Mediodorsal Thalamus Hypofunction Impairs Flexible Goal-Directed Behavior

Supplemental Information

Supplemental Methods & Materials

Virus Injection and Histology

The AAV2-hM4D-IRES-hrGFP (called AAV2-hM4D) adeno-associated virus was generated by inserting the hM4D sequence into the multiple cloning site of the pAAV-IRES-hrGFP vector (Agilent Technologies). In this virus, both hM4D and GFP expression are under the control of a cytomegalovirus promoter as previously described (1). The AAV2-hM4D as well as a control virus expressing only GFP (AAV2-hrGFP) were produced by Vector Biolabs (Eagleville, PA). Viral mice were injected bilaterally with 0.5 ml of AAV2-hM4D (1.04×10^{13} particles/ml) or control AAV2-hrGFP (1×10^{12} particles/ml). Viruses were pressure injected using a glass pipette (10-15 µm) into the mediodorsal thalamus (MD) (coordinates: A/P, -1.3 mm; M/L, ±0.35 mm; D/V, -3.2 mm). After each injection, the pipette remained *in situ* for 10 min to minimize leaking. Eight weeks-old mice were injected and then tested 4 weeks later for behavioral testing. After testing, mice were anesthetized with a xylazine/ketamine mix (100 mg/kg and 10 mg/kg) and then transcardially perfused successively with PBS and 4% PFA. Brains were then post-fixed 5 hours in 4% PFA. Fifty µm-thick coronal sections were cut on a vibratome and kept at 4°C in 0.4% PFA until mounting using Vectashield mounting media (Vector Laboratories). GFP autofluorescence was visualized with a microscope to assess for correct transgene expression within the MD. Of note, we found in our previous study that transgene expression was exclusively neuronal (1). In addition, due to the extremely low number of interneurons in the mouse MD (<<1%), virtually only relay neurons express the transgenes.

After histology, mice for which GFP was not visible or in an incorrect location were removed from the final analysis.

Behavioral Apparatus

The behavioral procedures were performed in sixteen matching operant chambers (model env-307w; Med-Associates, St. Albans, VT) enclosed in a sound- and light-attenuating cabinet equipped with an exhaust fan, which provided 72 dB background white noise inside the chamber. The internal dimensions of the experimental chamber were 22 × 18 × 13 cm, and the floor consisted of metal rods placed 0.87 cm apart. Each chamber was equipped with two retractable levers that could be extended to the left and right of a recessed food magazine. Attached to each food magazine were a pellet dispenser, used to deliver 20 mg grain-based food pellets (Bio-Serv, Frenchtown, NJ) and a liquid dipper used to deliver either evaporated milk or 20% sucrose solution depending on the experiment. An infrared photocell detector (4 mm from trough opening) was used to record head entries into the food magazine. Illumination was provided by a house light (model 1820; Med Associates) located on the wall opposite the magazine. Depending on the experiment, either a flashing or constantly illuminating house light was used to produce visual cues while tone (2 kHz; 80 dB) and white noise (80 dB) generators were used to produce the auditory cues. A set of two microcomputers running the Med-PC program (Med Associates) controlled all experimental events and recorded responses.

Discrimination and Reversal Learning Tasks

Dipper Training

Mice were first trained to consume the liquid reward (evaporated milk) from the dipper by placing them inside the chambers with the dipper in the raised position, providing access to a drop of evaporated milk. The dipper was retracted 10 s after the first head entry into the feeder

trough. A variable intertrial interval (ITI) ensued, followed by a new trial identical to the first. The session ended after 30 min or 20 dipper presentations. On the following day, mice received another session similar to the first, except that the dipper was raised for 8 s and then lowered independently of whether or not mice had made a head entry. The session ended after 20 dipper presentations.

Standard Lever Press Training

At the beginning of the session the lever was extended into the chamber, and lever presses were reinforced on a continuous reinforcement schedule. In this and all subsequent sessions, the reward consisted of raising the dipper for 5 s. The lever was retracted after every two reinforcements and then re-extended after a variable ITI (average 30 seconds). The session ended when the mouse earned 60 reinforcements or one hour elapsed. Mice continued receiving sessions like this until they earned 60 rewards in two consecutive sessions. Mice then moved to variable interval training (VI).

VI Training

In VI training, lever presses were not reinforced until after a variable interval (timed relative to the lever extension) had elapsed. Mice began on VI-2s schedule, meaning that the first lever press occurring on average 2 s after lever extension was reinforced. Each reinforcement was followed by a variable ITI (mean = 30 s, range = 110 s), during which the lever remained retracted, and then a new trial, signaled by the extension of the lever. When a mouse earned at least 40 rewards in one session, the VI schedule was changed in the next session. The VI schedules used were 2 s, 10 s, and 20 s. Once mice reached the criterion of 40 rewards in one session on the VI-20s schedule it was moved to the discrimination task.

Discrimination Task

During the discrimination task, the two groups of mice injected with either AAV2-hrGFP or AAV2-hM4D virus received daily i.p. saline injections before each session. Lever presses were reinforced on a VI-20s schedule when one of the two visual stimuli were present (S+) but lever presses in the presence of the other stimulus (S-) never led to reward. Each daily session was composed of 20 S+ and 20 S- trials. S+ and S- trials were randomly intermixed and separated by a variable ITI averaging 30 s during which the house light was turned off. The duration of S+ trials was 1 min. Thus, the average number of rewards earned during all S+ trials was 60. S- trials had a minimum duration of 1 min. During the S- mice were required to withhold lever presses for 20 s in order for the S- to end and the next ITI begin. The discrimination task ended after 7 sessions were completed.

Extinction Task

After the discrimination task, mice went through an extinction task. During this task, mice received either saline or CNO before each session. The extinction task was identical to the discrimination task except that lever presses during both the prior S+ and S- were not rewarded. Mice performed one session per day, and the extinction phase lasted for 6 days.

Reversal Task

After extinction, mice went through a reversal task. Mice received either saline or CNO before each reversal session. This task was identical to the discrimination phase except that the contingencies between the stimuli and the outcome were reversed so that the S+ became S- and the S- became S+.

Contingency Degradation

After the reversal task, mice were trained to press the lever on a random ratio (RR) schedule. To minimize interference from the prior tasks, mice were trained on the lever that had never been previously presented and the house light was off throughout the sessions. Mice performed 2 sessions under a RR5 [p(O/R) = 0.2] schedule followed by 5 sessions under a RR20 [p(O/R) = 0.05]. The session ended after mice obtained 20 rewards or after 30 min. Mice were then trained for two additional 20 min sessions of RR20 under saline or CNO treatment to obtain a baseline level of lever pressing and then were exposed to 5 consecutive sessions during which the response-outcome contingency was degraded. In these sessions reward was still delivered for lever pressing on the RR20 schedule, however the reward was also delivered non-contingently with the same probability [p(O/no R) = 0.05] in each second without a response. CNO or saline was administered before each degradation session.

Outcome Specific Devaluation Task

Dipper/Feeder Training

Mice were first trained to consume separately the two rewards (sucrose 20% solution and food pellets) from the reward receptacle in morning and afternoon sessions (morning session 10:00 am, afternoon session 3:00 pm) for 2 days. The training sessions where sucrose solution was the outcome were identical to that used for evaporated milk described above. For training sessions with food pellets as an outcome, individual food pellets were delivered on a random time (30 s) schedule.

Instrumental Training

After food magazine training, the mice received 11 days of instrumental training during which two actions (left and right lever press) were trained with different outcomes (pellets and

sucrose) in separate sessions each day; one-half of the mice earned pellets on the left lever during one session and sucrose for right lever presses in the other, whereas the rest of the mice earned each outcome on the opposite levers. The order of the training sessions was varied over the days. Each session ended when 20 outcomes were earned or when 30 min had elapsed. For the first 2 days, lever pressing was continuously reinforced. On days 3-5 mice earned rewards on: a RR5 schedule [p(O/R) = 0.2); on days 6-8 a RR10 [p(O/R) = 0.1] was in effect, and a RR20 [p(O/R) = 0.05] schedule was used on days 9–11.

Devaluation Test

After instrumental training, mice went to the first day of the outcome devaluation test. Mice received *ad libitum* access to one of the two outcomes, either pellets (15 g placed in a petri dish), or sucrose (15 ml in a drinking bottle), for 1 h in distinct feeding cages located in a room different from the test room. One-half of the mice in each action–outcome assignment received pellets and the remaining mice received sucrose. The mice were then given a 10 min choice extinction test in which both levers were available but no outcomes were delivered. After 1 day of re-exposure to the RR20 schedule, the same procedure was repeated except that mice that had been given *ad libitum* access to pellets now received sucrose, and mice that had been given *ad libitum* access to sucrose now received pellets. Mice received saline or CNO treatment just before pre-feeding during the devaluation test.

Outcome Specific Pavlovian-to-Instrumental Transfer (PIT)

Pavlovian Conditioning and Instrumental Training

After food magazine training as described for the outcome devaluation task, mice received 7 daily sessions of Pavlovian conditioning during which levers were retracted. Each session was 1 h long and consisted of presentations of the two conditioned stimuli (CSs) (i.e.,

tone or white noise), each consistently paired with either the sucrose or the pellet outcomes. Each CS lasted 2 min and was presented four times in a pseudorandom order with a variable ITI of 5 min. One-half of the animals received tone–pellet and noise–sucrose pairings, whereas the other half received the opposite CS–outcome pairings. The appropriate outcome was delivered during the tone or noise CS on a random-time 30 s schedule. Performance in Pavlovian conditioning was assessed using a Pavlovian elevation score [(head entries rate during CS) – (head entries rate during the pre-CS period)]. After Pavlovian training, the mice received 11 days of instrumental training as described for the outcome devaluation task. After Pavlovian training, the mice received 11 days of instrumental training during which two actions (left and right lever press) were trained with different outcomes (pellets and sucrose) as described for the outcome devaluation task.

Outcome Specific Pavlovian-to-Instrumental Transfer

After the final day of RR20 training, mice were given 2 days of Pavlovian-to-instrumental testing. During test sessions, both levers were inserted into the box, but no outcomes were delivered. Responding was extinguished on both levers for 8 min to reduce the baseline rate of performance after which each CS was presented four times over the next 40 min in the following order: noise–tone–tone–noise–tone–noise–noise–tone. Stimulus presentations lasted 2 min and were separated by a 3 min fixed ITI. A transfer score was calculated by measuring the difference in lever press responding in the presence versus absence of CS. The "Same" responses are lever presses on the lever that is paired with the same outcome as the CS. The "Different" responses are lever presses on the lever that is paired with the different outcome than the CS.

Statistics

Data are presented as means ± SEM. Two-way analysis of variance (ANOVA) with group and sessions as factors was carried out for the extinction task, the reversal learning task, the contingency degradation and for the Pavlovian conditioning phase of the PIT. For the outcome-specific PIT, two-way ANOVA with group and transfer as factors was carried out. The devaluation task was analyzed using a three-way ANOVA with group, value and time as factors.

Supplemental Reference

1. Parnaudeau S, O'Neill PK, Bolkan SS, Ward RD, Abbas AI, Roth BL *et al.* (2013): Inhibition of mediodorsal thalamus disrupts thalamofrontal connectivity and cognition. *Neuron* 77(6):1151-1162.