wagner & Brandon, 1989) speculated that in Pavlovian learning a predictive stimulus can enter into separate associations with sensory and emotional aspects of reward. More recent work has established that separate neural systems underlies these two forms of learning (e.g., Balleine & Killcross, 2006). However, there are

other aspects of reward that an organism can encode that can factor into learning (Delamater, 2012; Delamater & Oakeshott, 2007). Since Pavlov (1927) we have known that conditioned responding is temporally organized, and, indeed, there has been a considerable amount of research examining processes involved in interval timing (e.g., Addyman, French, & Thomas, 2016; Buhusi & Oprisan, 2013; De Corte & Matell, 2016; Kirkpatrick, 2014; Matell & Meck, 2004). We have, elsewhere, suggested that learning about the specific sensory properties of a reward, i.e., learning "what" it is, and learning to time its arrival, i.e., learning "when" it will occur, could entail two separate underlying learning systems with distinct, though overlapping, neural processes governing them (Delamater, Desouza, Rivkin, & Derman, 2014). However, recent research has also suggested that the occurrence of rewards in time are fundamentally encoded and that responding is a decision-based process based on various computations performed on the raw data stored within a temporal memory structure (e.g., Balsam & Gallistel, 2009; Gallistel & Balsam, 2014). Such an approach may predict that learning about "what" the reward is and "when" it will occur may recruit similar underlying psychological and neural processes that are not so distinct.

Research devoted to examining this issue has been scarce. If learning what a reward is and when it will occur entail separate systems, then variables affecting one should have little effect on the other. Balsam and his colleagues (Drew, Zupan, Cooke, Couvillon, & Balsam, 2005; Ohyama, Gibbon, Deich, & Balsam, 1999) showed that animals learn to time the arrival of reward from the onset of conditioning and retain this information throughout extinction as the conditioned response dissipates. These results suggest independence of the motivation to respond and learning to time the arrival of reward. However, a somewhat different conclusion comes from studies that have asked if interval timing effects are impacted by various manipulations of reward value. One popular method for assessing interval timing is known as the "peak procedure" (Roberts, 1981). In this task, animals are trained to press a lever for food reward after a specific amount of time has elapsed since the onset of a stimulus. Critically, responding is also assessed on non-rewarded probe trials in which the duration of the stimulus is usually 2-3 times that of rewarded trials. Responding averaged over these test trials gradually increases and peaks at a point centered close to the actual reward time before then falling off somewhat more gradually. This distribution of responding is taken as a strong indication that the animal has encoded when the reward is expected to occur.

There have been several reports suggesting that variables affecting reward value may have some impact on peak timing functions in this task, a result that would suggest the two processes to be interdependent (see Kirkpatrick, 2014 for a review). For instance, Roberts (1981) demonstrated that tests conducted while the animal is satiated can shift the entire peak response function to the right compared to when the animal is tested food deprived (also Galtress & Kirkpatrick, 2009; Galtress, Marshall, & Kirkpatrick, 2012). Interpretation of this result is complicated, though, by the additional observation that training the rat while sated and testing hungry also shifts the distribution to the right, as opposed to having the symmetrically opposite effect as would be expected if reward value interacts with reward timing in any straightforward way (Galtress et al, 2012).

Additional findings have shown a dissociation between manipulations affecting overall response rate and response timing. For instance, Meck (2006) found that caudate putamen lesions undermined response timing in a dual-peak procedure without undermining differences in overall levels of responding, whereas nucleus accumbens shell lesions eliminated peak rate response differences in this task without impacting temporal control. These data support the view that factors affecting response rate and timing are dissociable. Similarly, Ohyama et al. (2000) also showed that systemic injection of a dopamine antagonist immediately suppressed the overall rate of responding without immediately affecting temporal control in a peak timing task. Together these data suggest dissociations between the processes mediating response rate and temporal control in this task.

Another approach has been to examine the role of reward magnitude on interval timing in the peak procedure. It has been observed that peak distributions produced with small rewards are right-shifted to those produced with large rewards (Galtress & Kirkpatrick, 2009), though sometimes other aspects of the peak distribution such as initiation times are more sensitive than the actual peak of the distribution (Balci et al., 2010; Ludvig, Balci, & Spetch, 2011). The result suggests that timing itself may depend upon the magnitude of the anticipated reward.

Perhaps the most relevant type of experiment for the question of whether reward identity ("what") and temporal encoding ("when") involve dissociable or interactive systems uses the reward devaluation task. In this task, it is determined if devaluing a reward following a conditioning phase affects stimulus control when the stimulus is tested under extinction conditions. The logic of this test is that if a stimulus has associated with some specific feature of the reward and that feature has then independently been devalued (in the absence of the stimulus), subsequent testing under extinction conditions should reveal a decrement in performance. The effect is assumed to reflect the fact that the stimulus is capable of activating a specific representation of the now-devalued outcome (e.g., Delamater & LoLordo, 1991; Rozeboom, 1958). To date, there have only been two studies using this task in connection with a peak timing procedure in order to separately assess reward identity and timing, and the results have not been entirely consistent with one another. Galtress and Kirkpatrick (2009) observed that compared to a baseline training phase, the peak function was shifted to the right when testing occurred following reward devaluation training (where food intake was independently paired with a LiCl injection designed to establish an aversion to the food). Delamater et al (2014), however, reported that following devaluation the motivation to respond was reduced, but the timing of the peak function was not affected.

There were several procedural differences between the tasks used by Galtress and Kirkpatrick (2009) and Delamater et al (2014) that could be crucial. Delamater et al (2014) trained their animals in a Pavlovian task with two separate stimuli each paired with a qualitatively different reward, and then devalued one of these rewards through extensive selective devaluation training (one reward was paired repeatedly, and the other unpaired, with LiCl). Finally, the two stimuli were tested at the same time under extinction conditions. In contrast, Galtress and Kirkpatrick (2009) trained their rats in an instrumental task with a single stimulus and reward pair. They then tested their animals under extinction conditions after a limited amount of reward devaluation training (one or two food-LiCl pairings). Peak

functions shifted to the right in the Galtress and Kirkpatrick (2009) study, but this assessment depended upon a comparison of responding following devaluation to baseline responding during training sessions that also included rewarded trials. In the Delamater et al (2014) study, responding following reward devaluation was assessed in the same extinction session to one stimulus whose associated outcome had been devalued to another whose outcome had not been devalued. Under those conditions we did not find any evidence to suggest that reward devaluation impacted the peak timing function.

In order to further investigate the potential independence or interdependence of reward identity and timing the present studies extended these studies in two ways. First, because we think it is important to assess timing functions at the same time following a selective reward devaluation manipulation, we assessed the generality of our findings in a Pavlovian task to an instrumental peak timing task. Second, we assessed, perhaps more directly, the independence of "what" and "when" learning by assessing the effects of region-specific brain lesions on reward timing and selective devaluation effects. Other research has shown, convincingly, that the basolateral amygdala (BLA) is necessary for rats to encode sensory aspects of reward in Pavlovian devaluation (and other) tasks (e.g., Blundell, Hall, & Killcross, 2001; Corbit, 2005; Hatfield, Han, Conley, Gallagher, & Holland, 1996; Johnson, Gallagher, & Holland, 2009). Here, we ask whether pre-training BLA lesions might affect selective reward devaluation and reward timing performance differentially. If learning to encode the "what" and "when" of reward entails distinct systems, and "what" learning depends upon a functioning BLA, then such lesions should disrupt the selective reward devaluation effect but leave peak timing functions intact.

2. Methods

Experiment 1

2.1. Subjects—Subjects were 32 experimentally naïve Long-Evans rats, male (n=16) and female (n=16), that were bred at Brooklyn College and derived from Charles River laboratories. The study was run in two identical replications (n=16 per replication). Males' free-feeding weights ranged between 338 – 376 g in replication 1 and 489 – 608 g in replication 2, and females' weights ranged between 223 – 256 g in replication 1 and 240 – 305 g in replication 2. They were maintained at 85% of their free feeding weights in a colony room on a 14 hr LD cycle and housed in groups of 3-4 animals per cage in standard transparent plastic tub cages ($17 \times 8.5 \times 8$ in) with wood chip bedding. All experimental procedures were performed during the light phase of their light/dark cycle at the same time of day. All procedures were performed in accordance with the approved guidelines of the IACUC of Brooklyn College.

2.2. Apparatus—Eight identical operant conditioning chambers (BRS-Foringer, RC series) housed in a Med Associates sound-, and light-resistant shell were used for behavioral training and testing. The chambers measured 30.5 cm $\times$ 24.0 cm $\times$ 25.0 cm, and consisted of front and back walls were made of aluminum and remaining walls and ceiling were made of transparent Plexiglas. The chamber floor was made up of 0.60 cm diameter stainless steel rods spaced 2.0 cm apart. The recessed food magazine was located 1.2 cm centered above

the grid floor and measured $3.0 \times 3.6 \times 2.0$ cm. Two feeders were connected to this food magazine and could deliver different 45-mg food pellets (BioServ Purified (F0021), TestDiet (5TUM) that served as the reinforcers in these studies. Each reinforcement consisted of 2 individual pellet deliveries spaced 0.5 s apart. Head entries within the magazine were recorded by means of an infrared detector and emitter (Med Associates ENV-303HDA) enabling automatic recording. To the left and right of the magazine were two protruding levers that were covered with sheet metal when not in use. Mounted on the top back end of the operant chamber were two 28 volt, 2.8 W light bulbs covered by a translucent plastic sheet angled between the ceiling and the top portion of the rear aluminum wall. Activation of these rear house lights allowed for a white light stimulus to flash with equal on-off pulse durations at a frequency of 2 pulses per second. A Med Associates sonalert module (ENV-223AM) was mounted centrally on the outer side of the chamber ceiling to present a 2900 Hz tone stimulus, 6 dB above a background level of 74 dB (C weighting). A fan attached to the outer shell provided cross-ventilation within the shell as well as background noise. All experimental events were controlled and recorded automatically by an Intel-based PC and Med Associates interfacing equipment located within the same room.

2.3. Behavioral Procedures—Rats were initially magazine trained with both types of pellet rewards for two days. Both response levers were covered, and on each day one 20-min training session with one outcome was followed by a second 20-min session with the other outcome. The order of outcome presentations was counterbalanced within and across days. Each training session contained 20 pellets of one type delivered according to a random time 60-s schedule of reinforcement.

All rats were then trained to lever press on a continuous reinforcement schedule until 50 outcomes of each type were earned. Only the left lever was used in the study (and the other lever was permanently covered). Animals were first trained with one type of pellet reward (counterbalanced) and then again in another session with the other pellet reward.

Following initial lever training, all animals were given 20 days of training on the peak interval timing task. On reinforced trials a stimulus (flash or tone) was presented for 60 s, but only the first response occurring after 20 s was reinforced. The flashing light and tone stimuli signaled that different reward types (counterbalanced) could be earned. On non-reinforced probe trials the stimulus was presented for 60 s but reinforcement was withheld. Each training session contained 12 reinforced and 4 non-reinforced probe trials of each stimulus for a total of 32 trials per session. The inter-trial interval was increased from 40 s in the first two days to 80 s on the next two days to 120 s for the remaining days of training.

The reward devaluation procedure consisted of 6 cycles in which one of the rewards was paired with an IP injection of lithium chloride (LiCl, 0.3M, 1.5% body weight) on one day, and the other was presented on the next day without an injection. In each session, one of the rewards was delivered non-contingently as in magazine training. Both response levers were covered and the flash and tone stimuli were not presented throughout this phase. On days in which the animals were presented the devalued reward they were given LiCl injections immediately following the 20-min session. Rats were fed their daily food ration 1 hour after the LiCl treatment. On days in which the animals were presented with the non-devalued

reward, rats were given a 20-minute magazine training session after which they were immediately returned to their home cage.

There were three test sessions conducted under extinction conditions. Each session consisted of 8 non-reinforced probe trials of each 60-s stimulus (flash and tone). Lever responding was assessed in 1-s bins prior to and during each stimulus.

Experiment 2

2.4. Subjects—Subjects were 24 experimentally naïve Long-Evans rats (12 male and 12 female) housed and maintained as in Experiment 1. The free-feeding weights ranged from 422 – 514 g for the males and 283 – 346 g for the females.

2.5. Apparatus—The same apparatus was used as in Experiment 1.

2.6. Behavioral Procedures—The same general procedures were used as in Experiment 1 with the following exceptions. This experiment used a Pavlovian task, so the rats were not trained to lever press and both response levers were permanently covered throughout the study. In addition, there were 36 days of Pavlovian peak interval training, and during each reinforced trial the appropriate reward was delivered 20 s after stimulus onset. All other parameters were the same as in Experiment 1. The main data of interest consisted of examining magazine response rates in each 1-s bin prior to and during the stimuli on non-reinforced trials.

2.7. Surgical Procedures—Rats in Experiment 2 were given pre-training lesions targeting the basolateral amygdala. They were anesthetized with isoflurane (VetOne, ID, Cat. No. 13985-528-60) and placed in a stereotaxic frame (Stoelting). We made an incision into the scalp to expose the skull and adjusted the incisor bar to the level head position, using bregma and lambda as reference points. We intracranially injected rats using level-head coordinates that we derived from the stereotaxic atlas of Paxinos and Watson (2005): for rats under 400 g we injected at anteroposterior, -2.7 and -3.0 (two sites); lateral, ±5.2; ventral, -8.2 mm, for rats above 400 g we injected at anteroposterior, -2.7 and -3.0 (two sites); lateral, ±5.2; ventral, -8.7 mm. For rats in the BLA lesion group (n=16) we injected a total of 0.6 μL bilaterally (0.15 μL/site) of *N*-Methyl-D-aspartic acid (Tocris, PA, Cat. No. 0114,) in phosphate buffered saline (PBS) at a concentration of 20 μg/μL. We conducted the same surgical procedure for rats in Group Sham (n=8), however we injected PBS alone. We delivered all intracranial injections with a 2 μL Hamilton syringe (Hamilton, NV, cat. No. 1701,) at a rate of 0.1 μL/min and the needle tip was left in place for 5 min to allow diffusion of the toxin. After the final injection we covered the skull holes with bone wax and we sutured the incision closed. At the conclusion of the surgery we then injected rats with 0.09% physiological saline (s.c.), penicillin (100,000 U/kg, i.m, PenJect+B, Henry Schien, NY, Cat. No. 023203.) and buprenorphrine hydrochloride (0.05mg/kg, s.c., Buprenex, Butler Schien, OH, Cat. No.031919).

2.8. Histological Procedures—At the end of the experiment, we injected rats with a lethal barbiturate overdose and perfused transcardially with 0.9% saline followed by 10% formalin solution. We stored the brains in 10% formalin solution for 48 hr before being

transferred to a 30% sucrose solution. We allowed a few days to pass for brains to sink in the sucrose solution, before sectioning the brains. We sectioned coronal slices at 40 μm throughout the region of the BLA, and then mounted the slices on glass slides coated in a 4% gelatin solution. We stained the slices with cresyl violet. We examined slices for extent of lesion under a microscope and assessed placement of lesions with reference to the stereotaxic atlas of Paxinos and Watson (2005) and we compared lesioned brains relative to sham brains using several defining features including: gross morphological changes such as holes and tissue collapse; position and extent of gliosis and scarring and signs of neuronal cell body shriveling and loss.

3. Results

3.1. Statistical Analyses

The data were evaluated using several different indices of temporal control. The rate of responding in 1-s intervals during the stimuli (and pre stimuli periods) was recorded and compared between stimuli signaling devalued or non-devalued outcomes with dependent samples t tests. Temporal control was assessed by first transforming responding in each 1-s interval into a proportion of total response score. Cumulative response functions were then calculated for each animal and group averages are presented for an impression of how responding unfolded across the trial. The data points corresponding to when 25%, 50%, and 75% of the total responses made were calculated from these cumulative response functions. In addition, three other response timing measures were assessed. These included a measure of (1) the peak rate of responding (the maximum rate of responding shown in any 1-s interval), (2) the peak interval (the interval in which the peak rate was displayed), and (3) a "timing ratio" (the peak rate of responding divided by the overall rate during the stimulus). Following the procedures recommended by Rodger (1974), between stimulus statistical assessments were conducted using repeated measures ANOVAs performed on each group using a pooled error term, in addition to a between group main effects test. A Type 1 error rate of 0.05 was adopted as our test criterion in all cases.

3.2. Experiment 1 results

By the end of the training phase rats displayed normal peak functions to the two stimuli, and performance was matched in terms of responding to the stimulus whose associated outcome would subsequently be devalued or not. By the end of the reward devaluation phase all rats ceased consuming the food pellet that was paired with LiCl injections, but still consumed all of the pellets not paired with LiCl. The data of main interest concerned the results from the test phase, conducted under extinction conditions. Figure 1A shows responding averaged across test trials and across test sessions 1-3. Because there were no differences as a function of sex the data was collapsed across this variable as well. Lever press responding in the 20-s period prior to stimulus onset was uniformly low, but then it steadily increased once the stimuli were presented and peaked close to the previously reinforced interval of 20 s before responding slowly decreased thereafter. Of most importance, overall response levels were lower to the stimulus whose associated outcome had been devalued, $t(31) = 3.31$, $p = 0.002$. In order to assess whether the temporal distribution of responding differed between the two stimuli, i.e., whether the distribution shifted as a function of reward devaluation, the data in

each 1-s interval was expressed as a proportion of total responses for each animal and then the functions displaying the cumulative proportion of total responses across the interval were compared. Figure 1B shows that although overall lever press rates were lower to the stimulus whose associated outcome had been devalued, responding in the presence of both stimuli emerged similarly across time. The two cumulative response functions overlapped virtually identically. This indicates that the temporal distribution of responding to the two stimuli did not differ.

Additional analyses were performed on these data to corroborate the above findings. The peak rate of responding and the peak interval in which that rate occurred was determined for each subject in the presence of each stimulus. These data are shown in Figure 2A and 2B, respectively. The peak rate of responding was lower in the presence of the stimulus whose associated outcome had been devalued, t(31) = 3.69, p = 0.001, and the peak intervals did not differ between the two stimuli, t(31) = 1.14, p = 0.263. These data are consistent with the findings reported by Delamater et al. (2014) that used a Pavlovian peak training procedure, and extend those results to an instrumental task. They provide clear evidence that the outcome's identity and time of occurrence are encoded by the stimuli and that reward devaluation decreases the motivation to respond but has no impact on timing per se.

3.3. Experiment 2 results

3.3.1. Histological results—BLA lesions were considered valid if there was more than 75% neuronal loss and significant gliosis in target areas. There were 7 rats that met this criterion of the 16 given intracranial NMDA injections. Among these rats, 82.2% of the BLA was effectively lesioned with damage extending into 43.3% of the neighboring central amygdala. Characteristic damage was observed within the posterior BLA, lateral amygdala and the capsular central amygdala. The anterior and medial amygdala was largely spared. Figure 3 illustrates the maximum and minimum extent of the effective lesions in Group BLA. The remaining 9 animals given NMDA during surgery displayed either lesions outside of BLA region or had only unilateral lesions. Among these rats, only 16.1% of the BLA was judged to have been effectively lesioned with damage extending to 13.5% of the central amygdala.

3.3.2. Behavioral results—Since the patterns of data did not differ in animals given sham lesions and those given "off-target" lesions (i.e., judged by histological analysis to have a mostly intact BLA) the data were combined across these two subgroups for the purposes of the statistical analyses. This was treated as a separate control group (n=17). The effective lesion group (n=7) and the controls both acquired the peak timing task equivalently. Figure 4A and 4B shows responding on non-reinforced probe trials in 1-s intervals during pre-CS and CS periods averaged over the final 8 sessions of training. Responding was low in pre-CS periods and then rapidly increased upon stimulus onset at around 20 s in the presence of both CSs (i.e., the ones whose associated outcomes were scheduled to be devalued or not). Responding then gradually declined throughout the latter part of the stimuli. The overall rates of responding during the CSs did not differ between the groups, t(22) < 1, p > 0.05. In addition, although the peak rate of responding was numerically higher in the controls than in the lesion group (1.1 +/- 0.17 and 0.98 +/- 0.21 responses/s, respectively, +/-

SEM), the groups did not significantly differ, $t(22) < 1$, $p > 0.05$. The mean peak interval also did not differ between the two groups, $t(22) < 1$, $p > 0.05$ (means for Group BLA and Group Control, respectively, were 11.8 +/- 3.3 and 13.2 +/- 1.2). In order to examine the temporal patterning of responding more thoroughly, we assessed the cumulative proportion of total responding across time in both groups. Figure 4C shows that the two groups' cumulative response functions increased over the trial at comparable rates. This indicates that responses were distributed similarly across time in the two groups. The mean intervals in which 25%, 50%, and 75% of total responses were emitted did not differ between the groups, $t(22)s = 0.63$, 1.36, 1.80, $ps > 0.09$, respectively.

The data of most interest came from the test phase. The response rate distributions for the control animals and the effective BLA lesion groups are displayed in Figure 5A and 5B. Both groups displayed the characteristic peak function with responding steadily increasing above pre-CS levels and peaking near the expected time of reward and then gradually declining thereafter. However, whereas the sham and off-target lesion animals displayed a normal selective reward devaluation effect, the BLA lesioned animals did not. These data were analyzed by assessing overall responding (averaged across the entire trial) in the presence of the two CSs using a pooled error term (MSE = 24.288). Group Control displayed significantly lower responding to the CS whose associated outcome had been devalued, $F(1,22) = 11.60$, $p < 0.05$, whereas Group BLA animals did not, $F(1,22) = 0.00$, $p > 0.05$. Additional analyses were performed to assess the temporal distribution of responding in a more fine-grained manner.

The cumulative proportion of total responses was computed and compared between the two groups to determine if BLA lesions changed the temporal patterning of responding during the devaluation test sessions. Figure 5C shows that responding emerged across both stimuli very similarly for the two groups, although Group BLA tended to respond somewhat earlier during the stimuli (especially the CS signaling the devalued outcome). Figure 5D shows the intervals in which 25%, 50%, and 75% of the responses occurred to the CSs in the two groups. These data were analyzed by performing repeated measures ANOVAs across the stimulus factor (CS dev, CS not dev) for each group using pooled error terms at each level of the cumulative response factor (MSE = 7.106, 3.427, and 8.334 for 25%, 50%, and 75% respectively). In addition, a main effect of Group was also assessed at each level of the cumulative response factor. The only difference revealed by this analysis was that the mean interval in which 50% of the total responses were made was lower to the CS whose outcome had been devalued than to the CS whose outcome had not been devalued in Group BLA, $F(1,22) = 5.67$, $p < 0.05$. Nonetheless, the two groups did not differ at these three points along the cumulative response function. These data reveal that BLA lesioned and control animals generally responded similarly across time.

In order to assess between-group differences in the maximum of the response timing curve more fully, peak rate and peak intervals were calculated for each stimulus (Figure 5E and 5F). Whereas control rats displayed a lower peak rate of magazine responding to the CS whose associated outcome had been devalued than to the CS whose associated outcome had not been devalued, $F(1,22) = 5.85$, $p < 0.05$, lesioned rats failed to show this difference, $F(1,22) = 0.17$, $p > 0.05$ (pooled MSE = 0.037). In this study (unlike in Experiment 1),

control rats expressed this peak rate to the CS whose associated outcome had been devalued in an earlier interval than to the control CS, $F(1,22) = 4.81$, $p < 0.05$, whereas lesioned animals did not show this difference, $F(1,22) = .09$, $p > 0.05$ (pooled MSE = 78.829). In addition, the peak interval, combined across CSs, did not differ between the two groups, $F(1,22) = 0.06$, $p > 0.05$ (MSE = 71.705).

An additional analysis was performed on another measure of timing. For this measure, the maximum rate of responding within the stimulus was divided by the rate of responding averaged across the entire stimulus (Figure 5G). One subject in Group BLA had to be eliminated from this analysis because its overall response level was 0.25 responses per minute for CS devalued and 2.25 for CS not devalued. These low values dramatically skewed this subject's ratio scores compared to all other subjects. The analysis revealed a lower ratio in Group Control for the CS whose associated outcome had been devalued compared to the other CS, $F(1, 21) = 4.43$, $p < 0.05$ (pooled MSE = 0.267). This reflects the fact that these subjects responded less to the CS whose outcome had been devalued. However, this difference was not observed in Group BLA, $F(1,21) = 0.93$, and the two groups did not differ overall, $F(1,21) = 1.07$, $p < 0.05$ (MSE = 0.584).

Overall, the data from this study reveals that BLA lesions effectively eliminated the reward devaluation effect in this Pavlovian task, while largely sparing interval timing. There was some indication that responding emerged somewhat sooner in lesioned animals to the CS whose outcome had been devalued, but there were no further between group differences with any of the other indices of timing used here. Furthermore, although not apparent in the moment by moment response rate or cumulative proportion data, the peak interval occurred sooner in controls in the presence of the CS whose associated outcome had been devalued. While this effect was statistically reliable it is difficult to interpret because the entire temporal distribution of responding was very similar to both stimuli (as indicated in Figure 5A) and because the effect was not observed in Experiment 1.

4. Discussion

The main findings of the present studies were (1) that reward devaluation selectively decreased responding in the presence of stimuli signaling the devalued reward relative to stimuli signaling a non-devalued reward in both instrumental and Pavlovian peak procedures, (2) that the temporal organization of responding was not appreciably affected by reward devaluation, (3) that pre-training BLA lesions eliminated the selective reward devaluation effect, and (4) that BLA lesions had little to no impact on the temporal organization of behavior across the stimulus in the Pavlovian peak timing task. These findings add to our growing knowledge of how the brain encodes and learns about different features of reward.

The present studies were motivated by an interest in determining if learning the "what" and "when" aspects of reward might entail distinct systems. Our logic has been to examine whether variables that affect one aspect of learning might not impact the other. We explored this in two ways. We determined if the effects of reward devaluation might selectively diminish the motivation to respond to stimuli associated with the devalued reward, without

affecting the timing of responding. Second, we explored the effects of BLA lesions on not only the reward devaluation effect, but on reward timing as well. Our results are mostly consistent with the view that manipulations affecting behavior controlled by a representation of what the reward is selectively affects behavior controlled by that specific reward encoding, and have little to no impact on behavior controlled by a representation of when the reward is due to arrive. We observed this in both of our studies by showing that reward devaluation diminished responding selectively to the stimulus whose associated outcome had been devalued without affecting temporal control, and that BLA lesions undermined the reward devaluation effect without also affecting the temporal organization of responding.

Our results are both consistent and inconsistent with various findings in the literature. Our reward devaluation effect resembles results obtained by Ohyama et al (1999). They reported that following Pavlovian conditioning with a consistent CS-US interval in ring doves, responding peaked near the US time during extinction sessions in which the CS duration was increased to assess a peak interval response function. Of particular interest was the additional finding in this study that over repeated test sessions the peak of the response distributions did not change while overall levels of responding steadily decreased as a result of extinction. In our situation, the effect of reward devaluation was to selectively decrease responding to the stimulus signaling the devalued reward, and, as was true in the Ohyama et al (1999) study peak responding did not change. More recent research with rats, however, shows that the peak function during extinction can recalibrate to an increased CS duration (Drew, Walsh, & Balsam, 2017), as though the animals had learned during training that the US occurs after the CS is terminated and when a longer duration CS is tested in extinction the animals adjust their temporal expectation to match the new extended CS duration. This sort of process is unlikely to have played a role in our situation since we used the same long duration CS throughout the study.

Our results are, to some extent, inconsistent with Galtress and Kirkpatrick (2009) who found that following reward devaluation peak functions shifted to the right. Our two studies are difficult to directly compare because of various procedural differences. Nonetheless, we obtained little or no effect of reward devaluation on peak functions in both Pavlovian and instrumental procedures, while Galtress and Kirkpatrick (2009) did find such an effect using an instrumental task. Perhaps more importantly, we assessed the impact of reward devaluation on peak responding in the same test sessions to stimuli whose associated outcomes had been devalued or not. Galtress and Kirkpatrick (2009) assessed the effects of reward devaluation by comparing responding following a devaluation treatment to responding in baseline training sessions. We think this comparison is less than ideal because the conditions of testing are not identical to those of baseline sessions. The latter included reinforced trials intermixed with non-reinforced probe trials, whereas test sessions only included non-reinforced probe trials. Furthermore, there was no non-devalued control used in the Galtress and Kirkpatrick (2009) study, aside from the baseline comparison, and so it is not known whether response distributions may have changed spontaneously over time in the experiment or simply as a function of exposure to LiCl injections. We overcame these problems in the present studies by using an internal control stimulus whose associated outcome had not been devalued, and by assessing post-devaluation responding to these stimuli in the same test sessions under the same test conditions. Any non-specific changes in

response functions would apply equally to both our target and control stimuli. Under these conditions we found little evidence suggesting that devaluation impacted the timing functions to these stimuli. It is true that in Experiment 2 we observed the peak interval to the devalued CS to occur somewhat earlier than to the non-devalued CS, but the cumulative timing functions did not differ across time. Furthermore, we did not see a difference in Experiment 1, so the generality of this result is unclear. If anything, though, this minor effect is in the opposite direction to that reported by Galtress and Kirkpatrick (2009). Another noteworthy difference between our studies is that we gave a much more extensive amount of reward devaluation training (5 devaluation cycles) because we found that the aversions were weak or non-existent after only 1 or 2 food-LiCl pairings. Future work may be needed to resolve any discrepancies between our studies, but it is worth noting that because our studies used very different procedures these discrepancies do not constitute a failure to replicate.

Our findings are consistent with those in the literature showing that pretraining BLA lesions undermine the US devaluation effect and other effects that depend upon encoding specific features of reward (e.g., Blundell et al., 2001; Corbit & Balleine, 2005; Hatfield et al., 1996; Johnson et al., 2009). In Experiment 2 reported here, it may be noted that our brain lesions were fairly large and not uniquely specific to the BLA. Although, 82% of BLA tissue was damaged, a non-trivial amount (43%) of the central nucleus of the amygdala was also damaged. We think it is unlikely that the lesion effect we report here was caused by damage to the central nucleus of the amygdala because other research has shown that damage to this region had no effect on behavior controlled by outcome-specific encoding (e.g., Corbit & Balleine, 2005; Hatfield et al., 1996).

The present results extend prior findings by showing that while BLA lesions undermined the selective reward devaluation effect seen in our studies, it had little to no impact on temporal encoding. Lesioned animals displayed peak timing functions to the two CSs that were nearly identical to those functions seen in control animals. This result is consistent with the general notion that learning about the what and when aspects of reward entail separate systems. Our findings are, therefore, consistent with those of others, cited above, showing that the BLA is a critical structure for learning about the what aspects of reward. They go beyond those results by showing that this structure seems to play little to no role in learning to time the arrival of reward.

An alternative interpretation of our findings may be that BLA-lesioned animals are perfectly capable of encoding the "what" aspects of reward, but that devaluation performance, per se, was disrupted by BLA lesions. For instance, although the CS may activate a representation of its specific associated outcome, the selective devaluation effect also requires that the updated value of this outcome representation guides test performance. If BLA lesions disrupt this value updating process, as opposed to the actual selective outcome encoding itself, then our findings could be explained. While we have no data to refute this possibility it is worth noting that other results in the literature strongly point to an outcome encoding function of the BLA. For instance, BLA lesions, or suppression of BLA-OFC projections, have been shown to disrupt outcome-selective Pavlovian-to-instrumental transfer, the differential outcome effect, and also outcome-selective devaluation effects (see Blundell, Hall, & Killcross, 2001; Corbit & Balleine, 2005; Lichtenberg et al., 2017). Since all of

these effects depend upon outcome-specific encoding but only the selective devaluation effect additionally requires outcome value updating, the most parsimonious explanation is that the BLA is critical for outcome-selective encoding (Delamater, 2007).

Our results are somewhat at odds with other recent findings suggesting a role for the BLA in temporal encoding in fear conditioning. Díaz-Mataix, Ruiz Martinez, Schafe, Ledoux, & Doyère, (2013) demonstrated that a shift in the CS-US interval (from early trials to a "reactivation" trial) can induce a protein-synthesis dependent reconsolidation process within the amygdala. These findings suggest that from an early point in training the amygdala processes the specific CS-US interval used in fear conditioning and can detect when a different CS-US interval is used. Furthermore, Díaz-Mataix, Tallot, & Doyère (2014) review other literature suggesting involvement of the amygdala in interval timing (also Doyère & El Massioui, 2016). In view of our findings, it could be that (a) there is a fundamental difference in the role of the amygdala in timing during aversive and appetitive tasks, (b) that other brain structures may have taken over the functional loss in timing following pre-training lesions in our procedures, or (c) task-specific timing deficits are induced by amygdala manipulations. One such task-specific variable of potential significance is training with multiple target intervals. In our studies only one target interval was trained. If each of two stimuli, for instance, signaled reinforcement at distinct intervals, that could more effectively recruit BLA involvement. The Díaz-Mataix, Ruiz Martinez, Schafe, Ledoux, & Doyère, (2013) studies did use multiple intervals in their task. Future work will be required to provide more answers to these divergent findings.

Nonetheless, our finding of a dissociation between the effects of reward devaluation and BLA lesions on the motivation to respond in an outcome-specific manner and to time the arrival of food reward suggests that a process-dissociation may exist between learning about what and when aspects of reward. In order to make that point more forcefully, however, it would be necessary to investigate manipulations that impair interval timing to determine if outcome-specific encoding processes are similarly unaffected. Investigators have suggested that the dorsal striatum may play an especially critical role for temporal processing (e.g., Buhusi & Oprisan, 2013; Gu, van Rijn, & Meck, 2015; Hinton & Meck, 2004; Matell & Meck, 2004; Matell, Meck, & Nicolelis, 2003). This structure, therefore, is an interesting one to investigate in this context. While some investigators have found a role for dorsomedial striatum in goal-directed instrumental responding (Yin, Ostlund, Knowlton, & Balleine, 2005) a behavioral function that depends upon learning about what outcome is due to occur, the dorsolateral striatum appears not to depend upon such encoding (Yin, Knowlton, & Balleine, 2006). It remains unclear to what extent Pavlovian and instrumental outcome-specific and temporally-controlled responding might be affected by manipulations of these brain structures. But such research could shed additional light on how the brain learns about what and when aspects of reward.

One possibility is that learning about these different aspects of reward encoding depend upon completely dissociable neural networks, while another is that learning about multiple features of reward depend upon some common underlying neural network. Gallistel and Balsam (2014) and Balsam and Gallistel (2009) have suggested that learning can be best conceptualized, to a large if not exclusive extent, upon a decision-based process in which

temporal information plays a key role. It is not obvious, from this perspective, how non-temporal aspects of reward might factor into the control of behavior. It seems clear, though, from our data that learning to anticipate what reward will occur on a given trial appears to be distinct from learning about when that reward will occur. For example, learning about what reward will occur can be abolished without affecting learning when a reward will occur. This result would appear to support the view that learning of multiple reward features recruit separate, non-interacting, systems (see Delamater, et al, 2014). An organism's behavior, in some circumstances, though, could very well be controlled by an integrated representation of the various features of reward, both temporal and non-temporal. It remains an interesting research challenge to understand how learned behavior might be controlled by these different aspects of reward and how the brain enables it.

Acknowledgments

References

Addyman C, French RM, Thomas E. ScienceDirect Computational models of interval timing. Current Opinion in Behavioral Sciences. 2016; 8:140–146. DOI: 10.1016/j.cobeha.2016.01.004

Balci F, Ludvig EA, Abner R, Zhuang X, Poon P, Brunner D. Motivational effects on interval timing in dopamine transporter (DAT) knockdown mice. Brain Research. 2010; 1325:89–99. DOI: 10.1016/j.brainres.2010.02.034 [PubMed: 20167208]

Balleine BW, Killcross S. Parallel incentive processing: an integrated view of amygdala function. Trends in Neurosciences. 2006; 29(5):272–279. DOI: 10.1016/j.tins.2006.03.002 [PubMed: 16545468]

Balsam PD, Gallistel CR. Temporal maps and informativeness in associative learning. Trends in Neurosciences. 2009; 32(2):73–78. DOI: 10.1016/j.tins.2008.10.004 [PubMed: 19136158]

Blundell P, Hall G, Killcross S. Lesions of the basolateral amygdala disrupt selective aspects of reinforcer representation in rats. The Journal of Neuroscience. 2001; 21(22):9018–26. https://doi.org/http://www.jneurosci.org/content/21/22/9018. [PubMed: 11698612]

Buhusi CV, Oprisan SA. Time-scale invariance as an emergent property in a perceptron with realistic, noisy neurons. Behavioural Processes. 2013; 95:60–70. DOI: 10.1016/j.beproc.2013.02.015 [PubMed: 23518297]

Corbit LH. Double Dissociation of Basolateral and Central Amygdala Lesions on the General and Outcome-Specific Forms of Pavlovian-Instrumental Transfer. Journal of Neuroscience. 2005; 25(4):962–970. DOI: 10.1523/JNEUROSCI.4507-04.2005 [PubMed: 15673677]

De Corte BJ, Matell MS. ScienceDirect Interval timing, temporal averaging, and cue integration. Current Opinion in Behavioral Sciences. 2016; 8:60–66. DOI: 10.1016/j.cobeha.2016.02.004 [PubMed: 27175440]

Delamater AR. On the nature of CS and US representations in Pavlovian learning. Learning and Behavior. 2012; 40(1)doi: 10.3758/s13420-011-0036-4

Delamater AR, Desouza A, Rivkin Y, Derman R. Associative and temporal processes: A dual process approach. Behavioural Processes. 2014; 101doi: 10.1016/j.beproc.2013.09.004

Delamater AR, LoLordo VM. Event revaluation procedures and associative structures in Pavlovian conditioning. In: Dachowski L, Flaherty CF, editorsCurrent topics in animal learning: Brain, emotion, and cognition. NJ, Hillsdale: Erlbaum; 1991. 55–94.

Delamater AR, Oakeshott S. Learning about multiple attributes of reward in Pavlovian conditioning. Annals of the New York Academy of Sciences. 2007; 1104doi: 10.1196/annals.1390.008

Díaz-Mataix L, Ruiz Martinez RC, Schafe GE, Ledoux JE, Doyère V. Detection of a temporal error triggers reconsolidation of amygdala-dependent memories. Current Biology. 2013; 23(6):467–472. DOI: 10.1016/j.cub.2013.01.053 [PubMed: 23453952]

Díaz-Mataix L, Tallot L, Doyère V. The amygdala: A potential player in timing CS-US intervals. Behavioural Processes. 2014; 101:112–122. DOI: 10.1016/j.beproc.2013.08.007 [PubMed: 23973708]

Doyère V, El Massioui N. A subcortical circuit for time and action: Insights from animal research. Current Opinion in Behavioral Sciences. 2016; 8:147–152. DOI: 10.1016/j.cobeha.2016.02.008

Drew MR, Zupan B, Cooke A, Couvillon PA, Balsam PD. Temporal Control of Conditioned Responding in Goldfish. Journal of Experimental Psychology: Animal Behavior Processes. 2005; 31(1):31–39. DOI: 10.1037/0097-7403.31.1.31 [PubMed: 15656725]

Gallistel CR, Balsam PD. Time to rethink the neural mechanisms of learning and memory. Neurobiology of Learning and Memory. 2014; 108:136–144. DOI: 10.1016/j.nlm.2013.11.019 [PubMed: 24309167]

Galtress T, Kirkpatrick K. Reward value effects on timing in the peak procedure. Learning and Motivation. 2009; 40(2):109–131. DOI: 10.1016/j.lmot.2008.05.004

Galtress T, Marshall AT, Kirkpatrick K. Motivation and timing: Clues for modeling the reward system. Behavioural Processes. 2012; 90(1):142–153. DOI: 10.1016/j.beproc.2012.02.014 [PubMed: 22421220]

Gu BM, van Rijn H, Meck WH. Oscillatory multiplexing of neural population codes for interval timing and working memory. Neuroscience and Biobehavioral Reviews. 2015; 48:160–185. DOI: 10.1016/j.neubiorev.2014.10.008 [PubMed: 25454354]

Hatfield T, Han JS, Conley M, Gallagher M, Holland P. Neurotoxic lesions of basolateral, but not central, amygdala interfere with Pavlovian second-order conditioning and reinforcer devaluation effects. The Journal of Neuroscience. 1996; 16(16):5256–5265. https://doi.org/http://www.jneurosci.org/content/16/16/5256. [PubMed: 8756453]

Hinton SC, Meck WH. Frontal-striatal circuitry activated by human peak-interval timing in the supra-seconds range. Cognitive Brain Research. 2004; 21(2):171–182. DOI: 10.1016/j.cogbrainres.2004.08.005 [PubMed: 15464349]

Johnson AW, Gallagher M, Holland PC. The basolateral amygdala is critical to the expression of pavlovian and instrumental outcome-specific reinforcer devaluation effects. The Journal of Neuroscience. 2009; 29(3):696–704. DOI: 10.1523/JNEUROSCI.3758-08.2009 [PubMed: 19158296]

Kirkpatrick K. Interactions of timing and prediction error learning. Behavioural Processes. 2014; 101:135–145. DOI: 10.1016/j.beproc.2013.08.005 [PubMed: 23962670]

Konorski J. Integrative activity of the brain. Chicago: University of Chicago Press; 1967.

Ludvig EA, Balci F, Spetch ML. Reward magnitude and timing in pigeons. Behavioural Processes. 2011; 86(3):359–363. DOI: 10.1016/j.beproc.2011.01.003 [PubMed: 21244854]

Matell MS, Meck WH. Cortico-striatal circuits and interval timing: Coincidence detection of oscillatory processes. Cognitive Brain Research. 2004; 21(2):139–170. DOI: 10.1016/j.cogbrainres.2004.06.012 [PubMed: 15464348]

Matell MS, Meck WH, Nicolelis MAL. Interval timing and the encoding of signal duration by ensembles of cortical and striatal neurons. Behavioral Neuroscience. 2003; 117(4):760–773. DOI: 10.1037/0735-7044.117.4.760 [PubMed: 12931961]

Ohyama T, Gibbon J, Deich JD, Balsam ED. Temporal control during maintenance and extinction of conditioned keypecking in ring doves. Animal Learning & Behavior. 1999; 27(1):89–98. DOI: 10.3758/BF03199434

Roberts S. Isolation of an internal clock. Journal of Experimental Psychology: Animal Behavior Processes. 1981; 7(3):242–268. DOI: 10.1037/0097-7403.7.3.242 [PubMed: 7252428]

Paxinos G, Watson C. The rat brain in stereotaxic coordinates: The new coronal set. 5. San Diego: Elsevier Academic Press; 2005.

Rozeboom WW. "What is learned?" An empirical enigma. Psychological Review. 1958; 65:22–33. [PubMed: 13505979]

Wagner AR, Brandon SE. Evolution of a structured connectionist model of Pavlovian conditioning (AESOP). In: Klein SB, Mowrer RR, editorsContemporary learning theories: Pavlovian conditioning and the status of traditional learning theory. Hillsdale, NJ: Erlbaum; 1989. 149–189.

Yin HH, Knowlton BJ, Balleine BW. Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. Behavioural Brain Research. 2006; 166(2):189–196. DOI: 10.1016/j.bbr.2005.07.012 [PubMed: 16153716]

Yin HH, Ostlund SB, Knowlton BJ, Balleine BW. The role of the dorsomedial striatum in instrumental conditioning. European Journal of Neuroscience. 2005; 22(2):513–523. DOI: 10.1111/j.1460-9568.2005.04218.x [PubMed: 16045504]

Highlights

- Instrumental and Pavlovian peak timing tasks were used to assess learning of reward identity and timing.

- Responding for a devalued reward was selectively impaired but temporal control was maintained.

- Basolateral amygdala lesions undermined this selective reward devaluation effect but not interval timing.

- The data suggest that different neural systems contribute to reward identity and temporal learning.
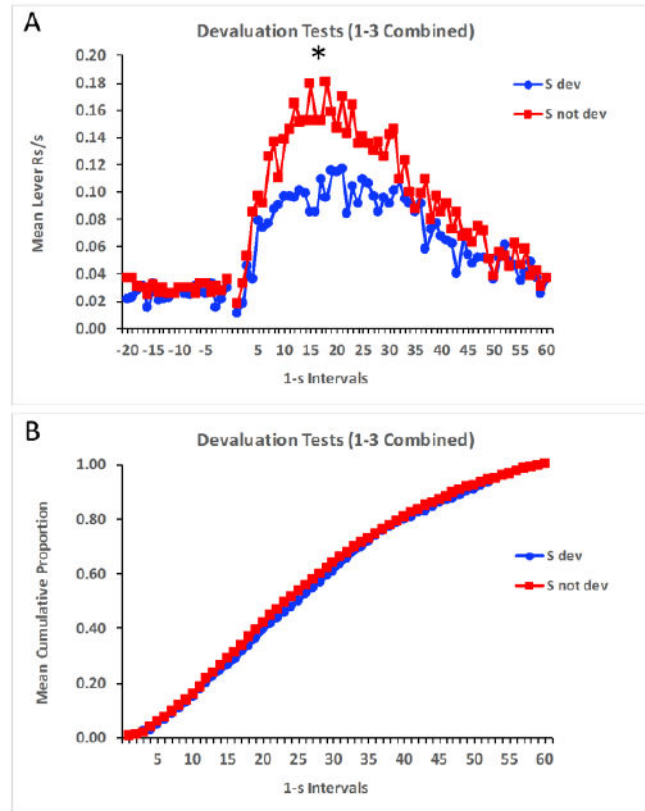
Figure 1.
Mean lever responses per second (A) and mean cumulative proportion of total responses (B) across time in the presence of the stimuli signaling devalued and not devalued rewards during the devaluation tests of Experiment 1. The * indicates significance at p < 0.05.
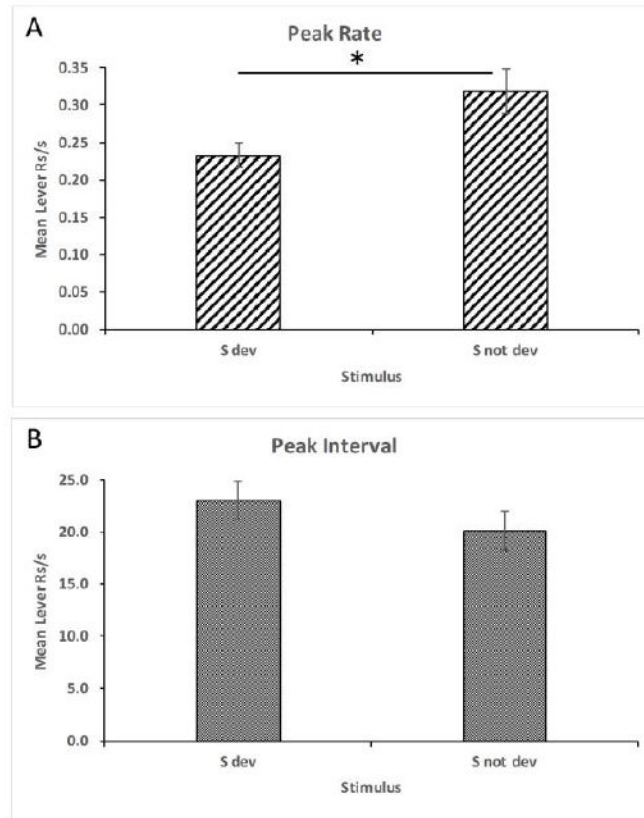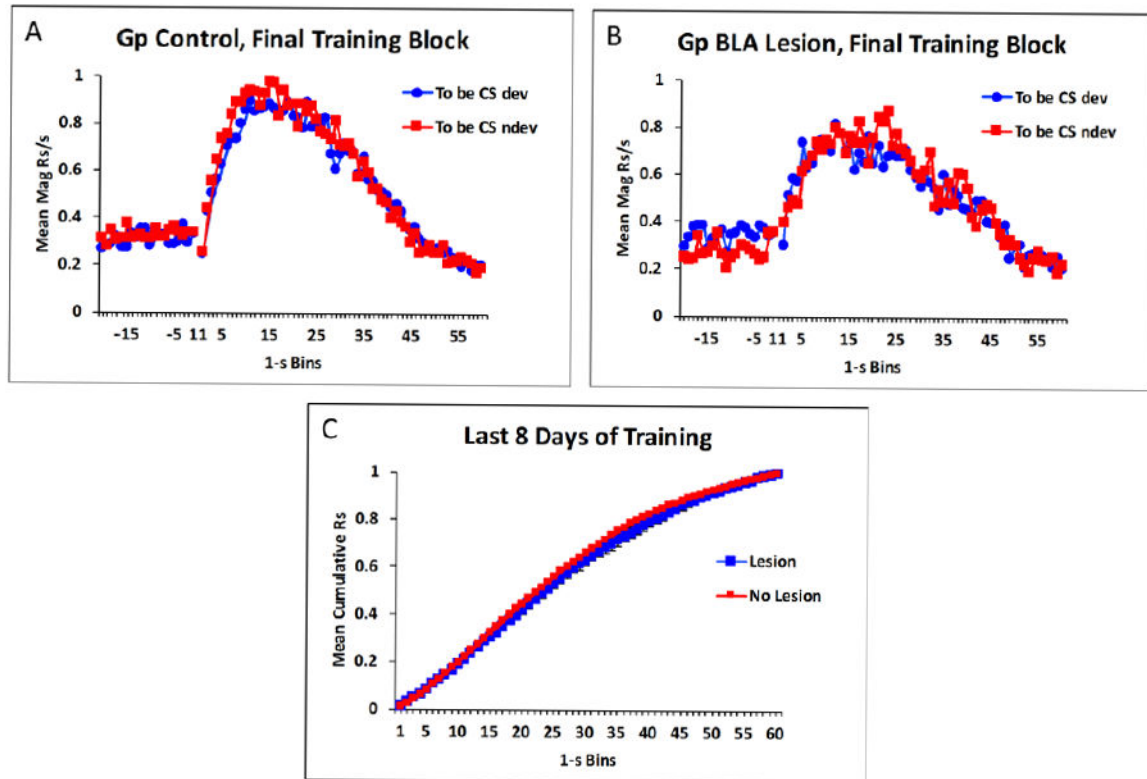
Figure 2.
Mean peak rate (A) and interval (B) of responding (+/- SEM) during devaluation tests in Experiment 1. The * indicates significance at $p < 0.05$.

-1.80 mm

-2.28 mm

-2.76 mm

-3.36 mm

-3.84 mm
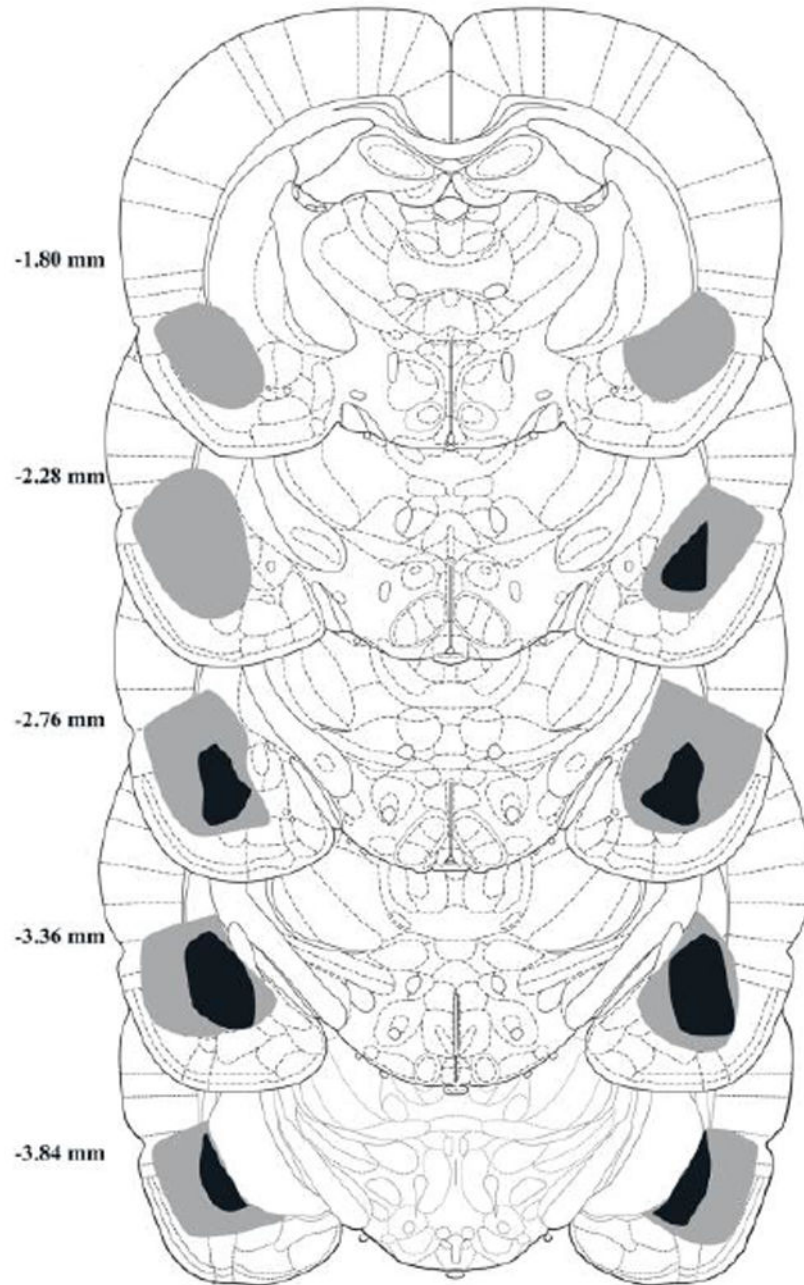
Figure 3.
Coronal sections showing minimum (black) and maximum (grey) lesioned areas in different A/P sections across animals judged to have effective BLA lesions (n=7) in Experiment 2.

Figure 4.
Mean magazine responses per second in the final 8 sessions of training in control rats (A), and in lesioned rats (B) in Experiment 2. Mean cumulative responding over time collapsed across the two conditioned stimuli (CS) in lesioned and control animals (C).
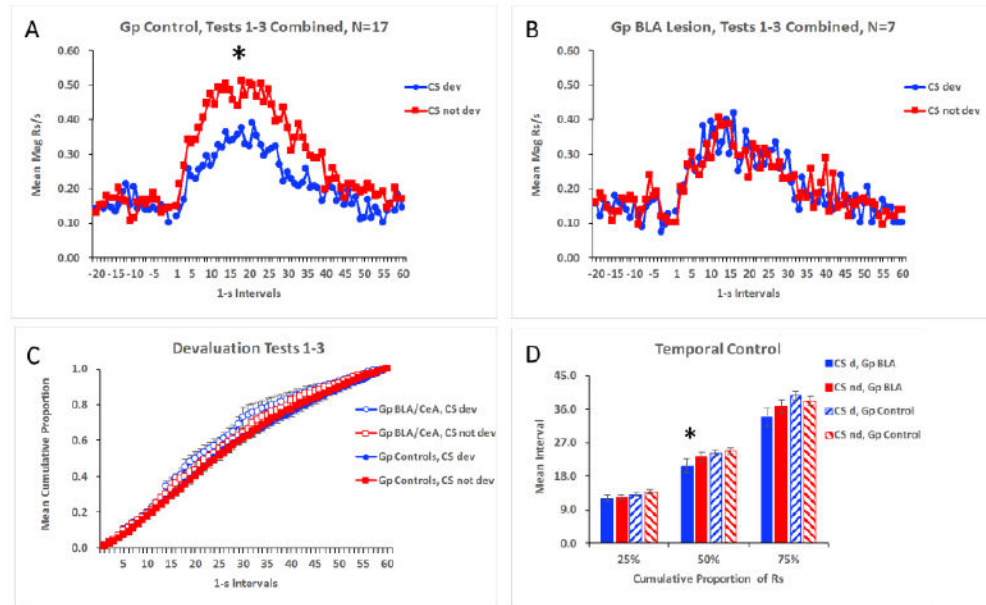
Figure 5.
Mean magazine responses per s during devaluation test sessions in Experiment 2 across 1-s intervals (both pre stimulus and during the stimulus) in the presence of the conditioned stimuli (CS) paired with the devalued or not devalued rewards for Groups Control (A) and BLA Lesion (B). Mean cumulative proporation of responding across the trial is shown for the two groups (C), and the interval in which 25%, 50%, and 75% of total responses were produced are displayed for the groups (D). Panels E and F and G show peak rate, peak interval, and the peak rate/average rate (Timing Ratio), respectively, in the presence of the two stimuli for the groups. * indicates p < 0.05.