RESEARCH ARTICLE SUMMARY

NEUROSCIENCE

A neuronal mechanism for motivational control of behavior

J. Courtin†*, Y. Bitterman†, S. Müller, J. Hinz, K. M. Hagihara, C. Müller, A. Lüthi*

INTRODUCTION: Acting to achieve goals underlies animals' behavioral flexibility and depends on the ability to motivate specific behaviors. Formal studies of goal-directed behavior are based on paradigms such as instrumental conditioning, in which subjects learn that specific actions result in a particular outcome. By definition, instrumental goal-directed actions are oriented toward specific outcomes and are sensitive to variations in outcome value and action-outcome contingency. However, the neuronal mechanisms that encode and maintain specific motivational control of self-paced behavior are poorly understood.

RATIONALE: Animal motivation to engage in goal-directed behavior is governed by a multitude of factors, ranging from the animal's general metabolic state to specific task parameters including outcome identity, outcome value, and action-outcome contingency. Using Ca^{2+} imaging, electrophysiology, and optogenetic manipulations, we examined the relationship between noncued, self-paced goal-directed actions and neuronal responses in the amygdala. Although the amygdala is a brain structure well known for encoding associative learning in the context of cue-triggered outcome-seeking behaviors, here, we specifically focused on neuronal correlates of essential parameters supporting goal-directed actions, both during action execution and reward consumption.

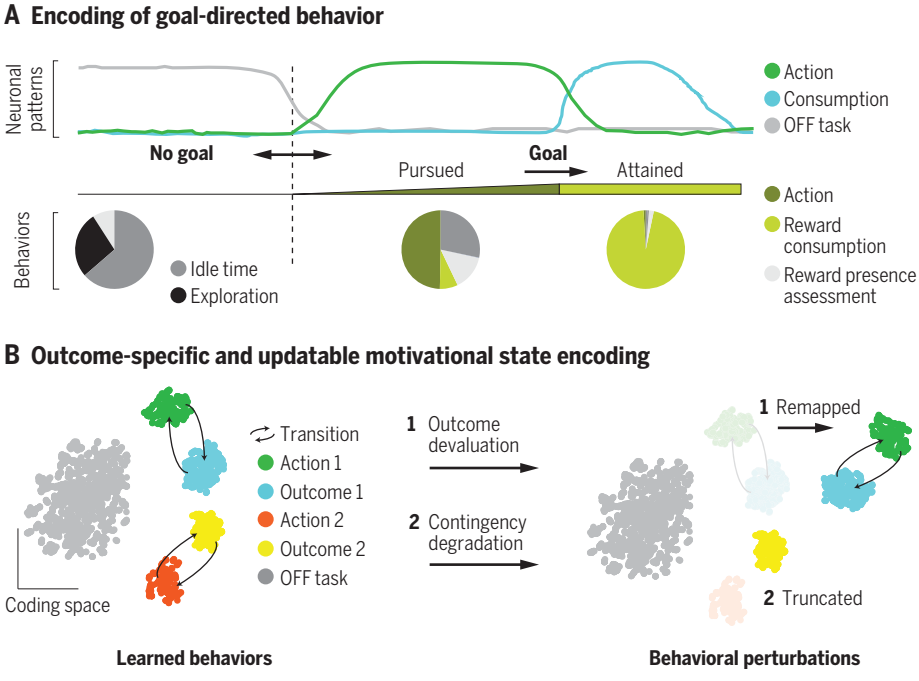
RESULTS: Mice were trained in a self-paced, instrumental goal-directed task in which one action led to the delivery of a sucrose reward and another action to the delivery of a milk reward, with no explicit stimuli signaling trial start or reward availability. Based on an automatically derived description of mouse behavior, we showed that after 5 days of training, mice exhibited stereotypical self-paced behavioral sequences, alternating between periods of action execution when mice pursued a goal and reward consumption when the goal was attained.

Using in vivo deep-brain Ca^{2+} imaging and electrophysiology, we found distinct outcome-specific ensembles of amygdala neurons activated during action-consumption sequences. These ensembles were spatially intermingled and broadcasted behaviorally relevant information to multiple striatum subregions. At the population level, distinct amygdala neuronal activity patterns tiled the entire action-consumption behavioral sequences. These internally generated action- and consumption-associated patterns maintained outcome-specific information along the session and across days.

Optogenetic inhibition of amygdala neurons during either action or consumption caused distinct behavioral phenotypes and indicated that action- and consumption-associated amygdala activity was necessary for the outcome-specific motivational control of goal-directed behavior.

Last, in agreement with experimental psychology models, we showed that action and consumption representations learned during training can be retrieved, truncated, or remapped upon manipulations of outcome value, such as after outcome devaluation or extinction or upon degradation of action-outcome contingency. More specifically, whereas outcome devaluation resulted in the immediate emergence of new action and outcome representations, action-outcome contingency degradation resulted in a gradual loss of the action representation with no effect on reward consumption representation.

CONCLUSION: We show here that the interplay between distinct outcome-specific and updatable neuronal activity patterns encodes and maintains specific motivational states necessary for the execution of adaptive, self-initiated action-consumption sequences. Amygdala neurons segregated behaviorally relevant information along different axes, such as outcome identity, value, and expectancy. Our study thus introduces the concept of specific motivation and extends the classic view that the amygdala controls outcome-seeking behavior by conferring value to outcome-predicting sensory stimuli. ■



A neuronal mechanism for noncued goal-directed behavior. (**A**) The interplay between distinct neuronal activity patterns encodes and maintains behaviorally relevant information at distinct time points along goal-directed action-consumption behavioral sequences. (**B**) Outcome-specific action and consumption patterns reflect the maintenance of a specific motivational state. Upon behavioral perturbations, action and consumption patterns differentially adapt. This adaptive code might be key to direct behavior toward a pursued goal.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: julien.courtin@fmi.ch (J.C.); andreas.luthi@fmi.ch (A.L.)
†These authors contributed equally to this work.
Cite this article as J. Courtin *et al.*, *Science* **375**, eabg7277 (2022). DOI: 10.1126/science.abg7277

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abg7277

RESEARCH ARTICLE

NEUROSCIENCE

A neuronal mechanism for motivational control of behavior

J. Courtin[1]*†, Y. Bitterman[1]†, S. Müller[1], J. Hinz[1,2], K. M. Hagihara[1,2], C. Müller[1], A. Lüthi[1,2]*

Acting to achieve goals depends on the ability to motivate specific behaviors based on their predicted consequences given an individual's internal state. However, the underlying neuronal mechanisms that encode and maintain such specific motivational control of behavior are poorly understood. Here, we used Ca^{2+} imaging and optogenetic manipulations in the basolateral amygdala of freely moving mice performing noncued, self-paced instrumental goal-directed actions to receive and consume rewards. We found that distinct neuronal activity patterns sequentially represent the entire action-consumption behavioral sequence. Whereas action-associated patterns integrated the identity, value, and expectancy of pursued goals, consumption-associated patterns reflected the identity and value of experienced outcomes. Thus, the interplay between these patterns allows the maintenance of specific motivational states necessary to adaptively direct behavior toward prospective rewards.

Maintaining motivational states allows animals to direct behavior to achieve desired goals (*1*). Formal studies of goal-directed behavior are based on paradigms such as instrumental conditioning, in which the subject learns through experience that specific actions result in a particular outcome. By definition, instrumental goal-directed actions are oriented toward specific outcomes and are sensitive to variations in outcome value and action-outcome contingency (*2*). However, current knowledge about the encoding of these parameters in the basolateral amygdala (BLA) is based on the use of cue-triggered outcome-seeking behaviors (*3*–*9*). Evidence of BLA involvement in noncued, self-paced instrumental goal-directed actions is mainly provided by system-level studies (*10*–*15*). Such studies suggest that the BLA confers specific motivational significance to goal-directed actions. However, the neuronal mechanisms encoding and maintaining such specific motivational states remain unknown.

Results

Self-paced action-consumption behavioral sequences

We trained food-restricted male mice in a self-paced instrumental goal-directed task in which one of two actions led to the delivery of a sucrose reward and the other to the delivery of a milk reward, with no explicit stimuli signaling trial start or reward availability. The number of actions required to obtain rewards increased over 5 training days on a variable ratio schedule of reinforcement [from constant reinforcement to variable ratio 5 (VR5); Fig. 1A]. During training, both instrumental actions were acquired at a similar rate (Fig. 1B). By day 5, performance of the actions was both hunger state dependent and goal directed, as demonstrated by its sensitivity to a change in hunger state, to specific satiety-induced outcome devaluation, and to action-outcome contingency degradation (Fig. 1, B and C). During each training session, mice were sequentially exposed to phases during which one or both instrumental actions was available or not (ON and OFF task phases, respectively). Using action and lick time stamps together with video tracking, we automatically derived a continuous description of mouse behavior during the entire session (Fig. 1D and fig. S1, A and B). We differentiated epochs of task-related behaviors such as instrumental actions, unrewarded lick and rewarded lick are epochs from epochs of non-task-related behaviors including idle times, when mice were in the task zone but did not perform any task-related behaviors, and context exploration idle times and context exploration are epochs, when mice ventured away from the task-related zone. Across training sessions, behaviors during ON task phases were increasingly composed of task-related behaviors. By contrast, behaviors during OFF task phases were stable across days (Fig. 1E and fig. S1, C and D).

By day 5, during ON task phases, mice repeatedly alternated between action and consumption periods. We thus defined a behavioral sequence as the time window between the first action of an action period and the last lick of the subsequent consumption period. Action periods, defined as the time from the first action to the last action of a behavioral sequence, included epochs of actions, unrewarded licks, and intermittent non-task-related behaviors, whereas consumption periods were exclusively composed of rewarded lick epochs (Fig. 1F). These action-consumption sequences were self-paced, as made evident by the occurrence of intermittent unrewarded lick epochs and the occasional performance of actions even after outcome delivery (fig. S2A). As described classically for variable ratio reinforcement schedules, mice showed constant action response rates with brief pauses after consumption periods (*16*). Action rate, duration of the distinct behavioral epochs, action to lick latency, and interbehavioral sequence intervals were stable across the session (Fig. 1, F and G, and fig. S2, B and C).

Specific activation of distinct BLA ensembles during action and consumption

To evaluate the neuronal correlates of goal-directed behavior in the BLA, we performed deep-brain Ca^{2+} imaging of BLA Ca^{2+}/calmodulin-dependent protein kinase 2 (CaMK2)–expressing principal neurons (PNs) using a miniaturized microscope and GCaMP6f as a Ca^{2+} indicator. The activity of single BLA PNs was extracted using constrained non-negative matrix factorization for microendoscopic data (CNMF-E) (see the materials and methods and fig. S3A; 113 ± 12 neurons per mouse, 905 total neurons from eight mice). On day 5, neuronal activity showed clear time-locked modulations to the different behavioral epochs during ON task phases (Fig. 1H). Across multiple action-consumption sequences, single neuron responses faithfully tracked the stereotypical behavioral switches between action and consumption periods (Fig. 2A). We therefore defined three functional types of BLA PNs. Action neurons (n = 137/905) displayed increased activity when mice were engaged in the performance of goal-directed actions (Fig. 2, A and B, and fig. S4, A to D). By contrast, consumption neurons (n = 170/905) exhibited increased activity time locked to rewarded lick epochs, when mice collected the reward. Finally, transition neurons (n = 39/905) showed increased activity during both action epochs and unrewarded lick or rewarded lick epochs (Fig. 2, A and B, and fig. S4, A to D). The three functional types of neurons were spatially intermingled, without clear anatomical organization either within each individual imaging field or along the anteroposterior (AP) or mediolateral (ML) BLA axes (fig. S3, B to F).

Of all BLA PNs recorded on day 5, 38% were significantly task modulated (n = 346/905 neurons from eight mice). In all three functional subsets, most neurons were outcome specific. That is, they discriminated between actions leading to, or the consumption of, a sucrose or milk reward. A small, but above chance level, proportion of neurons was not outcome specific (responding to milk and sucrose; $P < 0.001$, χ^2 test; Fig. 2C). Even though

[1]Friedrich Miescher Institute for Biomedical Research, CH-4058 Basel, Switzerland. [2]University of Basel, CH-4000 Basel, Switzerland.
*Corresponding author. Email: julien.courtin@fmi.ch (J.C.); andreas.luthi@fmi.ch (A.L.)
†These authors contributed equally to this work.

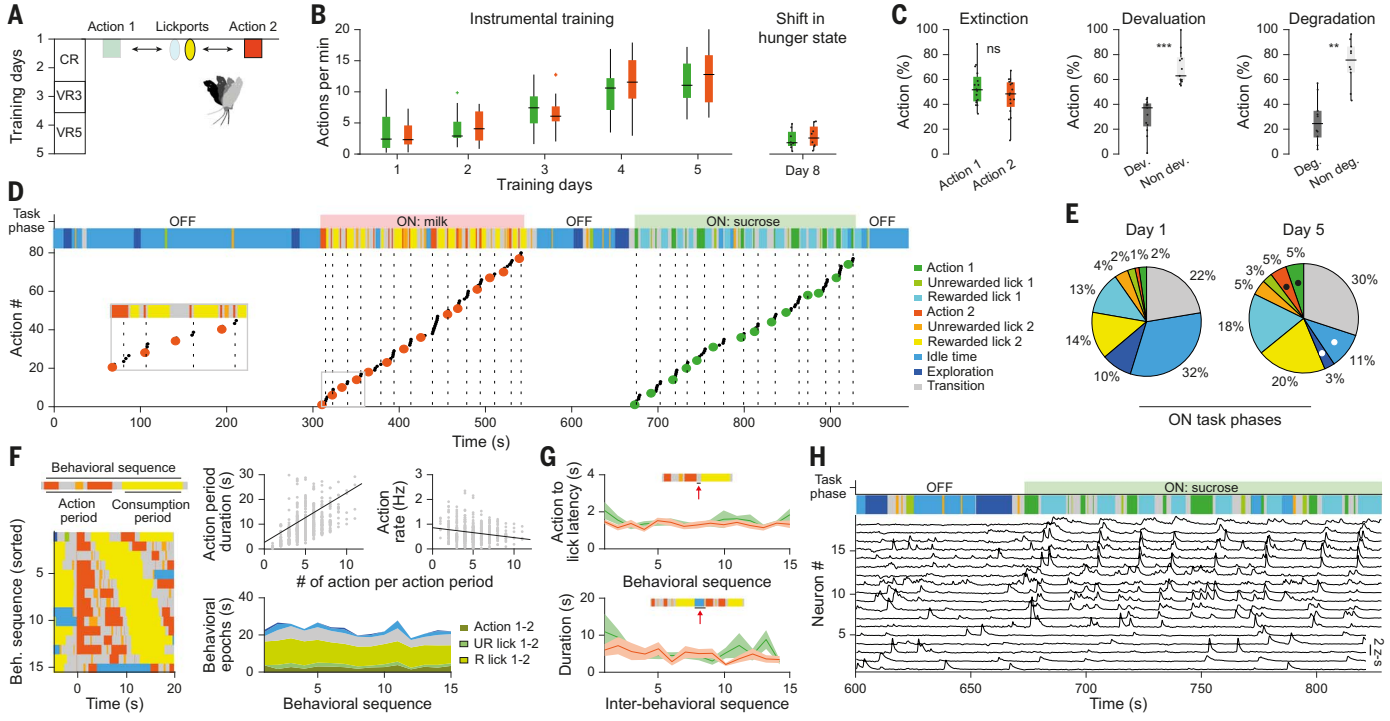


Fig. 1. Ca^{2+} imaging of BLA PNs during the learned, self-paced, instrumental goal-directed task. (**A**) Schematic of the behavioral context. Action availability is controlled and the required number of actions per outcome increases over training from CR to VR5. (**B**) Number of actions per minute over training (N = 16 mice in two cohorts) and during the free choice rewarded session in sated mice (N = 8 mice in two cohorts). (**C**) Action choice during the extinction test, satiety-induced outcome devaluation test, and action-outcome contingency degradation test (all tests were non-reinforced from two mice cohorts; N = 16, N = 13, and N = 10 mice, P = 0.72, P < 0.001, and P < 0.01, respectively; two-sided paired t tests comparing action choice). Box-and-whisker plots indicate median, interquartile, extreme data values, and outliers of the data distribution. (**D**) Task phases and behavioral epochs labeling during day 5 for one mouse. Black dots indicate individual actions; red and green dots, action periods onset (for milk and sucrose, respectively); vertical dashed lines, outcome delivery. Inset, initial behavioral sequences. Colors indicate the different behavioral epochs. (**E**) The proportion of time mice engaged in the different behavioral epochs during ON task phases on training days 1 and 5 (N = 16 mice; black or white dots denote significant increase or decrease in behavioral epochs duration, respectively; P < 0.01, two-sided Wilcoxon signed-rank tests comparing day 1 and 5 epoch durations). (**F**) Behavioral sequences aligned to action period onset (0 s) and sorted based on action period duration (left) while milk outcome was available. Duration of action periods (r = 0.21; P < 0.001) and action rate within action period (r = −0.07; P = 0.1) across action periods (top right). Mean duration of the distinct behavioral epochs (bottom right, N = 16 mice, day 5). (**G**) Action to lick latency and duration of the interbehavioral sequence intervals across behavioral sequences (N = 16 mice, day 5). Shading indicates SEM. (**H**) Ca^{2+} traces from 18 BLA CaMK2-expressing neurons simultaneously imaged during 4 min of day 5 [same mouse as in (D)]. Scale bar, 1 z-score. Colors at the top panels denote the different behaviors as in (D).

neuronal representations were outcome specific, behaviors of the two ON task phases were similar, suggesting comparable outcome values. Action and consumption neurons were oppositely modulated during ON task phases: Neurons activated during action were silent during consumption and vice versa (Fig. 2B). Accordingly, pairwise correlation and coactivation analysis of neuronal activity of action and/or consumption neurons indicated that coincident activity within each functional subset increased and coincident activity across functional subsets decreased during ON task phases (Fig. 2D). These patterns of coincidence were limited to the ON task phase, suggesting task-dependent functional interactions (fig. S4, E and F). We confirmed the results obtained with Ca^{2+} imaging using single-unit recordings (n = 78 neurons from four mice). Similar to the imaging results, 28% of all recorded neurons exhibited significant task-related modulation (P = 0.10 and χ^2 = 2.7, Friedman test). Correlation and co-firing analysis of pairwise neuronal activity between action and/or consumption neurons confirmed task-dependent functional interactions (fig. S5).

Because the activity of action and consumption neurons was outcome specific, we tested the performance of a linear decoder trained with the imaging dataset on the five-way distinction among action periods, consumption periods (milk or sucrose), and periods of non-task-related behavior during the OFF task phase. Decoder performance was high when trained with activity traces of all PNs (F1 score = 87.8 ± 3.2%) and dropped to chance levels when trained on shuffled neuronal data (19.7 ± 0.1%, P < 0.01, paired t test; Fig. 2F and fig. S4G). When subsets of PNs were selected, performance was higher for training with the activity of all task-modulated neurons (69.7 ± 6.2%) than for training with random selections of non-task-modulated neurons (57.3 ± 5.1%; see the materials and methods).

BLA activity patterns tile the entire action-consumption sequences

To investigate population-level representational principles and contrast them with our observations at the level of single neurons, we obtained population activity vectors by binning single neuron activity in 200-ms bins. Population activity vectors time locked to the beginning of action and consumption epochs showed regularity and specificity (pairwise Pearson's correlations action-action, r = 0.24 ± 0.02; consumption-consumption, r = 0.52 ± 0.02; action-consumption, r = −0.004 ± 0.01 for this example mouse; Fig. 3A). Accordingly, pairwise correlations between population activity vectors along the entire behavioral session mirrored both the different task phases and the behavioral epochs in action-outcome

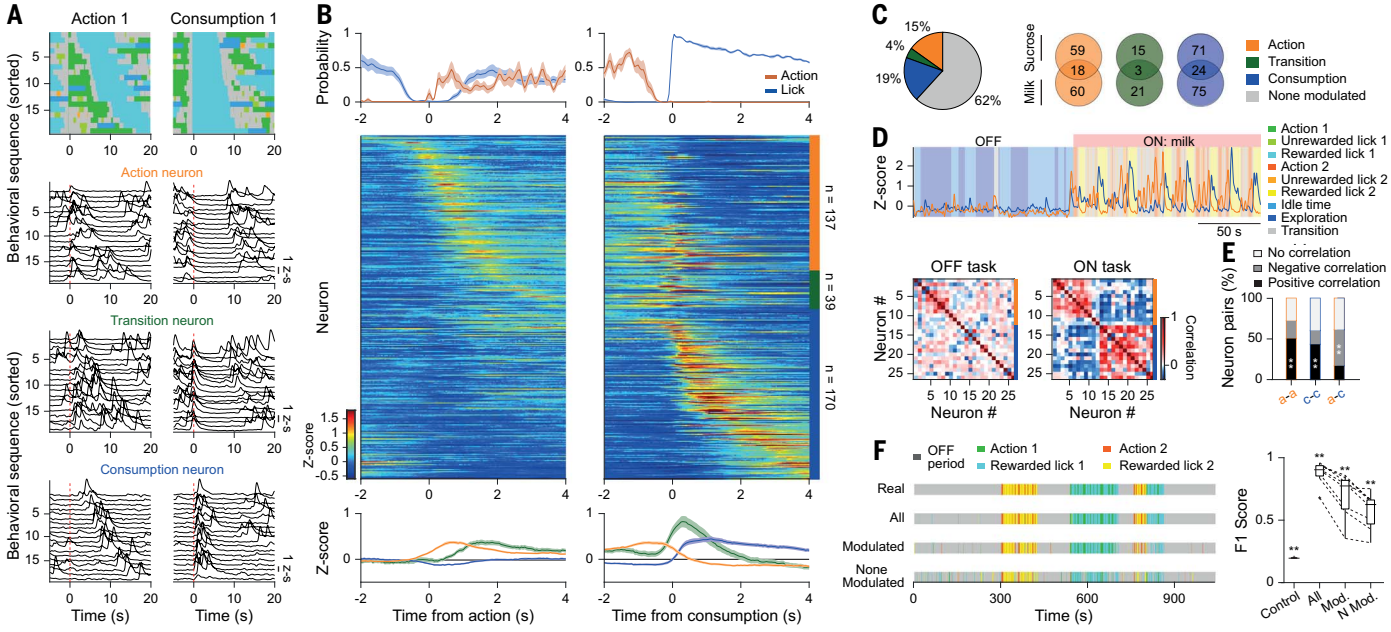


Fig. 2. Distinct sets of BLA PNs encode specific behavioral epochs. (**A**) Mouse behavior (top) and corresponding activity of three neurons (bottom) during 20 behavioral sequences when sucrose outcome was available. The behavioral sequences were aligned to action period onset (left) or to consumption period onset (right) and sorted by action or consumption period duration [colors at the top panels denote behavioral epochs as in (D)]. (**B**) Action and lick probability (top, around action and consumption periods onsets), individual neuron response of all task-modulated neurons (middle), and average response by functional group (bottom; n = 137 action, n = 39 transition, and n = 170 consumption neurons) aligned to action periods (left) and consumption periods (right) onsets (N = 8 mice in two cohorts, n = 905 neurons recorded on day 5). Shading indicates SEM. (**C**) Proportion (left) and outcome selectivity (right) of the task-modulated neurons. (**D**) Average activity of 10 action neurons (orange) and 16 consumption neurons (blue) during OFF and ON phases (top) aligned to the mouse behavior (background colors). Also shown is the pairwise correlation of activity for these neurons during the OFF and ON task phases (bottom). (**E**) Percentage of correlated neuron pairs during ON task phases on day 5 for action-action (a-a), consumption-consumption (c-c), and action-consumption (a-c) neuron pairs (**P < 0.01). (**F**) Performance of a linear decoder trained on the distinction between the five behavioral epochs depicted using responses of all PNs (n = 80 ± 10, mean ± SEM neurons per mouse), only the task-modulated neurons (Mod., n = 33 ± 4 neurons per mouse), or a random set of none-task-modulated neurons (N Mod., n = 33 ± 4 neurons randomly chosen to match the number of modulated neurons). Left, Example mouse decoder results compared with the real behavioral description (top). Right, Mean F score for all mice (N = 8, two-sided paired t test, **P < 0.01). Box-and-whisker plots indicate median, interquartile, extreme data values, and outliers of the data distribution.

behavioral sequences (Fig. 3B). To quantify the mapping between behavior and the population activity vector, we averaged the pairwise correlations between all population activity vectors in action and consumption epochs. Intrabehavioral epoch correlations were positive for outcome-specific action and consumption epochs and higher than interbehavioral epoch correlations (action-consumption) and OFF task correlations (Fig. 3C and fig. S6B). Throughout the session, the population activity vectors during action and consumption epochs remained positively correlated and distinct from OFF task population activity vectors (Fig. 3D and fig. S6C), suggesting that specific neuronal activity patterns robustly define distinct behavioral epochs in an outcome-specific manner. We then used dimensionality reduction [t-distributed stochastic neighbor embedding (t-SNE); see the materials and methods] with correlation as the distance measure to visually demonstrate the distinction between the neuronal activity patterns associated with the OFF task phase and the two outcome-specific ON task phases (fig. S6A). Mapping the population activity vectors associated with the different behavioral epochs to the embedded activity space revealed discrete clusters corresponding to distinct neuronal activity patterns occurring at specific behavioral epochs (Fig. 3E and fig. S6, B and C).

Defining action-associated neuronal activity patterns as all population activity vectors correlated with the population response occurring at the time of the first action of a behavioral sequence (P < 0.01; see the materials and methods), action-associated neuronal activity patterns were not restricted to action epochs but had already started before the first action of a behavioral sequence and lasted until reward detection (Fig. 3F). Indeed, the action-associated neuronal activity patterns that were maintained for several seconds throughout their respective action periods encompassed unrewarded lick and idle time epochs but not exploration epochs (Fig. 3G and fig. S6D). Therefore, BLA action-associated neuronal activity patterns persistently represent information about the pursued outcome from the initiation of the behavioral sequence until reward detection, not just during action performance. By tracking neurons across days, we observed that BLA action-associated neuronal activity patterns were reactivated upon re-exposure to the task during both non-reinforced and reinforced tests, but their activity was maintained only in the latter case (fig. S7). By contrast, consumption-associated neuronal activity patterns were restricted to consumption epochs and were not detected during unrewarded lick epochs, suggesting that consumption-associated neuronal activity patterns reflect reward detection rather than licking behavior (Fig. 3, F and G, and fig. S6D). Using single-unit recordings and inferred spikes extracted from deconvoluted signal using CNMF-E, we confirmed that these results were not merely reflecting slow Ca^{2+} dynamics (figs. S8 and S9).

Specific motivational control of goal-directed behavior

To test whether action- and consumption-associated BLA activity is necessary for instrumental goal-directed behavior, we inhibited the

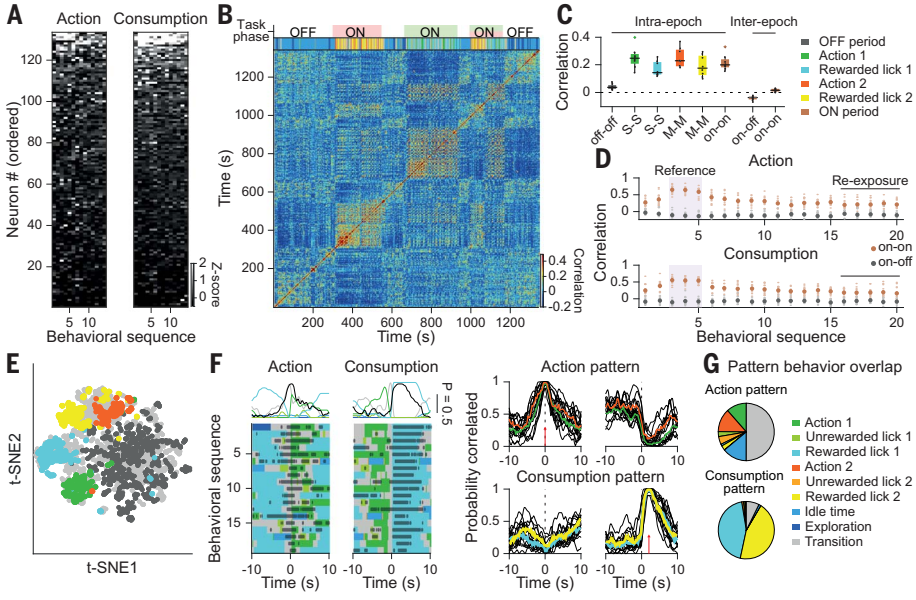


Fig. 3. Distinct BLA neuronal activity patterns maintain outcome-specific information along action-outcome behavioral sequences. (**A**) Population activity vectors from one mouse occurring at action period onset (left) and 2 s after the consumption period onset (right panel; bin size, 200 ms). Neurons ordered for each panel on the basis of their average activity over behavioral sequences. (**B**) Cross-correlation matrix of population activity vectors for the entire task. Same mouse as in (A) (bin size, 200 ms). Top, Task phases (red, milk; green, sucrose) and behavioral epochs labeling; colors are as in (G). (**C**) Intra- and interbehavioral epoch correlation between population activity vectors (N = 8 mice in two cohorts; colors denote the different behavioral epochs). Box-and-whisker plots indicate median, interquartile, extreme data values, and outliers of the data distribution. (**D**) Correlation between the reference activity vector and population activity vectors occurring at action (top) or rewarded lick (bottom) epochs along 20 behavioral sequences of day 5 session (N = 8 mice × two outcomes; the activity vectors selected as a mean reference for the correlation analysis are shaded). (**E**) 2D embedding (t-SNE) of all population activity vectors from (B) using correlation as a distance measure. (**F**) Left, Example of the maintenance of action (left)– or consumption (right)–locked activity patterns overlaid on behavior for 20 behavioral sequences. Black dots denote positive and significant correlation (P < 0.01) between population activity vectors and a reference activity pattern for each behavioral sequence (action: activity at action period onset; consumption: activity 2 s after consumption period onset). Probability of a correlated neuronal pattern and probability of behavioral epochs occurrences for each time bin is shown on the top (bin size, 200 ms; scale bars, probability of 0.5). Right, Probability of correlated action- or consumption-associated activity patterns locked to action and consumption period onsets (N = 8 mice × two outcomes). (**G**) Proportions of different behavioral epochs that occurred when the activity pattern is correlated with the action (top) or consumption (bottom) activity pattern.

activity of BLA PNs on day 5 using an optogenetic approach. We expressed archaerhodopsin (ArchT) or green fluorescent protein (GFP; control group) in PNs and delivered light (589 nm) in a closed-loop manner to independently inhibit neuronal activity during individual action or consumption periods (fig. S10). For each manipulation and each outcome, we compared behaviors between laser-OFF (behavioral sequences 1 to 8) and laser-ON periods (behavioral sequences 9 to 16) (Fig. 4A and fig. S11). Inhibition of PNs during action or consumption periods promoted non-task-related behaviors such as idle times and context exploration epochs, which consequently resulted in a general slowdown of instrumental performance (Fig. 4, B and C, and fig. S11, C and D). However, task-related behaviors such as instrumental actions, unrewarded lick epochs, and rewarded lick epochs were generally not affected by the two types of manipulation (Fig. 4D and fig. S11, E to K).

Inhibition of BLA PNs during either action or consumption periods caused markedly distinct behavioral phenotypes. Inhibition of BLA PNs during consumption periods selectively prolonged the interbehavioral sequence duration, i.e., action initiation of subsequent behavioral sequences was delayed. Consumption behavior was not affected (fig. S11, E to K). Inhibition of BLA PNs during action periods prolonged action period duration, latency to reward consumption, and interbehavioral sequence duration, thereby both extending the time to obtaining the reward and delaying action initiation (Fig. 4B and fig. S11D). No significant differences were observed when comparing behaviors during laser-OFF periods of both ON-task phases (all comparisons, P > 0.15), indicating that perturbing the neuronal activity when one of the rewards was available did not affect behavior in the subsequent phase, when the other reward was available. Moreover, all ArchT-expressing mice earned and consumed the maximum number of rewards.

During optogenetic inhibition of BLA PNs, mice did not show obvious aversive reactions, but rather engaged in non-task-related behaviors. Consistently, inhibition of BLA PNs did not result in real-time place avoidance (fig. S11L) or in mice switching preference to bigger rewards when they were allowed to choose between licking a 5% sucrose solution without laser stimulation and a 20% sucrose solution paired with laser (fig. S11M).

To confirm the causal role of action-associated activity in the motivational control of goal-directed behavior, we inhibited the activity of BLA PNs under non-reinforced conditions by delivering light (589 nm) in a closed-loop manner during individual action epochs (fig. S12). Inhibition of BLA PNs during a free choice non-reinforced test impaired instrumental performance (fig. S12, C to F). Moreover, inhibition of BLA PNs during specific satiety-induced outcome devaluation and action-outcome contingency degradation tests impaired both instrumental performance and action choice bias classically induced by both procedures (fig. S12, G to N).

Behavioral manipulations reveal an adaptive BLA code

To determine how BLA activity supporting goal-directed behaviors adapts to variations in outcome value and action-outcome contingency, we tracked how action- and consumption-associated activity was affected by manipulations of such behavioral parameters. First, we investigated the consequence of either specific satiety-induced outcome devaluation (six mice) or action-outcome contingency degradation (five mice) on BLA action-associated activity (fig. S13). During both non-reinforced tests, mice showed a clear bias in action choice toward the non-devaluated or non-degraded actions, and, concomitantly, only the activity associated with these actions was reactivated. This was evident at both the single-neuron level and the population level (fig. S13, C, D, G, and H).

We then simultaneously assessed how both action- and consumption-associated activity was affected by violating the expected outcome value and action-outcome contingency under reinforced conditions (Fig. 5, A and G, and fig. S14). After the completion of 5 training days, mice were first allowed to obtain eight rewards after the established VR5 schedule (referred to as the initial period), after which a perturbation was executed. During the violation of outcome value paradigm, outcomes

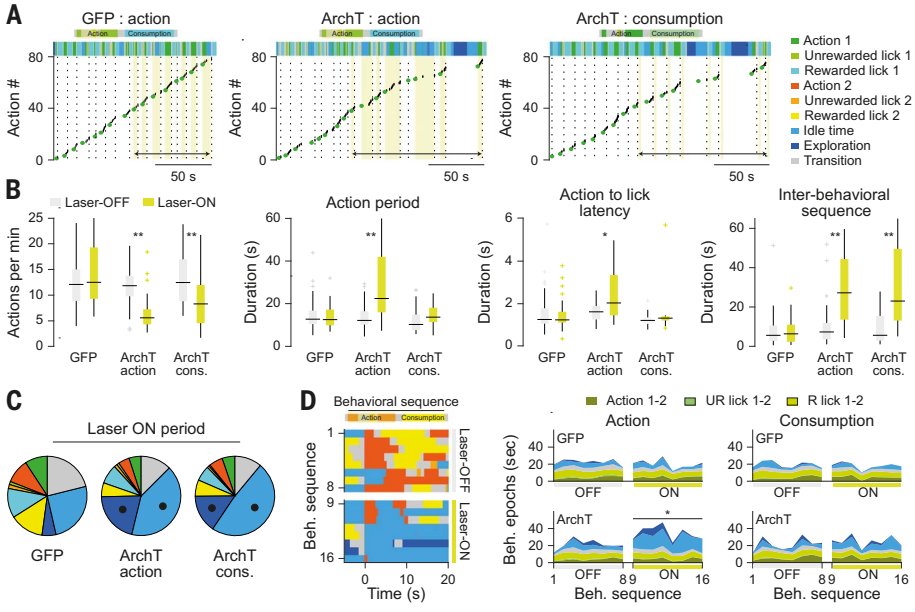


Fig. 4. BLA PNs activity is necessary for the maintenance of goal-directed actions. (**A**) Examples of the effects of optogenetic manipulations on goal-directed behaviors on day 5 for GFP- and ArchT-expressing mice. Black dots indicate individual actions; green dots, action period onset; vertical dashed lines, outcome delivery. Colors indicate behavioral epochs. Shaded yellow areas indicate laser deliveries; double-headed arrows, laser-ON periods (behavioral sequences 9 to 16). (**B**) Left to right, actions per minute, action period duration, action to lick latency, and interbehavioral sequence duration during laser-OFF (behavioral sequences 1 to 8, white bars) and laser-ON periods, respectively (yellow bars; two-sided paired *t* test for action per minute and two-sided Wilcoxon signed-rank tests for the other variables comparing laser-OFF and laser-ON periods, $*P < 0.05$; $**P < 0.01$). GFP, $N = 20$ mice in three cohorts × two outcomes ($N = 12$, $N = 8$ mice with laser stimulation during action periods and consumption periods, respectively, with data pooled). ArchT action, $N = 12$ mice in two cohorts × two outcomes (laser delivered during action periods); ArchT cons., $N = 8$ mice in two cohorts × two outcomes (laser delivered during consumption periods). Box-and-whisker plots indicate median, interquartile, extreme data values, and outliers of the data distribution. (**C**) Proportions of different behavioral epochs during laser-ON periods (color denote behavioral epochs as in (A); Black dots denote significant increase in behavioral epoch duration; $P < 0.01$, two-sided Wilcoxon signed-rank tests comparing laser-OFF and laser-ON periods; see raw values in fig. S11). (**D**) Mouse behavior during laser-OFF and laser-ON periods while milk outcome was available (left). Colors denote behavioral epochs as in (A). Also shown is the duration of the distinct behavioral epochs during the behavioral sequences for each sequence along laser-OFF and laser-ON periods (right, two-sided Wilcoxon signed-rank test).

9 to 16 were delivered on the usual VR5 schedule but were diluted by a factor of 4. This caused a disruption in the established goal-directed behavior without affecting consumption behavior. Mice exhibited significantly longer action period duration, interbehavioral sequence duration, and latency to reward consumption (Fig. 5, A and B) caused by increased non-task-related behaviors both inside and between behavioral sequences (fig. S14, A to C). Eventually, all mice obtained and consumed the maximum number of available outcomes.

We then tested the effect of the violation of the outcome value on the neuronal representations of action and consumption at both the single-neuron level and the population level (326 neurons from four mice). First, whereas action and consumption neurons detected during the initial period showed significantly reduced responses during the perturbed period (Fig. 5, C and D), additional consumption and action neurons emerged during the perturbed period (fig. S15, C and D). Similarly, immediately after the start of outcome value violation (i.e., after consumption of the first diluted reward) new action- and consumption-associated neuronal activity patterns emerged that replaced the action- and consumption-associated neuronal activity patterns of the initial period (Fig. 5, E and F, and fig. S15F). This effect was also observed when considering only neuronal activity patterns associated with the first or the last action in an action period, indicating that it was not a consequence of the increased action period duration (fig. S15G, and see comparison with neuronal activity pattern stability on day 5 in fig. S16).

During the violation of action-outcome contingency paradigm, the initial eight rewards were delivered contingent on the actions on the usual VR5 schedule. Subsequently, outcomes 9 to 20 were delivered noncontingently on the animal's actions. Whereas at the behavioral level, this test disrupted goal-directed behavior similarly to the violation of outcome value (Fig. 5, G and H, and fig. S14, D to F), the impact on BLA neuronal activity was different (434 neurons from five mice). Action neurons detected during the initial period showed a reduction of their responses during the entire perturbed period (Fig. 5I). However, unlike the immediate effect of violating outcome value, this change developed slowly after the start of the perturbation and was not accompanied by the emergence of additional action neurons (fig. S15H). In accordance with this finding, at the population level, the action-associated neuronal activity patterns gradually lost their correlation with the action-associated neuronal activity patterns of the initial period, but no new action-associated neuronal activity patterns emerged during the perturbed period (Fig. 5K and fig. S15J; also see the results for the first and last action periods in fig. S15L). In clear contrast to the impact of the violation of outcome value, consumption-associated activity during the initial period after violation of action-outcome contingency remained stable along the entire perturbed period (Fig. 5, J to L, and fig. S15K). As previously reported (*17*), we were able to detect a population of consumption neurons that gradually increased their activity over the course of unexpected reward deliveries (fig. S15I). However, the overall population activity pattern during consumption remained highly correlated throughout the entire session. Although changing outcome value resulted in the immediate emergence of new action and outcome representations, action-outcome contingency violation resulted in a gradual loss of the representation of the action with no emergence of a new stable action representation and no effect on reward consumption representation.

Discussion

Our findings support a crucial role for the BLA in the motivational control of goal-directed behavior. At the time of goal-directed action, BLA PNs integrated and encoded pursued outcome identity, pursued outcome value, and action-outcome contingency information. The maintenance of such prospective, outcome-specific, and updatable neuronal activity reflects a specific motivational state (*18*, *19*) that differs from a primary motivational state known to energize goal-directed actions in an unspecific manner (*1*, *20*). Furthermore, this state could be retrieved upon re-exposure to the task, but its maintenance depended on continued reinforcement. At the time of consumption, BLA PNs represented current outcome identity and value but not licking behavior. This is consistent with the observations that BLA neurons encode distinct reward features (*21*), including value (*22*), magnitude

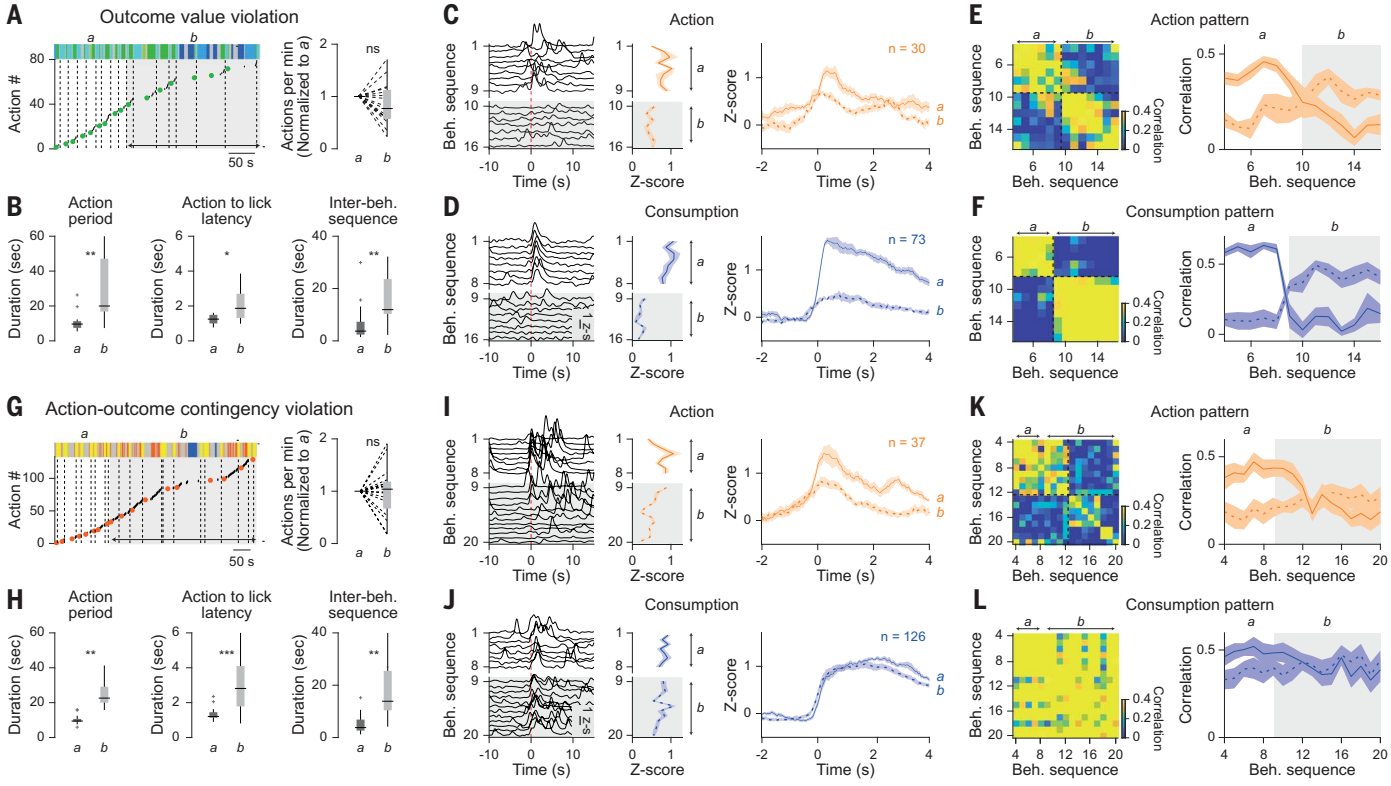


Fig. 5. BLA PNs adaptively control goal-directed actions. (**A**) Example of the effects of outcome value violation on goal-directed behaviors (left). Black dots indicate individual actions; green dots, action period onset; vertical dashed lines, outcome delivery (*a*, initial period, white; *b*, perturbed period, shaded area). Colors indicate behavioral epochs. Shown are actions per minute normalized to the initial period (right, $N = 8$ mice in two cohorts × two outcomes; $P > 0.15$, two-sided paired *t* test). (**B**) Action period duration (left), action to lick latency (middle), and interbehavioral sequence duration (right) during the initial and perturbed periods (*$P < 0.05$, **$P < 001$, two-sided Wilcoxon signed-rank tests). (**C**) Left, Action neuron activity aligned to action periods onset during the initial and perturbed (shaded area) periods. Middle, Mean activation per sequence of initial action neurons ($n = 30$, with four mice in two cohorts × two outcomes averaged over 1 s after action initiation; $P < 0.05$, two-sided paired *t* test comparing initial and perturbed period *z*-score values). Right, Average response of action (orange) and consumption (purple) neurons during the initial (full line) and perturbed (dashed line) periods. (**D**) Same as (C) but for initial consumption neurons aligned to consumption period onset ($n = 73$; averaged over 3 s after consumption initiation; $P < 0.001$, two-sided paired *t* test). Shading indicates SEM. (**E**) Left, Pairwise correlation between action activity vectors for one mouse. Dashed lines indicate the transition between neuronal patterns. Right, Dynamics of the correlation between action activity vectors ($N = 4$ mice × two outcomes), averaging correlations before (full line) and after (dashed line) the transition. (**F**) Same as (E) for consumption activity vectors. (**G** to **L**) Same as (A) to (F) but for action-outcome contingency violation ($N = 8$ mice in two cohorts for behavioral data; $N = 5$ mice in two cohorts for neuronal data). For (I), $n = 37$ action neurons ($P < 0.01$, two-sided paired *t* test). For (J), $n = 126$ consumption neurons ($P = 0.28$, two-sided paired *t* test). Box-and-whisker plots indicate median, interquartile, extreme data values, and outliers of the data distribution in all panels.

(*23*), and probability, (*17*) in a state-dependent manner. Such representations are thought to reflect the incentive value of outcomes (*24*). Optogenetic inhibition of BLA PNs during consumption delayed the initiation of action-consumption sequences without affecting consummatory behavior. Action and consumption PNs formed distinct functional populations showing opposite activity during action and consumption periods. This divergent feature was also observable at the population level, because action- and consumption-associated neuronal activity patterns were anticorrelated. Together, action- and consumption-associated activity integrated behaviorally relevant information at distinct time points along goal-directed action-consumption sequences. Upon behavioral perturbations, action- and consumption-associated activity differentially adapted. Although changing outcome value resulted in the remapping of both activities, affecting action-outcome contingency resulted in the degradation of action-associated activity only. This adaptive code might be a central component to direct behavior toward a pursued goal. Overall, these results indicate that the interplay between action- and consumption-associated neuronal activity patterns maintains specific motivational states necessary for the execution of self-initiated, goal-directed action-consumption sequences, and demonstrate that the BLA is a key neuronal substrate for the control of voluntary behavior.

In contrast to a large body of work on the contribution of the BLA in outcome-seeking behaviors in Pavlovian tasks or in tasks combining instrumental actions with Pavlovian cues that signal trial start and/or reward availability, execution of action consumption sequences in the present study was entirely self-paced. Our findings thus extend the classic model implicating BLA in controlling outcome-seeking behavior by conferring motivational significance to outcome-predicting sensory stimuli (*3*–*5*, *9*), showing encoding of the specific motivational state from before action initiation until reward attainment, consistent with findings from self-controlled behavior in monkeys (*18*). These results are consistent with BLA lesion, inactivation, or pharmacological studies that, despite notable discrepancies regarding the effect of pre- and post-training manipulations, support the view that noncued instrumental goal-directed action performance depends on BLA integrity (*12*).

We found that BLA PNs, at both the single-neuron level and the population level, differentiate

outcomes of similar value according to their identity (milk or sucrose). This distinction was encoded by distinct neuronal activity patterns during goal-directed action and consumption. Functional antagonism between BLA ensembles or circuits has been already observed in spatial exploration tasks (*25*) and in diverse Pavlovian tasks (*26*–*30*), where distinct BLA ensembles signal behaviors of opposite valence (avoidance and approach behaviors). Our results extend this view by showing that BLA population activity segregates behaviorally relevant information along different axes, such as outcome identity, value, and expectancy. Thereby, the BLA adaptively provides updated outcome-related information along the execution of goal-directed action-consumption sequences, with detailed representation of the motivational significance of action, suggesting a central role for the BLA in the circuit that supports the capacity for goal-directed action.

Although we found that most recorded BLA neurons were not significantly task modulated (*31*, *32*), we observed that neurons that do not show significant task-modulated activity carry outcome-specific information regarding the different periods of a behavioral sequence. Thus, as a population, BLA neurons form neuronal activity patterns that maintain specific action- and consumption-related information along the entire temporal extent of action-consumption behavioral sequences. The underlying neuronal mechanisms may involve precise interactions between specific functional subpopulations through intra-BLA recurrent inhibitory and excitatory circuits (*3*, *33*, *34*) under the control of neuromodulatory inputs (*35*, *36*). Moreover, BLA action and consumption neurons may be driven by differential external inputs from a variety of brain structures that process relevant information (*37*–*40*) and project to distinct targets, including the central nucleus of the amygdala (*41*), the nucleus accumbens (*42*), and the dorsomedial striatum (*43*). Action and consumption neurons could be found among BLA neurons projecting to the medial shell of the nucleus accumbens or to the posterior dorsomedial striatum (figs. S17 and S18). Although we cannot exclude that other BLA projections would be functionally more homogeneous, our data show that information encoded by BLA PNs along action-consumption behavioral sequences is broadcasted to distinct striatum subregions known to differentially control goal-directed behavior (*42*).

Materials and Methods

Animals

All animal procedures were performed in accordance with institutional guidelines at the Friedrich Miescher Institute for Biomedical Research and were approved by the Veterinary Department of the Basel-Stadt Canton. Male mice (C57BL/6JRccHsd, Envigo) were used throughout the study. We chose to do experiments with male mice because they are on average heavier than female mice, which facilitates the carrying of the electrode or gradient-index (GRIN) lenses implants during locomotion. Moreover, this allowed us to directly compare our present results with previous studies. Mice were individually housed for at least 2 weeks before starting behavioral paradigms and were kept in a 12-hour light–dark cycle. Mouse well-being was monitored throughout the experimental period. All behavioral experiments were conducted during the light cycle.

Surgical procedures

Eight-week-old mice were anesthetized using isoflurane (3 to 5% for induction, 1 to 2% for maintenance; Attane, Provet) in oxygen-enriched air (Oxymat 3, Weinmann) and placed in a stereotactic frame (model 1900, Kopf Instruments). Injections of buprenorphine (Temgesic, Indivior UK; 0.1 mg/kg body weight subcutaneously 30 min before anesthesia) and ropivacain (Naropin, AstraZeneca; 0.1 ml locally under the scalp before incision) were provided for analgesia. Postoperative pain medication included buprenorphine (0.1 mg/kg body weight in the drinking water overnight) and injections of meloxicam (Metacam, Boehringer Ingelheim; 1 mg/kg subcutaneously) for 3 days after surgery. Ophthalmic gel was applied to avoid eye drying (Viscotears, Bausch + Lomb). The body temperature of the experimental animal was maintained at 36°C using a feedback-controlled heating pad (FHC).

Surgeries for deep-brain Ca^{2+} imaging

AAV2/5.CaMK2.GCaMP6f (University of Pennsylvania Vector Core) was unilaterally injected into the BLA (300 nl) using a precision micropositioner (model 2650, Kopf Instruments) and pulled glass pipettes (tip diameter ~20 μm) connected to a Picospritzer III microinjection system (Parker Hannifin) at the following coordinates from bregma: anterior–posterior (AP): −1.6 mm; medial–lateral (ML): −3.35 mm; dorsal–ventral (DV): 4.2 mm. retroAAV.Ef1a.GCaMP6f (FMI Vector Core) was unilaterally injected either into the caudal dorsomedial striatum (DMS; 100 nl) at the following coordinates from bregma: AP: +0.5 mm; ML: −1 mm; DV: 2 mm or along rostro-caudal axis of the medial shell of the nucleus accumbens (mShell; 100 nl) at the following coordinates from bregma: AP: +1.3 or +0.8 mm; ML: −0.5 mm; DV: 4.2 to 4.4 mm. After injection, a track above the imaging site was cut with a sterile needle to aid the insertion of the GRIN lens (0.6 × 7.3 mm, GLP-0673, Inscopix). The GRIN lens was subsequently lowered into the brain using a micropositioner through the track (4.3 mm from brain surface) with a custom-built lens holder and fixed to the skull using ultraviolet light-curable glue (Henkel, Loctite 4305). Dental acrylic (Paladur, Heraeus) was used to seal the skull and attach a custom-made head bar for animal fixation during the miniature microscope mounting procedure. Mice were allowed to recover for 3 weeks after surgery before checking for GCaMP expression.

Surgeries for optogenetic sessions

AAV2/5.CaMK2.ArchT.GFP or AAV2/5.CaMK2.GFP (University of Pennsylvania Vector Core) was bilaterally injected into the BLA (200 nl per hemisphere, coordinates from bregma: AP: −1.6 mm; ML: −3.3 mm; DV: 4.2 mm). After injection, mice were bilaterally implanted with custom-made optic fiber connectors. Fiber tips were lowered to −3.9 mm below the brain surface using a micropositioner. Implants were fixed to the skull using cyanoacrylate glue (Ultra Gel, Henkel) and miniature screws (P.A. Precision Screws). Dental acrylic mixed with black paint (to avoid light spread) was used to seal the skull. Mice were allowed to recover for 3 weeks before behavioral training to ensure adequate virus expression.

Surgeries for in vivo electrophysiology

A set of mice was unilaterally implanted with a custom-built 32-wire electrode into the BLA using a micropositioner (coordinates from bregma: AP: −1.6 mm; ML: −3.3 mm; DV: 4.2 mm). Electrodes consisted of two bundles of 16 individually insulated, gold-plated nichrome wires (13 μm inner diameter, impedance 30 to 100 kΩ, Sandvik). Each wire in the bundles was attached to an 18-pin connector (Omnetics). Another set of mice was equipped with a custom-built optrode consisting of an optic fiber with an attached 16-wire electrode. Electrodes tips were cut at an angle to protrude ~300 to 500 μm beyond the fiber end. In this case, the surgical procedure was similar to that described above for the optogenetic section. Implants were fixed to the skull using cyanoacrylate glue (Ultra Gel, Henkel) and miniature screws (P.A. Precision Screws). Dental acrylic was used to seal the skull. Mice were allowed to recover for 2 weeks before recordings started.

Instrumental goal-directed behavioral context

The two behavioral contexts used in this study measured 26 cm L × 25 cm W × 40 cm H and each context was enclosed inside an acoustic foam isolated box. One context was equipped with two nose poke ports (left/right; H21-09M_CA, Coulbourn instruments) and the other one with two levers (left/right; ENV-312-2M, Med Associates). All *operanda* (two instrumental actions apparatus, left/right; and two lick ports, left/right) were located on the same wall. In each context, the two instrumental actions apparatus (left/right) were located at the extremes (5 cm from the wall limits). The lever edges and nose poke port centers were positioned 2 cm

from the floor. Lever extensions and retractions and nose poke port opening and closing (guillotine door) were remotely controlled. Custom-made lick ports were located at the middle of the wall, close to each other (4 cm between). The two lick ports were separated by a small barrier (to force mice to physically move to lick from one to another lick port). The lick ports were positioned 2 cm from the floor (fig. S1A). The lick port was composed of an empty cylinder (made of polyoxymethylene) positioned horizontally with an open window on the top where mice could access liquids (open window: ellipse of 6 × 3 mm). Liquids were delivered in a receptacle inserted in the cylinder (receptacle: half-ellipsoid of 6 × 3 × 2 mm). The receptacle was made of aluminum to measure tongue contacts through the analog input board of a neural recording data acquisition processor system (OmniPlex, Plexon). Lick onsets were inferred offline by detecting potential rise times. Each lick port allowed delivery of two outcomes through remotely controlled syringe pumps (PHM-107, Med Associates). A video camera recorded from above at 40 frames/s (fps) for video-tracking purposes using Cineplex Software (Plexon). All time stamps of camera frames, miniscope frames, nose pokes or lever presses, and analog signals from lick ports were recorded with the Omniplex neural recording data acquisition processor system at 40 kHz. Behavior, optogenetic, single-unit recordings, and miniscope were synchronized and controlled by a multi–input/output (I/O) processor (RZ6, Tucker Davis).

Instrumental goal-directed behavioral procedures

Food restriction

Mice were food restricted to 85% of their free-feeding body weight 4 days before and throughout the behavioral experiments and fed ~2 hours after their daily behavioral sessions with ~2.5 g of regular food.

Instrumental training

Instrumental training for the experimental set (N = 16 mice; N = 8 of 16 mice were equipped with a miniature microscope; Figs. 1 to 3 and figs. S1 to S9) was preceded with a session in which either sucrose (20%) or sweetened condensed milk (15%, Régilait) solutions were accessible, each at a fixed lick port (right and left lick port, respectively) upon licking (maximum duration 20 min or 20 of each outcome). The following day, instrumental training started with constant reinforcement (CR), in which outcomes were delivered after each action performed by the mouse. To speed up learning, only during this day was food dispensed in nose-poke ports or onto levers. Mice then started instrumental learning with 2 days of CR (without food in or onto actions, called day 1 and day 2). CR sessions were structured as follows: (i) 5 min without action or outcome availability (licking behavior was possible but not rewarded, OFF task); (ii) one of the two actions (left/right apparatus) was available for 30 min or until mice obtained 20 outcomes (milk/sucrose, ON task); (iii) both actions were not available for 2 min (OFF task); (iv) the second action (right/left apparatus) was available for 30 min or until mice obtained 20 outcomes (sucrose/milk, ON task); and (v) 2 min without action or outcome availability (OFF task). After CR sessions, mice went through variable action-outcome ratio training [variable response (VR)], first with average ratio 3 on day 3 (VR3, between 1 and 5 after normal distribution, $\mu = 3$, $\sigma = 1.5$) and average ratio 5 on days 4 and 5 (VR5, between 1 and 11 after normal distribution, $\mu = 5$, $\sigma = 2.5$). VR sessions were structured as follows: (i) 5 min without action or outcome availability (OFF task); (ii) one of the two actions (left/right apparatus) was available for 30 min or until mice obtained 15 outcomes (milk/sucrose, ON task); (iii) both actions were not available for 2 min (OFF task); (iv) the second action (right/left apparatus) was available for 30 min or until mice obtained 15 outcomes (sucrose/milk, ON task); (v) both actions were not available for 1 min (OFF task); (vi) both actions (left/right apparatus) were available for 30 min or until mice obtained 5 of each of the outcomes (milk/sucrose, ON task); and (vii) 2 min without action or outcome availability (OFF task). For CR and VR sessions, outcome order was counterbalanced.

Shift in hunger state test

To demonstrate the hunger-state dependency of instrumental actions, we exposed mice to a free choice rewarded test (N = 8 of 16 mice in the instrumental training group). A shift in hunger state was accomplished by interrupting food restriction from day 6 to day 8. The test was conducted on day 8 (Fig. 1B). The free choice reinforced test was structured as follows: (i) 5 min without action or outcome availability (OFF task); (ii) the two actions (left/right apparatus) were available for 30 min or until mice obtained 20 of each of the outcomes (milk/sucrose, ON task); and (iii) 2 min without action or outcome availability (OFF task). Outcomes were delivered on a VR5 basis.

Free choice extinction test

After instrumental training, we exposed mice to a free choice extinction (non-reinforced) test on day 6 (N = 16 mice in Fig. 1C; N = 8 mice in fig. S7; N = 20 mice in fig. S12), structured as follows: (i) 5 min without action or outcome availability (OFF task); (ii) the two actions (left/right apparatus) were available for 8 or 12 min (ON task; 8 min in fig. S12; 12 min in Fig. 1C and fig. S7); and (iii) 2 min without action or outcome availability (OFF task).

Free choice reinforced test

Two hours after the completion of the free choice extinction test, we exposed mice to a free choice reinforced test (N = 8 mice in fig. S7), structured as follows: (i) 5 min without action or outcome availability (OFF task); (ii) the two actions (left/right apparatus) were available for 30 min or until mice obtained 20 of each of the outcomes (milk/sucrose, ON task); and (iii) 2 min without action or outcome availability (OFF task). Outcomes were delivered on a VR5 basis.

Satiety-induced outcome devaluation

Instrumental actions are characterized as goal-directed if the actions are sensitive to variations in outcome value. After training, outcome devaluation was accomplished by pre-feeding mice with one of the two outcomes (the devaluated one) for 1 hour in home cage (N = 13 mice in Fig. 1C; N = 6 mice in fig. S13; N = 20 mice in fig. S12). The identity of the devaluated outcome was counterbalanced between mice. The experimenter was blinded to the devaluated outcome. Animals were randomly allocated to experimental groups and were later identified by unique markers for group assignment. Mice were then exposed on day 6 or day 8 to a free choice non-reinforced test (day 6 in fig. S12; day 8 in fig. S12), structured as follows: (i) 5 min without action or outcome availability (OFF task); (ii) the two actions (left/right apparatus) were available for 8 or 12 min (ON task; 8 min in fig. S12; 12 min in Fig. 1C and fig. S13); and (iii) 2 min without action or outcome availability (OFF task).

Action-outcome contingency degradation

Instrumental actions are characterized as goal directed if the actions are sensitive to variations in the action-outcome contingency. After instrumental training, action-outcome contingency degradation was accomplished by unpairing one of the two actions from its respective outcome (degraded action; N = 10 mice in Fig. 1C; N = 5 mice in fig. S13; N = 20 mice in fig. S12). The identity of the degraded action-outcome contingency was counterbalanced between mice. This training phase was conducted on day 6 and day 7, with the same structure as day 5 VR5 training (see above). Mice then were exposed on day 8 to a free choice non-reinforced test, structured as follows: (i) 5 min without action or outcome availability (OFF task); (ii) the two actions (left/right apparatus) were available for 8 or 12 min (ON task; 8 min in fig. S12; 12 min in Fig. 1C and fig. S13); and (iii) 2 min without action or outcome availability (OFF task).

Optogenetic sessions

Mice were trained to obtained two outcomes as described earlier. For optogenetic manipulations performed on day 5 (N = 40 mice;

Fig. 4 and figs. S10 and S11), sessions were structured as follows: (i) 5 min without action or outcome availability (OFF task); (ii) one of the two actions (left/right apparatus) was available for 30 min or until mice obtained 16 outcomes (milk/sucrose, ON task); (iii) 2 min without action or outcome availability (OFF task); (iv) the second action was available for 30 min or until mice obtained 16 outcomes (ON task); and (v) 2 min without action or outcome availability (OFF task). For optogenetic manipulations performed during non-reinforced tests (N = 40 mice; fig. S12), sessions were structured as follows: (i) 5 min without action or outcome availability (OFF task); (ii) the two actions (left/right apparatus) were available for 8 min (ON task); and (iii) 2 min without action or outcome availability (OFF task). Laser delivery strategies are described below.

Outcome value violation

Outcome value violation session differs from the specific satiety-induced outcome devaluation classically used to characterize goal-directed behavior. Nevertheless, the decrease in outcome value was achieved during this session by diluting outcomes, as evident by the reduction in instrumental behavior, and by the clear preference for nondiluted outcomes mice showed during the two-sucrose choice challenge (see below). On the basis of these observations, we call this session outcome value violation. After mice were trained to obtain two outcomes as described earlier (N = 8 mice; N = 4 of 8 mice were equipped with a miniature microscope), they underwent an outcome value violation session on day 6 (Fig. 5 and figs. S14 and S15), structured as follows: (i) 5 min without action or outcome availability (OFF task); (ii) one of the two actions was available for 30 min or until mice obtained 16 outcomes (ON task); (iii) 2 min without action or outcome availability (OFF task); (iv) the second action was available for 30 min or until mice obtained 16 outcomes (ON task); and (v) 2 min without action or outcome availability (OFF task). During ON task phases, mice could obtain up to eight outcomes (called the initial period), and then outcomes 9 to 16 were diluted by a factor of 4 (called the perturbed period). During initial and perturbed periods, outcomes were delivered on a VR5 schedule.

Action-outcome contingency violation

After mice were trained to obtain two outcomes as described earlier (N = 8 mice; N = 5 of 8 mice were equipped with a miniature microscope), they underwent a retraining session on day 7 (in the same conditions as on day 5) followed by an action-outcome contingency violation session on day 8 (N = 3 mice; one mouse was excluded from analysis because it bit the miniature microscope cable at the beginning of the session). An extra group of N = 2 mice that were submitted to an action-outcome contingency violation session on day 6 was then added to the analysis (Fig. 5 and figs. S14 and S15). Action-outcome contingency violation session was structured as follows: (i) 5 min without action or outcome availability (OFF task); (ii) one of the two actions was available for 30 min or until mice obtained 20 outcomes (ON task); (iii) 2 min without action or outcome availability (OFF task); (iv) the second action was available for 30 min or until mice obtained 20 outcomes (ON task); and (v) 2 min without action or outcome availability (OFF task). During ON task phases, mice were allowed to obtain eight outcomes delivered on a VR5 schedule (called the initial period) and then outcomes 9 to 20 were delivered noncontingently to action (called the perturbed period).

Instrumental goal-directed behavior analysis

Behavioral characterization

The instrumental training session was divided into OFF task phases (no action or outcome availability) and ON task phases (one or both actions available), as described in detail above. We collected action and lick time stamps together with mouse position tracking using Bonsai software (*44*). We then used these reading to automatically derive a continuous description of the mouse behavior during the entire session. We used nine descriptors differentiating epochs of task-related behaviors: action, unrewarded lick, and rewarded lick epochs from epochs of non-task-related behaviors: idle times and context exploration epochs, in an outcome-specific manner. Action epochs were defined as the time bin of a single action or series of actions (consecutive actions not interrupted by other behavioral epochs). Rewarded lick epochs were defined as the first lick bouts (series of two or more consecutive licks, with inter-lick interval <1 s) that occurred after an outcome delivery. Unrewarded lick epochs were defined as all lick bouts not defined as rewarded lick epochs (see above). Idle times epochs were defined as time windows when mice when mice were at the task zone (mouse center of mass located at ~8 cm (1/3 of the context) from the wall including instrumental actions and lick ports) but did not perform any task-related behaviors (actions/licks) for >4 s. Exploration epochs were defined as time windows when mice were away from the task zone. Transition epochs were defined as all time bins not allocated to any of the different behavioral epochs defined above. We further defined behavioral sequences consisting of action and outcome consumption periods. The action period was defined as the time from the first action until the last action between consumption epochs and included epochs of instrumental actions or unrewarded lick and transition epochs. Consumption periods are identical to the rewarded lick epochs.

Behavioral quantification

We used different parameters to quantify goal-directed behavior. Action rate (actions per minute; total number of actions divided by the total ON task phase duration) was used as the quantifier of the general instrumental performance. Action to lick latency (in seconds; time between the last action of an action period and the first rewarded lick that followed) was used to quantify the latency to collect outcome. Interbehavioral sequence interval (in seconds; time between the last lick of a consumption period and the subsequent action) was used to quantify the latency to initiate a new behavioral sequence after a consumption period. We used these three parameters to estimate mice motivation to obtain outcomes. We used the continuous description of the mouse behavior to quantify the duration of behavioral sequences, action periods, consumption periods, and the different behavioral epochs.

Ca^{2+} imaging using the miniature microscope

Miniature microscope imaging

Four weeks after surgery, mice were head-fixed to check GCaMP expression using a miniature microscope (nVista HD, Inscopix Inc.). If the expression level was sufficient, mice were anesthetized with isoflurane (3 to 5% for induction, 1 to 2% for maintenance; Attane, Provet) in oxygen-enriched air (Oxymat 3, Weinmann) to fix the miniature microscope baseplate (BLP-2, Inscopix) on top of the cranium implant using blue-light-curable glue (Vertise Flow, Kerr). The miniature microscope was then detached, the base plate was capped with a base plate cover (Inscopix), and the mouse was returned to its home cage. During the 3 days preceding instrumental training, mice were habituated to head fixation (~5 min), followed by a free exploration session in home cage (~10 min). This daily procedure allowed mice to habituate to miniature microscope mounting and carrying. It allowed us to select mice with a high number of neurons and better signal quality (>50 neurons, N = 19 of 24 mice included in the analysis) and to set the miniature microscope light-emitting diode (LED) power and electronic focus (for the nVista3 system). The miniature microscope was mounted immediately before each behavioral session by head fixation. Imaging data were acquired using nVista HD software (Inscopix Inc.) at a frame rate of 20 Hz (exposure time, 50 ms) with an LED power of 0.6 to 0.8 mW/mm^2 (excitation irradiance at objective front surface), an analog gain of 1 to 4, and a field of view of 1080 × 1080 pixels (~648 × 648 μm

with the nVista2 system) and 1280 × 800 pixels (~768 × 480 μm, nVista3 system). Imaging frame time stamps were recorded with the Omniplex neural recording data acquisition processor system.

Motion correction

Imaging files produced by the nVista2 or nVista3 systems were imported into Matlab 2017b using custom code. Two regions of interest (ROIs) manually selected by the investigators were used to correct for translational motion. To address background noise during motion correction, we subtracted a Gaussian blurred image on a frame-by-frame basis from the raw imaging data. We used Fourier fast transform–based Image registration (*45*) to register the first 100 frames on one ROI, and then used the median image of these 100 frames as a template to register the rest of the frames in the imaging session until correction was lower than a user-defined threshold (<0.01 average pixel shift on average across all frames). The procedure was then repeated on the resulting motion-corrected movie using the second user-defined ROI to ensure minimal nonrigid motion. We applied calculated shifts to the raw movie and used the raw, motion-corrected movies for the extraction of Ca^{2+} traces and all subsequent analysis.

Ca^{2+} trace extraction

We processed each session independently and used CNMF-E, a recently developed algorithm based on nonnegative matrix factorization (*46*), for automatic extraction of Ca^{2+}traces. We then excluded from the analysis ROIs based on anatomy (ROI size, shape, or vicinity to the edge of lens), low signal-to-noise ratio or large overlap in signal and spatial location with other neurons (>60% spatial overlap and >0.6 Pearson's correlation between the traces across the entire session). Linear trends across an entire session were removed from the Ca^{2+} traces and further calculations were performed on the *z*-scores of the detrended traces.

Registration of neuron identities across imaging session

After the individual sessions were extracted and sorted, taking day 5 as a reference, we performed automatic neuron alignment between sessions using centroid and shape-matching algorithms (*47*). To ensure that the correct alignment visual inspection was performed to assess whether the imaged neuron was the same across days, the ROI shape and location in the field of view had to be consistent across sessions.

Histology

After the completion of behavioral experiments, mice were deeply anesthetized with urethane (2 g/kg of body weight intraperitoneally) and transcardially perfused with PBS followed by 4% paraformaldehyde in PBS. Brains were removed and postfixed overnight at 4°C. Then, 80-μm coronal brain slices containing the BLA were cut with a vibratome (VT1000 S, Leica) and stored in PBS. Slices were washed for 10 min in PBS, given a 5-min exposure to 4′,6-diamidin-2-phenylindol (DAPI, 1:10,000, Sigma-Aldrich), and then washed 3× for 15 min each in PBS. Slices were mounted on glass slides, coverslipped, and imaged using an Axioscan Z1 slide scanner (Carl Zeiss AG), equipped with a 10× air objective (Plan-Apochromat 10×/0.45). Mice were excluded post hoc if the GRIN lens was not placed in BLA ($N = 3$ mice excluded).

Imaging plan location and anatomical reconstruction

We assessed the spatial location of each neuron within the imaging field and along AP or ML BLA axes (fig. S3). The position of the center of each GRIN lens was matched against a mouse brain atlas (*48*), and the resulting coordinates were used to infer the location of the center of mass of each ROI (or neuron) along the AP or ML BLA axes and to determine the relative concentration of the distinct functional neuronal populations, as previously described (*49*).

Freely moving electrophysiology

Single-unit recordings

Mice were habituated to the head-stage (Plexon) connection procedure by handling and short head restraining for at least 3 days before experiments. The head stages were connected to a 32-channel preamplifier (gain ×100, bandpass filter 150 Hz to 9 kHz for unit activity and 0.7 Hz to 170 Hz for field potentials, Plexon). Spiking activity was digitized at 40 kHz, bandpass filtered from 250 Hz to 8 kHz, and isolated by time-amplitude window discrimination and template matching using the Omniplex neural recording data acquisition processor system.

Single-unit analysis

Single-unit spike sorting was performed using Offline Spike Sorter software (Plexon). Principal component scores were calculated for unsorted waveforms and plotted in a three-dimensional principal component space; clusters containing similar valid waveforms were manually defined. A group of waveforms was considered to be generated from a single neuron if the waveforms formed a discrete, isolated cluster in the principal component space and did not contain a refractory period of <1 ms, as assessed using autocorrelogram analyses. To avoid analysis of the same neuron recorded on different channels, we computed cross-correlation histograms. If a target neuron presented a peak of activity at a time that the reference neuron fired, only one of the two neurons was considered for further analysis. Putative inhibitory interneurons were separated from putative excitatory projection neurons using an unsupervised cluster algorithm based on Ward's method. In brief, the Euclidian distance was calculated between all neuron pairs based on the three-dimensional space defined by each neuron's average half-spike width (measured from trough to peak), the firing rate, and the area under the hyperpolarization phase of the spike (*50*). Further analyses were performed as for Ca^{2+} traces using *z*-score–transformed spike trains binned in 50-ms bins.

Histology

Before transcardial perfusion with 4% paraformaldehyde in PBS, an electrolytic lesion was made at the electrode tip by applying 0.1 μA for 10 s to one of the electrode wires in deeply anesthetized mice to mark the recording site. Brains were postfixed in 4% paraformaldehyde for at least 2 h at 4°C and cut in 80-μm coronal slices using a vibratome (VT1000S). Sections containing the BLA were washed three times for 10 min in PBS, exposed to DAPI (1:10,000, Sigma-Aldrich) for 10 min, and washed again three times for 10 min in PBS before mounting on glass slides. To verify the electrode placement, sections were scanned with an Axioscan Z1 slide scanner (Carl Zeiss AG), equipped with a 10× air objective (Plan-Apochromat 10×/0.45). Electrode placements were matched against a mouse brain atlas (*48*).

Optogenetic experiments

The experimenter was blinded to each animal's experimental cohort (virus condition, GFP, or ArchT). Animals were randomly allocated to experimental groups and were later identified by unique markers for group assignment. Before behavioral experiments, all mice were habituated to the optical fiber connection procedure by handling and brief head restraining for at least 3 days. On the experimental days, implanted fibers (fiber numerical aperture of 0.48, fiber inner core diameter of 200 μm; Thorlabs) were connected to a custom-built laser bench (Life Imaging Services) with custom fiber patch cables. An acousto-optic modulator (AA Opto-Electronic) controlled the laser intensity (MGL-F-589, 589-nm wavelength, CNI Lasers). Laser power at the fiber tip was measured before every subject with an optical power and energy meter (PM100D, ThorLabs) and adjusted to reach an irradiance value of ~4 mW at fiber tip.

Closed-loop optogenetic sessions

To inhibit BLA PNs during action periods, the laser was switched on from the first action until the last action of each action period. To inhibit BLA PNs during consumption periods,

the laser was switched on from the first lick after outcome delivery until the last rewarded lick of each consumption period. Licks were detected online, and rewarded licks were defined as the first lick bout (series of two or more licks, with inter-lick intervals <1 s) after outcome delivery. Maximal laser duration was fixed at 20 s. Laser was controlled by a multi-I/O processor (RZ6, Tucker Davis). The laser was switched on during individual action or consumption periods from the ninth to the 16th behavioral sequences, called the laser-ON period. The first to the eighth behavioral sequences were called the laser-OFF period (Fig. 4 and figs. S10 and S11). For extinction tests, to inhibit BLA PNs during action epochs the laser was switched on from the first action until the last action of each action epoch. Maximal laser duration was fixed at 4 s. Laser was delivered for each action epochs during the first 4 min of the test (laser-on period). No laser was delivered during the last 4 min of the test (laser-off period; fig. S12).

Two-sucrose choice challenge

Two-sucrose choice challenge was performed in a behavioral context measured 26 cm L × 25 cm W × 40 cm H. The context contained two lick ports located on the same wall and separated by 10 cm. For 12 min, food restriction mice freely chose between 5 and 20% sucrose solutions delivered from different lick ports. Licks triggered outcome delivery at slow rate (2 μl/s). Mice showed sustained licking bouts during the 12 min of the test (fig. S11M). The behavioral session was divided in two phases: In the first 6 min, the laser was off and in the last 6 min, each lick bouts of the 20% sucrose solution was paired with light. Laser was controlled by a multi-I/O processor (RZ6, Tucker Davis).

Real-time place avoidance paradigm

The real-time place avoidance experiment was performed in a behavioral context composed of two compartments (20 × 20 cm each) connected by an alleyway (5 × 5 cm). The two compartments differed in shape (circular and square) and visually (gray Plexiglas with black horizontal stripes or green Plexiglas). Mouse position was tracked outline using Cineplex software (Plexon). On day 1, mice were allowed to freely explore the entire behavioral context during a 12-min pre-exposure period. After pre-exposure, the compartment in which the mice spent the most time was designated as the most visited compartment. On day 2, mice were submitted to a 12-min test session during which light pulses (1-s pulse width, repeated at 1 Hz) were delivered while the mice occupied the most visited compartment but not when they occupied the less-visited compartment (fig. S11L). The laser was controlled by a multi-I/O processor (RZ6, Tucker Davis).

Histology

Mice were transcardially perfused (as described above) and optical fibers removed. Brains were postfixed in 4% paraformaldehyde for at least 2 h at 4°C and cut into 80-μm coronal slices using a vibratome (VT1000S). Sections containing the BLA were immediately mounted on glass slides and coverslipped. To verify the specificity of viral expression and fiber tip placement, sections were scanned with an Axioscan Z1 slide scanner (Carl Zeiss AG), equipped with a 10× air objective (Plan-Apochromat 10×/0.45). Fiber tip placements were matched against a mouse brain atlas (*48*). Mice were excluded from the analysis post hoc if they did not show bilateral expression of the virus, if virus expression (cell bodies expressing GFP) was detected outside of the BLA, or if they did not exhibit correct fiber placement (<300 μm away from the BLA; figs. S10 to S12). Twenty-four of 104 mice were excluded.

Data analysis

We used custom routines written in MATLAB (The MathWorks) to perform all data analyses.

Identification of task-modulated neurons

We identified, independently for each outcome, three functional types of neurons showing significant and consistent increase activity at action and/or consumption times: action neurons, consumption neurons, and transition neurons. We first collected the Ca^{2+} traces in a time window around action/consumption time stamps (see below). Neurons were classified as task modulated if the maximum of their average response in a more restricted reference time window fulfilled three criteria: (i) it exceeded a threshold on the maxima obtained for each neuron by a bootstrapping procedure using the same time window around random time stamps along the session (number of random time stamps matched the number of real time stamps, 1000 iterations, threshold set at the 99 percentile); (ii) it exceeded the average response outside of the restricted reference window; (iii) it exceeded the upper limit of the 95% confidence interval around zero (a *z*-score value of 0), computed using Student's *t* distribution with the empirical mean and standard error of the maxima of individual traces. To characterize BLA PNs as action neurons, we used a time window of 2 s before and 4 s after the onset of one of the following actions: (i) the first action in an action period (after a rewarded lick epoch), (ii) the last action before an unrewarded lick, or (iii) the last action in an action period. This allowed us to capture neurons locked to the distinct actions in action periods. In terms of behavior, these windows isolate the stereotypical behavioral switches between actions and licks (Figs. 1F and 2A and fig. S3A). The restricted reference time window used for action neurons started 0.5 s before action onsets and lasted the duration of the shortest action epoch (determined per mouse; no shorter than 0.8 s). To characterize BLA PNs as consumption neurons, we used a time window of 2 s before and 8 s after the start of consumption periods (see the description of behavior in this time window in Fig. 2A and fig. S3B). The restricted reference time window used for consumption neurons started 0.1 s before consumption onset and lasted the duration of the shortest consumption epoch (determined per mouse). BLA PNs characterized as both action and consumption neurons were classified as transition neurons. During outcome value and action-outcome contingency violation sessions, task-modulated neurons were identified by taking into account time stamps of behavioral events occurring either during the initial or the perturbed periods.

Coincident activity between task-modulated neurons

For Ca^{2+} data, we used correlation and coactivation analysis between pairs of action and/or consumption neurons in high-activity frames (*z*-scored traces values <2 were filtered to 0) to assess the intrafunctional subset and interfunctional subset coincident activity independently for OFF and ON task phases. Correlation between neuron pairs was evaluated by using Pearson's correlation coefficient (r). We determined the proportion of noncorrelated, positively correlated, and negatively correlated neuron pairs ($P < 0.01$) for action-action, consumption-consumption, and action-consumption pairs. Then, we determined whether the proportions of correlated neurons were different from those obtained for random neuronal pairs (proportion bigger than the 99% percentile of proportions obtained with random pairs). The coactivation index for each pair of action and/or consumption neurons was calculated as the ratio of co-occurring high-activity frames to the total number of high-activity frames (*51*). For single-unit data, correlation and co-firing analyses between pairs of action and/or consumption neurons were performed after binning the spike train in 50-ms bins.

Decoder

We trained a linear discriminant analysis (LDA) classifier with fivefold cross-validation on the distinction between five behavioral epochs: action periods (left/right), consumption periods (milk/sucrose), and periods of non-task-related behavior during the OFF task phase (unrewarded licks epochs during the OFF task phase were excluded). Neuronal responses and behavior were binned in 200-ms bins, and bins with other behavioral descriptions were excluded from the analysis.

To address the unbalanced design (more non-task-related epochs) we used the F score calculated from the confusion matrices as a measure of the decoder performance. We repeated the analysis with the behavioral labels shuffled in relation to the real neuronal data (control, 100 repetitions) using only the responses of the task-modulated neurons (modulated) or random selections of non-task-modulated neurons with a sample size that matched the number of task-modulated neurons for each animal (non-modulated, 100 repetitions).

Pairwise correlation between population activity patterns

We binned the single neuron activity in 200-ms bins to obtain a single population activity vector per time point and quantified the similarity between two activity vectors by the pairwise Pearson's correlation coefficient (r). For the mean correlation analysis presented in Fig. 3C, we calculated intrabehavioral epoch mean correlations by first collecting the activity vectors in all the epochs of a specific behavior (two action epochs and two rewarded lick epochs) and then calculating the average pairwise Pearson's correlation between them (Fig. 3C). The mean on-on intra-epoch correlation was then calculated by averaging the four specific behavior intra-epoch correlations per mouse. Inter-epoch mean correlations were calculated per mouse by averaging the pairwise correlation between activity vectors during OFF task and ON task (off-on) or between on activity vectors during different ON task behaviors (on-on).

To track the consistency of the activity patterns along the session, we used as reference patterns the mean activity in two selections of time points: 24 OFF task time bins (every 10 s in the 4 min before the first ON task phase) and all time bins in three action/consumption epochs (behavioral sequences 3 to 5). We then calculated the pairwise correlations between the activity vectors during each action-consumption epoch and these two reference patterns (Fig. 3D).

To estimate the persistence of action- and consumption-associated patterns along the behavioral sequence and their relationship to behavior (Fig. 3, F and G), we selected as a reference for each sequence a representative action and consumption activity vector (time of the first action in an action epoch to follow action pattern consistency along the action period or 2 s after the first lick in a consumption period to differentiate from unrewarded lick epochs). We then marked all bins in the sequence with activity vectors that had a positive correlation with the reference (the lower bound of the 99% confidence interval of the Pearson's correlation >0). The same procedure was performed with single-unit recordings and inferred spikes after binning the spike trains in 50-ms bins (figs. S8 and S9).

Transitions in action- and consumption-associated patterns were detected by k-means clustering (n clusters = 2) of the mean pattern correlation matrices obtained by first averaging the pairwise correlations between the mean activity vectors during epochs in different sequences and then averaging the resulting correlation matrices (N = 4/5 mice, for outcome value/action-outcome contingency violation sessions, respectively, × 2 outcomes).

Dimensionality reduction

We used tSNE (tsne.m function from the MATLAB's Statistics Toolbox with perplexity = 60 and learning rate = 600) with correlation as the distance measure to visually demonstrate the distinct clusters of neuronal activity patterns and their relation to behavior.

Statistics

Statistical analysis was performed with MATLAB (The MathWorks). Normality of the data was assessed using Shapiro-Wilk test and either parametric (paired t) or nonparametric (Wilcoxon signed-rank or rank-sum) tests were used. In the figures, box-and-whisker plots indicate median (vertical line), interquartile (horizontal thick line), extreme data values (horizontal thin line), and outliers (plus sign) of the data distribution. No statistical methods were used to predetermine sample sizes, but our sample sizes are similar to those generally used in the field. Statistical tests are mentioned in the figure legends (ns indicates $P > 0.05$; $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$).

REFERENCES AND NOTES

1. A. Dickinson, B. Balleine, Motivational control of goal-directed action. *Anim. Learn. Behav.* **22**, 1–18 (1994). doi: 10.3758/BF03199951
2. B. W. Balleine, The meaning of behavior: Discriminating reflex and volition in the brain. *Neuron* **104**, 47–62 (2019). doi: 10.1016/j.neuron.2019.09.024; pmid: 31600515
3. P. H. Janak, K. M. Tye, From circuits to behaviour in the amygdala. *Nature* **517**, 284–292 (2015). doi: 10.1038/nature14188; pmid: 25592533
4. K. M. Tye, P. H. Janak, Amygdala neurons differentially encode motivation and reinforcement. *J. Neurosci.* **27**, 3937–3945 (2007). doi: 10.1523/JNEUROSCI.5281-06.2007; pmid: 17428967
5. G. D. Stuber *et al*., Excitatory transmission from the amygdala to nucleus accumbens facilitates reward seeking. *Nature* **475**, 377–380 (2011). doi: 10.1038/nature10194; pmid: 21716290
6. J. J. Paton, M. A. Belova, S. E. Morrison, C. D. Salzman, The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* **439**, 865–870 (2006). doi: 10.1038/nature04490; pmid: 16482160
7. G. Schoenbaum, A. A. Chiba, M. Gallagher, Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nat. Neurosci.* **1**, 155–159 (1998). doi: 10.1038/407; pmid: 10195132
8. F. Ambroggi, A. Ishikawa, H. L. Fields, S. M. Nicola, Basolateral amygdala neurons facilitate reward-seeking behavior by exciting nucleus accumbens neurons. *Neuron* **59**, 648–661 (2008). doi: 10.1016/j.neuron.2008.07.004; pmid: 18760700
9. A. Servonnet, G. Hernandez, C. El Hage, P. P. Rompré, A. N. Samaha, Optogenetic activation of the basolateral amygdala promotes both appetitive conditioning and the instrumental pursuit of reward cues. *J. Neurosci.* **40**, 1732–1743 (2020). doi: 10.1523/JNEUROSCI.2196-19.2020; pmid: 31953370
10. S. B. Ostlund, B. W. Balleine, Differential involvement of the basolateral amygdala and mediodorsal thalamus in instrumental action selection. *J. Neurosci.* **28**, 4398–4405 (2008). doi: 10.1523/JNEUROSCI.5472-07.2008; pmid: 18434518
11. D. A. Simmons, D. B. Neill, Functional interaction between the basolateral amygdala and the nucleus accumbens underlies incentive motivation for food reward on a fixed ratio schedule. *Neuroscience* **159**, 1264–1273 (2009). doi: 10.1016/j.neuroscience.2009.01.026; pmid: 19344638
12. R. Chesworth, L. Corbit, "The contribution of the amygdala to reward-related learning and extinction," in *The Amygdala: Where Emotions Shape Perception, Learning and Memories*, B. Ferry, Ed. (IntechOpen, 2017); https://www.intechopen.com/chapters/54509.
13. B. W. Balleine, A. S. Killcross, A. Dickinson, The effect of lesions of the basolateral amygdala on instrumental conditioning. *J. Neurosci.* **23**, 666–675 (2003). doi: 10.1523/JNEUROSCI.23-02-00666.2003; pmid: 12533626
14. S. E. Rhodes, E. A. Murray, Differential effects of amygdala, orbital prefrontal cortex, and prelimbic cortex lesions on goal-directed behavior in rhesus macaques. *J. Neurosci.* **33**, 3380–3389 (2013). doi: 10.1523/JNEUROSCI.4374-12.2013; pmid: 23426666
15. A. W. Johnson, M. Gallagher, P. C. Holland, The basolateral amygdala is critical to the expression of pavlovian and instrumental outcome-specific reinforcer devaluation effects. *J. Neurosci.* **29**, 696–704 (2009). doi: 10.1523/JNEUROSCI.3758-08.2009; pmid: 19158296
16. C. B. Ferster, B. F. Skinner, *Schedules of Reinforcement* (Appleton-Century-Crofts, 1957).
17. M. R. Roesch, D. J. Calu, G. R. Esber, G. Schoenbaum, Neural correlates of variations in event processing during learning in basolateral amygdala. *J. Neurosci.* **30**, 2464–2471 (2010). doi: 10.1523/JNEUROSCI.5781-09.2010; pmid: 20164330
18. I. Hernádi, F. Grabenhorst, W. Schultz, Planning activity for internally generated reward goals in monkey amygdala neurons. *Nat. Neurosci.* **18**, 461–469 (2015). doi: 10.1038/nn.3925; pmid: 25622146
19. W. Schultz, Multiple reward signals in the brain. *Nat. Rev. Neurosci.* **1**, 199–207 (2000). doi: 10.1038/35044563; pmid: 11257908
20. Y. Niv, D. Joel, P. Dayan, A normative perspective on motivation. *Trends Cogn. Sci.* **10**, 375–381 (2006). doi: 10.1016/j.tics.2006.06.010; pmid: 16843041
21. H. Nishijo, T. Ono, H. Nishino, Single neuron responses in amygdala of alert monkey during complex sensory stimulation with affective significance. *J. Neurosci.* **8**, 3570–3583 (1988). doi: 10.1523/JNEUROSCI.08-10-03570.1988; pmid: 3193171
22. A. Fontanini, S. E. Grossman, J. A. Figueroa, D. B. Katz, Distinct subtypes of basolateral amygdala taste neurons reflect palatability and reward. *J. Neurosci.* **29**, 2486–2495 (2009). doi: 10.1523/JNEUROSCI.3898-08.2009; pmid: 19244523
23. M. A. Bermudez, W. Schultz, Reward magnitude coding in primate amygdala neurons. *J. Neurophysiol.* **104**, 3424–3432 (2010). doi: 10.1152/jn.00540.2010; pmid: 20861431
24. B. W. Balleine, S. Killcross, Parallel incentive processing: An integrated view of amygdala function. *Trends Neurosci.* **29**, 272–279 (2006). doi: 10.1016/j.tins.2006.03.002; pmid: 16545468
25. J. Gründemann *et al*., Amygdala ensembles encode behavioral states. *Science* **364**, eaav8736 (2019). doi: 10.1126/science.aav8736; pmid: 31000636
26. P. Namburi *et al*., A circuit mechanism for differentiating positive and negative associations. *Nature* **520**, 675–678 (2015). doi: 10.1038/nature14366; pmid: 25925480
27. J. Kim, M. Pignatelli, S. Xu, S. Itohara, S. Tonegawa, Antagonistic negative and positive neurons of the basolateral amygdala. *Nat. Neurosci.* **19**, 1636–1646 (2016). doi: 10.1038/nn.4414; pmid: 27749826
28. X. Zhang, B. Li, Population coding of valence in the basolateral amygdala. *Nat. Commun.* **9**, 5195 (2018). doi: 10.1038/s41467-018-07679-9; pmid: 30518754
29. P. Kyriazi, D. B. Headley, D. Pare, Multi-dimensional coding by basolateral amygdala neurons. *Neuron* **99**, 1315–1328.e5 (2018). doi: 10.1016/j.neuron.2018.07.036; pmid: 30146300
30. W. Zhang *et al*., Functional circuits and anatomical distribution of response properties in the primate amygdala. *J. Neurosci.* **33**, 722–733 (2013). doi: 10.1523/JNEUROSCI.2970-12.2013; pmid: 23303950
31. X. Zhang, J. Kim, S. Tonegawa, Amygdala reward neurons form and store fear extinction memory. *Neuron* **105**, 1077–1093.e7 (2020). doi: 10.1016/j.neuron.2019.12.025; pmid: 31952856

32. C. Herry *et al*., Switching on and off fear by distinct neuronal circuits. *Nature* **454**, 600–606 (2008). doi: 10.1038/nature07166; pmid: 18615015
33. S. Duvarci, D. Pare, Amygdala microcircuits controlling learned fear. *Neuron* **82**, 966–980 (2014). doi: 10.1016/j.neuron.2014.04.042; pmid: 24908482
34. S. Krabbe *et al*., Adaptive disinhibitory gating by VIP interneurons permits associative learning. *Nat. Neurosci.* **22**, 1834–1843 (2019). doi: 10.1038/s41593-019-0508-y; pmid: 31636447
35. A. A. Grace, S. B. Floresco, Y. Goto, D. J. Lodge, Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends Neurosci.* **30**, 220–227 (2007). doi: 10.1016/j.tins.2007.03.003; pmid: 17400299
36. J. A. da Silva, F. Tecuapetla, V. Paixão, R. M. Costa, Dopamine neuron activity before action initiation gates and invigorates future movements. *Nature* **554**, 244–248 (2018). doi: 10.1038/nature25457; pmid: 29420469
37. M. Malvaez, C. Shieh, M. D. Murphy, V. Y. Greenfield, K. M. Wassum, Distinct cortical-amygdala projections drive reward value encoding and retrieval. *Nat. Neurosci.* **22**, 762–769 (2019). doi: 10.1038/s41593-019-0374-7; pmid: 30962632
38. K. Lavi, G. A. Jacobson, K. Rosenblum, A. Lüthi, Encoding of conditioned taste aversion in cortico-amygdala circuits. *Cell Rep.* **24**, 278–283 (2018). doi: 10.1016/j.celrep.2018.06.053; pmid: 29996089
39. S. L. Parkes, B. W. Balleine, Incentive memory: Evidence the basolateral amygdala encodes and the insular cortex retrieves outcome values to guide choice between goal-directed actions. *J. Neurosci.* **33**, 8753–8763 (2013). doi: 10.1523/JNEUROSCI.5071-12.2013; pmid: 23678118
40. K. Yoshida, M. R. Drew, M. Mimura, K. F. Tanaka, Serotonin-mediated inhibition of ventral hippocampus is required for sustained goal-directed behavior. *Nat. Neurosci.* **22**, 770–777 (2019). doi: 10.1038/s41593-019-0376-5; pmid: 30988523
41. J. Kim, X. Zhang, S. Muralidhar, S. A. LeBlanc, S. Tonegawa, Basolateral to central amygdala neural circuits for appetitive behaviors. *Neuron* **93**, 1464–1479.e5 (2017). doi: 10.1016/j.neuron.2017.02.034; pmid: 28334609
42. G. Hart, B. K. Leung, B. W. Balleine, Dorsal and ventral streams: The distinct role of striatal subregions in the acquisition and performance of goal-directed actions. *Neurobiol. Learn. Mem.* **108**, 104–118 (2014). doi: 10.1016/j.nlm.2013.11.003; pmid: 24231424
43. L. H. Corbit, B. K. Leung, B. W. Balleine, The role of the amygdala-striatal pathway in the acquisition and performance of goal-directed instrumental actions. *J. Neurosci.* **33**, 17682–17690 (2013). doi: 10.1523/JNEUROSCI.3271-13.2013; pmid: 24198361
44. G. Lopes *et al*., Bonsai: An event-based framework for processing and controlling data streams. *Front. Neuroinform.* **9**, 7 (2015). doi: 10.3389/fninf.2015.00007; pmid: 25904861
45. M. Guizar-Sicairos, S. T. Thurman, J. R. Fienup, Efficient subpixel image registration algorithms. *Opt. Lett.* **33**, 156–158 (2008). doi: 10.1364/OL.33.000156; pmid: 18197224
46. P. Zhou *et al*., Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data. *eLife* **7**, e28728 (2018). doi: 10.7554/eLife.28728; pmid: 29469809
47. G. Corder *et al*., An amygdalar neural ensemble that encodes the unpleasantness of pain. *Science* **363**, 276–281 (2019). doi: 10.1126/science.aap8586; pmid: 30655440
48. G. Paxinos, K. Franklin, K. Franklin, *The Mouse Brain in Stereotaxic Coordinates* (Academic, ed. 2, 2001).
49. B. Engelhard *et al*., Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature* **570**, 509–513 (2019). doi: 10.1038/s41586-019-1261-9; pmid: 31142844
50. J. Courtin *et al*., Prefrontal parvalbumin interneurons shape neuronal activity to drive fear expression. *Nature* **505**, 92–96 (2014). doi: 10.1038/nature12755; pmid: 24256726
51. N. Karalis *et al*., 4-Hz oscillations synchronize prefrontal-amygdala circuits during fear behavior. *Nat. Neurosci.* **19**, 605–612 (2016). doi: 10.1038/nn.4251; pmid: 26878674

ACKNOWLEDGMENTS

We thank all members of the A. Lüthi laboratory for comments and helpful discussions; P. Argast, P. Buchmann, and all staff of the FMI Animal Facility for outstanding technical assistance; the FMI IT department for support with data storage; and the Facility for Imaging and Microscopy at the FMI, in particular S. Bourke. **Funding:** This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant no. 669582 to A.L.) and by SNSF Ambizione grant PZ00P3_180057 to J.C. **Author contributions:** J.C. designed the experiments. J.C. performed the experiments with help from S.M. and C.M. J.H. and K.M.H. helped with deep-brain imaging data extraction. J.C. and Y.B. analyzed behavior and deep-brain imaging data. J.C., Y.B., and A.L. wrote the paper. All authors contributed to the interpretation of the data and commented on the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All processed data and scripts needed to evaluate the conclusions in the paper are available at: https://data.fmi.ch/PublicationSupplementRepo/ and https://github.com/fmi-basel/1Photon_Analysis.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abg7277
Figs. S1 to S18
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

24 January 2021; resubmitted 16 August 2021
Accepted 5 November 2021
10.1126/science.abg7277

A neuronal mechanism for motivational control of behavior

J. CourtinY. BittermanS. MüllerJ. HinzK. M. HagiharaC. MüllerA. Lüthi

Science, 375 (6576), eabg7277. • DOI: 10.1126/science.abg7277

The amygdala and goal-directed actions

Almost everything we do in our daily lives is goal directed. The brain can maintain a motivational state to direct actions to achieve desired outcomes. Using deep-brain calcium imaging, electrophysiology, and optogenetics in mice, Courtin *et al*. observed that at the time of goal-directed action, basolateral amygdala principal neurons integrate and encode pursued outcome identity, pursued outcome value, and action-outcome contingency information. At the time of consumption, basolateral amygdala neuronal firing represents current outcome identity and value. Together, action- and consumption-associated activity integrate behaviorally relevant information at distinct time points along goal-directed action-consumption sequences. —PRS