Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Esther Seo April 8, 2022

Investigating behavioral flexibility controlled by the melanocortin-4 receptor in the dorsal striatum

by

Esther H. Seo

Shannon L. Gourley
Adviser

Psychology

Shannon L. Gourley

Adviser

Donna L. Maney

Committee Member

Leila E. Rieder

Committee Member

Yuk F. Cheong

Committee Member

2022

Investigating behavioral flexibility controlled by the melanocortin-4 receptor in the dorsal striatum

By

Esther Seo

Shannon Gourley
Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Department of Psychology

2022

Abstract
Investigating behavioral flexibility controlled by the melanocortin-4 receptor in the dorsal striatum

By Esther Seo

Flexible behavior refers to the ability to update behavior in response to external or internal stimuli – to integrate new information into old patterns of behavior. One brain region necessary for behavioral flexibility is the dorsal striatum. The dorsal striatum is divided into two subregions in rodents: the dorsomedial striatum (DMS) and dorsolateral striatum (DLS). Neural activity in the DMS and DLS is directly correlated with whether an action is considered goal-directed or habitual, respectively. The factors in the dorsal striatum responsible for coordinating behavioral flexibility remain incompletely understood. One possible factor is the melanocortin-4 receptor (MC4R), a receptor directly implicated in the regulation of behavioral flexibility. MC4R controls the direction of excitatory neural activity in the DMS and DLS such that MC4R activation decreases neural activity in the dorsal striatum. Administration of a systemic MC4R agonist would be expected to impact MC4R activity simultaneously in the DMS and DLS. However, neural activity in the DMS and DLS is not concurrent and is dependent on the present action (action->outcome or stimulus->response). We found that systemic administration of an MC4R agonist during a period when mice express flexible responding interferes with DMS-dependent behavior, resulting in a bias towards inflexible, habit-like behavior. Meanwhile, we found that systemic administration of an MC4R agonist during the expression of habit blocked DLS-dependent behavior, resulting in goal-directed action selection, when habitual behavior would otherwise occur. These findings could provide novel therapeutic insight for treating neuropsychiatric disorders characterized by maladaptive habit, such as substance use disorders.

Investigating behavioral flexibility controlled by the melanocortin-4 receptor in the dorsal striatum

By

Esther Seo

Shannon Gourley
Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Department of Psychology

2022

Acknowledgements

Thank you to Dr. Shannon Gourley for her mentorship throughout this year; to Dr. Leila Rieder, Dr. Donna Maney, and Dr. Yuk Fai Cheong for their endless support and feedback. Special thank you to Elizabeth Heaton for her guidance, mentorship, humor, and compassion.

And thank you to my family and friends without whom none of this would be possible. Finally, I would like to give all glory to God.

Table of Contents

List of Figures

Introduction

Flexible behavior describes the ability to update behavior in response to external or internal stimuli – to integrate new information into old patterns of behavior. For instance, taking a detour when your typical route is blocked by construction describes flexible behavior. In contrast, many neuropsychiatric disorders are characterized by inflexible, habit-like behaviors that can contribute to poor health outcomes (McKim et al., 2021). For example, individuals with substance use disorder (SUD) may defer to inflexible habits (substance use) at the expense of flexible goal-directed action (rehabilitation). Thus, studying this flexible toggling between goal-directed action and habit may be crucial in developing treatments of neuropsychiatric disorders, including SUD.

One brain region necessary for behavioral flexibility is the dorsal striatum. The dorsal striatum is divided into two subregions in rodents: the dorsomedial striatum (DMS) and dorsolateral striatum (DLS). Excitatory neural activity in the DMS facilitates goal-directed behavior, which is motivated by the anticipation of an outcome (Yin, 2005). Goal-directed behavior allows one to flexibly update an action in response to new information about the outcome. In rats, lesion of the DMS ablates the ability of rats to flexibly update their action to reflect their experience (Yin, 2005). In contrast, the DLS regulates habitual behavior that is motivated by the presentation of a stimulus, rather than an expected outcome. Rats with lesions of the DLS readily engage in flexible, goal-directed responding in an outcome devaluation task, obstructing their ability to form habits (Yin et al., 2004). Thus, the dorsal striatum serves as a potential target for investigating behavioral flexibility.

Neural activity in the DMS and DLS is directly correlated with whether an action is considered goal-directed or habitual. Although an action may initially start goal-directed, the activity may transition into a habit. For example, the first time one drives to work, one has to consciously navigate an unfamiliar route to arrive on time. While driving this unfamiliar route, neural activity is expected to be elevated in the DMS based on previous literature (Gremel and Costa, 2013). After this action has been repeated multiple times over a consistent period, the route become familiar, and as a consequence, the action of driving to this location becomes a habit. As goal-directed actions transition into habits, neural activity decreases in the DMS and increases in the DLS, reflecting a flexible shift between these two brain regions in relation to the familiarity of the action (Gremel and Costa, 2013). Taking advantage of this well-studied balance between DMS and DLS activity, as it corresponds to behavior, could serve as a strategy by which to understand the molecular mechanisms controlling *maladaptive* actions, such as aberrant habit formation or adherence.

The factors in the dorsal striatum responsible for coordinating behavioral flexibility remain incompletely understood. One possible factor is the melanocortin-4 receptor (MC4R), which is agonized by α-melanocyte stimulating hormone (α-MSH) (Anderson et al., 2016). MC4R is expressed throughout the brain, and it has been extensively studied in the hypothalamus, as it plays a role in energy homeostasis (Anderson et al., 2016). MC4R is also expressed within both the DMS and DLS, primarily on dopamine-1-receptor-medium spiny neurons (D1R-MSNs), where the receptor regulates GluA2-AMPAR localization at the cell membrane (Lim et al., 2012). In the dorsal striatum, the activation of MC4R on D1R-MSNs leads to the reduction of GluA2-AMPAR (Lim et al., 2012). Likewise, decreasing striatal MC4R activity leads to an increase in excitatory, glutamatergic signaling in the striatum (Xu et al.,

2013). Thus, activation of MC4R would lead to a decrease in neural activity in the dorsal striatum, whereas a reduction of MC4R would lead to an increase of dorsal striatal neural activity. Therefore, MC4R controls neural activity in the DMS and DLS.

Recent research has directly implicated dorsal striatal MC4R in the regulation of behavioral flexibility. The activity of *Mc4r*+ MSNs in the DMS – the subregion required for goal-directed behavior – is necessary to express and sufficient to induce flexible, goal-directed action (Heaton, unpublished). MC4R levels in the DMS are directly correlated with behavioral flexibility, such that increased MC4R is associated with decreased flexibility (Allen et al., 2022). Accordingly, reducing *Mc4r* in the DMS via viral vector mediated gene knockdown enhances the ability of mice to modify behavior when rewards are absent or devalued (Allen et al., 2022). Similarly, the activity of *Mc4r*+ MSNs in the DLS – the subregion required for habit – is sufficient for the expression of inflexible, habit-like behavior (Heaton, unpublished). Reducing *Mc4r* in the DLS promotes bias towards familiar, routine behaviors, thus promoting the function of this striatal subregion in habitual behavior (Heaton, unpublished). MC4R therefore serves as a prime target for studying the effects of dorsal striatal neuropeptide systems on behavioral flexibility.

Current Research

The studies discussed above were conducted using chemogenetics, specifically using Gi- and Gq-DREADDs; however, MC4R agonists exist and are well-tolerated. Administration of a systemic MC4R agonist would be expected to simultaneously impact central MC4R activity, including in the DMS and DLS. However, neural activity in the DMS and DLS is not concurrent and is dependent on the present action (action->outcome or stimulus->response) (Gremel &

Costa, 2013). As discussed, DMS activity is critical during the expression of flexible, goal-directed action. During this period, the DMS is "online" and the DLS is "offline." We predict that systemic administration of an MC4R agonist during a period when mice express flexible responding interferes with DMS-dependent behavior, will result in a bias towards inflexible, habit-like behavior. By the same reasoning, we predict that systemic administration of an MC4R agonist during the expression of habit (when the DMS is "offline" and the DLS is "online") will block DLS-dependent behavior, resulting in goal-directed action selection, when habitual behavior would otherwise occur.

To test these predictions, we used setmelanotide, a synthetic MC4R agonist that was approved in November 2020 by the FDA to treat melanocortin-related monogenic obesity. Previous research suggests setmelanotide impacts energy homeostasis (Anderson et al., 2016). However, little is known of potential effects of setmelanotide on behaviors controlled by brain regions such as the dorsal striatum. This study is the first to directly test whether systemic setmelanotide administration impacts behavioral flexibility.

In sum, we predicted that these experiments would demonstrate that it is possible to manipulate flexible and inflexible behavior via systemic administration of an MC4R agonist. Our results may provide novel therapeutic insight for treating neuropsychiatric disorders characterized by maladaptive habits, such as SUD.

Methods

Animals

All experiments were conducted in mice homozygous for a 'floxed' *Mc4r* gene (*Mc4r*-flox) maintained on a mixed C57BL/6J-129S1/SvlmJ background (Jackson Labs). In absence of

Cre-recombinase, the 'floxed' *Mc4r* gene maintains typical function, essentially rendering mice in the following experiments indistinguishable from wild-type counterparts (Sohn et al., 2013). Mice were weaned from the dam at or soon after postnatal day 21 and housed in single-sex cages with siblings or unrelated mice of the same age. Mice were maintained on a 14-h light cycle (0700 on) and provided food and water *ad libitum*. Procedures were approved by the Emory University IACUC and carried out in accordance with the recommendations of the *Guide for the Care and Use of Laboratory Animals*.

Setmelanotide preparation and administration

Setmelanotide hydrochloride (Creative Peptides Inc.) was dissolved in saline + 0.1% bovine serum albumin (BSA). Mice were injected with 0.1 mL per weight (g) vehicle (1% BSA) or setmelanotide (0.2 or 1.0 mg/kg) intraperitoneally (i.p.).

Evaluation of setmelanotide dosage on food intake and locomotion

Behaviorally active doses of setmelanotide are defined by their ability to reliably reduce food intake in mice (Collet et al., 2017). However, previous literature reported that high doses of setmelanotide may also impact locomotion in mice, which would confound behavioral testing (Bischof et al., 2016). Thus, it was necessary to conduct a pilot study to identify a dose of setmelanotide that was both behaviorally active (reduced food intake) but did not interfere with gross locomotion for use in future experiments.

Locomotion and free feeding were measured in large, open locomotion chambers. Each locomotion chamber was equipped with 16 horizontally placed photobeams. Total number of photobeam breaks was used as a measure of gross locomotion. All mice were exposed to all

three drugs across three test days in a counterbalanced fashion. First, mice (n = 8) were placed in the center of the chamber and allowed to habituate for an hour. Then, the mice were injected with one of three drug conditions: 1 mg/kg setmelanotide, 0.2 mg/kg setmelanotide, or vehicle. Before returning the mice to their chambers, 4g of vivarium chow was scattered across the locomotion chamber below the locomotor beams. We expected setmelanotide to reach peak efficacy in the dorsal striatum after 90 minutes post-injection (Collet et al., 2017). Thus, upon their re-entrance into the locomotion chambers, we continued to record their behavior for 2 hours post-injection. At the end of the session, we recorded the amount of food eaten.

Instrumental conditioning: Short training

Instrumental conditioning was conducted in a separate cohort of mice (n = 12). Mice were trained to nose poke for pellets using a "short training" procedure. First, mice were trained to respond on two apertures for a food reward using a fixed ratio (FR1) schedule of reinforcement. Mice could earn up to 30 reinforcers per aperture (60 reinforcers/session), and training sessions ended either when mice acquired all 60 reinforcers or when 70 minutes had passed. Following FR1 training, mice were shifted to a random interval 30-second schedule of reinforcement (RI30) for 2 days to enhance response rates.

Once responding was stable, one aperture was occluded, and the response-outcome contingency associated with the remaining aperture was rendered null: responding at the available aperture no longer predicted food pellet delivery and instead, pellets were delivered non-contingently at a rate yoked to the animal's reinforcement rate from the last day of training ("noncontingent aperture"). As a control, responding at the other aperture during a subsequent session remained reinforced ("reinforced aperture"). On the final day, both apertures were made

available during a choice test conducted in extinction. Mice that preferentially responded on the aperture associated with the intact contingency were considered goal-directed – sensitive to response-reward contingency. In contrast, mice that responded equally on both apertures were considered habitual – deferring to familiar routines. Following "short training", we expect a typical control mouse to exhibit goal-directed behaviors that are sensitive to changes in instrumental contingencies.

Mice went through two rounds of testing for response flexibility. In the first round, mice received an i.p. injection of either 0.2 mg/kg setmelanotide or vehicle 90 minutes prior to exposure to the noncontingent aperture. This setmelanotide dose was found to be behaviorally active but did not impact gross locomotion in our pilot test. Since setmelanotide decreases food intake as seen in previous literature (Bischof et al., 2016) and in our experiments, we controlled for effects on food consumption in the reinforced vs. noncontingent condition by administering 0.2 mg/kg setmelanotide in all mice 90 minutes prior to exposure to the reinforced aperture. After completing the first choice test, mice went through several "refresher" sessions where responding at the aperture was again rewarded as normal to reinvigorate response rates. Once response rates were again stable, mice entered the second round of testing. This time, 90 minutes prior to exposure to the noncontingent and reinforced apertures, mice received an injection of the drug condition that was not experienced in the first round of testing (e.g., a mouse receives an injection of 0.2 mg/kg setmelanotide during the first round, and a vehicle injection in the second round). Response rates during the second choice test were recorded. The order of drug injection was counterbalanced across mice.

Instrumental conditioning: Extended training

In a separate cohort (n = 12), mice were trained using an "extended training" procedure. As above, mice were first trained to respond on two apertures for a food reward using a FR1 schedule of reinforcement. Once responding was stable, mice were shifted to 4 days of RI30 followed by 10 days of RI60. Extensive experience with RI schedules reduces sensitivity to instrumental contingencies (Yin et al., 2006). Thus, typical mice were expected to exhibit habitual behavior at the following choice test.

Following extended training, mice were tested in two rounds of our task of response flexibility. Behavioral testing and drug administration were conducted as above under "short training."

Statistical analyses

For measures of locomotion and free-feeding, two-tailed statistical analyses with $a = 0.05$ were performed using the Statistical Package for the Social Sciences (SPSS). The number of locomotion beam breaks and food intake were analyzed by repeated measures analysis of variance (ANOVA) with drug condition as the within-subjects factor. Following main effects, post-hoc t-tests were applied.

For tests of behavioral flexibility, response rates and response preference ratios were compared by 2-factor repeated measures analysis of variance (ANOVA), with drug condition and aperture as within-subjects (repeated measure) factors. Following interactions or for the purposes of planned comparisons, paired subjects post hoc t-tests were applied, as appropriate. One-sample t-tests were used to determine if ratios of response rates at the reinforced versus noncontingent aperture were significantly different from 1. In some cases, trending effects (defined as $p > .05$ but $<.14$) will be discussed.

Differences in sex were not analyzed because the sample size was too small to conduct analyses. Sex differences will be analyzed with the addition of replicates.

Results

Setmelanotide Dosage Pilot Study

This experiment was conducted to investigate the dose-dependent effect of setmelanotide on food intake. We administered 0.2 mg/kg setmelanotide, 1.0 mg/kg setmelanotide, or vehicle 90 minutes prior to the beginning of the experiment. This experiment was conducted in three rounds with each mouse experiencing each of the drug conditions. The order of drug condition was administered in a counterbalanced fashion. We found that food intake significantly differed as a function of the drug administered ($F(2, 14) = 9.48, p = .002$). Post-hoc comparisons revealed that there was a significant decrease in food intake in the 0.2 mg/kg condition ($t(7) = 3.39, p = .012$) and the 1.0 mg/kg condition ($t(7) = 3.18, p = .016$) in comparison to the control (Figure 1A). Thus, the 0.2 mg/kg and 1.0 mg/kg doses were sufficient to reduce food intake.

This experiment also investigated the effect of setmelanotide on locomotion in a dose-dependent manner. Gross locomotion did not differ significantly across the three drug conditions ($F(2, 14) = 3.61, p = .055$; Figure 1B). This indicated that both the 0.2 mg/kg and 1.0 mg/kg setmelanotide doses had no effects on locomotion. From the food intake and locomotion data, it was concluded that both the 0.2 mg/kg and 1.0 mg/kg doses were behaviorally active and eligible for our experiments. We chose the 0.2 mg/kg dose to minimize off-target effects of an excessive dose of setmelanotide.

To determine whether setmelanotide administration impacted operant response rates, we trained mice in operant conditioning chambers to respond for the delivery of a food reward on a

FR1 schedule (Figure 1C). Then, we administered 0.2 mg/kg setmelanotide or vehicle 90 minutes prior to an FR1 session. We conducted a paired samples t-test to determine if there was a significant difference in responding at apertures pre- *vs*. post-administration of setmelanotide. There was no significant difference in responding following setmelanotide administration compared to the condition prior to drug administration ($t(12) = 1.47, p = .166$; Figure 1D).

Short Training

The purpose of short training was to train the mice to exhibit bias towards goal-directed behavior. The mice were trained using a FR1 schedule of reinforcement until responding was stable. Then, mice were further trained using a RI30 schedule of reinforcement for two days. Response rates increased across training ($F(10, 110) = 49.49, p < .001$; Figure 2A).

Following training, the contingency at one aperture was rendered null while the contingency at the other aperture remained reinforced. The interaction between noncontingent *vs*. reinforced condition and drug condition was trending ($F(1, 11) = 19.31, p = .067$; Figure 2B). Planned comparisons revealed that control mice favored the reinforced aperture ($t(11) = 2.39, p = .036$; Figure 2B), whereas setmelanotide blocked that preference ($t(11) = -0.37, p = .717$; Figure 2B). This pattern leads to the conclusion that the mice in the control condition exhibited a bias towards goal-directed behavior, while setmelanotide may block behavioral flexibility, causing mice in the setmelanotide condition to defer to habit-based behavior. However, we are not rejecting the null hypothesis because the interaction between the noncontingent *vs*. reinforced condition and drug condition was not significant.

The next day, response preference was measured in a choice test conducted in extinction where we investigated whether the mice were able to update their behavior according to their prior learning during contingency degradation. There was no main effect of drug on response

rates between the control and setmelanotide conditions ($F(1, 11) = 2.30, p = .158$). We found a main effect of aperture ($F(1, 11) = 5.42, p = .040$), likely driven by high responding at the reinforced aperture in the control group. There was a trending interaction between drug and aperture ($F(1, 11) = 2.55, p = .138$). Planned comparisons between the reinforced vs. noncontingent aperture for each group showed that there was a significant increase in responding at the reinforced aperture compared to the noncontingent aperture in the control condition ($t(11) = 3.33, p = .007$; Figure 2C), but there was no difference between the reinforced and noncontingent apertures in the setmelanotide condition ($t(11) = 1.27, p = .232$; Figure 2C). Again, it is important to note that we are not rejecting the null hypothesis.

Moreover, we conducted a one sample t-test against the test value of 1 on the response ratio between responses at the reinforced versus degraded apertures during the choice test. The response ratio of the control condition significantly differed from 1 ($t(11) = 3.09, p = .010$; Figure 2D), while the response ratio of the setmelanotide condition did not differ from 1 ($t(11) = 1.77, p = .104$; Figure 2D). Thus, during this choice test, the responding of mice in the control condition was significantly higher on the reinforced aperture compared to the noncontingent aperture, while the mice in the setmelanotide condition showed no preference in their choice of aperture. This indicated a bias towards goal-directed behavior when exposed to the control condition. On the other hand, mice failed to show a preference and instead deferred to habit-like behavior when exposed to the setmelanotide condition. However, there was no difference between groups in responding at reinforced or noncontingent apertures ($t(11) = 1.45, p = .176$). *These results are preliminary, and replications are underway*.

Extended Training

The purpose of extended training was to train the mice to exhibit biases towards habitual behavior. As above, mice were trained to respond on two apertures for the delivery of a food pellet on an FR1 schedule. When responding was stable, the schedule was shifted to 4 days of RI30 training, followed by 10 days of RI60 training. Response rates increased across training ($F(22, 242) = 23.19, p < .001$; Figure 3A).

As above, the contingency at one aperture was next rendered null, while the contingency at the other aperture remained intact, and response rates during the noncontingent *vs*. reinforced sessions were collected. When we analyzed response rates as mice first encountered these changes in contingencies, we found that there was no main effect of drug ($F(1, 11) = 0.26, p = .617$). However, there was a main effect of aperture ($F(1, 11) = 16.47, p = .002$), likely driven by lower response rates at the degraded aperture in the setmelanotide condition. The interaction between noncontingent *vs*. reinforced condition and drug condition was not significant during contingency degradation ($F(1, 11) = 0.99, p = .342$). Thus, we unexpectedly did not induce habitual behavior in the control condition and groups did not differ.

The next day, response preference was measured in a choice test conducted in extinction. There was no main effect of drug on response rates ($F(1, 11) = 0.15, p = .710$). There was a main effect of aperture on response rates ($F(1, 11) = 11.82, p = .006$; Figure 3C). There was no significant interaction between the drug and aperture conditions ($F(1, 11) = 0.03, p = .865$). A one sample t-test against the test value of 1 using the same response ratio calculated similarly as in short training revealed that both the response ratio of the control condition ($t(11) = 2.77, p = .018$; Figure 3D) and the setmelanotide condition significantly differed from 1 ($t(11) = 2.28, p = .044$; Figure 3D). Thus, at the choice test, both experimental groups were unexpectedly biased

towards goal-directed behavior, which will be discussed below. *Again, these results are preliminary in nature, and replications are underway.*

Discussion

The dorsal striatum is a brain region that influences behavioral flexibility: activity in the DMS drives goal-directed behavior and activity in the DLS drives habit. Within the dorsal striatum, activating MC4R decreases the excitability of MSNs (Lim et al., 2012), while reducing MC4R enhances glutamatergic excitability (Xu et al., 2013). To date, the role of MC4R in striatal-dependent behavioral flexibility has been primarily investigated via viral vector-mediated gene knockdown and chemogenetics (Allen et al., 2022). This experiment is the first to test whether systemic administration of setmelanotide, a MC4R agonist, biases animals towards goal-directed or habitual behaviors.

Setmelanotide blocked bias formation towards goal-directed behavior in mice exposed to short training

Mice were first trained according to schedules of reinforcement that induce biases towards goal-directed behavior (Figure 2A). Upon administering 0.2 mg/kg setmelanotide i.p., however, goal-directed action strategies were disrupted, and mice deferred instead to habit-based behavior (Figure 2B). When mice perform goal-directed behavior, the DMS is online and neural activity is increased (Gremel and Costa, 2013). Mice in the control condition exhibited a bias towards goal-directed behavior in the operant conditioning chambers (Figure 2B and Figure 2C), and thus, we can theorize that the DMS was online during this task. The administration of an MC4R agonist such as setmelanotide activates the MC4R on the D1R-MSNs (Collet et al.,

2017), which decreases the expression of GluA2-AMPARs that control excitatory neural activity. Thus, we theorized that administering setmelanotide would decrease neural activity within the brain region that is active and relevant during a given task. The mice in the control condition exhibited a bias towards goal-directed behavior, and thus, we theorized that the DMS was active during this task. Upon exposure to setmelanotide, mice failed to develop goal-sensitive action strategies and instead, deferred to habitual behavior (Figure 2B and Figure 2C). In accordance with previous literature, this suggests that there was a decrease in activity within the DMS after setmelanotide administration (Gremel & Costa, 2013). Furthermore, we believe that this change in neural activity in the DMS occurred as a result of reduced expression of GluA2-AMPAR (Lim et al., 2012), as MC4R was agonized by setmelanotide.

Does setmelanotide have potential for blocking habitual behavior?

We next attempted to train mice to exhibit a bias towards habitual behavior using extended RI training, which is expected to correspond with an increase in DLS activity (Gremel and Costa, 2013). When we measured response rates when mice first encountered a violation in familiar contingencies, mice exposed to setmelanotide flexibly inhibited a behavior that was not reinforced, suggesting the activity in the DLS was reduced. The strongest effect of the drug was seen *during* this initial experience with contingency degradation (Figure 3B), when mice were learning about the contingency of one aperture, and while setmelanotide was active. However, the control and setmelanotide conditions unexpectedly did not differ; thus, we unfortunately did not appear to induce habitual behavior, making it difficult to draw conclusions about the effects (if any) of setmelanotide on this behavior.

Reinforcing this notion, at the choice test, there was also a main effect of aperture (Figure 3C). This means that regardless of drug condition, there was a significantly higher response rate at the reinforced aperture in comparison to the noncontingent aperture, which suggests that all groups exhibited a bias towards goal-directed behavior. This contradicts our expectation that mice in the extended training cohort would exhibit a bias towards habitual behavior. This is further supported by the one-sample t-test where the ratio of responses in both conditions significantly differed from the test value of 1. It is interesting to note, though, that setmelanotide did not obstruct response flexibility in this experiment, suggesting that the effects in Figure 2 reflect a vulnerability of response flexibility to MC4R agonism when mice are relatively new to a task and are still learning task parameters, relative to when the task is highly familiar. In future experiments, more extensive training will be used to intensify habitual behavior.

Implications of setmelanotide in humans regarding behavioral flexibility

Setmelanotide is an FDA approved drug used to treat monogenic obesity and has been extensively researched in both rodent models and humans (Bischof et al., 2016; Clement et al., 2020; Markham, 2021). The present research used setmelanotide in rodents to study the effects of MC4R activation, this time on striatal-dependent decision-making behavior. The findings of the current research have translational impact in patients with neuropsychiatric disorders characterized by routine, inflexible behaviors, such as SUD. As seen in the current research, administering setmelanotide to mice biased towards habitual behavior enabled them to flexibly update their behavior according to on-going experience, thus enriching goal seeking. This outcome suggests that setmelanotide administration during a period of heightened bias towards

habitual behavior in humans may open a window of opportunity for intervention and facilitation of goal-directed behavior.

We also find that control of action flexibility by setmelanotide occurs at the lowest dose used in previous research using mice models (Bischof et al., 2016). This pattern suggests that a similar change in behavior may occur in low dose of setmelanotide in humans, as MC4R is also apparent within the human brain and previous translational research shows a similarity in MC4R function between mice models and humans (reviewed in Fatima et al., 2021). In conjunction with the setmelanotide working on hypothalamic systems to combat monogenic obesity, setmelanotide may also act in the dorsal striatum to reduce bias towards habit-based eating.

Conclusion

In summary, we report that an MC4R agonist, setmelanotide, obstructs striatal-dependent decision-making behavior. Drug effects depended on the mouse's training history, such that it obstructed goal-directed behavior following training with schedules of reinforcement that promote behavior flexibility, and it obstructed habitual behavior following training with schedules of reinforcement that induce automatized habits. This research has implications for humans, as setmelanotide was recently approved to treat obesity; our findings suggest that it may impact decision-making behavior.

References

Allen, A. T., Heaton, E. C., Shapiro, L. P., Butkovich, L. M., Yount, S. T., Davies, R. A., . . . Gourley, S. L. Inter-individual variability amplified through breeding reveals control of reward-related action strategies by Melanocortin-4 Receptor in the dorsomedial striatum. *Communications Biology*, *5(116)*. https://doi.org/10.1038/s42003-022-03043-2

Anderson, E. J. P., Çakir, I., Carrington, S. J., Cone, R. D., Ghamari-Langroudi, M., Gillyard, T., . . . Litt, M. J. (2016). Regulation of feeding and energy homeostasis by α-MSH. *Journal of Molecular Endocrinology*, *56*(4), 157-174. https://doi.org/10.1530/JME-16-0014

Balleine, B. W., Delgado, M. R., & Hikosaka, O. (2007). The Role of the Dorsal Striatum in Reward and Decision-Making. *Journal of Neuroscience*, *27*(31), 8161-8165. https://doi.org/10.1523/JNEUROSCI.1554-07.2007

Balleine, B. W., & O'Doherty, J. P. (2009). Human and Rodent Homologies in Action Control: Corticostriatal Determinants of Goal-Directed and Habitual Action. *Neuropsychopharmacology*, *35,* 48-69. https://doi.org/10.1038/npp.2009.131.

Bischof, J. M., Van Der Ploeg, L. H., Colmers, W. F., & Wevrick, R. (2016). Magel2-null mice are hyper-responsive to setmelanotide, a melanocortin 4 receptor agonist. *British Journal of Pharmacology*, *173*(17), 2614-2621. https://doi.org/10.1111/bph.13540

Collet, T. H., Dubern, B., Mokrosinski, J., Connors, H., Keogh, J. M., Mendes de Oliveira, E., Henning, E., Poitou-Bernert, C., Oppert, J. M., Tounian, P., Marchelli, F., Alili, R., Le Beyec, J., Pépin, D., Lacorte, J. M., Gottesdiener, A., Bounds, R., Sharma, S., Folster, C., Henderson, B., … Van der Ploeg, L. (2017). Evaluation of a melanocortin-4 receptor (MC4R) agonist (Setmelanotide) in MC4R deficiency. *Molecular Metabolism, 6*(10), 1321-1329. https://doi.org/10.1016/j.molmet.2017.06.015.

Fatima, M. T., Ahmed, I., Fakhro, K. A., & Akil, A. S. A. S. (2022). Melanocortin-4 receptor complexity in energy homeostasis,obesity and drug development strategies. *Diabetes, Obesity & Metabolism*, *24*(4), 583-598. https://doi.org/10.1111/dom.14618.

Gremel, C. M., & Costa, R. M. (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nature Communications*, *4*, 2264. https://doi.org/10.1038/ncomms3264

Lim, B. K., Huang, K. W., Grueter, B. A., Rothwell, P. E., & Malenka, R. C. (2012). Anhedonia requires MC4R-mediated synaptic adaptations in nucleus accumbens. *Nature*, *487*(7406), 183-189. https://doi.org/10.1038/nature11160

Markham, A. (2021). Setmelanotide: First Approval. *Drugs*, *81*(3), 397-403. https://doi.org/10.1007/s40265-021-01470-9

McKim, T. H., Dove, S. J., Robinson, D. L., Fröhlich, F., & Boettiger, C. A. (2021). Addiction history moderates the effect of prefrontal 10-Hz transcranial alternating current stimulation on habitual action selection. *Journal of Neurophysiology*, *125*(3), 768-780. https://doi.org/10.1152/jn.00180.2020

Pin, X., Grueter, B. A., Britt, J. K., McDaniel, L., Huntington, P. J., Hodge, R., . . . Pieper, A. A. (2013). Double deletion of melanocortin 4 receptors and SAPAP3 corrects compulsive behavior and obesity in mice. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(26), 10759-10764. https://doi.org/10.1073/pnas.1308195110

Sohn, J. W., Harris, L. E., Berglund, E. D., Liu, T., Vong, L., Lowell, B. B., . . . Elmquist, J. K. (2013). Melanocortin 4 receptors reciprocally regulate sympathetic and parasympathetic preganglionic neurons. *Cell*, *152*(3), 612-619. https://doi.org/10.1016/j.cell.2012.12.022

Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *European Journal of Neuroscience*, *19*(1), 181-189. https://doi.org/10.1111/j.1460-9568.2004.03095

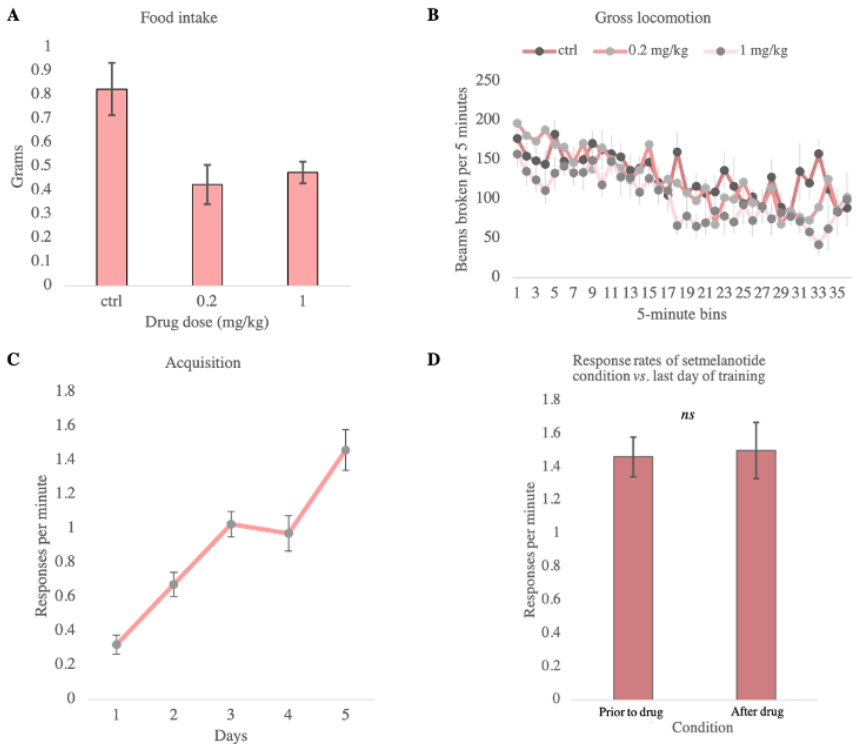
Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2005). Blockade of NMDA receptors in the dorsomedial striatum prevents action–outcome learning in instrumental conditioning. *European Journal of Neuroscience*, *22*(2), 505-512. https://doi.org/10.1111/j.1460-9568.2005.04219.

Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2006). Inactivation of dorsolateral striatum enhances sensitivity to changes in the action–outcome contingency in instrumental conditioning. *Behavioural Brain Research*, *166*(2), 189-196. https://doi.org/10.1016/j.bbr.2005.07.012

Yin, H. H., Ostlund, S. B., Knowlton, B. J., & Balleine, B. W. (2005). The role of the dorsomedial striatum in instrumental conditioning. *European Journal of Neuroscience*, *22*(2), 513-523. https://doi.org/10.1111/j.1460-9568.2005.04218.

Figure 1

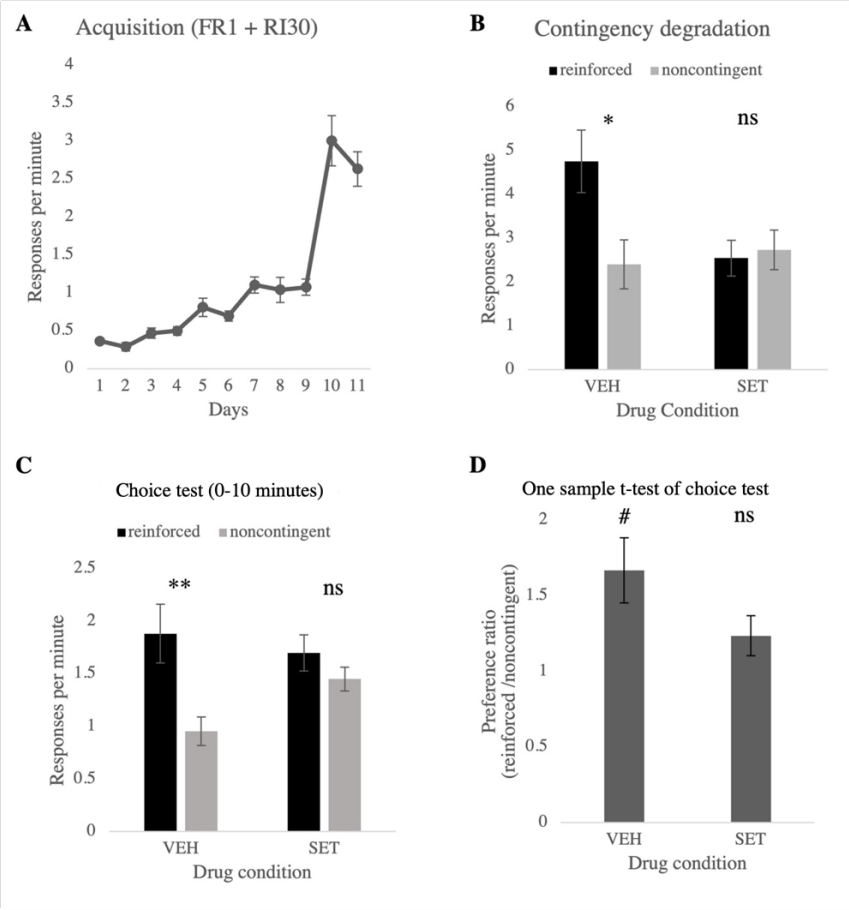
Setmelanotide Dosage Pilot



(A) Decreased food intake in 0.2 mg/kg and 1.0 mg/kg conditions compared to the control condition. (B) No significant difference in gross locomotion. (C) Acquisition (FR1 + RI30) shows significant difference between the first day of training *vs*. the last day of training. (D) Paired samples t-test of ratio of responses of the setmelanotide condition versus the last day of training. * = $p < .05$; ns = not significant

Figure 2

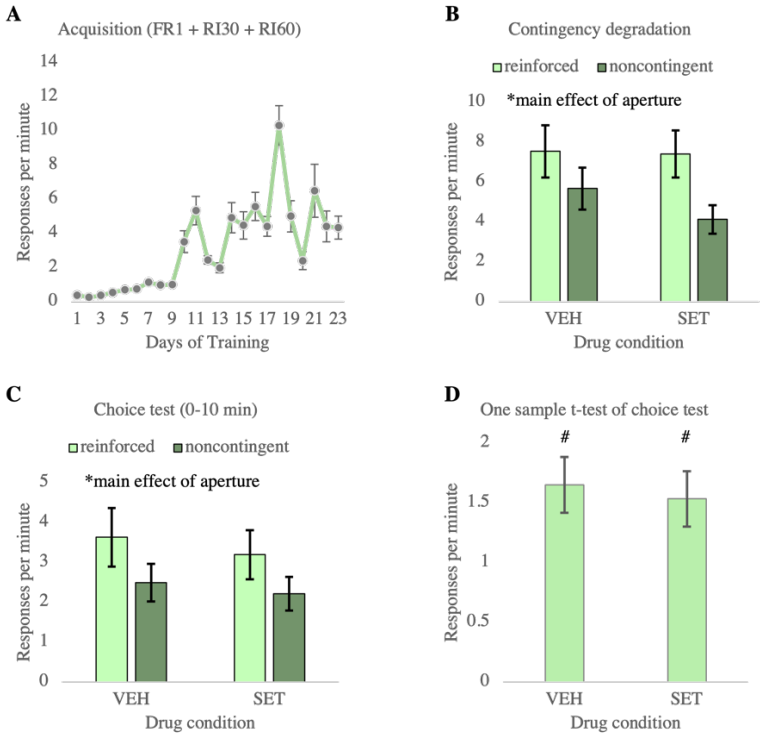
Short Training



(A) Acquisition (FR1 + RI30) shows significant difference between first and last day of training. (B) Contingency degradation results in significant difference between response rates at the reinforced *vs*. noncontingent apertures in vehicle condition and no significant difference between response rates at the reinforced *vs*. noncontingent apertures in setmelanotide condition. (C) 0–10-minute choice test reveals similar pattern of behavior as seen in the contingency degradation results. (D) One sample t-test against test value of 1 of preference ratios during choice test which shows mice responded significantly higher on the reinforced aperture *vs*. the noncontingent aperture in the vehicle condition and not the setmelanotide condition. * = $p < .05$; ** = $p < .01$; # = $p < .05$ for a one sample t-test with a test value of 1; ns = not significant

Figure 3

Extended Training



(A) Acquisition (FR1 + RI60) shows significant increase in response rates between the first day of training *vs*. the last day of training. (B) There was an unexpected main effect of aperture during contingency degradation, which indicates that a bias towards goal-directed behavior was induced in both the vehicle and setmelanotide conditions. (C) There was an unexpected main effect of aperture during the 0–10-minute choice test, which indicates that a bias towards goal-directed behavior was observed in both drug conditions. This might indicate a decreased vulnerability to MC4R agonism when the operant is highly familiar. (D) One sample t-test against test value 1 of preference ratios during choice test shows that mice in both drug conditions preferred the reinforced aperture, further confirming the mice in both drug conditions exhibited a bias towards goal-directed behavior. ** = $p < .01$; # = $p < .05$ for a one sample t-test with a test value of 1