Disruption of model-based decision making by silencing of serotonin neurons in the dorsal raphe nucleus

Yu Ohmura[1*], Kentaro Iwami[2], Srikanta Chowdhury[3,#], Hitomi Sasamori[1], Chiaki Sugiura[1], Youcef Bouchekioua[1], Naoya Nishitani[1], Akihiro Yamanaka[3], and Mitsuhiro Yoshioka[1]

[1] Department of Neuropharmacology, Faculty of Medicine and Graduate School of Medicine, Hokkaido University, N15 W7 Kita-ku, Sapporo 060-8638, Japan

[2] Hokkaido University School of Medicine, N15 W7 Kita-ku, Sapporo 060-8638, Japan

[3] Department of Neuroscience II, Research Institute of Environmental Medicine (RIEM), Nagoya University, Nagoya 464-8601, Japan

#Present Address: Department of Medicine, Naomi Berrie Diabetes Center, Columbia University, 1150 St. Nicholas Avenue, New York, NY 10032, USA

*Corresponding author/Lead Contact: Yu Ohmura

E-mail: yohmura@med.hokudai.ac.jp; gwd0701@yahoo.co.jp

Twitter handle : @OhmuraYu

SUMMARY (207/250 words)

Adapting to changing environmental conditions requires a prospective inference of future actions and their consequences, a strategy also known as model-based decision making [1-3]. In stable environments, extensive experience of actions and their consequences leads to a shift from a model-based to a model-free strategy; whereby behavioral selection is primarily governed by retrospective experiences of positive and negative outcomes. Human and animal studies, where subjects are required to speculate about implicit information and adjust behavioral responses over multiple sessions, point to a role for the central serotonergic system in model-based decision making [4-8]. However, to directly test a causal relationship between serotonergic activity and model-based decision making, phase-specific manipulation of serotonergic activity is needed in a one-shot test, where learning by trial and error is neutralized. Moreover, the serotonergic origin responsible for this effect is yet to be determined. Herein, we demonstrate that optogenetic silencing of serotonin neurons in the dorsal raphe nucleus, but not in the median raphe nucleus, disrupts model-based decision making in lithium-induced outcome devaluation tasks [9-11]. Our data indicate that the serotonergic behavioral effects are not due to increased locomotor activity, anxiolytic effects, or

working memory deficits. Our findings provide insights into the neural mechanisms underlying neural weighting between model-free and model-based strategies.

RESULTS AND DISCUSSION

Validation of transgenic mice expressing Archaerhodopsin T (ArchT) in central serotonergic neurons

TPH staining and ArchT–EYFP expression in coronal slices co-localize to the cell bodies of the dorsal raphe nucleus (DRN) and median raphe nucleus (MRN) (Figures 1A and 1B). One coronal brain slice was used from each bi-transgenic mouse (n = 4). In total, 1057 DRN and 235 MRN cells were counted. Quantitative analysis of the MRN and DRN revealed that over 85% of TPH-positive cells expressed ArchT–EYFP (Figure 1B; DRN: n = 925/1052; MRN: n = 198/231) and that almost all ArchT–EYFP-positive cells expressed TPH (Figure 1B; DRN: n = 925/930; MRN: n = 198/202), indicating that serotonergic neurons selectively expressed ArchT–EYFP in the raphe nuclei in these bi-transgenic mice.

Ex vivo electrophysiology revealed that yellow light illumination induced outward currents (Figures 1C and 1D) and membrane hyperpolarization (Figures 1E and 1F) in ArchT–EYFP-positive neurons. Spontaneous firing of DRN and MRN neurons was

completely inhibited by 0.03 mW/mm^2 yellow light; further, all ArchT–EYFP-positive neurons were inhibited by yellow light illumination (Figures 1G and 1H).

Insert Figure 1 About Here

Silencing serotonin neurons in the DRN reverses lithium devaluation-induced response suppression in the outcome devaluation task

A direct method for assessing model-based vs. model-free decision making is to devaluate the outcome value offline; that is, in a distinct phase from the test being evaluated [12, 13]. Conventionally, the outcome is devalued by intraperitoneal injection of lithium chloride after free access to the reward. In subsequent phases, inhibition of responses to obtain the reward is considered an index of model-based decision making (Figure 2A). In the operant training used in the present study, the schedule of reinforcement was a fixed ratio 1 (FR1). The lights of the central three holes were turned on throughout the task and these holes were never closed, i.e., free operant. We implanted optic fibers into the DRN or MRN of mice and examined whether the inhibition of responses is altered by silencing serotonin neurons in each nucleus.

Representative fiber placements are shown in Figure S1A. Although yellow lights were applied continuously for the 15 min probe test and 15 min post-test food consumption test, we did not observe a clear loss of EYFP expression (Figure S1B). Furthermore, we speculate that serotonergic function was intact during these tests because *in vivo* microdialysis data showed that continuous yellow light application (30 min) to the dorsal raphe nucleus decreased extracellular levels of serotonin in the dorsal striatum of bi-transgenic mice. In addition, the serotonin levels of bi-transgenic mice were comparable to those of littermate controls after light application was stopped (Figure S1C), indicating that continuous light application at this intensity (~1 mW) will not activate or damage serotonin neurons.

Lithium-induced taste aversion gradually reduced the consumption of reward pellets when lithium was injected soon after pellet consumption, whether or not mice were bi-transgenic, and this reduction persisted to the end of the experiment (Figures 2B and 2D). However, when lithium was injected 6 hours after pellet consumption (i.e., unpaired), mice failed to develop taste aversion (Figures 2B and 2D); this ensured that mice associated lithium-induced sickness with the pellets and that lithium itself did not suppress the responses of the mice. In this probe test, we observed a tendency of suppressed responses in the paired injection group of control littermates compared to

the baseline except for one outlier (Figures S2C and S2D), indicating that mice used a model-based strategy. However, the suppression of nose-poke responses was reversed by inhibiting the firing of the DRN serotonin neurons in the paired injection group of bi-transgenic mice (Figure 2C), while the suppression of food consumption persisted during the post-test consumption test (Figure 2B), indicating that silencing serotonin neurons in the DRN disrupts model-based decision making. In contrast to the DRN, the suppression of nose-poke responses was not reversed in bi-transgenic mice that received yellow light stimulation to the MRN (Figure 2E), while *in vivo* yellow light application to the MRN could decrease extracellular levels of serotonin (Figure S1D). Unfortunately, the reward consumption of bi-transgenic mice in the unpaired group was larger than that of other groups even before the aversive conditioning (Figure 2D). Although these mice might have been more motivated to food, it is unlikely that the difference affected the results of the probe test because we used the baseline normalized data for statistical analysis.

Note that nose-poke responses increased up to 140-200% in the group of mice that received unpaired lithium injections (Figures 2C and 2E). In other words, 140-200% is the baseline in the probe test, and the reduction from baseline in the control group (Figures 2C and 2E) demonstrated that responses were suppressed, indicating that

model-based decision making occurs in these mice. Moreover, pre-test light application for 15 min did not affect subsequent responses during the probe test (Figure S2A), suggesting that continuous light application does not damage neurons.

Because we used limited instrumental training to avoid developing habitual responses, mice had few opportunities to learn the task contingencies. To confirm that mice actually learned that poking their noses into holes leads to reward delivery, we conducted a non-contingency training where the reward was delivered at semi-random intervals, regardless of nose poke responses. In this non-contingency control experiment, the number of responses was significantly lower than in the experiment with response-reward contingencies (Figure S2B), indicating that our limited instrumental training was sufficient for developing goal-directed responses and preventing the manifestation of habitual responses.

It should be noted that we used baseline normalized data for statistical analysis because individual differences were so large. For reference, the actual number of responses is shown in Figures S2C–S2F. Although some outliers were observed, the directions and trends were almost the same as those in the normalized data.

Thus, we demonstrated that the DRN, but not the MRN, has a pivotal role in model-based decision making (Figure 2). However, serotonin neurons in the DRN project to

many brain regions including the striatum and several cortical areas [14, 15]. A primate study demonstrated that firing rates in the lateral prefrontal cortex were related to transitive inference [16]. Although it is unclear if rodents have a prefrontal cortex, the prelimbic cortex in rodents seems to be homologous to the lateral prefrontal cortex in primates [17] and may be involved in model-based decision making [18]. Furthermore, a human study showed that the BOLD signal in the dorsomedial striatum was increased under model-based conditions [4]. Moreover, several studies have shown that the orbitofrontal cortex plays a critical role in decision making using model-based representations [10, 19, 20]. Interestingly, these brain regions have dense serotonergic neuronal invasions from the DRN, which is also one of the origins for serotonergic projections to the forebrain [14]. Thus, the prelimbic cortex, dorsomedial striatum, and orbitofrontal cortex are promising candidates for serotonergic modulation from the DRN and for modulating model-based decision making.

Insert Figure 2 About Here

Locomotor activity, not anxiety or working memory, are facilitated by silencing serotonin neurons in the DRN

To assay for potential confounding results with the functional changes induced by our manipulations (e.g., increased locomotor activity and anxiolytic effects), we examined the effects of silencing serotonin neurons on these functions. Silencing serotonin neurons in the DRN of bi-transgenic mice increased the total distance traveled in the open field test compared to control littermates (Figure 3A), consistent with previous findings [21, 22]. However, it did not increase locomotor activity in the elevated plus-maze test (Figure 3B). Optogenetic inhibition of serotonin neurons in the DRN had no effects on anxiety-like behaviors in the open filed test (Figure 3A) or the elevated plus-maze test (Figure 3B). Thus, altered anxiety levels cannot explain the disinhibition of nose-poke responses in the outcome devaluation task while increased locomotor activity could account for these results.

Insert Figure 3 About Here

Another potential confounding factor is working memory deficits, as this can disrupt model-based decision making [23]. Importantly, the serotonergic system is involved in working memory [24], while the suppression of the DRN has no effect [25]. We showed that silencing serotonin neurons in the DRN of bi-transgenic mice did not alter the number of entries into the arms of the maze (Figures S3A and S3B) or the percentage of spontaneous alternations (Figures S3A and S3C) in the Y-maze test. Thus, working memory deficits were not induced by silencing serotonin neurons; congruent with a previous study [25]. These results indicate that working memory deficits cannot explain the disruption of model-based decision making in the present study. However, as the Y-maze test is known to be insensitive to minor impairments of working memory [26], we cannot completely discount the possibility that silencing serotonin neurons caused minor impairments to working memory. However, we can at least conclude that the Y-maze test in our setup detected the impairments in working memory induced by scopolamine (Figures S3D and S3E).

Silencing serotonin neurons in the DRN disinhibits responses to a previously preferred hole in the reward-specific outcome devaluation task

Because response inhibition or disinhibition in the outcome devaluation task (Figure 2A) could be influenced by locomotor activity, we developed a new outcome devaluation task where only one of two different rewards was devalued (Figure 4A). A reduction in the response ratio to a previously preferred hole was regarded as model-based decision making; we used this ratio as an index of biased responses to one of two holes to exclude the effects of altered locomotor activity. Increased locomotor activity was expected to increase responses to both holes similarly, without altering their ratio. To confirm that mice learned that poking their noses into holes results in a reward delivery, we performed a non-contingency control experiment. We demonstrated that the number of responses was significantly lower than in the experiment with response-reward contingencies (Figure S4A).

Lithium-induced taste aversion gradually reduced the consumption of preferred reward pellets regardless of genotype, but did not reduce the consumption of the unpreferred reward (Figures 4A and 4B). In the probe test, the response ratio to preferred/unpreferred holes was genotype-dependent (Figures 4C). For control littermates, the response ratio depended on training/probe sessions and converged at 50% in the probe test, indicating that the mice used a model-based strategy. For bi-transgenic mice, biased responses were maintained by silencing serotonin neurons in the

DRN, indicating that silencing serotonin neurons in the DRN disrupts model-based decision making. However, we do not know whether these bi-transgenic mice searched for food more than control littermates because we did not count the number of magazine entries, regrettably.

Silencing serotonin neurons in the DRN did not reverse the suppression of food consumption in the post-test consumption test (Figure 4B), indicating that silencing serotonin neurons in the DRN left aversive memory unaltered. Furthermore, silencing serotonin neurons in the DRN did not alter responses to the central hole (Figure 4D), indicating that silencing serotonin neurons in the DRN did not alter responses to the inactive hole. In other words, the disinhibition of responses was selective for cues that are associated with a goal.

Insert Figure 4 About Here

Thus, it is unlikely that the observed effects were due to overall disinhibition of nose-poke responses. Rather, we observed that the inhibition of DRN serotonin neurons specifically disinhibited responses toward the hole associated with a devalued outcome; strongly suggesting that this manipulation disrupted the retrieval of updated values for

the outcome from predictive/indirect cues. Alternatively, inhibition of DRN serotonin neurons may leave the retrieval of the updated outcome values intact, but impair the ability to combine independently acquired event memories to mentally simulate/make new predictions.

It should be noted that the mice displayed model-free strategy behavior, in addition to model-based strategy behavior, even after they acquired taste aversion (Figures 2 and 4). If the mice used only a model-based strategy, the responses to the devalued hole would completely disappear; however, substantial responses remained, and thus it appears that mice balance both strategies. Although this is a possibility, the serotonergic system could modulate weighting between model-free and model-based strategies without altering memory function itself, as previously suggested [5, 27].

A previous study demonstrated that tryptophan depletion in healthy subjects impaired model-based learning in reward-seeking conditions whereas tryptophan depletion facilitated model-based learning in punishment-avoiding conditions [5]. In the present study, mice might regard the task as a reward-seeking condition. Alternatively, tryptophan depletion may induce non-serotonergic effects on the kynurenine pathway [28], thus facilitating model-based learning in punishment-avoiding conditions in a serotonin-independent manner. Another human study demonstrated that chronic

selective serotonin reuptake inhibitor (SSRI) treatment in patients with OCD impaired model-based learning during reward-seeking tasks [6], which appeared to be inconsistent with our results; however, the effects of SSRI on extracellular serotonin levels are long-lasting. Chronic SSRI treatment could have various consequences including desensitization of serotonin receptors and neurotrophic changes [29-31]. Indeed, these secondary effects could dampen model-based learning. Alternatively, the phasic activity of serotonin neurons may be required to facilitate model-based decision making or learning. If this is the case, tonic activation induced by SSRI should disrupt model-based decision making or learning. Future studies are needed to record the activity of serotonin neurons in the DRN during model-based decision making with precise time resolution. If these data become available, we could address this issue by fully taking advantage of the optogenetic temporal specificity; we can inhibit serotonin neurons at specific times when activity will be enhanced for decision making.

A common feature of these previous studies is that they assessed model-based decision making indirectly by testing online behavioral adjustments to outcome manipulations. An online outcome devaluation has the advantage of measuring behavioral adaptations trial by trial, and the data can be fit to a Q-learning model [1]. However, the offline outcome devaluation and one-shot tests in the present study would

require more prospective inferential processes, where learning by trial and error is neutralized, resulting in a more direct measurement of model-based decision making. In theory, this would make learning-phase specific manipulations possible, which we took advantage of in the present study. To integrate online and offline methods with our findings, we need to alternate between them, a strategy we intend to employ during future studies.

Conclusions

In summary, we found that the inhibition of serotonin neurons in the DRN, but not in the MRN, disrupted model-based decision making. It would be difficult to adapt to these changing/complicated situations using only model-free strategies [1, 2]. Previous studies have shown that patients with substance abuse [32], schizophrenia [33], obsessive-compulsive disorder [6], or under chronic stress [30] manifest impaired model-based learning. Moreover, serotonergic drugs have been used for treating several psychiatric disorders. Elucidating the neural mechanisms underlying model-based decision making will provide insights for treating these common impairments.

ACKNOWLEDGEMENTS

The authors thank JAM Post (https://www.jamp.com/index.cfm) for the English language review. This work was supported by the Takeda Science Foundation and a Grant-in-Aid for Scientific Research on Innovative Areas awarded to Y.O. [JSPS KAKENHI Grant Number: JP19H04976]. We would like to thank Tomoko Furukawa, Aki Tanimori, and Yuki Kumakura, for help with breeding the transgenic mice.

AUTHOR CONTRIBUTIONS

Conceptualization, Y.O.; Investigation, Y.O., K.I., S.C., H.S., C.S., and Y.B.; Writing – Original Draft, Y.O.; Writing – Review & Editing, Y.B., N.N., A.Y., and M.Y.; Funding Acquisition, Y.O.; Resources, A.Y.; Supervision, M.Y.

DECLARATION OF INTEREST

The authors declare no conflicts of interest.

Titles and Figure Legends

Figure 1. Validity of Tph2-tTA::tetO-ArchT–EYFP bi-transgenic mice.

(A) Mice expressing central serotonergic neuron-specific ArchT were obtained by crossing a tetO-ArchT BAC transgenic mouse line with a Tph2-tTA line. **(B)** TPH staining and ArchT–EYFP expression in coronal slices display co-localization in cell bodies of the DRN and MRN. One coronal brain slice was used from each bi-transgenic mouse (n = 4). In total, 1057 DRN and 235 MRN cells were counted. **(C)** Photocurrent data from whole-cell voltage-clamp recordings of 5-HT neurons at -60 mV hold during 2 sec of continuous yellow light illumination with different intensities (see methods). **(D)** Summary of data in **C** (DRN: n = 6, and MRN neurons: n = 7). **(E)** Whole-cell current-clamp recordings from 5-HT neurons in the raphe nucleus. Light-evoked hyperpolarized potentials during continuous application (2 sec) of yellow light with different intensities. **(F)** Summary of data in **E** (DRN (n = 8) and MRN neurons (n = 7)). **(G)** Representative traces showing that the spontaneous firing of both DRN and MRN neurons was completely inhibited by 0.03 mW/mm^2 of yellow light. **(H)** Summary of data in **G** (DRN: n = 8, $F_{1.20,8.42} = 19.96$, $P = 0.0014$; and MRN neurons: n = 8, $F_{1.17,8.19} = 30.86$, $P < 0.001$). The data are presented as the means ± S.E.M.

Figure 2. Silencing serotonin neurons in the DRN, but not in the MRN, reversed lithium devaluation-induced response suppression in the outcome devaluation task. **(A)** Schematic illustrations for the primary part of the outcome devaluation task procedure. Either bi-transgenic mice or control littermates received continuous yellow light applications to the DRN or MRN during the probe test and post-test consumption tests. **(B)** The paired lithium injection group gradually reduced their consumption of reward pellets over days, and this taste aversion persisted to the end of the experiment irrespective of genotype, while the unpaired lithium injection group did not show taste aversion (days × paired/unpaired interaction effect, $F_{1.89,70.05}$= 29.95, P < 0.001) **(C)** Yellow light-induced silencing of serotonin neurons in the DRN of bi-transgenic mice disinhibited the nose-poke responses compared to control littermates (genotype × paired/unpaired interaction effect, $F_{1, 37}$ = 4.49, P = 0.041; control vs. bi-transgenic in paired group, P = 0.0079). **(D)** The paired lithium injection group gradually reduced the consumption of reward pellets over days, and the taste aversion persisted to the end of the experiment irrespective of genotype, while the unpaired lithium injection group did not develop taste aversion (days × paired/unpaired interaction effect, $F_{2.12,44.47}$ = 32.59, P < 0.001). **(E)** Yellow light-induced silencing of serotonin neurons in the MRN of bi-transgenic mice did not disinhibit nose pokes compared to control littermates (genotype × paired/unpaired

interaction effect, $F_{1, 21} = 0.51$, $P = 0.48$; a main effect of paired/unpaired, $F_{1, 21} = 6.20$, $P = 0.021$). The data are presented as the means ± S.E.M. CL, control littermates. BI, bi-transgenic. See also Figures S1, S2 and S3.

Figure 3. Effects of inhibiting serotonin neurons on locomotor activity and anxiety-like behavior.

(A) Bi-transgenic (n = 9) or littermate control (n = 10) mice received continuous yellow light applications to the DRN during the 10 min of the open field test. Silencing serotonin neurons of bi-transgenic mice did not alter the time spent in the central arena ($t_{17}= 0.02$, $p = 0.98$) but increased the total distance traveled compared to control littermates ($t_{17} = 2.26$, $p = 0.037$). **(B)** Bi-transgenic (n = 9) or littermate control (n = 10) mice received continuous yellow light application to the DRN during the 5 min of the elevated plus-maze test. Optogenetic inhibition of serotonin neurons did not affect the total distance traveled ($t_{17} = 1.18$, $p = 0.26$) or the time spent in open arms ($t_{17} = 1.12$, $p = 0.28$). The data are presented as the means ± S.E.M. *$P < .05$.

Figure 4. Silencing serotonin neurons in the DRN reversed lithium devaluation-induced response suppression in the reward-specific outcome devaluation task.

(A) Schematic illustrations for the primary part of the partial outcome devaluation task procedure. Either bi-transgenic mice (n = 11) or control littermates (n = 10) received continuous yellow light application to the DRN during the probe test and post-test consumption tests. **(B)** Lithium-induced taste aversion gradually reduced their consumption of preferred reward pellets, independent of genotype (an interaction effect of day × preferred/unpreferred, $F_{2,76} = 58.68$, $P < 0.001$). **(C)** Yellow light application to the DRN of control littermates did not disinhibit the nose-poke responses to the devalued hole but did in bi-transgenic mice (an interaction effect of genotype × preferred/unpreferred, $F_{1,38} = 5.11$, $P = 0.03$). **(D)** Yellow light application to the DRN did not affect responses to the central hole. The data are presented as the means ± S.E.M. $^{\#}P < .05$, unpreferred vs. preferred. $^{*}P < .05$. See also Figures S4.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Yu Ohmura (yohmura@med.hokudai.ac.jp).

Materials availability

This study did not generate new unique reagents. Transgenic mice are available from RIKEN BRC(RRID:IMSR_RBRC05842, IMSR_RBRC05846).

Data and code availability

Original/source data for figures in the paper is available (DOI: 10.17632/6wkmndk4ss.1).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animal Models

We used adult male and female C57BL/6N mice or transgenic mice on a C57BL/6N background and aged >56 days prior to the start of behavioral experiments. We used male mice only for the lithium-induced outcome devaluation tasks because female mice show faster habit formation [11], making it difficult to observe model-based decision making. Indeed, our preliminary study indicated that we could not measure model-based decision making in our assays with female mice (data not shown). All mice were housed in groups prior to starting behavioral experiments or surgery. C57BL/6N mice were supplied from Nippon SLC Co. Ltd (Hamamatsu, Japan). The animal rooms were under an alternating light-dark cycle (light from 7 p.m. to 7 a.m.) at 25 ± 2°C and a relative humidity of 40–50%. All tests were performed during the dark period. The treatment of animals complied with the Guidelines for the care and use of Laboratory Animals of the Animal Research Committee of Hokkaido University.

We created ArchT-expressing animals in central serotonergic neurons by crossing the tetracycline operator (tetO)-ArchT–EYFP BAC transgenic mouse line (RRID:IMSR_RBRC05842) with a Tph2-tTA BAC transgenic mouse line (RRID:IMSR_RBRC05846). The generation of tetO-ArchT BAC transgenic mice and Tph2-tTA BAC transgenic mice have been described in previous reports [22, 34, 35]. These mice were backcrossed to the C57BL/6N strain for more than six generations to ensure genetic homogeneity in our experiments.

METHOD DETAILS

Immunohistochemistry

Immunohistochemistry and confocal microscopy were performed as described previously [36] with minor changes. In brief, after intracardial perfusion with 4% paraformaldehyde in PBS at pH 7.2, brains were post-fixed overnight, placed in 0.1 M PB containing 20% sucrose, and sectioned at a thickness of 50 µm. The primary antibodies used were mouse anti-GFP (1:1 000; 012-20461, WAKO, Osaka, Japan) and sheep anti-tryptophan hydroxylase (TPH; 1:1 000; AB1541, Millipore, Burlington, MA, USA). The secondary antibodies used were donkey anti-mouse Alexa 488 (Invitrogen, Carlsbad, CA, USA) and donkey anti-sheep indocarbocyanine (Cy3; Jackson ImmunoResearch, PA, USA). Images were captured with a confocal laser-scanning microscope (FV1000, Olympus, Tokyo, Japan).

Acute Slice *In Vitro* Electrophysiology

Brain slice preparation and patch-clamp recording methods were modified from previously described protocols [37]. Briefly, Tph2-tTA::tetO-ArchT mice of either sex, aged 2-3 months, were decapitated and brains were quickly isolated and chilled in ice-cold cutting solution (in mM: 110 K-gluconate, 15 KCl, 0.05 EGTA, 5 HEPES, 26.2 $NaHCO_3$, 25 glucose, 3.3 $MgCl_2$, and 0.0015 (±)-3-(2-Carboxypiperazin-4-yl)propyl-1-phosphonic acid), then gassed with 95% O_2 and 5% CO_2. Coronal brain slices of 300-µm thickness containing raphe nuclei were generated using a vibratome (VT-1200S; Leica, Wetzlar, Germany), and were incubated in bath solution (in mM: 124 NaCl, 3 KCl, 2 $MgCl_2$, 2 $CaCl_2$, 1.23 NaH_2PO_4, 26 $NaHCO_3$, and 25 glucose) gassed with 95% O_2 and 5% CO_2 at 35°C for at least 60 min. Slices were then incubated in the same incubation chamber at room temperature for another 30 min, for recovery.

Patch-clamp recordings were performed in neurons that expressed EYFP. Brain slices were perfused with bath solution gassed with 95% O_2 and 5% CO_2 using a peristaltic pump (Dynamax; Rainin, Oakland, CA, USA) at 1.5 mL/min in a recording chamber (RC-26G; Warner Instruments, Hamden, CT, USA) installed on an upright fluorescence microscope stage (BX51WI; Olympus, Tokyo, Japan). An infrared camera (C3077-78; Hamamatsu Photonics, Hamamatsu, Japan) was installed in the fluorescence microscope along with an electron-multiplying charge-coupled device camera (Evolve 512 delta; Photometrics, Tucson, AZ, USA); both images were displayed on separate monitors. Recordings were conducted using glass micropipettes (4-6 M

USA). Output signals were low-pass filtered at 5 kHz and digitized at a 10-kHz sampling rate. Patch-clamp data were then recorded through an analog-to-digital (AD) converter (Digidata 1550A; Molecular Devices) using pClamp 10.2 software (Molecular Devices). Yellow light (575 ± 12.5 nm) was generated by a light-emitting diode (Spectra light engine; Lumencor, Beaverton, OR, USA) and was guided to the microscope stage with a 1-cm diameter optical fiber via the 40× objective lens. Yellow light intensity (in %) of 1, 2, 5, 10, 20, 40, and 80 indicates light power density (in mW/mm^2) of 0.01, 0.02, 0.03, 0.06, 0.23, 0.47, and 0.98, respectively. Whole-cell voltage-clamp recordings were performed in the presence of synaptic blockers (400 µM picrotoxin, 50 µM AP-5, and 20 µM CNQX) in bath solution and a sodium channel blocker (1 mM QX-314) in a KCl-based internal solution. Light-induced sustained currents and hyperpolarized membrane potentials were measured just before light cessation. Yellow light of 5% intensity (0.03 mW/mm^2) was applied during cell-attached recordings for 10 sec.

In Vivo Microdialysis

We conducted microdialysis with HPLC-ECD, as previously described [22]. Samples from the dorsal striatum or ventral hippocampus were collected every 15 min, and two consecutive samples were averaged for each phase: pre, light, and post phases. During the light phase, continuous yellow light was applied to the DRN or MRN for 30 min.

Surgical Procedure

Mice were anesthetized with isoflurane and fixed in a stereotaxic frame (Narishige, Tokyo, Japan). For light applications to the DRN or MRN, an optic fiber with a mirror tip at 45° (MA45; Doric Lenses, Quebec, Canada) was implanted with coordinates 4.5 mm posterior to the bregma, 1.0 mm lateral to the midline, and 2.3 (DRN) or 3.7 (MRN) mm ventral to the dura [38]. After surgery, ointments containing antibiotics and steroids (Dolmycin, Zeria Pharmaceutical Co., Ltd., Tokyo, Japan; Kenalog, Bristol-Myers Squibb, New York, NY, USA) were applied to the wounds. A piece of jelly containing carprofen (MediGel, Clear H_2O, Portland, ME, USA) was placed in the home cage, and the mice were housed individually and allowed to recover for 7 days prior to behavioral experiments.

In Vivo Light Illumination Procedure

For light applications to the DRN or MRN, yellow (575 nm) light was generated by a SPECTRA 2-LCR-XA light engine (Lumencor, Beaverton, OR, USA), and the light intensity at the fiber tip was approximately 1 mW/mm^2. A fiber-optic rotary joint (Doric Lenses, Quebec, Canada) was used for unrestricted in vivo illumination. The light was controlled via TTL pulses driven by a stimulator (Nihon Kohden, Tokyo, Japan).

Verification of Optical Fiber Placements

After the experiments were completed, optical fiber placement was verified under a microscope. In brief, after intracardial perfusion with 4% paraformaldehyde in PBS at pH 7.2, brains were post-fixed overnight in 0.1 M PB containing 20% sucrose. Coronal sections (of 50 µm thickness) were cut on a cryostat and mounted onto slides. After drying, the sections were stained with toluidine blue, and cannula placements were verified under a microscope according to the brain atlas [38]. Data from mice with incorrect placements were excluded from all analyses.

Outcome Devaluation Task

We used aluminum operant chambers measuring W22 × D26 × H18 cm (Med Associates Inc., St. Albans, VT, USA). The curved rear wall of each chamber contained nine holes. Each hole had an infrared photocell beam to detect nose-poke responses, and a yellow LED light was located behind each hole. Every other hole was sealed so that only the three centrally positioned ports were accessible. A food magazine was located on the opposite wall of the chamber, and a house light was located at the top of this wall. The food magazine had a yellow LED light located on its ceiling. The apparatus was controlled by a computer program written in the MED-PC programming language (Med Associates Inc., St. Albans, VT, USA).

Mice were individually housed, and each mouse underwent 1 min of handling once a day for 3 consecutive days. After completing the handling acclimatization, a 9-day procedure was used for the outcome devaluation task as described below.

Day 1: Usual food (CE-2; CLEA JAPAN Inc., Tokyo, Japan) was removed from the home cage and each mouse was allowed to eat 0.4 g of reward pellets (20 mg each, dustless precision pellets F0071, Bio-serv, Frenchtown, NJ, USA) in their home cages for a day to habituate the mouse to this food.

Day 2: Each mouse was moved into the operant chamber and instrumental training was started. A reward pellet was always delivered to the food magazine 1 min after the mouse entered the magazine during a 45-min session. After the session, each mouse was allowed to eat usual food (CE-2; CLEA JAPAN Inc., Tokyo, Japan) in the home cage for 3 hours.

Days 3 – 4: In the operant chamber, the mouse was required to nose poke in any of the 3 holes to obtain a reward pellet during a daily session. Each session was conducted for 30 min or until the number of nose pokes became more than 8, whichever came first. All three holes were illuminated for the session. In a non-contingency control group, the reward was delivered at semi-random intervals (112, 225, or 338 sec) regardless of nose poke response. The intervals were chosen based on the number of obtained rewards in a response-reward contingency group so

that the number of rewards was almost the same between groups. After the session, each mouse was allowed to eat usual food (CE-2) in the home cage for 3 hours.

Day 5: A session was conducted in the operant chamber for 15 min without limiting the number of nose pokes. If the number of nose pokes was less than 5, the mouse was omitted from the study. In the non-contingency control group, the reward was delivered at semi-random intervals (30, 60, or 90 sec) regardless of the nose poke response. After the session, each mouse was allowed to eat usual food (CE-2) in the home cage for 3 hours.

Days 6 – 8: In the operant chamber, the mouse was confined to an opaque plastic box, but was still able to access the food magazine. The magazine was filled with 2 g of reward pellets and the mouse was allowed to eat them for 15 min. Immediately after (paired) or 6 hours after (unpaired) removing the mouse from the chamber, the mouse received an intraperitoneal injection of lithium chloride (LiCl, 0.15 M) at a volume of 40 ml/kg. This dose of LiCl should induce illness and taste aversion of reward pellets [9]. Three hours after the session, each mouse was allowed to eat usual food (CE-2) in the home cage for 3 hours.

Day 9: The mouse was placed in the operant chamber without the plastic box again as conducted on Day 5, but nose pokes were not rewarded. The number of nose pokes into holes during a 15-min probe test was counted. We calculated the ratio of nose-poke responses in the probe test to that on Day 5 and regarded the response inhibition as model-based decision making. Subsequently, mice received a 15-min post-test consumption test in the operant chamber without the plastic box. The magazine was filled with 2 g of reward pellets and the mouse was allowed to eat them for 15 min. A group of mice received continuous yellow light throughout the probe test and post-test consumption test to examine whether the inhibition of serotonin neurons disrupts model-based decision making or taste aversion. Another group of mice received yellow light in their home cages before the probe test for 15 min and during the post-test consumption test to determine if the light damages neurons and if this damage affects behavior during the probe test. After completing the light application in their home cages, we waited for 10 min before starting the probe test because ArchT-induced inhibition will last for 5 to 10 min even after the light application is stopped [35].

Open Field Test

The acrylic box (30 × 30 × 30 cm^3) was covered by rough-surfaced polypropylene sheets. The illumination of the room was set to 150 lux. The behavior of each mouse was monitored by a CCD

camera over the 10-min testing period and automatically analyzed using a software package (LimeLight, Actimetrics, USA). The total distance traveled was used as a measure of locomotor activity. The time spent in the central arena (10 cm × 10 cm square) was analyzed to assess anxiety-like behavior.

Elevated Plus-Maze Test

Mice experienced the open field test previously were used. The elevated plus-maze test was performed as described preciously [22]. In brief, the apparatus was made of wood and consisted of two open arms (25 × 5 cm) and two closed arms (25 × 5 cm) that extended from the central platform (5 × 5 cm). The closed arms were surrounded by 20-cm-high side walls. The maze was elevated 40 cm above the floor, and the illumination of the room was set to 200 lux. The behavior of each mouse was monitored by a CCD camera during a 5-min testing period and automatically analyzed using a software package (LimeLight). The total distance traveled and the number of total entries/exits to/from each arm were used as measures of locomotor activity. The time spent in the open arms was used as a measure of anxiety-related behavior, because mice typically avoid open arms [39].

Reward-Specific Outcome Devaluation Task

The apparatus for the reward-specific outcome devaluation task was adapted from that used in the previous outcome devaluation task; in the present study, two plastic food magazines were used to receive only one kind of reward each. One magazine was white-colored and associated with chocolate-flavored pellets (20 mg each, dustless precision pellets F05301, Bio-serv) while the other one was black-colored and associated with sucrose pellets (20 mg, dustless precision pellets F0071, Bio-Serv). These plastic magazines were designed using Tinkercad software and built using a 3D printer (Creator Pro, Apple Tree Co., Ltd.).

Mice were individually housed, and each mouse underwent 1 min of handling once a day for 3 consecutive days. After completing the handling acclimatization, a 5-day procedure was used for a reward-specific outcome devaluation task as described below.

Day 1: Daily food (CE-2; CLEA JAPAN Inc., Tokyo, Japan) was removed from the home cage and each mouse was allowed to eat 0.8 g of reward pellets, including 0.4 g of sucrose pellets (F0071, Bio-serv) and 0.4 g of chocolate flavored pellets (F05301, Bio-serv) in their home cage for a day to habituate mice to these foods.

Day 2: Each mouse received an instrumental training session inside the operant chamber. A reward pellet was always delivered to each food magazine (i.e., two pellets in total) 1 min after the mouse entered both magazines during a 45-min session. Subsequently, each mouse received

a second operant training session in the operant chamber. They were required to nose poke into any of the two-sided holes to obtain a reward pellet. A nose poke into a side hole always resulted in the delivery of a sucrose pellet, while a nose poke into the other side hole always resulted in the delivery of a chocolate flavored pellet. Nose pokes into the central hole were also counted, but kept inactive throughout the experiment (i.e., never resulted in a reward delivery). The operant training session ended after 6 hours or if the total number of nose pokes into the side holes exceeded 50. The two side holes were illuminated during the sessions while the central hole had no light. We calculated the ratio of nose-poke responses into each side hole for each mouse to estimate the hole/reward preference of each animal. In a non-contingency control group, the reward was delivered at semi-random intervals (107, 215, or 323 sec) regardless of nose poke response. The operant training session ended after 166 min or if the total number of nose pokes into the side holes exceeded 50. The intervals and the session time limit were chosen based on the number of obtained rewards in a response-reward contingency group so that the number of rewards was almost the same between groups.

Day 3 – 4: The mouse was confined into an opaque plastic box put inside the operant chamber, such that the animal had access to only one of the two food magazines. The magazine was filled with 2 g of chocolate flavored or sucrose reward pellets and the mouse was allowed to eat them for 15 min. Immediately after removing the mouse from the chamber, it received an intraperitoneal injection of 40ml/kg of saline or lithium chloride (LiCl, 0.15 M). The devalued reward was that for which the mouse developed a preference during the Day 2 operant training session, whereas the other type of reward was not devalued. The same session was repeated twice with a 6-hour interval, but the type of reward and injection solution were counterbalanced. For example, mice that first received sucrose pellets and saline injection during the first session, received chocolate flavored pellets and LiCl injection during the second session 6 hours later.

Day 5: The mouse was tested in the operant chamber without the plastic box in the same manner as during the Day 2 operant training session, but now nose-poke responses did not result in reward delivery. The number of nose pokes to each hole was counted during a 30-min probe test. We calculated the ratio of nose pokes into each hole and compared the ratio in the probe test with the ratio obtained during Day 2 operant training. The reduction of the response ratio to a previously preferred hole was regarded as model-based decision making. We used the ratio as an index of biased response to one of two holes, to exclude the effects of altered locomotor activity. Increased locomotor activity is expected to increase the responses to both holes similarly without altering their ratio. In addition, we calculated the ratio of nose pokes into the central hole during the probe test to that during the training session, and compared the ratio between the genotypes.

A high ratio in only bi-transgenic mice would suggest that our optogenetic manipulation disinhibited responses toward cues previously inhibited by any other reason. Conversely, no difference in the ratio between groups would suggest that our optogenetic manipulation selectively affected cues previously associated with goal-directed behaviors. After the probe test, mice received a 15-min post-test consumption test in the operant chamber without the plastic box. A magazine was filled with 2 g of sucrose pellets while the other magazine was filled with 2 g of chocolate flavored pellets. Mice were allowed to freely consume pellets for 15 min. Mice received continuous yellow light throughout the probe test and post-test consumption test to examine whether the inhibition of serotonin neurons disrupts model-based decision making or taste aversion.

Y-Maze Test

The Y-maze test was performed as described previously [40]. In brief, the apparatus was made of wood and consisted of three arms (10 cm-wide, 45 cm-length, and 35 cm-high-walls) radiating out from a center platform (10 × 10 × 10 cm triangle). The illumination of the room was set to 20 lux. The behavior of each mouse was monitored by a CCD camera during an 8-min testing period. Entry into an arm was defined as when the hind paws of the mouse were completely within the arm. Spontaneous alternation was counted when the mouse entered all three arms in overlapping triplet sets. The percentage of alternation was calculated as (successive triplet sets/total number regarded as working memory impairment [41].

QUANTIFICATION AND STATISTICAL ANALYSIS

Electrophysiological data were analyzed using one-factor repeated-measures ANOVA. Most behavioral parameters in each group (DRN or MRN) were analyzed separately using a two-factor ANOVA. For two-group comparisons, Student's t-test was used. If Levene's test showed significant results, Welch's t-test was used instead of the Student's t-test. In cases where a within-subject design was used, paired t-tests were used. We used a one-sample Wilcoxon test when all values in one of the two groups were identical.. Time-course changes of food consumption were analyzed using a two-factor (or a three-factor) mixed-design ANOVA with the day as the within-subject factor and the genotype as the between-subject factor (and preferred/unpreferred reward as a between-subject factor). When there were significant interactions, simple main effects were calculated. If Mauchlys sphericity test was significant, the Greenhouse–Geisser correction was used. All data are expressed as the mean ± standard error of the mean (S.E.M.). The alpha

level was set at 0.05 for all comparisons. All statistical procedures were conducted using SPSS (version 23.0) or GraphPad Prism (version 9.0).

References

1. Daw, N.D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat Neurosci *8*, 1704-1711.
2. Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? Neural Netw *12*, 961-974.
3. McDannald, M.A., Takahashi, Y.K., Lopatina, N., Pietras, B.W., Jones, J.L., and Schoenbaum, G. (2012). Model-based learning and the contribution of the orbitofrontal cortex to the model-free world. Eur J Neurosci *35*, 991-996.
4. Fermin, A.S., Yoshida, T., Yoshimoto, J., Ito, M., Tanaka, S.C., and Doya, K. (2016). Model-based action planning involves cortico-cerebellar and basal ganglia networks. Sci Rep *6*, 31378.
5. Worbe, Y., Palminteri, S., Savulich, G., Daw, N.D., Fernandez-Egea, E., Robbins, T.W., and Voon, V. (2016). Valence-dependent influence of serotonin depletion on model-based choice strategy. Mol Psychiatry *21*, 624-629.
6. Voon, V., Baek, K., Enander, J., Worbe, Y., Morris, L.S., Harrison, N.A., Robbins, T.W., Ruck, C., and Daw, N. (2015). Motivation and value influences in the relative balance of goal-directed and habitual behaviours in obsessive-compulsive disorder. Transl Psychiatry *5*, e670.
7. Boulougouris, V., and Robbins, T.W. (2010). Enhancement of spatial reversal learning by 5-HT2C receptor antagonism is neuroanatomically specific. J Neurosci *30*, 930-938.
8. Matias, S., Lottem, E., Dugue, G.P., and Mainen, Z.F. (2017). Activity patterns of serotonin neurons underlying cognitive flexibility. Elife *6*.
9. Gourley, S.L., Swanson, A.M., Jacobs, A.M., Howell, J.L., Mo, M., Dileone, R.J., Koleske, A.J., and Taylor, J.R. (2012). Action control is mediated by prefrontal BDNF and glucocorticoid receptor binding. Proc Natl Acad Sci U S A *109*, 20714-20719.
10. Pickens, C.L., Saddoris, M.P., Setlow, B., Gallagher, M., Holland, P.C., and Schoenbaum, G. (2003). Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. J Neurosci *23*, 11078-11084.
11. Quinn, J.J., Hitchcott, P.K., Umeda, E.A., Arnold, A.P., and Taylor, J.R. (2007). Sex chromosome complement regulates habit formation. Nat Neurosci *10*, 1398-1400.

12. Dickinson, A. (1985). Actions and Habits - the Development of Behavioral Autonomy. Philos T Roy Soc B *308*, 67-78.
13. Balleine, B.W., and ODoherty, J.P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology *35*, 48-69.
14. Muzerelle, A., Scotto-Lomassese, S., Bernard, J.F., Soiza-Reilly, M., and Gaspar, P. (2016). Conditional anterograde tracing reveals distinct targeting of individual serotonin cell groups (B5-B9) to the forebrain and brainstem. Brain Struct Funct *221*, 535-561.
15. Azmitia, E.C., and Segal, M. (1978). An autoradiographic analysis of the differential ascending projections of the dorsal and median raphe nuclei in the rat. J Comp Neurol *179*, 641-667.
16. Pan, X., Fan, H., Sawa, K., Tsuda, I., Tsukada, M., and Sakagami, M. (2014). Reward inference by primate prefrontal and striatal neurons. J Neurosci *34*, 1380-1396.
17. Uylings, H.B., Groenewegen, H.J., and Kolb, B. (2003). Do rats have a prefrontal cortex? Behav Brain Res *146*, 3-17.
18. Hart, G., and Balleine, B.W. (2016). Consolidation of Goal-Directed Action Depends on MAPK/ERK Signaling in Rodent Prelimbic Cortex. J Neurosci *36*, 11974-11986.
19. Jones, J.L., Esber, G.R., McDannald, M.A., Gruber, A.J., Hernandez, A., Mirenzi, A., and Schoenbaum, G. (2012). Orbitofrontal cortex supports behavior and learning using inferred but not cached values. Science *338*, 953-956.
20. Lucantonio, F., Caprioli, D., and Schoenbaum, G. (2014). Transition from model-based to model-free behavioral control in addiction: Involvement of the orbitofrontal cortex and dorsolateral striatum. Neuropharmacology *76 Pt B*, 407-415.
21. Correia, P.A., Lottem, E., Banerjee, D., Machado, A.S., Carey, M.R., and Mainen, Z.F. (2017). Transient inhibition and long-term facilitation of locomotion by phasic optogenetic activation of serotonin neurons. Elife *6*.
22. Ohmura, Y., Tanaka, K.F., Tsunematsu, T., Yamanaka, A., and Yoshioka, M. (2014). Optogenetic activation of serotonergic neurons enhances anxiety-like behaviour in mice. Int J Neuropsychopharmacol *17*, 1777-1783.

23. Otto, A.R., Raio, C.M., Chiang, A., Phelps, E.A., and Daw, N.D. (2013). Working-memory capacity protects model-based learning from stress. Proc Natl Acad Sci U S A *110*, 20941-20946.
24. Cowen, P., and Sherwood, A.C. (2013). The role of serotonin in cognitive function: evidence from recent studies and implications for understanding depression. J Psychopharmacol *27*, 575-583.
25. Warburton, E.C., Harrison, A.A., Robbins, T.W., and Everitt, B.J. (1997). Contrasting effects of systemic and intracerebral infusions of the 5-HT1A receptor agonist 8-OH-DPAT on spatial short-term working memory in rats. Behav Brain Res *84*, 247-258.
26. Klinkenberg, I., and Blokland, A. (2010). The validity of scopolamine as a pharmacological model for cognitive impairment: a review of animal behavioral studies. Neurosci Biobehav Rev *34*, 1307-1350.
27. Doya, K. (2002). Metalearning and neuromodulation. Neural Netw *15*, 495-506.
28. Young, S.N. (2013). Acute tryptophan depletion in humans: a review of theoretical, practical and ethical aspects. J Psychiatry Neurosci *38*, 294-305.
29. Gray, N.A., Milak, M.S., DeLorenzo, C., Ogden, R.T., Huang, Y.Y., Mann, J.J., and Parsey, R.V. (2013). Antidepressant treatment reduces serotonin-1A autoreceptor binding in major depressive disorder. Biol Psychiatry *74*, 26-31.
30. Yamauchi, M., Tatebayashi, T., Nagase, K., Kojima, M., and Imanishi, T. (2004). Chronic treatment with fluvoxamine desensitizes 5-HT2C receptor-mediated hypolocomotion in rats. Pharmacol Biochem Behav *78*, 683-689.
31. Lyttle, K., Ohmura, Y., Konno, K., Yoshida, T., Izumi, T., Watanabe, M., and Yoshioka, M. (2015). Repeated fluvoxamine treatment recovers juvenile stress-induced morphological changes and depressive-like behavior in rats. Brain Res *1616*, 88-100.
32. Voon, V., Derbyshire, K., Ruck, C., Irvine, M.A., Worbe, Y., Enander, J., Schreiber, L.R., Gillan, C., Fineberg, N.A., Sahakian, B.J., et al. (2015). Disorders of compulsivity: a common bias towards learning habits. Mol Psychiatry *20*, 345-352.
33. Culbreth, A.J., Westbrook, A., Daw, N.D., Botvinick, M., and Barch, D.M. (2016). Reduced model-based decision-making in schizophrenia. J Abnorm Psychol *125*, 777-787.
34. Tanaka, K.F., Matsui, K., Sasaki, T., Sano, H., Sugio, S., Fan, K., Hen, R., Nakai, J., Yanagawa, Y., Hasuwa, H., et al. (2012). Expanding the repertoire of

optogenetically targeted cells with an enhanced gene expression system. Cell Rep *2*, 397-406.

35. Tsunematsu, T., Tabuchi, S., Tanaka, K.F., Boyden, E.S., Tominaga, M., and Yamanaka, A. (2013). Long-lasting silencing of orexin/hypocretin neurons using archaerhodopsin induces slow-wave sleep in mice. Behav Brain Res *255*, 64-74.

36. Shikanai, H., Yoshida, T., Konno, K., Yamasaki, M., Izumi, T., Ohmura, Y., Watanabe, M., and Yoshioka, M. (2012). Distinct neurochemical and functional properties of GAD67-containing 5-HT neurons in the rat dorsal raphe nucleus. J Neurosci *32*, 14415-14426.

37. Chowdhury, S., and Yamanaka, A. (2016). Optogenetic activation of serotonergic terminals facilitates GABAergic inhibitory input to orexin/hypocretin neurons. Sci Rep *6*, 36039.

38. Franklin, K.B.J., and Paxinos, G. (2008). The mouse brain in stereotaxic coordinates. 3rd Edition. (Amsterdam ; Boston: Academic Press,).

39. Carobrez, A.P., and Bertoglio, L.J. (2005). Ethological and temporal analyses of anxiety-like behavior: the elevated plus-maze model 20 years on. Neurosci Biobehav Rev *29*, 1193-1205.

40. Hamadate, N., Yamaguchi, T., Sugawara, A., Togashi, H., Izumi, T., Yoshida, T., Ohmura, Y., and Yoshioka, M. (2010). Liposome-encapsulated hemoglobin ameliorates impairment of fear memory and hippocampal dysfunction after cerebral ischemia in rats. J Pharmacol Sci *114*, 409-419.

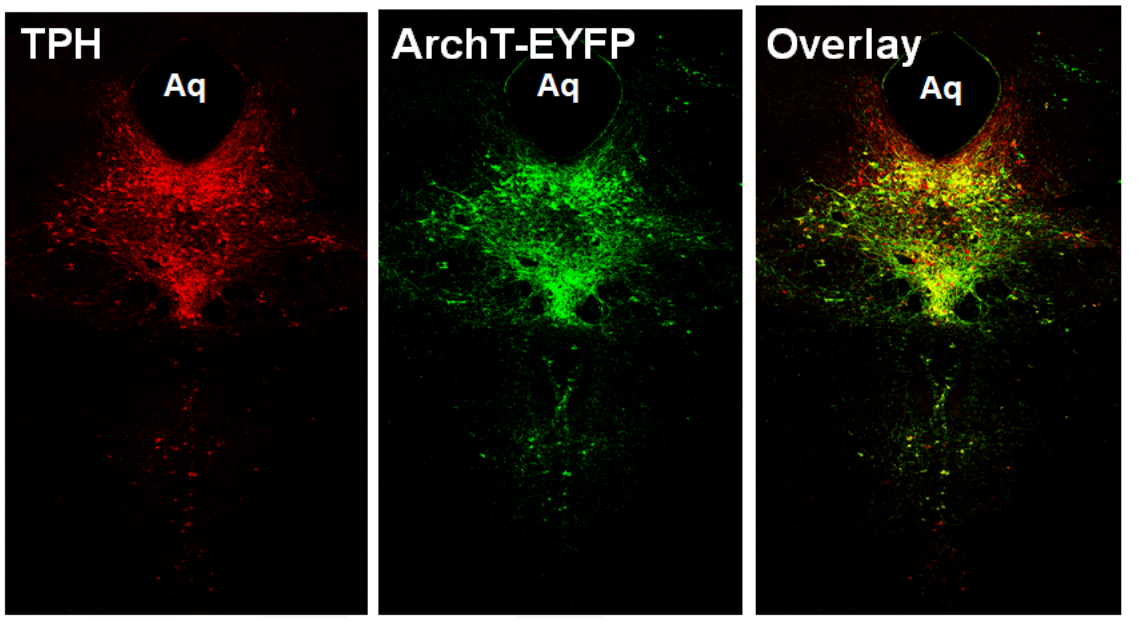
41. Sarter, M., Bodewitz, G., and Stephens, D.N. (1988). Attenuation of scopolamine-induced impairment of spontaneous alteration behaviour by antagonist but not inverse agonist and agonist beta-carbolines. Psychopharmacology (Berl) *94*, 491-495.

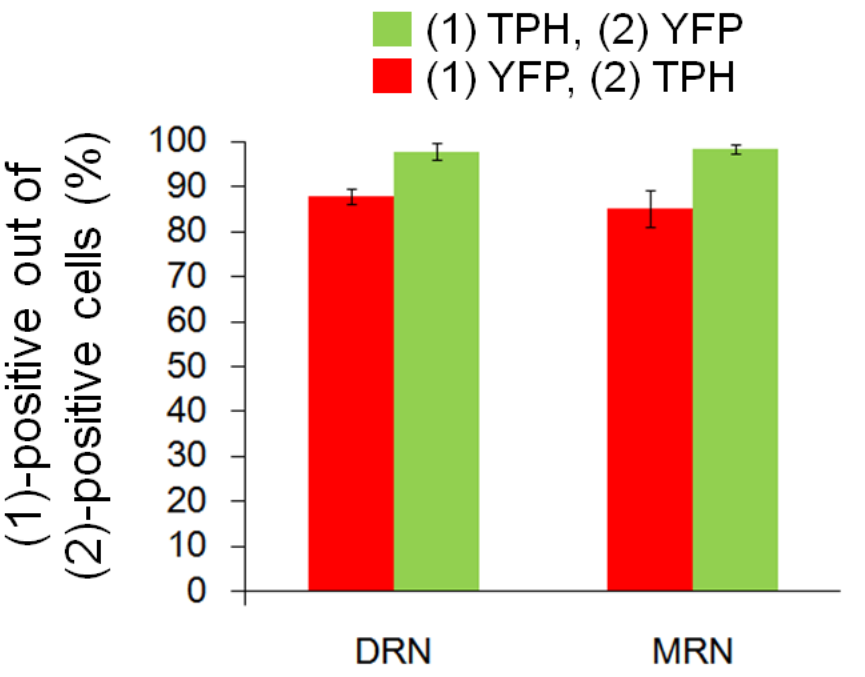
A

Tph2-tTA × tetO-ArchT-EYFP → Bi-transgenic

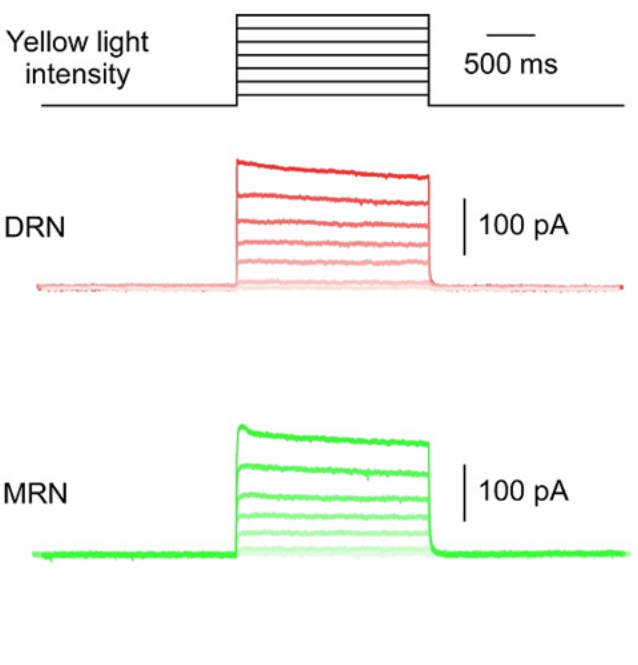
Serotonergic neuron-specific ArchT-EYFP expression

B

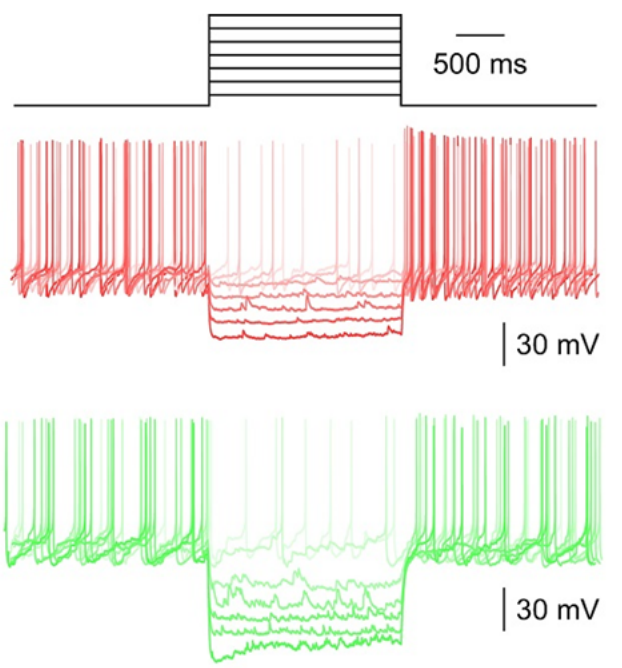




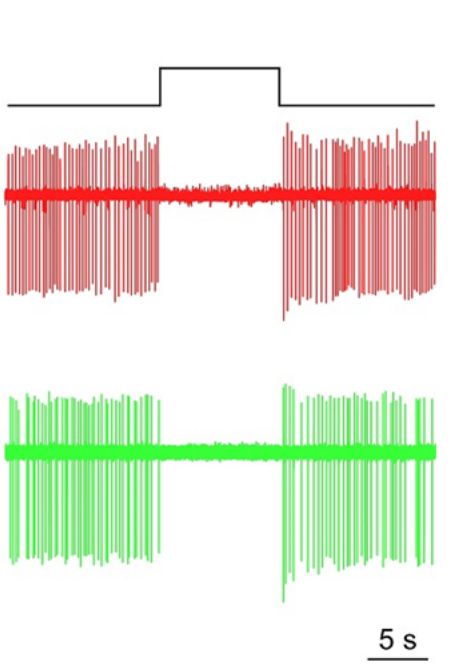
C



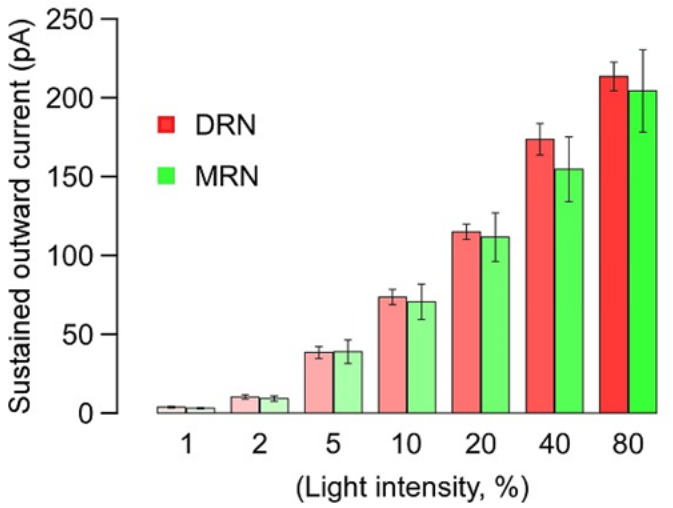
E



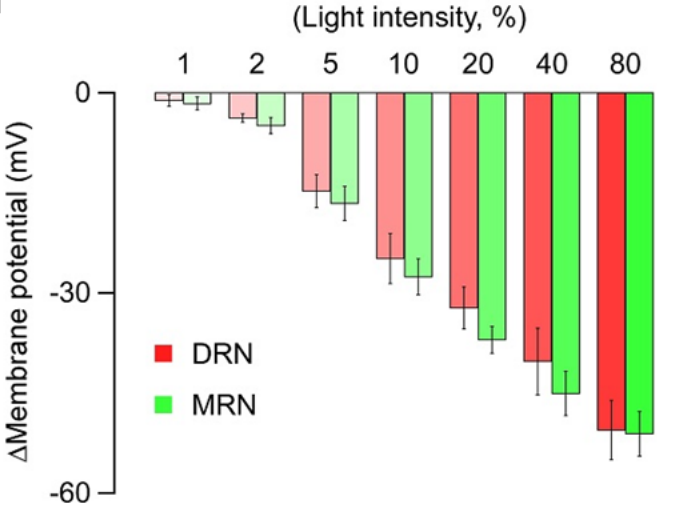
G



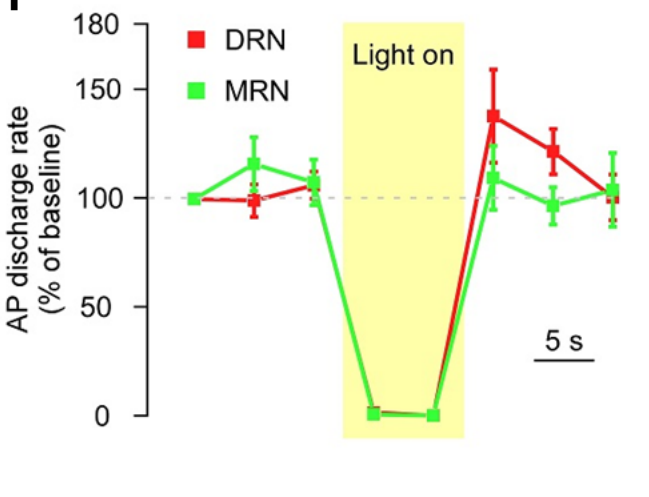
D



F



H



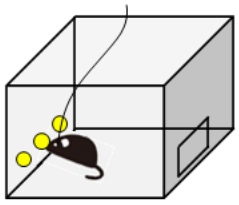
A

Day 3 – 5

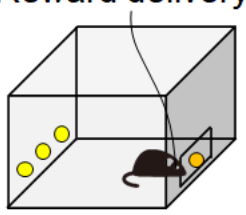
No light

Operant training

Nose poke to lit holes



Reward delivery

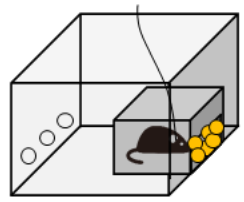


Day 6 – 8

No light

Taste aversion

Reward consumption in a small box



Paired

LiCl injection soon after the consumption

Unpaired

LiCl injection 6 hours after the consumption

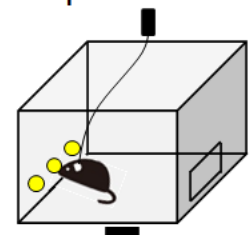
Day 9

Yellow light

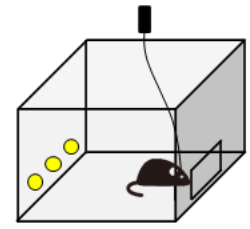
Probe test

(extinction)

Nose poke to lit holes

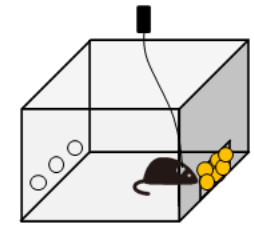


No reward



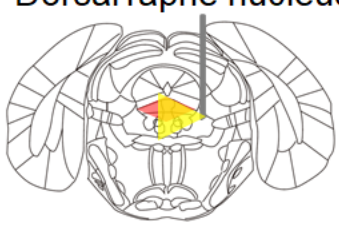
Post test

Reward consumption

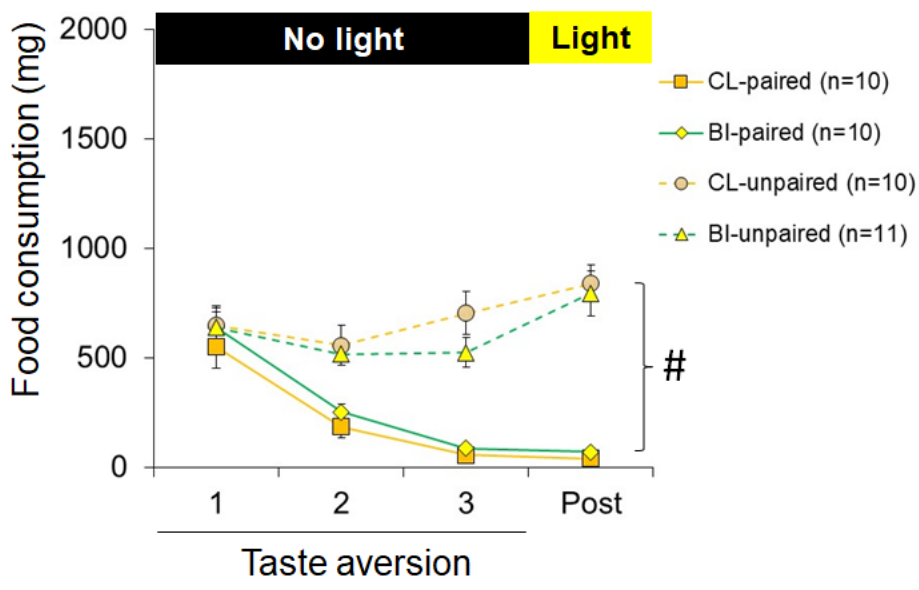


B

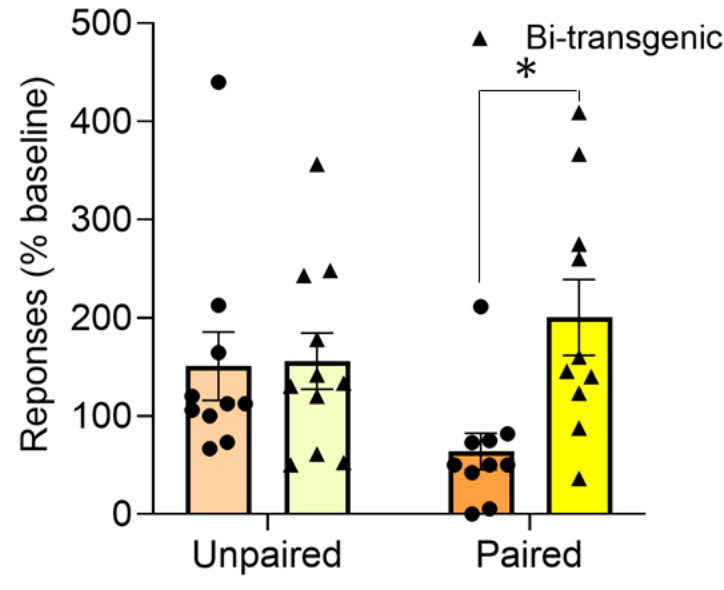
Dorsal raphe nucleus



Bregma -4.5 mm

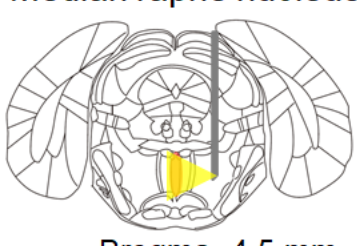


C

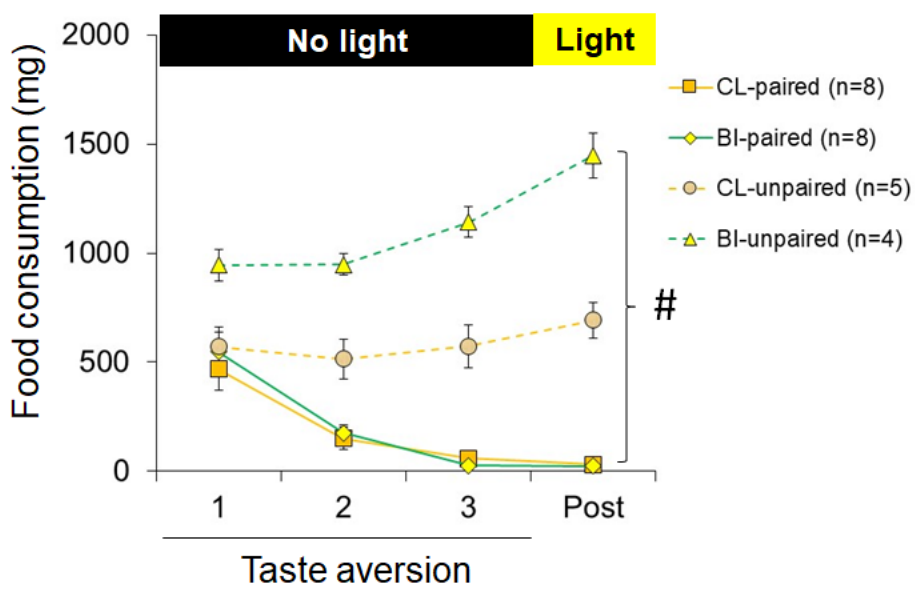


D

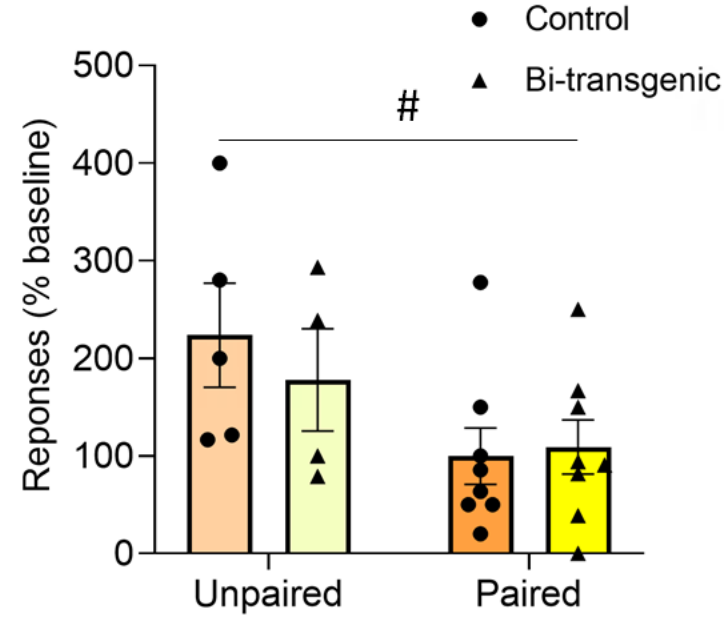
Median raphe nucleus



Bregma -4.5 mm

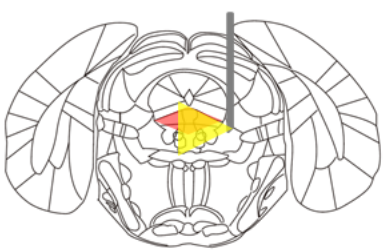


E



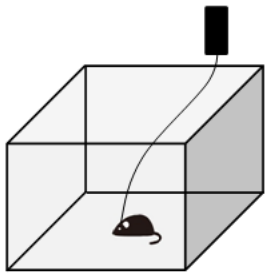
A

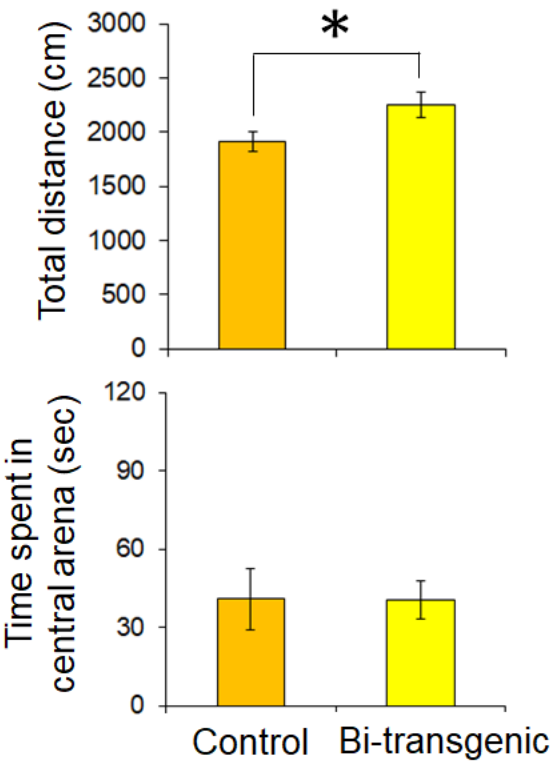
Dorsal raphe nucleus



Bregma -4.5 mm

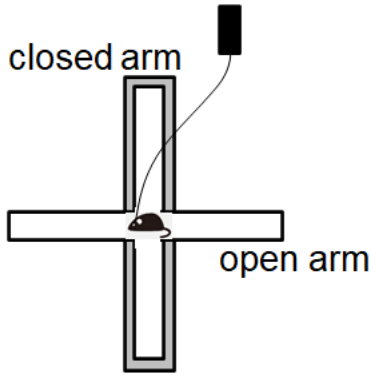
Open field test

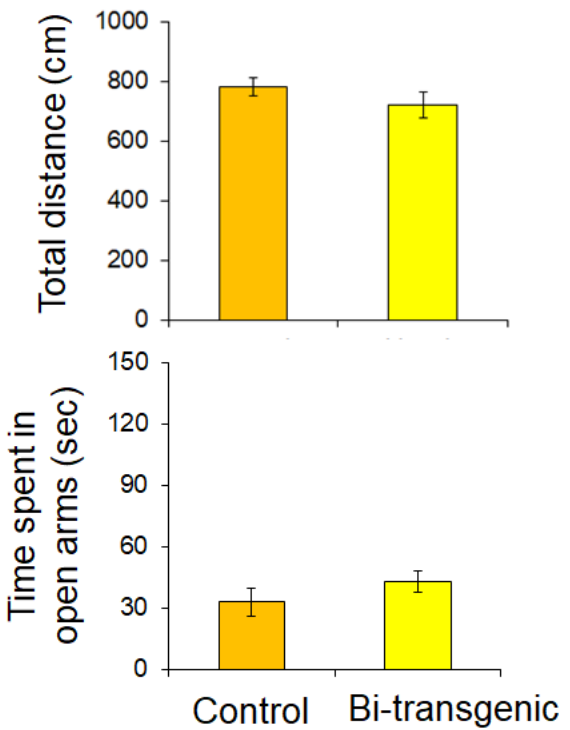




B

Elevated plus maze test





A

Day 2 (Operant training)

No light

Nose poke to right hole

Reward A is delivered

Nose poke to left hole

Reward B is delivered

Nose poke to central hole

No reward

Day 3 – 4 (taste aversion)

No light

Reward A consumption in a small box

Reward B consumption in a small box

LiCl injection

Saline injection

Preferred reward is devalued

Day 5 (test)

Yellow light

Probe test (extinction)

Nose poke to holes

No reward

Post test

Rewards A and B consumption

B

Dorsal raphe nucleus

Bregma -4.5 mm

Food consumption (mg)

1000

800

600

400

200

0

No light

Yellow light

#

1

2

Post test

Taste aversion

bi-transgenic unpreferred

bi-transgenic preferred (devalued)

control unpreferred

control preferred (devalued)

C

Responses to side holes (% total responses)

150

100

50

0

Unpreferred

Preferred (devalued)

*

Training

Probe

Control

Bi-transgenic

D

Responses to central hole (%, probe/training)

500

400

300

200

100

0

Control

Bi-transgenic