

A quadruple dissociation of reward-related behaviour in mice across excitatory inputs to the nucleus accumbens shell

Erin B. Lind ^{1,2}, Brian M. Sweis ^{1,2,3}, Anders J. Asp^{1,4}, Manuel Esguerra^{1,2}, Keelia A. Silvis^{1,2}, A. David Redish ^{1,2} & Mark J. Thomas ^{1,2}✉

The nucleus accumbens shell (NAcSh) is critically important for reward valuations, yet it remains unclear how valuation information is integrated in this region to drive behaviour during reinforcement learning. Using an optogenetic spatial self-stimulation task in mice, here we show that contingent activation of different excitatory inputs to the NAcSh change expression of different reward-related behaviours. Our data indicate that medial prefrontal inputs support place preference via repeated actions, ventral hippocampal inputs consistently promote place preferences, basolateral amygdala inputs produce modest place preferences but as a byproduct of increased sensitivity to time investments, and paraventricular inputs reduce place preferences yet do not produce full avoidance behaviour. These findings suggest that each excitatory input provides distinct information to the NAcSh, and we propose that this reflects the reinforcement of different credit assignment functions. Our finding of a quadruple dissociation of NAcSh input-specific behaviours provides insights into how types of information carried by distinct inputs to the NAcSh could be integrated to help drive reinforcement learning and situationally appropriate behavioural responses.

¹Department of Neuroscience, University of Minnesota, 6-145 Jackson Hall, 321 Church St SE, Minneapolis, MN 55455, USA. ²Medical Discovery Team on Addiction, University of Minnesota, 3-432 McGuire Translational Research Facility, 2001 6th St SE, Minneapolis, MN 55455, USA. ³Department of Psychiatry, Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, 1 Gustave L. Levy Pl, New York, NY 10029, USA. ⁴Rehabilitation Medicine Research Center, Department of Physical Medicine and Rehabilitation, Mayo Clinic, 200 First St SW, Rochester, MN 55905, USA. ✉email: tmthomas@umn.edu

Goal-directed behaviour involves the integration of multiple cognitive, emotional, and motivational processes to coordinate the appropriate execution of environment- and situation-specific behaviours. The nucleus accumbens (NAc) is a key limbic-motor interface implicated in the integration of information which drives behavioural responses to motivationally relevant stimuli^{1–4}. While both the shell (NAcSh) and core subregions are involved in reward valuation^{5–9}, the NAcSh subregion is of particular interest as it receives input from both cortical and limbic brain regions and is well situated to receive, integrate, and respond to information about both appetitive and aversive stimuli^{10,11}. Additionally, alterations in excitatory transmission in NAcSh have been implicated in both adaptive and maladaptive motivated behaviour, and maladaptive learning and experience-dependent plasticity at excitatory synapses within the NAcSh are thought to underlie a range of psychiatric disorders including depression, addiction, and schizophrenia^{12–15}. Despite this, relatively little is known regarding how individual glutamatergic inputs to the NAcSh may guide reward valuation in goal-directed behaviour, or how these different inputs may contribute to the acquisition or expression of adaptive or maladaptive behaviours.

The NAcSh is primarily composed of medium spiny GABAergic neurons (MSNs) that receive excitatory input from multiple source nuclei including the medial prefrontal cortex (mPFC), ventral hippocampus (vHPC), basolateral amygdala (BLA), and paraventricular thalamus (PVT)^{11,16,17}. These source nuclei are most often implicated in regulating emotional processing and the expression of approach/avoidance behaviour, although the direction of these effects is sometimes equivocal, particularly within the BLA and PVT^{18–23}. Optogenetic approaches to selectively activate mPFC, vHPC, BLA, and PVT excitatory projections to the NAcSh during real time place preference or instrumental self-stimulation assays indicate each of these pathways are involved in emotional and motivational valence. Contingent stimulation of PVT → NAcSh inputs reduces or has variable effects on real time place preference^{19,24,25} but can also support lever-based self-stimulation²⁵. In contrast, contingent stimulation of mPFC → NAcSh, vHPC → NAcSh, or BLA → NAcSh inputs consistently support both real time place preference and instrumental responding, indicating that stimulation of these inputs is rewarding^{25–27}. Given the finding that mPFC → NAcSh, vHPC → NAcSh, and BLA → NAcSh inputs all have positive valence, it has been proposed that amount, rather than source, of excitatory drive to the NAcSh is what is most relevant for motivated behaviour²⁸. However, given the diverse and often complex roles of the mPFC, vHPC, and BLA in affective, motivational, and cognitive function^{29–33}, it is possible that input-specific differences in reward-directed behaviour were simply not detectable using classic paradigms, or that other aspects of motivated behaviour, such as reward valuation, reinforcement learning, and/or computational decision-making processes may be involved.

It remains unclear what, if any, specific information is provided by different excitatory inputs to the NAcSh or how this may regulate specific aspects of reward valuation or decision-making during foraging behaviour^{9,34}. Thus, in the present study we utilized a spatially dependent optogenetic self-stimulation task to investigate strategies employed by mice to obtain or avoid that optogenetic stimulation. We found that self-stimulation of individual excitatory inputs to the NAcSh resulted in distinct behavioural patterns, which we propose relate to the reinforcement of different credit assignment functions.

Results

Approach. We used optogenetic methods to selectively activate different excitatory inputs to the NAcSh (Fig. 1a, Supplementary

Fig. 1a). An adeno-associated virus (AAV2) was used to express channelrhodopsin2 (ChR2) or control fluorophore (eYFP) under a Ca^{2+} /calmodulin-dependent protein kinase IIa (CaMKIIa) promoter to drive eYFP or ChR2-eYFP expression in glutamatergic neurons in either the medial prefrontal cortex (mPFC), ventral hippocampus (vHPC), basolateral amygdala (BLA), or paraventricular thalamus (PVT) (Fig. 1b, Supplementary Fig. 1b). Optical fibers were then implanted bilaterally above eYFP+ or ChR2+ terminals in the NAcSh to allow for selective activation of mPFC → NAcSh, vHPC → NAcSh, BLA → NAcSh, or PVT → NAcSh pathways (Fig. 1c, Supplementary Fig. 1c). To identify potential input-specific differences in behavioural strategies exhibited during reinforcement learning, we assessed optogenetic self-stimulation behaviour in an open-field spatial task (Fig. 1d; Supplementary Fig. 2). The spatial arena used for this task consisted of an open square box (20" x 20") with four spatially restricted and contextually distinct zones, one in each corner (Fig. 1e). Mice were allowed to freely explore during an initial baseline session. Acquisition of self-stimulation behaviour was assessed the next day by pairing entry into one of these corner zones with optogenetic stimulation of either mPFC → NAcSh, vHPC → NAcSh, BLA → NAcSh, or PVT → NAcSh inputs. A reversal test was conducted the following day by switching the location of the active zone to confirm the reward valence of each input as well as the behavioural strategy that was observed during acquisition.

To identify strategies the mice exhibited while engaging with stimulation-paired environments during testing, we employed stimulation parameters that were contingent on the mouse's behaviour but also allowed mice flexibility in how they received optogenetic self-stimulation within the spatial arena. Entry into a stimulation-paired zone triggered a 465 nm blue-light LED for up to 5 s (30 Hz, 5 ms pulse width), followed by a 15 s non-reinforced "timeout" period. Mice could terminate stimulation early by exiting the active zone. If mice received the full 5 s stimulation, they could remain in the zone throughout the timeout period to gain additional bouts of stimulation without taking any further action. They could also bypass the timeout period by exiting and re-entering the zone to trigger another bout of stimulation. Mice could freely vary across these options throughout the test (Fig. 1f, Supplementary Fig. 2). Using this approach, we found that eYFP control mice explored the spatial arena similarly across refection groups during the baseline, acquisition, and reversal sessions, suggesting that viral surgery and light delivery into the NAcSh alone did not impact behaviour in the spatial arena (Supplementary Fig. 3). Thus, data from eYFP mice was pooled for subsequent analysis ($n = 12$ mice). Initially, separate one-way ANOVAs were used to assess behaviour across the four corner zones during baseline (Supplementary Fig. 4), acquisition (Fig. 2a), and reversal testing (Fig. 3a). These analyses revealed that while eYFP and ChR2+ mice exhibited similar exploratory behaviour during baseline testing, stark behavioural differences emerged when mPFC ($n = 7$), vHPC ($n = 7$), BLA ($n = 11$), or PVT ($n = 6$) inputs to the NAcSh were selectively activated upon active zone entry during acquisition and reversal sessions.

Expression of different behavioural strategies during acquisition in ChR2 + mice.

During acquisition, ChR2 +, but not eYFP control mice, showed alterations in behaviour across the different corner zones during optogenetic stimulation; however, the type of alteration differed between the different ChR2+ input pathways (Fig. 2). For example, both mPFC → NAcSh ($F_{3,24} = 14.7956$, $p < 0.0001$, $\eta^2 = 0.65$) and vHPC → NAcSh ($F_{3,24} = 11.4987$, $p < 0.0001$, $\eta^2 = 0.59$) stimulation produced real-time place preferences and increased the time mice spent in stimulation-paired corners (Fig. 2d), resulting in a significant amount of optogenetic

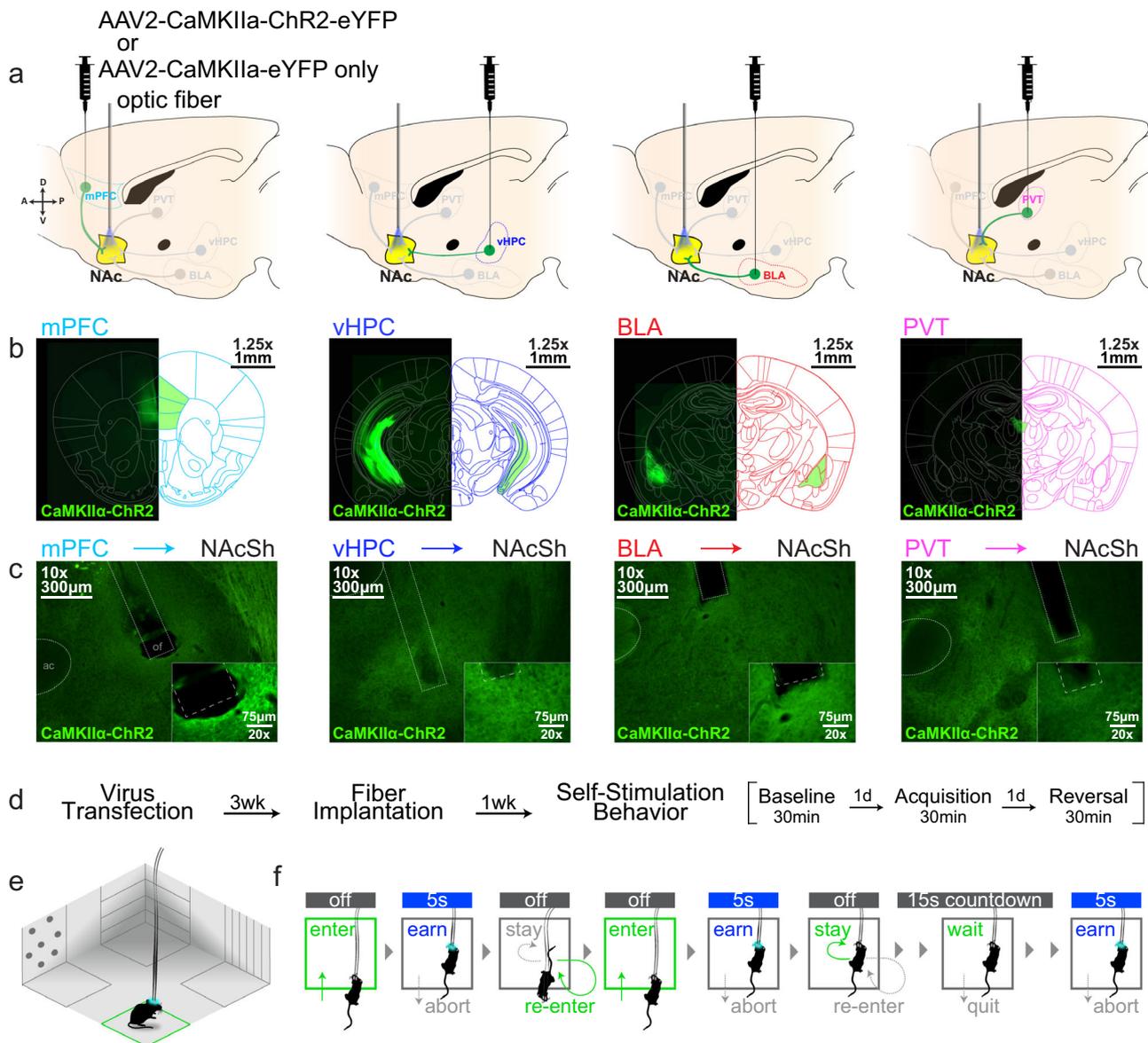


Fig. 1 Neural pathway targets and optogenetic self-stimulation behaviour task. **a** Viral vector injections and optic fiber implants targeted specific excitatory inputs to the NAcSh. **b** Representative images (4x) from coronal slices (40 μ m) showing viral localization and spread. **c** Representative images of input-specific ChR2+ terminal regions in NAcSh (10x, insets at 20x). The optical fiber (of) and anterior commissure (ac) are depicted with straight and rounded dashed lines, respectively. **d** Experimental timeline. Each session was 30 min long with one session per day. **e** Optogenetic stimulation was inactive during baseline, active in one zone during acquisition, and active in a different zone during reversal (30 Hz, 5 ms pulse width, 5 s max stimulation per bout). **f** Illustration of reward-related decisions during the task. A mouse could exit and re-enter the active zone to bypass a 15 s “timeout” period, wait in the active zone during the timeout for additional stimuli, or any combination of these two options.

stimulation of these two inputs being self-administered (Fig. 2e, $F_{3,24} = 9.7834$, $p = 0.0002$, $\eta^2 = 0.55$ and $F_{3,24} = 13.0889$, $p < 0.0001$, $\eta^2 = 0.62$, respectively). In contrast, BLA \rightarrow NAcSh mice showed only non-significant trends for elevated active zone-directed behaviour as assessed by these measures (Fig. 2d, $F_{3,40} = 2.0105$, $p = 0.1279$ and Fig. 2e, $F_{3,40} = 2.1827$, $p = 0.1051$). While PVT \rightarrow NAcSh mice also showed altered time spent (Fig. 2d, $F_{3,20} = 5.8616$, $p = 0.0048$, $\eta^2 = 0.47$) and optogenetic stimulation received (Fig. 2e, $F_{3,20} = 4.9177$, $p = 0.0102$, $\eta^2 = 0.42$) during the acquisition session, active zone-directed behaviour in these mice was reduced rather than increased, suggesting that these mice were instead exhibiting some degree of real-time place avoidance.

We next examined whether contingent optogenetic self-stimulation was associated with alterations in the number of

entries mice made into stimulation-paired zones (Fig. 2c, f). In eYFP control mice, exposure to light in the active zone did not impact how many entries mice made across the corner zones. In contrast, ChR2+ mice showed input-specific alterations in zone entry behaviour. Exposure to contingent stimulation of mPFC \rightarrow NAcSh inputs altered the number of zone entries made ($F_{3,24} = 3.7727$, $p = 0.0238$, $\eta^2 = 0.32$) and increased active relative to inactive zone entries, indicating that these mice were obtaining rewarding stimulation by repeatedly exiting and re-entering the active stimulation zone. Alternatively, this zone entry effect was not found in vHPC \rightarrow NAcSh mice ($F_{3,24} = 0.5509$, $p = 0.6524$) suggesting that these mice were preferentially obtaining rewarding stimulation by waiting in the zone for subsequent stimulations rather than leaving and re-entering. Exposure to BLA \rightarrow NAcSh stimulation also had no effect on the

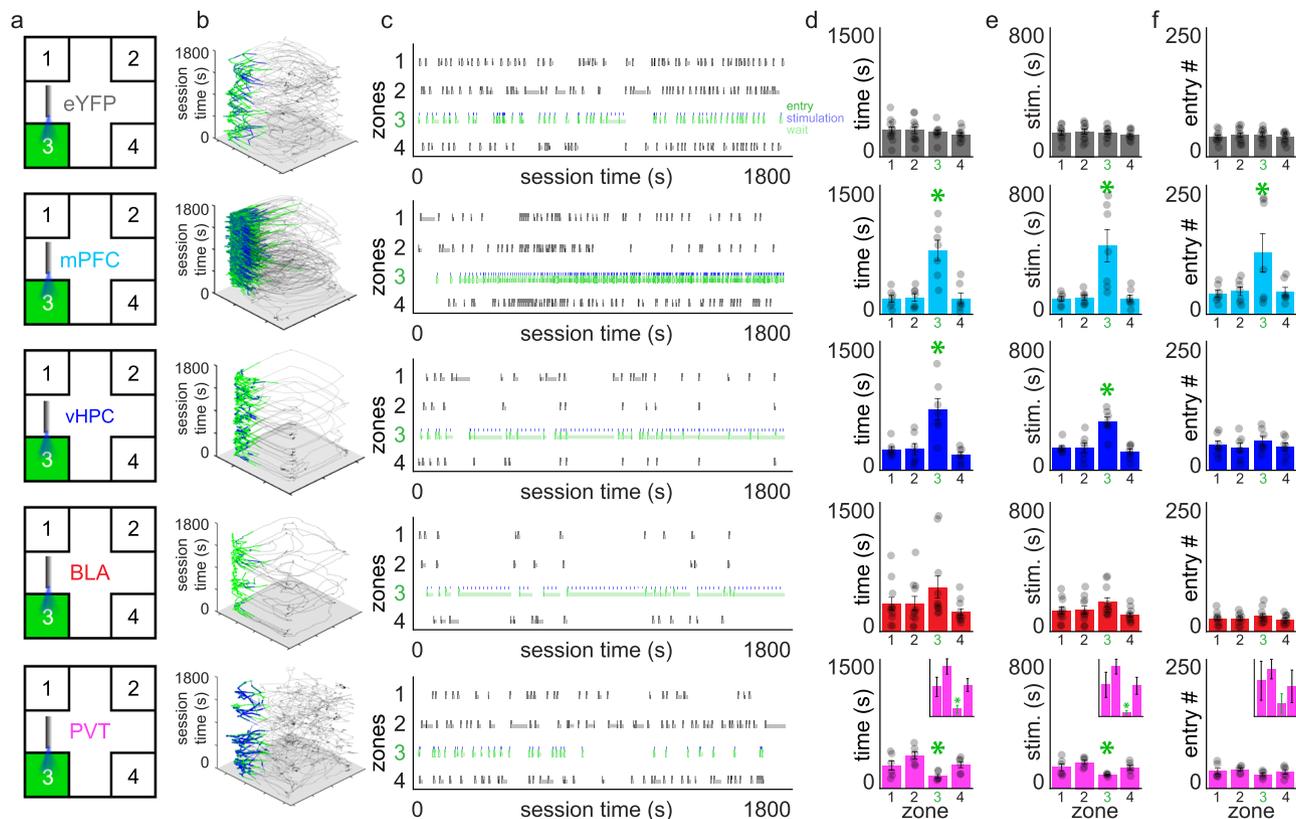


Fig. 2 Self-stimulation of different NAcSh inputs results in input-specific behavioural strategies. Different NAcSh input groups are depicted by different colours throughout: eYFP controls (gray, $n = 12$), mPFC \rightarrow NAcSh (cyan, $n = 7$ mice), vHPC \rightarrow NAcSh (blue, $n = 7$ mice), BLA \rightarrow NAcSh (red, $n = 11$ mice), and PVT \rightarrow NAcSh (pink, $n = 6$ mice). **a** Mice explored the spatial arena during a 30 min acquisition session with one active (3; green) and three inactive (1,2,4) corner zones. **b** x-y position tracking for representative mice throughout the session (green = active zone position; blue = stimulation position). **c** Raster plots showing individual zone entries over time in representative mice (from b). Green ticks indicate active zone entries, blue ticks indicate stimulation onset, green shading indicates time spent waiting during the timeout in the active (full height shading) and inactive (half height shading) zones. **d** Total time spent across the four corner zones. Active-zone time was increased in mPFC \rightarrow NAcSh and vHPC \rightarrow NAcSh mice and decreased in PVT \rightarrow NAcSh mice (inset presented for clarity). **e** Total stimulation time (real or mock) received in each corner zone. Stimulation time was increased in mPFC \rightarrow NAcSh and vHPC \rightarrow NAcSh mice and decreased in PVT \rightarrow NAcSh mice. **f** Total entries made into the different corner zones. Active-zone entries were only increased (vs. inactive zones) in mPFC \rightarrow NAcSh mice. Despite the reduced stimulation received in PVT \rightarrow NAcSh mice, these mice did not avoid entry into the active zone. Group data are expressed as mean \pm SEM. Main zone effects within each group and metric were identified using one-way ANOVAs. * $p < 0.05$ active vs. all other inactive zones, Student's post hoc comparisons.

number of entries made across the corner zones during acquisition testing ($F_{3,40} = 0.7286$, $p = 0.5410$). Interestingly, despite the observed reductions in time spent and stimulation received in the active zone in PVT \rightarrow NAcSh mice, these mice failed to show significant reductions in the number of entries they made across the corner zones during the acquisition session ($F_{3,20} = 0.9875$, $p = 0.4187$).

Input-specific differences in behaviour exhibited during reversal testing in Chr2+ mice. When the location of the active zone was changed during reversal testing (Fig. 3), eYFP control mice showed similar behaviour across the spatial arena as was found during the baseline and acquisition sessions, whereas input-specific behavioural differences were again detected in Chr2+ mice. While behaviour in mPFC \rightarrow NAcSh mice was preferentially directed at the active zone during acquisition, this active zone-directed behaviour was not fully recapitulated during reversal as indicated by the lack of significant zone effects for the time in zone (Fig. 3d, $F_{3,24} = 1.6413$, $p = 0.2062$), stimulation time (Fig. 3e, $F_{3,24} = 2.0868$, $p = 0.1286$), and zone entry (Fig. 3f, $F_{3,24} = 1.5569$, $p = 0.2257$) metrics, raising the possibility that in addition to being rewarding, stimulation of this input may have impacted behaviour in these mice in a way that led to an

inconsistent ability to maintain reward-directed behaviour. In contrast, vHPC \rightarrow NAcSh mice showed full reversal of reward-directed behaviour indicated by increased time spent (Fig. 3d, vHPC: $F_{3,24} = 16.0947$, $p < 0.0001$, $\eta^2 = 0.67$) and stimulation received (Fig. 3e, vHPC: $F_{3,24} = 21.1208$, $p < 0.0001$, $\eta^2 = 0.73$) in the reversal zone, without altering entries made across the corner zones (Fig. 3f, $F_{3,24} = 1.1030$, $p = 0.3672$). Interestingly, the effects of BLA \rightarrow NAcSh stimulation mice were more apparent in reversal than during acquisition testing, as these mice now exhibited significant differences in both time spent (Fig. 3d, $F_{3,40} = 2.9637$, $p = 0.0435$, $\eta^2 = 0.18$) and optogenetic stimulation received (Fig. 3e, $F_{3,40} = 3.1728$, $p = 0.0344$, $\eta^2 = 0.33$) across the corner zones. For PVT \rightarrow NAcSh mice, time in zone (Fig. 3d, $F_{3,20} = 5.2418$, $p = 0.0078$, $\eta^2 = 0.44$), stimulation received (Fig. 3e, $F_{3,20} = 3.2676$, $p = 0.0427$, $\eta^2 = 0.33$), and zone entry metrics (Fig. 3f, $F_{3,20} = 0.6659$, $p = 0.5828$) were consistent with the acquisition session, and mice shifted their avoidance behaviour towards the new active zone.

Pathway and contingency-dependent effects of stimulation on locomotor behaviour. Although we found clear differences in reward-related behaviours described above, movement around the arena (Fig. 2b, Fig. 3b) and entry patterns (Fig. 2c, Fig. 3c)

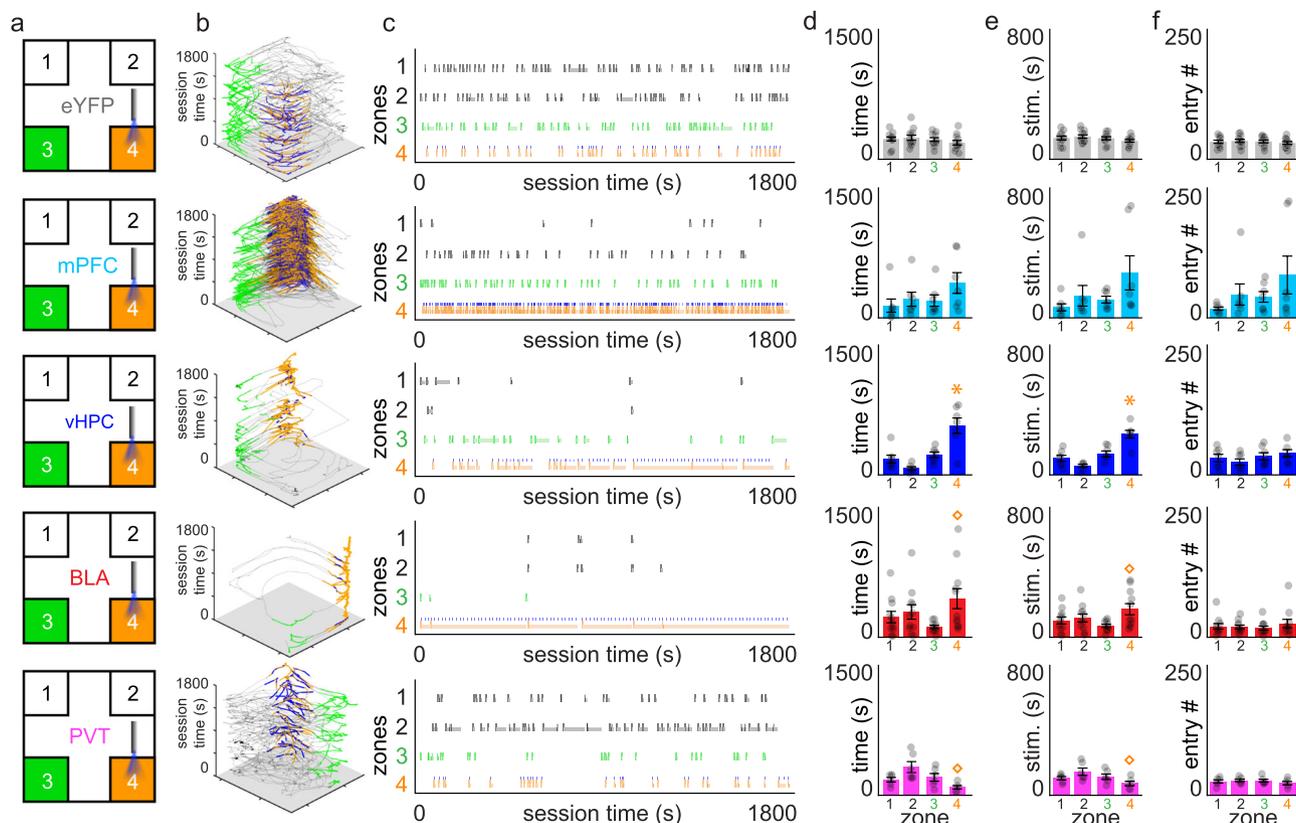


Fig. 3 Expression of NAcSh input-specific behavioural strategies during reversal testing. Different NAcSh input groups are depicted by different colours throughout: eYFP controls (gray, $n = 12$ mice), mPFC \rightarrow NAcSh (cyan, $n = 7$ mice), vHPC \rightarrow NAcSh (blue, $n = 7$ mice), BLA \rightarrow NAcSh (red, $n = 11$ mice), and PVT \rightarrow NAcSh (pink, $n = 6$ mice). **a** Following acquisition, mice explored the testing arena with a new active zone (4; orange). Inactive zones included the previously active zone (3; green) and zones 1 and 2 (black). **b** x-y position tracking for representative mice throughout the session (orange = position in active zone; blue = position during stimulation; green = position in previously active zone). **c** Raster plots showing individual zone entries over time in representative mice (same mice as in **b**). Orange shading indicates time spent remaining in the active zone. **d** Total time spent across the four corner zones. **e** Total stimulation received (real or mock) received. **f** Total entries made into the different corners. Group data are expressed as mean \pm SEM. Main zone effects within each group and metric were identified using one-way ANOVAs. $\diamond p < 0.05$ main effect across zones, $*p < 0.05$ active vs. all other inactive zones, Student's post-hoc tests.

during the acquisition and reversal sessions did not appear consistent across the different Chr2+ inputs, so it is possible that alterations in general motor activity may underlie some of the behaviours being exhibited by mice in the spatial arena. To more directly examine this, we first assessed distance traveled during behaviour sessions across both input pathways and behaviour sessions using a mixed-model ANOVA with subject as a random factor. This analysis revealed a pathway \times session interaction (Fig. 4a, $F_{8,76} = 3.1330$, $p = 0.0041$) indicating that activation of some, but not all pathways impacted motor behaviour during testing. Post-hoc comparisons within input pathways revealed that motor activity in mPFC \rightarrow NAcSh was increased during self-stimulation sessions relative to the baseline session, whereas activity in BLA-NAcSh mice was decreased. In contrast, neither vHPC \rightarrow NAcSh nor PVT \rightarrow NAcSh exhibited alterations in locomotor activity during testing, suggesting that behaviours in these mice were not confounded by non-specific motor effects.

To further identify whether exposure to optogenetic stimulation produced non-specific effects on locomotor activity, we also assessed locomotor behaviour in a separate cohort of mice that received passive optogenetic stimulation in a contextually distinct spatial arena that was not contingent on any behaviour. Analysis here revealed a significant interaction between input pathway \times stimulation period (Fig. 4b, $F_{8,56} = 5.6410$, $p < 0.0001$), suggesting again that stimulation of some, but not all pathways impacted motor behaviour. However, in this case only mPFC \rightarrow NAcSh

mice exhibited significant increases in motor activity during passive stimulation. Post-hoc comparisons indicated that mice increased their activity during the 5 min passive stimulation period and this effect that was reduced, but still apparent, after stimulation was discontinued. Together these data indicate a pathway-specific disconnect between locomotor activity exhibited in response to contingent versus non-contingent stimulation.

Waiting in and leaving the stimulation-paired active zone in mPFC \rightarrow NAcSh mice. As mPFC \rightarrow NAcSh mice exhibited increased motor activity in conjunction with either contingent or non-contingent stimulation, it is possible that the repeated entries made into the active corner zones during the self-stimulation sessions were merely a consequence of non-specific changes in locomotor activity. For example, it is possible that activation of mPFC \rightarrow NAcSh inputs transiently increased locomotor activity, causing the mice to leave the active zone during the stimulation bout, then returning to it after this effect wore off when stimulation was discontinued upon zone exit. If this were the case, we would expect mPFC \rightarrow NAcSh mice to be unable to wait in the stimulation-paired zone past the 5 s. stimulation bout, and instead would leave the zone (i.e., have an abort event) at a consistent duration after the onset of stimulation. However, analysis of waiting and leaving behaviour indicated this was not the case. A two-way ANOVA analysis of wait time (i.e., time spent in the zones during the timeout period) across sessions in

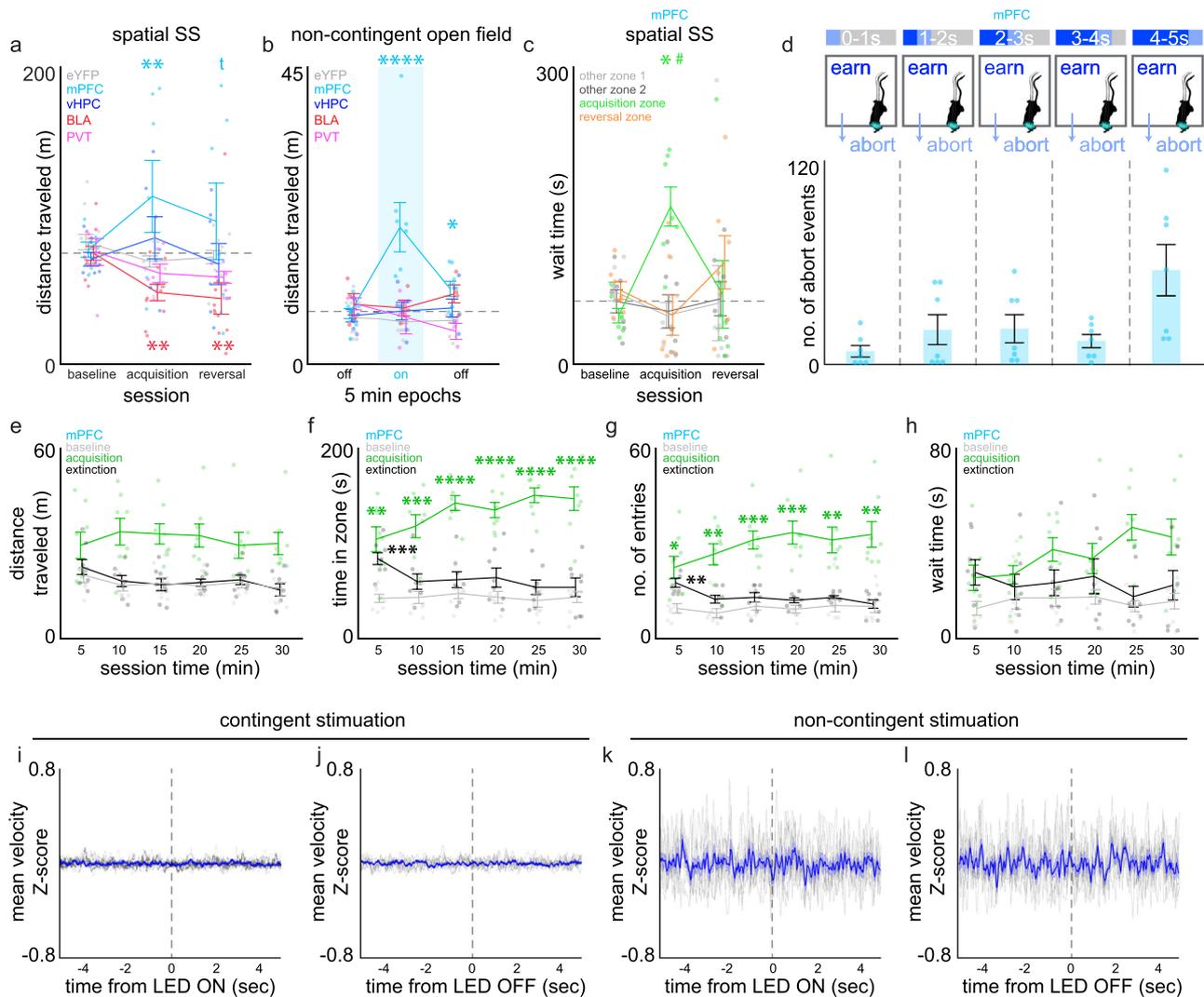


Fig. 4 Disconnects between locomotor activity and reward-directed behaviour in mPFC → NAcSh mice. **a** Total distance traveled (m) during spatial self-stimulation (SS) sessions in eYFP (gray, $n = 12$ mice), mPFC → NAcSh (cyan, $n = 7$), vHPC → NAcSh (blue, $n = 7$), BLA → NAcSh (red, $n = 11$), and PVT → NAcSh (pink, $n = 6$) mice. Contingent stimulation (30 Hz) increased total distance traveled in mPFC → NAcSh and decreased it in BLA → NAcSh mice. **b** Total distance traveled (m) during 5 min epochs in eYFP (gray, $n = 12$), mPFC → NAcSh (cyan, $n = 8$), vHPC → NAcSh (blue, $n = 6$), BLA → NAcSh (red, $n = 4$), and PVT → NAcSh (pink, $n = 3$) mice. Non-contingent stimulation (30 Hz, 5 s on, 5 s off during on epoch highlighted in blue) increased distance traveled, but only in mPFC → NAcSh mice. **c** Total time spent waiting in the corner zones during the timeout period across contingent self-stimulation sessions in mPFC → NAcSh mice. **d** The number of active zone abort stimulation decisions made by mPFC → NAcSh mice during acquisition expressed across progressive 1 s time bins within the 5 s stimulation duration. **e–h** Behaviour exhibited in the across (5 min bins) baseline (gray), acquisition (green), and extinction sessions (black) in mPFC → NAcSh mice ($n = 8$). **e** Total distance traveled (m) across the entire arena, (**f**) Time spent in the stimulation-paired active zone, (**g**) Entries made into the stimulation-paired active zone, (**h**) Time spent waiting in active zones during the timeout period. **i, j** Velocity Z-scores during contingent stimulation onset (**i**) or offset (**j**) during acquisition in mPFC → NAcSh mice. **k, l** Velocity Z-scores during non-contingent stimulation onset (**k**) or offset (**l**) during acquisition in mPFC → NAcSh mice. Note that differences in velocity amplitudes between contingent and non-contingent conditions is likely related to differences in apparatus dimensions. Data is expressed as mean \pm SEM. Analysis was done using mixed-model ANOVAs (**a, b**), two-way ANOVA (**c**), one-way ANOVA (**d**), two-way repeated measures ANOVAs (**e–h**), or paired two-tailed *t*-tests (**i, j**) $tp < 0.1$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$ versus baseline session (contingent) or first off epoch (non-contingent). $\#p < 0.05$ versus other corner zones, within same session using Student's (**a, b**) or Tukey's (**c, f, g**) post hoc comparisons.

these mice revealed a significant session \times zone interaction on wait times (Fig. 4c, $F_{6,72} = 3.1649$, $p = 0.0082$, $\eta^2 = 0.25$). Post-hoc comparisons further revealed that these elevations in wait time occurred selectively within the active zone during the acquisition session, despite locomotor activity being elevated at this time. Thus, mice were indeed capable of remaining in the active zone beyond the initial 5 s of stimulation. A one-way ANOVA analysis of the number of abort events across 1 s time bins during the 5 s stimulation bouts also indicated that these events occurred across varying times points after the initiation of

stimulation (Fig. 4d, $F_{4,30} = 4.3049$, $p = 0.0072$, $\eta^2 = 0.36$). Taken together, these data indicate that activation of mPFC → NAcSh mice inputs was producing purposeful changes in locomotor activity that were associated with the repeated actions being taken in these mice across the self-stimulation sessions.

Preservation of behaviour in mPFC → NAcSh mice when stimulation is discontinued. To further confirm that locomotor activity was associated with repeated actions in mPFC → NAcSh

mice, we also examined acquisition and extinction of behavioural strategies over time in a separate cohort of mice ($n = 8$). Two-way repeated measures ANOVAs were used to analyze overall locomotor behaviour and active zone-specific behaviour over time during these sessions (Fig. 4e–h). We found a significant effect of session on locomotor activity (Fig. 4e, $F_{2,21} = 11.07$, $p = 0.0005$, $\eta^2 = 0.88$), with locomotor activity being elevated during the acquisition compared to baseline or extinction sessions. We also found significant session \times time interaction effects for time spent in the active zone (Fig. 4f, $F_{10,105} = 3.887$, $p = 0.0002$, $\eta^2 = 0.37$) and active zone entries (Fig. 4g, $F_{10,105} = 2.722$, $p = 0.0052$, $\eta^2 = 0.26$), with a trend for interaction on active zone wait times (Fig. 4h, $F_{10,105} = 1.877$, $p = 0.0565$). In contrast to the enhanced locomotor activity, these measures developed over the course of the session, and post hoc comparisons indicated that elevation in the time in zone and zone entry metrics were still expressed early during extinction testing, then extinguished over time. Finally, we used SLEAP³⁵ pose and position analysis to identify any velocity changes that were directly associated with stimulation onset and/or offset during active (Fig. 4i, j, Supplementary Fig. 5) or passive stimulation (Fig. 4k, l, Supplementary Fig. 5) and found no evidence for any consistent alterations in velocity that were temporally paired with either stimulation onset or stimulation offset. Taken together, these data suggest that activation of mPFC \rightarrow NAcSh inputs were reinforcing purposeful actions made by the mice rather than simply producing non-specific elevations in locomotor activity.

BLA \rightarrow NAcSh mice are sensitive to time investments made in the active zone. Given discrepancies between the expression of place preferences and locomotor effects during testing between vHPC \rightarrow NAcSh mice and BLA \rightarrow NAcSh mice, we also sought to better distinguish behaviours across these two inputs by examining another behavioural metric afforded by this foraging paradigm: relationships with the passage of time. Because both vHPC \rightarrow NAcSh mice and BLA \rightarrow NAcSh appeared willing to remain in the active zone during the timeout period to receive subsequent stimulations, we assessed the relative probability these mice would stay in the active zone during the timeout period as a function of time already waited (p(stay), Supplementary Fig. 6a,b). While vHPC \rightarrow NAcSh mice only showed elevations in this metric during reversal testing relative to the baseline session, this time-dependent tendency to remain in the active zone specifically during the timeout period was consistently increased during both acquisition and reversal in BLA \rightarrow NAcSh mice. Further, this metric was significantly elevated in BLA \rightarrow NAcSh compared to vHPC \rightarrow NAcSh during both acquisition and reversal sessions (Supplementary Fig. 6c). This relative increase in sensitivity to time investments (i.e., ‘sunk costs’) was also seen in detrended curves and resulting slopes when data was pooled across inactive (baseline) and active (acquisition and reversal) sessions (Supplementary Fig. 6d, e). Thus, unlike the vHPC \rightarrow NAcSh mice who appeared to develop a more straightforward place preference, the decision of BLA \rightarrow NAcSh mice to remain in the reward-associated context in the absence of any ongoing stimulation was associated with a relatively stronger resistance to leave with the passage of time.

Quadruple dissociation of behaviours across NAcSh inputs.

Finally, to compare behaviour more directly across NAcSh inputs, we pooled behaviour metrics in the active zones across both acquisition and reversal tests and compared them against pooled metrics in the inactive zones during these tests (Fig. 5, Supplementary Fig. 8; baseline values are depicted for comparison but were not included in the analysis). Mixed-model ANOVA

analysis revealed a significant zone (active or inactive) \times brain pathway interaction for time in zone (Fig. 5a, $F_{3,213} = 13.4753$, $p < 0.0001$), stimulation time (Fig. 5b, $F_{3,213} = 19.2294$, $p < 0.0001$), wait time (Fig. 5c, $F_{3,213} = 10.8662$, $p < 0.0001$), total stimulations (Fig. 5d, $F_{3,213} = 16.1441$, $p < 0.0001$), entry stimulations (Fig. 5e, $F_{3,213} = 18.1867$, $p < 0.0001$), stay stimulations (Fig. 5f, $F_{3,213} = 5.3269$, $p = 0.0015$), and p(stay) metrics (Fig. 5g, $F_{3,224} = 31.0202$, $p < 0.0001$) indicating that cross-pathway differences occurred primarily in the active zones. Post-hoc comparisons of behaviour in the active stimulation zones across pathways for these metrics (Supplementary Fig. 7) clearly showed PVT \rightarrow NAcSh mice spent less time in the stimulation-paired zones compared to mPFC \rightarrow NAcSh, vHPC \rightarrow NAcSh, or BLA \rightarrow NAcSh mice (Fig. 5a). Despite spending similar amounts of time in the active zone, we found input-specific differences in the amount of optogenetic stimulation received (Fig. 5b) and timeout-related wait times (Fig. 5c) across mPFC \rightarrow NAcSh, vHPC \rightarrow NAcSh, and BLA \rightarrow NAcSh inputs. While mPFC \rightarrow NAcSh mice tended to self-administer optogenetic stimulation by repeatedly entering the active zone (Fig. 5e), vHPC \rightarrow NAcSh and BLA \rightarrow NAcSh mice were more likely to wait in the active stimulation zone to obtain additional stimulation (Fig. 5c, f). Only BLA \rightarrow NAcSh mice showed increased sensitivity to the passage of time spent waiting in the active zone during timeout periods relative to the other inputs (Fig. 5g). Notably, the input-specific differences in behaviour we observed were still apparent, albeit weaker, when assessed in separate groups of mice at lower frequencies (Supplementary Fig. 8), strongly suggesting that source, rather than amount, of excitatory input to the NAcSh mediated the effects we observed in this study.

Discussion

In the present study we found that optogenetic stimulation of individual excitatory inputs to the NAcSh results in the expression of distinct behavioural outputs in a contextually based spatial self-stimulation task that extends beyond the simple rewarding/aversive dichotomy. We further provide evidence that input-specific changes in locomotor activity during self-stimulation sessions reflect more purposeful alterations in activity rather than non-specific response to the optogenetic stimulation, consistent with previously demonstrated disconnects across NAcSh inputs in stimulation-related locomotor activity and depression-like behaviour³⁶. This quadruple dissociation of reward-related strategies highlights how integration of multiple excitatory inputs to the NAcSh guides motivated behaviour and response selection.

The source and target nuclei investigated in the present study comprise an integrative network that constructs outcome predictions and guides the selection of goal-directed behaviour^{16,37}. Within this network, the ventral striatum serves as a key site of convergent input from cortical and limbic regions that is critical for integration of information and regulation of behavioural outputs^{16,38}. However, the question remains what specific information may be relayed by the individual inputs to drive the behaviours we observed. One intriguing possibility is that individual NAcSh inputs could assist in such information integration by providing different valuation information to the NAcSh related to specific credit assignment functions. Credit assignment is a retrospective process that influences future predictions about outcomes that then guides decision-making and response selection. These functions play a critical role in reward valuation, reinforcement learning, computational decision-making processes, and goal-directed behaviour³⁹. In foraging animals, these functions can contribute to ongoing cost-benefit analyses and help hone behaviour towards a survival-favorable outcome by reinforcing behaviours that increase an animal’s ability to obtain

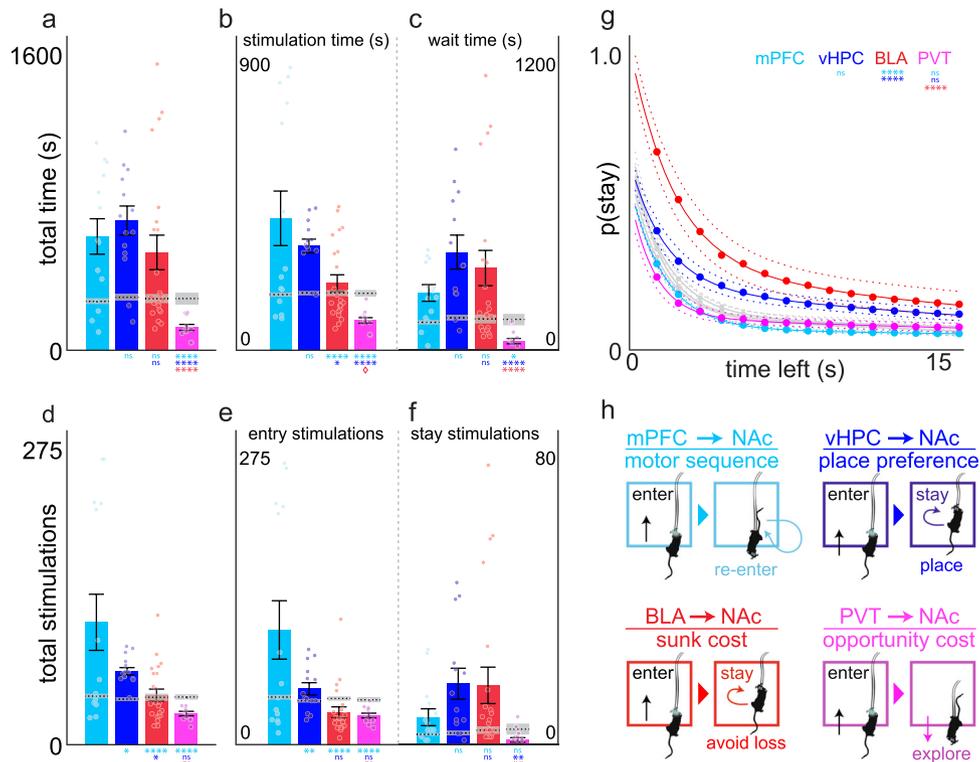


Fig. 5 Direct comparison of optogenetically-induced behavioural strategies across NAcSh inputs. Different NAcSh input groups are depicted by different colours throughout: mPFC → NAcSh (cyan, $n = 7$ mice), vHPC → NAcSh (blue, $n = 7$ mice), and BLA → NAcSh (red, $n = 11$ mice), and PVT → NAcSh (pink, $n = 6$ mice). **a** Total time spent in active stimulation zones during testing. mPFC → NAcSh, vHPC → NAcSh, and BLA → NAcSh mice spent similar amounts of time in active zone, while PVT → NAcSh mice spent relatively less time in the active zones. **b, c** Time spent being actively stimulated (stimulation time, **b**) versus waiting during the timeout period (wait time, **c**). The relative amount of time spent being stimulated versus waiting differed across inputs. **d** The total number of stimulation bouts was significantly higher in mPFC → NAcSh mice compared to other groups, which reflects the high number of stimulations by direct zone entry (**e**). In contrast, vHPC → NAcSh and BLA → NAcSh mice earned higher numbers of stay stimulations by prolonged lingering during the timeout period (**f**). **g** BLA → NAcSh mice were more likely to remain in the active zone throughout the timeout period, distinguishing it from other groups. **h** Proposed credit assignment functions for each set of NAcSh excitatory inputs. Figures represent pooled data (mean ± SEM) from active corner zones during acquisition and reversal sessions. Mixed-model ANOVAs were used to identify pathway × zone (pooled active or inactive) interactions and Tukey's post hoc tests (outlined in Supplementary Fig. 8) were used for between-pathway comparisons within pooled active zones. Inactive zone data are depicted in Supplementary Fig. 9. Dashed lines represent pooled data from the baseline session and are presented for comparison but were not included in the analysis. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$ versus other pathways as specified.

rewards and/or avoid danger. Furthermore, both incentive salience⁴⁰ and Pavlovian action^{41,42} reinforcement learning paradigms involve the credit assignment problem⁴³—given a recent positive or negative experience, how should one weigh sets of convergent stimuli when assigning credit?

The ventromedial mPFC plays an important role in executive control and is involved in reward representation, value-based decision making, action selection, response inhibition, attention, task switching, and habit formation^{16,44–47}. The infralimbic cortical area we targeted in the present study, which preferentially innervates the NAcSh¹⁶, has been further implicated in response-outcome encoding^{47,48} and synaptic plasticity at infralimbic inputs to the NAcSh selectively impacts re-evaluative, but not deliberative, aspects of decision-making processes⁹. While both the mPFC and ventral striatum have previously been identified as being involved in credit assignment processes^{49–51}, to our knowledge this is the first time that direct connections between the mPFC and NAcSh have specifically been implicated. In our case, the behaviour exhibited by mPFC → NAcSh mice is most consistent with this input assigning value to recent actions. These mice made repeated entries into the stimulation-paired zone during acquisition but also showed more variable active zone entry behaviours during reversal testing, consistent with work indicating that the mPFC, and particularly the infralimbic

area targeted here, is important for strategy shifting but is less involved in reversal learning⁵². If mPFC → NAcSh stimulation promoted recent action sequences, we would predict that mice would exhibit repeated active-zone entry during acquisition testing as entering the stimulation-paired zone was the most recent action exhibited prior to stimulation being delivered. This repeated entry into the original acquisition zone would presumably persist initially during reversal, but when this action was not reinforced by further stimulation, mice would likely take different subsequent exploratory search approaches in the arena, and these varying motor sequences would then be reinforced when mice entered the new stimulation-paired reversal zone. Indeed, when mice were given an extinction instead of a reversal session, they did initially preserve their entry behaviour towards the previously active zone, but this behaviour quickly dissipated in the absence of further stimulation. This latter finding is consistent with the role of the infralimbic cortex in learning new stimulus-reward associations during extinction training⁵³. The reinforcement of action-based credit assignment functions through the NAcSh may also provide a way that the infralimbic cortex can promote goal-directed response vigor⁵⁴ while also suppressing unwanted actions⁵⁵.

In addition to assigning credit to beneficial actions, animals must be able to link specific actions and outcomes with associated

environmental stimuli. The hippocampus is a critical for encoding such contextual information about the environment⁵⁶ and the vHPC subregion is known to be important for spatial navigation, context-based associative learning, and emotional and affective processing⁵⁷. The vHPC has also been implicated in value-based decision making⁵⁸ as well as credit assignment processes⁵⁹, although which outputs of the vHPC are critical and what specific valuation information is provided is less clear. The spatial arena we utilized in the present study was contextually rich in that each corner had distinct borders and contextual markers, so these features were the primary ones the animals had to use to identify the location of the stimulation-paired corner. Compared to the other inputs, vHPC → NAcSh stimulation produced the most consistent and selective place preferences for these stimulation-paired corners during both acquisition and reversal testing. These mice also lingered in the stimulation-paired corners during the timeout period in both acquisition and reversal sessions and were more likely to obtain additional bouts of stimulation compared to mPFC → NAcSh mice. Together, these findings indicate that instead of assigning credit to recent actions, vHPC → NAcSh mice were assigning value to the stimulation-paired context itself. Such context-based credit assignments would be consistent with the proposed role of vHPC inputs to the nucleus accumbens in model-based spatial navigation, goal-directed behaviour, and spatial reversal learning^{33,60–64} and would additionally provide a way for vHPC inputs to influence intertemporal choice, cost-benefit decision-making, and approach-avoidance behaviour^{58,65,66}.

The BLA is important for incentive and motivational processing, associative learning, behavioural flexibility, outcome-specific representations, intertemporal choice, and cost-benefit decision-making processes^{31,67}. While BLA projections to the NAc core subregion are important for outcome devaluation, projections to the NAcSh are more involved in outcome-specific Pavlovian-to-instrumental transfer and associative learning that pairs environmental cues and reward-related outcomes^{67–69}. Notably, in our task we did not pair temporally discrete cues with optogenetic stimulation, so this may be one reason why BLA → NAcSh mice self-administered less optogenetic stimulation and were relatively slower to show reward-directed behaviour compared to mPFC → NAcSh and vHPC → NAcSh inputs. However, it is more likely that these effects were primarily driven by the increased sensitivity to time investments that we identified in BLA → NAcSh mice. The assignment of credit to time investments (sunk costs) would also more readily explain why BLA → NAcSh mice showed reductions in locomotor activity, but only when stimulation was contingent on the animal's behaviour. While there is limited evidence for BLA involvement in credit assignment processes⁷⁰, such credit assignments would be very consistent with an alteration in intertemporal choice and an aversion to leaving a reward-associated environment that has been described in foraging animals and humans^{71–78}. Such credit assignments could also explain how BLA → NAcSh stimulation could impact both spatial and temporal aspects of cost-benefit decision-making processes^{79–81}.

Interestingly, only PVT → NAcSh mice showed reductions rather than elevations in active zone-directed behaviour during self-stimulation sessions. However, they did not avoid entering the zone entirely, suggesting that PVT → NAcSh stimulation produced a more complex interaction with the behavioural task that prevented the expression of full avoidance behaviour. Given the heterogeneous nature of the PVT and its implications in arousal and both appetitive and aversive behaviour^{22,23,82}, this complexity may reflect heterogeneity of these projections at either the pre- or post-synaptic level. For example, activity in anterior PVT projections is more associated with appetitive behaviour,

whereas projections from posterior PVT are more often associated with aversive behaviour²³. While we chose to target coordinates that have previously been shown to have dense projections to the NAcSh, produce real-time place aversion upon optogenetic stimulation, and undergo plasticity after exposure to addictive drugs²⁴, it has been also been demonstrated that optogenetic activation of more anterior PVT regions promotes appetitive behaviour^{21,83}. Given the spread of our viral injections across both anterior and posterior regions of the PVT (Supplementary Fig. 1), it is possible we are accessing both types of pathways. On the post-synaptic side, because PVT → NAcSh projections synapse onto both dopamine receptor type 1 (D1+) and type 2 (D2+) containing MSNs²⁵, contingent stimulation of PVT → NAcSh inputs may affect both post-synaptic cell types, resulting in both appetitive and aversive-like responding. However, as all PVT mice showed some level of avoidance of the stimulation-paired zone, these possibilities cannot fully explain the behaviour exhibited by these mice. Alternatively, another credit assignment function, opportunity cost, could explain the behaviour we observed in PVT → NAcSh mice. Activation of this credit assignment function would decrease the value of remaining in the stimulation-paired corner and increase the value of exploring other options. The potential opportunity costs to increase exploratory behaviour seems particularly likely given the low risk and low opportunity cost of doing so in our spatial task^{84,85}. However, after exploring the other corner zones and finding no reward opportunities there, mice would presumably return to the stimulation-paired zone to reassess the situation, which is consistent with our observations (Fig. 2c, f; Fig. 3c, f).

Given the present data, we propose a potential pathway mechanism for how different excitatory inputs provide convergent valuation information to the NAcSh to guide behavioural strategies during reinforcement learning (Fig. 5h). We propose that each excitatory input to the NAcSh reinforces different credit assignment functions, with mPFC → NAcSh assigning value to recent actions; vHPC → NAcSh inputs assigning value to spatial contexts, BLA → NAcSh inputs assigning value to time invested (sunk costs), and PVT → NAcSh inputs assigning value to opportunities. While credit assignment processes in reinforcement learning are typically discussed in relation to striatal dopamine function⁸⁶, our data strongly suggest that striatal glutamate release at individual NAcSh inputs is also critically involved. While specific local pathway effects and network-wide dynamics remain to be elucidated, the convergence of such information in the NAcSh has important implications for both reinforcement learning and value-based decision-making processes and provides mechanistic insight into how behaviours could be selected and adapted during foraging. Given that different experiences will access these different inputs, and that these pathways could be separately affected by diseases or trauma, our findings provide insight regarding how plasticity at different NAcSh inputs contributes to both adaptive and maladaptive learning found across multiple psychiatric disorders.

Methods

Subjects. A total of 125 adult male C57Bl/6J mice (Jackson Laboratories) were used. Mice were ~6 weeks of age (18–22 g) at the beginning of study and were housed in a temperature and humidity-controlled vivarium under a 12-h light-dark cycle (lights on at 0600). All mice were habituated to the vivarium for at least 5 days before undergoing any surgical procedures. Mice were group housed until fiber implant surgery at which time they were single housed. This approach was done to minimize stress due to single housing as well as to preserve the integrity of the implants. All surgical and experimental procedures were approved by the University of Minnesota Institutional Animal Care and Use Committee and followed guidelines of the American Association for the Accreditation of Laboratory Animal Care.

Experimental design. The overall experimental design is illustrated in Fig. 1. Mice underwent one surgery to inject a viral vector that expresses the blue-light activated Channelrhodopsin (ChR2) under a CamKII α promoter to drive expression in glutamatergic projection neurons terminating in the NAcSh, followed by a second surgery to implant an optical fiber above ChR2+ terminal regions in the NAcSh. After recovery, mice underwent three consecutive days of behavioural testing to examine pathway-specific self-stimulation behaviour.

Surgical procedures. Mice were anesthetized with a ketamine/xylazine cocktail (100 and 10 mg/kg, IP) and were given bilateral injections of AAV2-CamKII α -eYFP (control virus) or AAV2-CamKII α -ChR2(H134R)-eYFP (University of North Carolina Vector Core) using a 5 μ l Hamilton syringe with a 29-gauge needle. Injections were targeted to the infralimbic cortex (mPFC: 0 $^\circ$, A/P +1.8, M/L +/−0.4, D/V −3.1 from Bregma), ventral hippocampus (vHPC: 0 $^\circ$, A/P −3.08, M/L +/−2.9, D/V −4.25 from Bregma), basolateral amygdala (BLA: 0 $^\circ$, A/P −1.3, M/L +/−3.1, D/V −4.9 from Bregma), or paraventricular nucleus of the thalamus (PVT: 4 $^\circ$, A/P −1.2, M/L +/−0.1, D/V −3.2 from Bregma). Viral volume was 0.5 μ l/side, except for PVT injections which were 0.25 μ l/side. PVT injections were given bilaterally to mimic bilateral injection parameters of the other brain regions and but were angled to avoid hitting the midline ventricular areas. Coordinates were identified using from a mouse brain atlas⁸⁷ and were consistent with previous research that targeted these regions for assessment of real-time place preference or instrumental self-stimulation procedures^{24,26,27,88}. Approximately 3–4 weeks after virus injection, a second surgery was used to place custom-made optical fiber implants (200 μ m core, 0.66 NA fiber; 230 μ m ID ferrules, ThorLabs). Fibers were implanted bilaterally directly above the nucleus accumbens shell (NAcSh: 14 $^\circ$, A/P +1.5, M/L +/−1.63, D/V −4.1 from Bregma/skull). Fibers were secured to the skull using stainless steel machine screw anchors (0.0625") and Geristore dental acrylic. When not in use, fibers were kept capped (Precision Fiber Products) to prevent damage. Mice were given 5–10 days of recovery before behavioural testing. We chose to target the NAcSh subregion given that it is involved in reward valuation, is a key region where cortical and limbic inputs converge^{11,17,89,90}, and that plasticity in this region has been implicated in multiple psychiatric disorders^{12,14,15,38}. Optogenetic stimulation of glutamatergic mPFC²⁷, vHPC²⁷, and BLA^{26,27,88} inputs to the NAcSh subregion have also been demonstrated to induce real-time place preferences and reinforce instrumental behaviour, whereas excitatory PVT inputs to the NAcSh cause real-time place avoidance²⁴, but also can promote instrumental responding²⁵. Furthermore, as the NAcSh is a site thought to be critical for integration of information³⁸, we wanted to further explore whether different sources of glutamate input to this subregion matter for reward-directed behaviour²⁸.

Optogenetic self-stimulation task. To examine optogenetic self-stimulation behaviour, we developed an open-field arena (20'x20") with isolated corners zones containing different contextual cues (e.g., triangles, dots, horizontal or vertical lines). This approach allowed mice to discriminate between the different corner zones and provided space in the arena where mice could explore that was free of contextual marker and/or stimulation consequences. Mice underwent three consecutive days of testing as described below.

Baseline. Mice were initially habituated to the spatial environment during a 30 min baseline session. In this session, mice were allowed to freely explore the spatial arena, but entry into any of the corner zones was without consequence. The amount of time spent in each zone during this baseline was measured. Neither the most preferred nor the least preferred zone was chosen as the active acquisition or reversal zone for subsequent self-stimulation tests to avoid artificial increases or decreases in time spent in the zone that were unrelated to the optogenetic stimulation itself. Thus, the 2nd or 3rd preferred zones were chosen as active zones and their designation as "acquisition" or "reversal" zones was counterbalanced across experimental groups. Zone assignments for the two other corner zones were also counterbalanced prior to acquisition testing.

Acquisition. Separate groups of mice were assessed for self-stimulation behaviour at a particular frequency (30 Hz, $n = 64$, 20 Hz, $n = 30$, or 10 Hz, $n = 31$) during a 30 min acquisition test. PlexBright LEDs (Plexon) were used to deliver blue light (465 nm, 5 ms pulse width, 10–15 mW, ~4–6 mW/mm² at target tissue) bilaterally through patch cables (200 μ m, 0.66 NA, Plexon) that were connected to brain implants using light shielded zirconia sleeves (1.25 mm OD, Plexon). A single corner zone was designated as the active stimulation zone and entry into this zone triggered an active LED, whereas entry into the other 3 "non-active" zones triggered an inactive LED (mock stimulation). Optogenetic stimulation in the active zone ended immediately upon zone exit, or after a maximum of 5 s if a mouse remained in the active zone. A 15 s timeout period was initiated after 5 s of stimulation after which another train of stimulation was initiated. Thus, mice could bypass the 15 s timeout period by exiting and re-entering the active zone or wait the duration of the 15 s timeout periods to receive additional stimulation. This approach allowed mice to titrate stimulation levels as well as provided mice with different strategy options for stimulation. It also allowed us to collect additional behavioural data beyond the total

time spent in each corner zone including the actual amount of optogenetic stimulation received (real or mocked stimulation time), the number of zone entries that were made, the time spent lingering in a zone during the timeout, and the number of times mice received more than one stimulation train per entry (stay stimulations).

Reversal. Mice received a 30 min reversal test to further examine the expression of the behavioural strategy observed during acquisition testing. The test had the same parameters as during acquisition, except that the previously active zone (acquisition zone) was deactivated, and a different zone was activated (reversal zone).

Extinction. To identify whether behaviours exhibited during acquisition maintained after discontinuation of contingent stimulation in the spatial arena, a separate cohort of mPFC \rightarrow NAcSh ($n = 8$) mice underwent baseline and acquisition sessions as described above, but instead of reversal testing they were given a single 30 min extinction session where stimulation availability was discontinued.

Non-contingent passive stimulation. Mice were assessed for locomotor activity in response to passive (30 Hz, 5 ms pulse width) optogenetic stimulation in a distinct open-field apparatus which consisted of a beige rectangular (22 x 42 x 20 cm) box with corn cob type bedding on the bottom and no contextual cues available. eYFP ($n = 12$), mPFC \rightarrow NAcSh ($n = 8$), vHPC \rightarrow NAcSh ($n = 6$), BLA \rightarrow NAcSh ($n = 4$), and PVT \rightarrow NAcSh ($n = 3$). Mice were initially habituated to the apparatus for 30 min. They then underwent three stimulation periods (5 min each, off-on-off), with optogenetic stimulation being delivered during the middle period (5 s on, 5 s off). Stimulation parameters were set to provide more continual stimulation than was received in the spatial active self-stimulation task, but still mimic stimulation bout timing parameters set in this task. This spacing also allowed us to directly assess locomotor velocity changes that were temporally paired with either onset or offset of passive stimulation.

Histology. After behavioural testing, mice were deeply anesthetized with pentobarbital (Fatal Plus, 390 mg/ml) and transcardially perfused with 1x phosphate buffered saline (PBS) followed by 4% paraformaldehyde. Brains were removed, placed in vials containing 4% paraformaldehyde overnight at 4 $^\circ$ C, and were then transferred to a 20% (1 day) and then 30% sucrose solution until brains had sunk and ready for slicing. A sliding microtome (Leica Biosystems) was used to cut 30 μ m coronal sections, which were subsequently slide mounted and imaged with an epifluorescent microscope (Leica) to identify location of ChR2 virus and optical fiber placement. Only data from mice with correct virus and fiber placement were used for analysis.

Statistics and reproducibility. Behavioural data was collected with ANY-maze (Stoelting Co.) and processed with ANY-maze, Matlab, or Python. JmpPro13/15 or GraphPad Prism9 was used for graphing and data analysis. Dependent variables included time spent in corner zones, time receiving actual or mock stimulation (active vs inactive zones, stimulation time), number of zone entries made (number of entry stimulations), wait time (time spent in corner zone when not being stimulated), number of stay stimulations (times mice received more than one stimulation per entry into a zone, i.e., stayed in zone through entire timeout period), number of abort events, distance traveled (m), velocity (z-scores), and probability of remaining in the zone as a function of time invested (p(stay)). One-way analysis of variance (ANOVA) with corner zone as a fixed factor was used to assess within-pathway behaviour across the four corner zones within each testing session (baseline, acquisition, reversal). A one-way ANOVA with time bin as a fixed factor was used to assess abort events in the acquisition zone during acquisition in mPFC \rightarrow NAcSh mice. Two-way ANOVAs with zone and session as fixed factors were utilized for within-pathway comparisons of wait time behaviour. Two-way repeated measures ANOVAs with session and time as fixed factors and time as the repeated measure were used to assess behaviour over time and across sessions for extinction experiments. Mixed-model ANOVAs with pathway and session/epoch as fixed factors and subject as a random factor were used for across-pathway and session/epoch comparisons of locomotor behaviour. Paired t -tests were used to analyze z-score velocity data averaged 5 s before and after stimulation onset and/or offset. Mixed-model ANOVAs with pathway, session, and time left as fixed factors and subjects as a random factor were used for across-pathway and session comparisons of p(stay) values in vHPC and BLA mice. Mixed-model ANOVAs with pathway and zone type as fixed factors and subject as a random factor were used also for across-pathway comparisons of pooled behavioural metrics across zones (active or inactive) during active stimulation sessions (i.e., acquisition and reversal, baseline data presented for comparison, but not included in these analyses). Significant main or interaction effects were followed by either Student's t -tests (for a priori comparisons) or Tukey's post hoc tests. Significance level for main and interaction effects was set at $p < 0.05$. Effect sizes were calculated as follows: $\eta^2 = SS_{\text{effect}}/SS_{\text{total}}$.

Reporting summary. Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Source data for the main and supplementary figures is provided as Supplementary Data 1. All related raw data and processing codes are available upon reasonable request.

Received: 25 March 2021; Accepted: 5 January 2023;

Published online: 30 January 2023

References

- Calhoun, G. G. & O'Donnell, P. Closing the gate in the limbic striatum: prefrontal suppression of hippocampal and thalamic inputs. *Neuron* **78**, 181–190 (2013).
- Mogenson, G. J., Jones, D. L. & Yim, C. Y. From motivation to action: functional interface between the limbic system and the motor system. *Prog. Neurobiol.* **14**, 69–97 (1980).
- Mulder, A. B., Hodenpjl, M. G. & Lopes da Silva, F. H. Electrophysiology of the hippocampal and amygdaloid projections to the nucleus accumbens of the rat: convergence, segregation, and interaction of inputs. *J. Neurosci.* **18**, 5095–5102 (1998).
- O'Donnell, P. & Grace, A. A. Physiological and morphological properties of accumbens core and shell neurons recorded in vitro. *Synapse* **13**, 135–160 (1993).
- Rescorla, R. A. & Wagner, A. R. in *Classical conditioning II* (eds Black, A. H. & Prokasy, W. F.) 64–99 (Appleton-Century-Crofts, 1972).
- Camara, E., Rodriguez-Fornells, A., Ye, Z. & Münte, T. F. Reward networks in the brain as captured by connectivity measures. *Front Neurosci.* **3**, 350–362 (2009).
- van der Meer, M. A. & Redish, A. D. Expectancies in decision making, reinforcement learning, and ventral striatum. *Front Neurosci.* **4**, 6 (2010).
- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P. & Dolan, R. J. Model-based influences on humans' choices and striatal prediction errors. *Neuron* **69**, 1204–1215 (2011).
- Sweis, B. M., Larson, E. B., Redish, A. D. & Thomas, M. J. Altering gain of the infralimbic-to-accumbens shell circuit alters economically dissociable decision-making algorithms. *Proc. Natl Acad. Sci. USA* **115**, E6347–E6355 (2018).
- Kelley, A. E. Functional specificity of ventral striatal compartments in appetitive behaviors. *Ann. N. Y. Acad. Sci.* **877**, 71–90 (1999).
- Salgado, S. & Kaplitt, M. G. The nucleus accumbens: a comprehensive review. *Stereotact. Funct. Neurosurg.* **93**, 75–93 (2015).
- Russo, S. J. & Nestler, E. J. The brain reward circuitry in mood disorders. *Nat. Rev. Neurosci.* **14**, 609–625 (2013).
- Luthi, A. & Luscher, C. Pathological circuit function underlying addiction and anxiety disorders. *Nat. Neurosci.* **17**, 1635–1643 (2014).
- Hearing, M., Graziane, N., Dong, Y. & Thomas, M. J. Opioid and psychostimulant plasticity: targeting overlap in nucleus accumbens glutamate signaling. *Trends Pharm. Sci.* **39**, 276–294 (2018).
- Turner, B. D., Kashima, D. T., Manz, K. M., Grueter, C. A. & Grueter, B. A. Synaptic plasticity in the nucleus accumbens: lessons learned from experience. *ACS Chem. Neurosci.* **9**, 2114–2126 (2018).
- Sesack, S. R. & Grace, A. A. Cortico-basal ganglia reward network: microcircuitry. *Neuropsychopharmacology* **35**, 27–47 (2010).
- Li, Z. et al. Cell-type-specific afferent innervation of the nucleus accumbens core and shell. *Front Neuroanat.* **12**, 84 (2018).
- Beyeler, A. et al. Divergent routing of positive and negative information from the amygdala during memory retrieval. *Neuron* **90**, 348–361 (2016).
- Do-Monte, F. H., Minier-Toribio, A., Quinones-Laracuente, K., Medina-Colon, E. M. & Quirk, G. J. Thalamic regulation of sucrose seeking during unexpected reward omission. *Neuron* **94**, 388–400.e384 (2017).
- Millan, E. Z., Kim, H. A. & Janak, P. H. Optogenetic activation of amygdala projections to nucleus accumbens can arrest conditioned and unconditioned alcohol consummatory behavior. *Neuroscience* **360**, 106–117 (2017).
- Cheng, J. et al. Anterior paraventricular thalamus to nucleus accumbens projection is involved in feeding behavior in a novel environment. *Front Mol. Neurosci.* **11**, 202 (2018).
- Barson, J. R., Mack, N. R. & Gao, W. J. The paraventricular nucleus of the thalamus is an important node in the emotional processing network. *Front Behav. Neurosci.* **14**, 598469 (2020).
- McGinty, J. F. & Otis, J. M. Heterogeneity in the paraventricular thalamus: the traffic light of motivated behaviors. *Front Behav. Neurosci.* **14**, 590528 (2020).
- Zhu, Y., Wienecke, C. F., Nachtrab, G. & Chen, X. A thalamic input to the nucleus accumbens mediates opiate dependence. *Nature* **530**, 219–222 (2016).
- Lafferty, C. K., Yang, A. K., Mendoza, J. A. & Britt, J. P. Nucleus accumbens cell type- and input-specific suppression of unproductive reward seeking. *Cell Rep.* **30**, 3729–3742.e3723 (2020).
- Stuber, G. D. et al. Excitatory transmission from the amygdala to nucleus accumbens facilitates reward seeking. *Nature* **475**, 377–380 (2011).
- Britt, J. P. et al. Synaptic and behavioral profile of multiple glutamatergic inputs to the nucleus accumbens. *Neuron* **76**, 790–803 (2012).
- Tye, K. M. Glutamate inputs to the nucleus accumbens: does source matter? *Neuron* **76**, 671–673 (2012).
- Euston, D. R., Gruber, A. J. & McNaughton, B. L. The role of medial prefrontal cortex in memory and decision making. *Neuron* **76**, 1057–1070 (2012).
- Strange, B. A., Witter, M. P., Lein, E. S. & Moser, E. I. Functional organization of the hippocampal longitudinal axis. *Nat. Rev. Neurosci.* **15**, 655–669 (2014).
- Wassum, K. M. & Izquierdo, A. The basolateral amygdala in reward learning and addiction. *Neurosci. Biobehav. Rev.* **57**, 271–283 (2015).
- Goodroe, S. C., Starnes, J. & Brown, T. I. The complex nature of hippocampal-striatal interactions in spatial navigation. *Front Hum. Neurosci.* **12**, 250 (2018).
- Stoianov, I. P., Pennartz, C. M. A., Lansink, C. S. & Pezzulo, G. Model-based spatial navigation in the hippocampus-ventral striatum circuit: A computational analysis. *PLoS Comput Biol.* **14**, e1006316 (2018).
- Stephens, D. W. Decision ecology: foraging and the ecology of animal decision making. *Cogn. Affect. Behav. Neurosci.* **8**, 475–484 (2008).
- Pereira, T. D. et al. SLEAP: A deep learning system for multi-animal pose tracking. *Nat. Methods* **19**, 486–495 (2022).
- Bagot, R. C. et al. Ventral hippocampal afferents to the nucleus accumbens regulate susceptibility to depression. *Nat. Commun.* **6**, 7062 (2015).
- Pezzulo, G. Goals reconfigure cognition by modulating predictive processes in the brain. *Behav. Brain Sci.* **37**, 154–155 (2014).
- Goto, Y. & Grace, A. A. Limbic and cortical information processing in the nucleus accumbens. *Trends Neurosci.* **31**, 552–558 (2008).
- Sutton, R. S. *Temporal credit assignment in reinforcement learning.* (University of Massachusetts Amherst, 1984).
- Berridge, K. C. & Robinson, T. E. Liking, wanting, and the incentive-sensitization theory of addiction. *Am. Psychol.* **71**, 670–679 (2016).
- Dayan, P. & Balleine, B. W. Reward, motivation, and reinforcement learning. *Neuron* **36**, 285–298 (2002).
- Dayan, P. & Berridge, K. C. Model-based and model-free Pavlovian reward learning: reevaluation, revision, and revelation. *Cogn. Affect. Behav. Neurosci.* **14**, 473–492 (2014).
- Sutton, R. S. & Barto, A. G. *Reinforcement learning: An introduction* (MIT press, 2018).
- Wang, Q. et al. Distributed value representation in the medial prefrontal cortex during ontemporal choices. *J. Neurosci.* **34**, 7522 (2014).
- Domenech, P. & Koechlin, E. Executive control and decision-making in the prefrontal cortex. *Curr. Opin. Behav. Sci.* **1**, 101–106 (2015).
- Hiser, J. & Koenigs, M. The multifaceted role of the ventromedial prefrontal cortex in emotion, decision making, social cognition, and psychopathology. *Biol. Psychiatry* **83**, 638–647 (2018).
- Maisson, D. J. N. et al. Choice-relevant information transformation along a ventrodorsal axis in the medial prefrontal cortex. *Nat. Commun.* **12**, 4830 (2021).
- Barker, J. M., Glen, W. B., Linsenbardt, D. N., Lapish, C. C. & Chandler, L. J. Habitual behavior is mediated by a shift in response-outcome encoding by infralimbic cortex. *eneuro* **4**, ENEURO.0337–0317.2017 (2017).
- Lim, S.-L., O'Doherty, J. P. & Rangel, A. The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *J. Neurosci.* **31**, 13214–13223 (2011).
- Akaishi, R. & Hayden, B. Y. A spotlight on reward. *Neuron* **90**, 1148–1150 (2016).
- Lim, D. H., Yoon, Y. J., Her, E., Huh, S. & Jung, M. W. Active maintenance of eligibility trace in rodent prefrontal cortex. *Sci. Rep.* **10**, 18860 (2020).
- Rich, E. L. & Shapiro, M. L. Prelimbic/infralimbic inactivation impairs memory for multiple task switches, but not flexible selection of familiar tasks. *J. Neurosci.* **27**, 4747–4755 (2007).
- Nett, K. E. & LaLumiere, R. T. Infralimbic cortex functioning across motivated behaviors: can the differences be reconciled? *Neurosci. Biobehav. Rev.* **131**, 704–721 (2021).
- Riveros, M. E., Forray, M. I., Torrealba, F. & Valdés, J. L. Effort displayed during appetitive phase of feeding behavior requires infralimbic cortex activity and histamine H1 receptor signaling. *Front. Neurosci.* <https://doi.org/10.3389/fnins.2019.00577> (2019).
- Capuzzo, G. & Floresco, S. B. Prelimbic and infralimbic prefrontal regulation of active and inhibitory avoidance and reward-seeking. *J. Neurosci.* **40**, 4773–4787 (2020).
- Smith, D. M. & Bulkin, D. A. The form and function of hippocampal context representations. *Neurosci. Biobehav. Rev.* **40**, 52–61 (2014).
- Fanselow, M. S. & Dong, H. W. Are the dorsal and ventral hippocampus functionally distinct structures? *Neuron* **65**, 7–19 (2010).

58. Schumacher, A., Vlassov, E. & Ito, R. The ventral hippocampus, but not the dorsal hippocampus is critical for learned approach-avoidance decision making. *Hippocampus* **26**, 530–542 (2016).
59. Duncan, K., Doll, B. B., Daw, N. D. & Shohamy, D. More than the sum of its parts: a role for the hippocampus in configural reinforcement learning. *Neuron* **98**, 645–657. e646 (2018).
60. Pennartz, C. M., Ito, R., Verschure, P. F., Battaglia, F. P. & Robbins, T. W. The hippocampal-striatal axis in learning, prediction and goal-directed behavior. *Trends Neurosci.* **34**, 548–559 (2011).
61. Ciochi, S., Passecker, J., Malagon-Vina, H., Mikus, N. & Klausberger, T. Selective information routing by ventral hippocampal CA1 projection neurons. *Science* **348**, 560–563 (2015).
62. Barker, J. M., Bryant, K. G. & Chandler, L. J. Inactivation of ventral hippocampus projections promotes sensitivity to changes in contingency. *Learn Mem.* **26**, 1–8 (2019).
63. Avigan, P. D., Cammack, K. & Shapiro, M. L. Flexible spatial learning requires both the dorsal and ventral hippocampus and their functional interactions with the prefrontal cortex. *Hippocampus* **30**, 733–744 (2020).
64. Cernotova, D., Stuchlik, A. & Svoboda, J. Roles of the ventral hippocampus and medial prefrontal cortex in spatial reversal learning and attentional set-shifting. *Neurobiol. Learn Mem.* **183**, 107477 (2021).
65. McHugh, S. B., Campbell, T. G., Taylor, A. M., Rawlins, J. N. & Bannerman, D. M. A role for dorsal and ventral hippocampus in inter-temporal choice cost-benefit decision making. *Behav. Neurosci.* **122**, 1–8 (2008).
66. Abela, A. R. & Chudasama, Y. Dissociable contributions of the ventral hippocampus and orbitofrontal cortex to decision-making with a delayed or uncertain outcome. *Eur. J. Neurosci.* **37**, 640–647 (2013).
67. Keefer, S. E., Gyawali, U. & Calu, D. J. Choose your path: divergent basolateral amygdala efferents differentially mediate incentive motivation, flexibility and decision-making. *Behav. Brain Res* **409**, 113306 (2021).
68. Shiflett, M. W. & Balleine, B. W. At the limbic-motor interface: disconnection of basolateral amygdala from nucleus accumbens core and shell reveals dissociable components of incentive motivation. *Eur. J. Neurosci.* **32**, 1735–1743 (2010).
69. Corbit, L. H. & Balleine, B. W. The general and outcome-specific forms of Pavlovian-instrumental transfer are differentially mediated by the nucleus accumbens core and shell. *J. Neurosci.* **31**, 11786–11794 (2011).
70. Chau, B. K. et al. Contrasting roles for orbitofrontal cortex and amygdala in credit assignment and learning in macaques. *Neuron* **87**, 1106–1118 (2015).
71. Nonacs, P. State dependent behavior and the marginal value theorem. *Behav. Ecol.* **12**, 71–83 (2001).
72. Hayden, B. Y., Pearson, J. M. & Platt, M. L. Neuronal basis of sequential foraging decisions in a patchy environment. *Nat. Neurosci.* **14**, 933–939 (2011).
73. Wikenheiser, A. M., Stephens, D. W. & Redish, A. D. Subjective costs drive overly patient foraging strategies in rats on an intertemporal foraging task. *Proc. Natl Acad. Sci. USA* **110**, 8308–8313 (2013).
74. Blanchard, T. C. & Hayden, B. Y. Monkeys are more patient in a foraging task than in a standard intertemporal choice task. *PLoS One* **10**, e0117057 (2015).
75. Carter, E. C., Pedersen, E. J. & McCullough, M. E. Reassessing intertemporal choice: human decision-making is more optimal in a foraging task than in a self-control task. *Front Psychol.* **6**, 95 (2015).
76. Constantino, S. M. & Daw, N. D. Learning the opportunity cost of time in a patch-foraging task. *Cogn. Affect Behav. Neurosci.* **15**, 837–853 (2015).
77. Carter, E. C. & Redish, A. D. Rats value time differently on equivalent foraging and delay-discounting tasks. *J. Exp. Psychol. Gen.* **145**, 1093–1101 (2016).
78. Sweis, B. M. et al. Sensitivity to “sunk costs” in mice, rats, and humans. *Science* **361**, 178–181 (2018).
79. Peck, C. J., Lau, B. & Salzman, C. D. The primate amygdala combines information about space and value. *Nat. Neurosci.* **16**, 340–348 (2013).
80. Amir, A., Lee, S. C., Headley, D. B., Herzallah, M. M. & Pare, D. Amygdala signaling during foraging in a hazardous environment. *J. Neurosci.* **35**, 12994–13005 (2015).
81. Orsini, C. A., Trotta, R. T., Bizon, J. L. & Setlow, B. Dissociable roles for the basolateral amygdala and orbitofrontal cortex in decision-making under risk of punishment. *J. Neurosci.* **35**, 1368–1379 (2015).
82. Kirouac, G. J. Placing the paraventricular nucleus of the thalamus within the brain circuits that control behavior. *Neurosci. Biobehav Rev.* **56**, 315–329 (2015).
83. Labouebe, G., Boutrel, B., Tarussio, D. & Thorens, B. Glucose-responsive neurons of the paraventricular thalamus control sucrose-seeking behavior. *Nat. Neurosci.* **19**, 999–1002 (2016).
84. Eccard, J. A. & Liesenjohann, T. The importance of predation risk and missed opportunity costs for context-dependent foraging patterns. *PLoS One* **9**, e94107 (2014).
85. Gruber, A. J., Thapa, R. & Randolph, S. H. Feeder approach between trials is increased by uncertainty and affects subsequent choices. *eneuro* **4**, ENEURO.0437–0417.2017 (2017).
86. Deserno, L. et al. Dopamine enhances model-free credit assignment through boosting of retrospective model-based inference. *eLife* **10**, e67778 (2021).
87. Paxinos, G. & Franklin, K. B. *The Mouse Brain in Stereotaxic Coordinates* (Academic press, 2019).
88. Tye, K. M. et al. Amygdala circuitry mediating reversible and bidirectional control of anxiety. *Nature* **471**, 358–362 (2011).
89. Brog, J. S., Salyapongse, A., Deutch, A. Y. & Zahm, D. S. The patterns of afferent innervation of the core and shell in the “accumbens” part of the rat ventral striatum: immunohistochemical detection of retrogradely transported fluoro-gold. *J. Comp. Neurol.* **338**, 255–278 (1993).
90. Voorn, P., Vanderschuren, L. J., Groenewegen, H. J., Robbins, T. W. & Pennartz, C. M. Putting a spin on the dorsal-ventral divide of the striatum. *Trends Neurosci.* **27**, 468–474 (2004).

Acknowledgements

This work was supported by grants from the National Institute on Drug Abuse (R01 DA019666, R01 DA041808, K02 DA035459 and P30 DA048742) and the Breyer-Longden Family Foundation. Thank you to the University of Minnesota MnDRIVE Optogenetics Core for access to optogenetic equipment and resources and to Ethan Huffington, Cynthia Zheng, Sonal Nagpal, Lucie Ozbek, and Megan Brickner for technical assistance making optical fibers and performing histology.

Author contributions

E.B.L., A.J.A., and M.J.T. designed the experiments; E.B.L., A.J.A., and K.A.S. performed surgeries; E.B.L. and A.J.A. ran the behavioural experiments; E.B.L., A.J.A., and K.A.S. performed histology; E.B.L., B.M.S., and M.E. prepared the figures and analyzed the data; E.B.L., B.M.S., A.D.R. and M.J.T. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-023-04429-6>.

Correspondence and requests for materials should be addressed to Mark J. Thomas.

Peer review information *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editor: Karli Montague-Cardoso.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023