Multi-dimensional Coding by Basolateral Amygdala Neurons

Highlights

- The type and timing of CRs is a major determinant of BLA activity
- The CS responses of BLA cells are distinct from the activity related to CRs
- Few BLA cells encode valence but it can be decoded at the population level
- Most BLA neurons concurrently encode multiple task features

Authors
Pinelopi Kyriazi, Drew B. Headley, Denis Pare

Correspondence
dbh60@newark.rutgers.edu (D.B.H.), pare@andromeda.rutgers.edu (D.P.)

In Brief
Kyriazi et al. show that most BLA neurons concurrently encode multiple task features, including the sensory properties of CSs and the behaviors they elicit. This leads to multiple task dimensions being represented at the population level.
Multi-dimensional Coding by Basolateral Amygdala Neurons

Pinelopi Kyriazi,1,3 Drew B. Headley,2,3,* and Denis Pare2,4,*
1Behavioral and Neural Sciences Graduate Program, Rutgers State University, Newark, NJ 07102, USA
2Center for Molecular and Behavioral Neuroscience, Rutgers State University, Newark, NJ 07102, USA
3These authors contributed equally
4Lead Contact
*Correspondence: dbh60@newark.rutgers.edu (D.B.H.), pare@andromeda.rutgers.edu (D.P.)
https://doi.org/10.1016/j.neuron.2018.07.036

SUMMARY

Conditioned appetitive and aversive responses (CRs) are thought to result from the activation of specific subsets of valence-coding basolateral amygdala (BLA) neurons. Under this model, the responses of BLA cells to conditioned stimuli (CSs) and the activity that drives CRs are closely related. We tested the strength of this correlation using a task where rats could emit different CRs in response to the same CSs. At odds with this model, the CS responses and CR-related activity of individual BLA cells were separable. Moreover, while the incidence of valence-coding cells did not exceed chance, at the population level there was similarity between valence coding for CSs and CRs. In fact, both lateral and basolateral neurons concurrently encoded multiple task features and behaviors. Thus, conditioned emotional behaviors may not depend on the recruitment of single cells that explicitly encode individual task variables but from multiplexed representations distributed across the BLA.

INTRODUCTION

The basolateral complex of the amygdala (BLA) is necessary for the acquisition and expression of conditioned responses (CRs) to stimuli (CSs) that predict aversive or rewarding outcomes (Ambroggi et al., 2008; Fanselow and Poulos, 2005; LeDoux, 2000). In the lateral nucleus (LA), the main station of the amygdala for sensory inputs (LeDoux, 2000; Sah et al., 2003), conditioning enhances the efficacy of synapses that convey CS information (McKernan and Shinnick-Gallagher, 1997; Rumpel et al., 2005; Tye et al., 2008), resulting in an increased CS responsiveness of target neurons (Genud-Gabai et al., 2013; Ghosh and Chattarji, 2015; Maren and Quirk, 2004; Quirk et al., 1995). Within the LA and basolateral nuclei (BL), largely separate populations of cells acquire responses to appetitive or aversive stimuli (Belova et al., 2008; Burgos-Robles et al., 2017; Sangha et al., 2013; Lee et al., 2017; however see Shabel and Janak, 2009) and they project to different downstream structures, which in turn mediate distinct behavioral responses (Beyeler et al., 2016; Namburi et al., 2015). Together these findings suggest that the changes in CS responsiveness displayed by BLA cells following appetitive or aversive conditioning constitute potentiated sensory responses that can automatically trigger approach or defensive CRs through distinct neuronal outputs.

However, several observations are inconsistent with this rigid serial mechanism. For instance, BLA cells can be recruited into memory traces based on increased excitability irrespective of their underlying function (Han et al., 2007; Yiu et al., 2014). Furthermore, a recent Ca²⁺ imaging study reported that at the population level, the changes of CS representation induced by fear conditioning involve a qualitative shift toward the representation of the unconditioned stimulus (US; Grewe et al., 2017). Yet another study revealed that the activity of BL neurons is not necessarily related to the sensory properties of the CS but is closely tied to the conditioned reward-seeking behavior (Lee et al., 2016). Last, animals display different defensive behaviors (Pelham and Kim, 2016) depending on threat proximity (Perusini and Fanselow, 2015) and a strict serial mechanism cannot account for such flexibility.

As the previous studies highlight, current conditioning paradigms typically limit animal behavior to a single CR (e.g., freezing), which can give the impression that CSs automatically trigger CRs when in fact it is the only option the experimental conditions allow. Since conditioning increases the likelihood that CSs will elicit a particular CR, it is difficult to disentangle whether training-induced alterations in BLA activity are related to the valence or sensory properties of the CSs, the behavior they elicit, or a mixture thereof.

To overcome these limitations, we developed a novel paradigm, the Risk-Reward Interaction (RRI) task, where rats can respond to a light CS in different ways depending on where the CS is presented. Rats learn that in some positions the CS signals reward availability (CS-R) and in others an impending footshock (CS-S). The footshock can be avoided passively or actively, depending on the rats’ position in relation to the CS. Behavioral freezing is also observed, allowing us to compare neuronal activity related to different defensive behaviors. Using this task, we found that LA and BL neurons encode heterogeneous mixtures of sensory and behavioral variables.

RESULTS

In the RRI task (Figure 1), light CSs indicated different outcomes based on their location. When the CS appeared below one of
three shock sectors (CS-S; Figure 1A, red), it signaled an impending footshock in that sector. When a CS was presented behind the left or right wall (CS-R; Figure 1A, blue), it indicated that a water reward would be delivered at that location. The rats’ position and head direction were monitored throughout the task.

After training, rats (n = 8) displayed different defensive and appetitive behaviors during the RRI task. On CS-S trials, if rats were on the shock sector at trial onset, they generally avoided the shock by moving to one of the unlit sectors (Figure 1C1). In 19% of such active avoidance trials, animals showed behavioral freezing prior to avoiding the shock. Alternatively, if rats were already on one of the unlit sectors at the start of the CS-S, they then typically displayed passive avoidance; that is, they stayed away from the lit sector until the trial ended (Figure 1C2). During appetitive trials, rats ran to the reward port shortly after CS-R onset. This reward approach behavior (Figure 1C3) was then followed by reward anticipation, defined as rats placing their two front paws onto the reward port and waiting there until reward delivery.

During the training phase, rats acquired active avoidance faster than reward seeking (Figure 1B; two-way ANOVA, Behavior Type × Session, F(1,104) = 33.32, p = 8.15 × 10⁻⁶). Passive avoidance remained stable throughout training (Figure S1A). As detailed in Figures S1C and S1D, conditioned behaviors varied significantly in their duration (one-way ANOVA, F(4,5259) = 31.74, p = 3.62 × 10⁻⁶) and latency from CS onset (one-way ANOVA, F(4,5259) = 7238.91, p = 0). Once rats performed ≥ 80% of the trials correctly (Figure S1B), they were implanted with silicon probes aimed at the BLA (Figures S2A–S2C and S3).

Activity of Principal BLA Cells during CSs

We recorded 344 cells in LA and 305 cells in BL while rats performed the RRI task. Cells were classified as putative principal neurons (PNs; LA n = 264; BL n = 212) or fast-spiking interneurons (ITNs; LA n = 31; BL n = 45) based on spike width and mean firing rate (Figures S2D and S2E). To assess how the proportion of cells with CS-responses in the RRI task compared to previous reports on appetitive and aversive tasks, we computed the proportion of cells significantly responsive during the first second after CS onset. We found that ~20% of LA and ~35% of BL cells significantly increased or decreased their firing rates in response to the CS-Rs or CS-Ss (Figures 2A–2B, top row; rank-sum tests, p < 0.005), consistent with prior studies (Beyeler et al., 2016; Burgos-Robles et al., 2017; Lee et al., 2016, 2017; Namburi et al., 2015; Sangha et al., 2013; Shabel and Janak, 2009). When we expanded the test window from 1 to 10 s (the CS period prior to the US), the distribution of cells responsive during the CS-Rs and CS-Ss changed significantly in both LA (Figures 2A and 2B; chi-square tests: CS-Rs, χ² = 32.59, p < 0.00001; CS-Ss, χ² = 20.48, p = 0.000036) and BL (Figures 2A and 2B; CS-Rs, χ² = 50.83, p < 0.00001; CS-Ss, χ² = 16.36, p = 0.00028).

To determine whether this pattern of responsiveness was specific to the sensory modality of the CSs we used, a subset of 37 BL cells were recorded in a rat that had been trained to seek rewards in response to auditory CS-Rs (white noise) presented above the left or right reward port and avoid shocks signaled by light CSs, as above. In keeping with the close parallels between the anatomical pathways that relay auditory and visual information to the BLA (LeDoux et al., 1985; LeDoux, 1993; Linke et al., 1999), the proportion of cells excited by the auditory CS-Rs observed in this sample of BL cells (19%) was similar to that seen with light CS-Rs (13%; χ² = 1.02, p > 0.31). Furthermore, the proportion of cells responsive to both the auditory CS-R and light CS-S (5%) was nearly identical to the proportion of cells responsive to both light CS-Rs and light CS-Ss (4%; χ² = 0.21, p > 0.64).

Also consistent with prior studies, our sample included cells whose firing rates increased selectively during the CS-Rs or CS-Ss. These cells, respectively termed R-Cells (Figure 2C; LA, 5%; BL, 10%) and S-Cells (Figure 2D; LA, 19%; BL, 20%), showed similar changes in activity during CSs of the opposite valence: if a cell was responsive to one CS-R (or one CS-S), it responded similarly to the other CS-R (or CS-S) but displayed no response or an inhibition during CSs of the opposite valence. Furthermore, these responses were independent of head direction (three-way ANOVA for stimulus-excited cells, Head Direction × Stimulus Type × Region, F(12,712) = 1.05, p = 0.4039) and stimulus proximity at CS onset (three-way ANOVA for stimulus-excited cells, Stimulus Proximity × Stimulus Type × Region, F(2,278) = 1.4, p = 0.2487). However, most CS-responsive cells, hereafter termed mixed cells (LA, 40%; BL, 53%),
displayed marked between-cells as well as trial-to-trial variations in the late part of their CS-related activity, making it difficult to classify them into separate groups. Two examples of mixed cells are depicted in Figures 2E and 2F. The first shows a transient inhibition at the onset of all CS types followed by an excitation during some CS-R and CS-S trials (Figure 2E). The second increased its firing rate at the onset of all CSs but displayed a late persistent excitation only during the CS-Rs (Figure 2F).

Activity of Principal BLA Cells in Relation to CRs

We reasoned that the trial-to-trial fluctuations seen in the late part of CS responses could be related to variations in the type of CS presented. This hypothesis was tested by analyzing the activity of principal BLA cells during the CS-Rs and CS-Ss. The data were presented in bar graphs showing the proportion of LA or BL cells with significantly altered firing rates (rank-sum test, p < 0.005) during the first second of the CSs (top) or their entire duration (bottom). Figure 2 illustrates the activity of R-cells (C) and S-cells (D) whose firing rates increased selectively during the CS-Rs or CS-Ss, respectively. Mixed cells (E and F) displayed marked trial-to-trial variations in the late part of their CS-related activity.

Figure 2. Activity of BLA Neurons during the CS-Rs and CS-Ss

(A) CS-R. (B) CS-S. Proportion of LA or BL cells with significantly altered firing rates (rank-sum test, p < 0.005) during the first second of the CSs (top) or their entire duration (bottom).

(C–F) Representative examples of cells that showed significant changes in firing rates during one or more of the CSs. R-Cells (C) and S-Cells (D) are neurons whose firing rates increased selectively during the CS-Rs or CS-Ss, respectively. Mixed cells (E and F) displayed marked trial-to-trial variations in the late part of their CS-related activity. Ticks, individual spike times. Thick lines, Z-scored averages of firing rates. Vertical dashed lines, onset of CS-Rs (blue) or CS-Ss (red). Abbreviations: CS-R, reward-predicting conditioned stimulus; CS-S, shock-predicting conditioned stimulus; US, unconditioned stimulus. Related to Figures S2 and S3.
or timing of the CRs. To test this, we computed peri-event time histograms (PETHs) of unit activity referenced to the onset of CRs instead of CSs, revealing that the activity of many PNs in LA (30.6%) and BL (31.1%) significantly increased in relation to one or more CRs relative to their firing rate immediately preceding behavior onset (rank-sum test, p < 0.005; see STAR Methods). CR-activated cells were found among the three groups described above (R-cells, S-cells, and mixed cells). In many of these cells (LA, 56 of 81; BL, 38 of 66), CR-locked activity occurred despite the absence of significant short-latency CS responses. Two examples of such cells are shown in Figures 3A and 3B. The first cell selectively increased its firing rate at the onset of active avoidance (Figure 3A1) and remained at baseline during passive avoidance trials (Figure 3A2). Note the absence of increased spiking at the onset of the light stimulus (CS-S), indicated by the orange lines in Figure 3A. The second cell did not change its firing rate during reward approach (Figure 3B1) but subsequently displayed a marked and persistent activation whose onset coincided with that of reward anticipation (Figure 3B2).

To test whether these CR-related increases in firing rates in fact represent delayed sensory responses to the light stimuli, we compared the activity of all putative CR-coding cells in three conditions: when rats emitted different CRs in response to the same type of CS, on correct versus error trials, or when rats emitted the same CR at different times with respect to CS onset.

For the first analysis, we took advantage of the fact that on aversive trials, rats could avoid the US actively or passively, depending on their location with respect to the CS-S. Thus, we averaged the activity of all cells that, individually, showed significant increases in firing rates during active avoidance (hereafter termed AA-cells for simplicity; LA, n = 49; BL, n = 32) and compared their

Figure 3. Behavioral Correlates of Unit Activity
Arrows and arrowheads mark the onset of behaviors and CSs, respectively. (A and B) Individual examples of principal neurons that strongly increase their firing rates in relation to (A1) active avoidance (AA) but not passive avoidance (PA; A2) or (B2) reward anticipation (RAnt) but not reward approach (RA, B1). Vertical dashed lines, onset of reference behavior. Ticks, individual spike times (as many trials as rows of ticks are shown). Thick lines, Z-scored averages of firing rates. Yellow ticks, onset of CS. Cyan ticks, end of reference behavior. (C) Comparison between Z-scored averaged firing rate ± SEM of principal cells (n = 81) during AA (red) versus PA (blue), referenced to behavior onset (C1) or CS onset (C2). (D) Comparison between Z-scored averaged firing rate ± SEM of principal cells during correct (black) and error (red) CS-S (D1; n = 68) or CS-R (D2; n = 53) trials. (E and F) Relation between behavior onset (arrows) and unit activity for CS-S (E) and CS-R trials (F). Individual principal cells are shown in E1 and F1 (Z-scored average ± SEM of multiple trials). Z-scored average ± SEM of all available principal cells are shown in E2 and F2. Abbreviations: CS-R, reward-predicting conditioned stimulus; CS-S, shock-predicting conditioned stimulus.

1318 Neuron 99, 1315–1328, September 19, 2018
activity on aversive trials that called for an active (red, Figure 3C) or passive avoidance (blue, Figure 3C) response. Whether the activity of AA-cells was referenced to the onset of CRs (Figure 3C1) or CS-Ss (Figure 3C2), they showed significantly lower firing rates during trials that called for a passive instead of an active avoidance response (paired t tests; CR-locked responses, t(29) = 3.737, p = 0.002; CS-locked responses, t(29) = 4.44, p = 0.0001). However, it is possible that on CS-S trials calling for an active versus passive avoidance response, rats actually experienced a different CS. At odds with this interpretation, however, we noticed that in 29% of trials where the rats were outside the shock sector, they nevertheless generated behavioral freezing in response to the CS-S, indicating that on those trials, they perceived the CS-S and interpreted it as threatening even though they were not in the shock sector. Furthermore, when we restricted our unit analysis to those passive avoidance trials when the rats froze, we still observed that AA cells fired at markedly higher rates during CS-S trials calling for an active than a passive avoidance response (Figure S4A). Yet, it remains that the rats’ position relative to the CS-S differs on trials that call for an active versus passive avoidance. Thus, strictly speaking the CS-Ss are not identical.

To further test whether this contrast resulted from the rats’ differing proximity to the light stimuli on active versus passive avoidance trials, we next compared the averaged activity of all the AA cells recorded during sessions with one or more error trials, where the rats failed to actively avoid the shock (46 LA cells, 22 BL cells, 18 sessions, 137 error trials). During the first 5 s following CS onset, by which time active avoidance responses are typically completed, AA cells showed significantly higher firing rates on correct than error trials (Figure 3D1; black, correct; red, error), even though their position relative to the light stimuli was the same in both cases (paired t test t(10) = 8.497, p = 6.92 × 10^{-6}).

However, it could be argued that on error trials, rats failed to pay attention to or recognize the CS-S, causing reduced CS-evoked neural responses in these trials. Incompatible with this interpretation, in 61% of error trials (where rats failed to avoid the shock), rats nevertheless displayed behavioral freezing, again indicating that not only had they perceived the CS-S, but also that they had correctly interpreted it as threatening. In fact, the incidence of freezing on error trials (61%) was significantly higher than on all CS-S trials combined (14%; χ² = 304.48, p < 0.00001), suggesting that the reason why rats did not emit the correct response (active avoidance) on many error trials is because they failed to stop freezing. Furthermore, when we restricted our unit analyses to those trials where rats had paid attention to the CS-S and interpreted it correctly (that is when they froze), AA cells still had significantly higher firing rates on correct than error trials during the first 5 s after CS-S onset (Figure S4B).

Furthermore, consistent with a recent report (Lee et al., 2016), an analogous phenomenon was observed in cells with significant increases in firing rates during the CS-Rs (31 LA cells; 22 BL cells, 14 sessions, 167 error trials). In these cells, while the early sensory-related part of the response to the CS-R remained when rats failed to approach and anticipate reward delivery, the late component was nearly abolished (Figure 3D2; paired t test t(19) = 17.345, p = 4.17 × 10^{-13}). However, rats might have incorrectly processed positional cues on error trials and as a result could not identify the responses that the CS-S or CS-R called for.

Additional evidence of the behavioral dependence of unit activity was obtained by comparing the time course of firing rates on trials where rats emitted the correct CRs at different latencies with respect to CS onset. Figure 3E1 shows an example AA cell whose activity increased when the avoidance behavior was emitted at different latencies across trials (black arrows). This dependence of cell activity on behavior latency was also evident when all AA cells (n = 81) were considered (Figure 3E2). Similar results were obtained with RA cells (n = 44; Figures 3F1 and 3F2). While the variations in BL activity as a function of CR latency suggest that some or all of the spikes BL cells generate long after CS onset are related to the CRs, it remains difficult to draw boundaries between CS- and CR-related activity. In fact, it could be argued that even late components are sensory responses whose time course varies from trial to trial. However, since there are no apparent reasons for these variations, it seems more parsimonious to ascribe late response components to variations in CR latency.

**Generalized Linear Model Untangles Cell Activity**

Even though the above suggests that the activity of BLA neurons is related to both stimuli and behaviors, the temporal overlap between CSs and associated CRs constitutes a major obstacle when determining what variables LA and BL cells actually encode with PETHs. To circumvent these limitations, we defined the CS component of spiking as time locked to the presentation of the CS, and the CR component as spiking around the onset of the CR. Using these definitions, we fit the spiking of individual cells using a group Lasso generalized linear model (GLM) with 10-fold cross validation (see STAR Methods and Figures S5–S7). This method is not based on PETHs, but takes full advantage of trial-to-trial variations in the type, timing, and duration of the variables of interest (stimuli, evoked, and spontaneous behaviors) to infer which one(s) neurons encode. It is used to fit the spiking of individual cells for the duration of the task. Importantly, this type of GLM allows for dimensionality reduction in correlated data and encourages sparsity when identifying the task variables that are related to cell activity (Breheny and Huang, 2015; Tibshirani, 1996; Yuan and Lin, 2006).

To test whether the GLM’s coding estimates had face validity, we first compared the model’s output to the results of PETHs referenced to different variables. For instance, standard analyses had revealed that the cell shown in Figure 4A (same as in Figure 2F) had short-latency (behavior-independent) responses to all CSs and that it increased its firing rate in relation to reward anticipation and, more weakly, during active avoidance. The model correctly captured this complex response profile (Figure 4A, CSs row 1; reward anticipation row 3, columns 4–5; active avoidance, row 3, column 1).

Although the model estimates were generally consistent with the results of the standard analyses, there were some discrepancies. These likely resulted from the GLM’s ability to disentangle coding for concurrent variables that PETHs could not isolate. A prime example of this is the encoding of movement...
speed versus active CRs (reward approach and active avoidance). This is illustrated in Figure 4B, which shows a principal BLA cell that had been identified as an AA and RA cell with PETHs. In order for the model to assess whether the increases in firing rates seen during active CRs were, in sum or in part, actually due to the coding of locomotion, we had the GLM fit the cell’s activity to the rat’s movement speed throughout the task (including the inter-trial intervals). The model estimated that part of the spiking observed during active avoidance and reward approach trials was related to speed (Figure 4B, row 5). However, since speed did not fully account for the cell’s increases in firing rates, the model attributed the rest of the response to the active CRs. Most importantly, the model-estimated spiking combined across all task-related variables closely matched observed spiking (Figure 4B, bottom row). See Figure S6 for more example cells fit by the GLM and Figure S7 for a characterization of the model’s estimation accuracy.

**Task-Related Activity in the BLA and Striatum**

Next, we used the GLM output to characterize coding of task variables by LA and BL neurons at the population level. For comparison, we carried out the same analyses in striatal neurons (n = 208) recorded above the amygdala. In the three structures, R62% of presumed principal cells (Figure 5A1) and R65% of interneurons (Figure 5A2) encoded two or more task variables. The number of encoded task features was distributed similarly across structures in principal cells and interneurons (Kullbeck-Leibler tests for permuted distributions: all percentiles within 95th two-sided cut-offs; see STAR Methods).
To study regional differences in the coding of task variables, we computed the proportion of cells with excitatory and inhibitory modulations (respectively, red and blue bars in Figures 5B1 and 5C1) and their average absolute magnitude (Figures 5B2 and 5C2). These analyses revealed a remarkable degree of regional similarity in the encoding of task features. In the three regions, >48% of principal cells (Figure 5B1) and >71% of interneurons (Figure 5C1) showed excitatory or inhibitory modulations by all task features with the exception of freezing in the striatum (41% of PNs, 37% of ITNs). A three-way ANOVA on the magnitude of the modulations revealed that PNs (Figure 5B2) had stronger modulations than interneurons (Figure 5C2) across all task features ($F_{(1,3785)} = 71.52, p = 0$). Furthermore, an interaction effect between feature type and structure ($F_{(7,3785)} = 2.47, p = 0.0158$) indicated that striatal cells were more strongly modulated by speed than LA and BL cells (Figure 5B2; Tukey-Kramer test, STR versus LA $p = 5.91 \times 10^{-6}$, STR versus BL $p = 5.91 \times 10^{-6}$), consistent with the role of the striatum in motor control (Brown and Robbins, 1989; Kravitz et al., 2010; Reading et al., 1991). Finally, LA cells were most strongly driven by freezing and this modulation was significantly higher than in striatal but not BL neurons (Figure 5B2; Tukey-Kramer test, LA versus STR $p = 0.0198$, LA versus BL $p = 0.0892$).

It was previously reported that BLA cells encode valence during appetitive and aversive tasks (Belova et al., 2008; Beyeler et al., 2016; Lee et al., 2017; Namburi et al., 2015; Sangha et al., 2013). In these prior studies, however, the definition of valence was based on CS responsiveness, raising the question of the degree to which BLA cells encode valence in both the stimulus and behavioral dimensions. Presumably, if a cell is encoding valence, it should not only respond in a selective manner to the cues that elicit positive or aversive outcomes but also show corresponding changes in activity in relation to the conditioned behaviors they elicit. Thus, we defined valence as an excitatory response to either an appetitive or aversive CS, and at least one behavior of the same valence, plus an inhibition or no change in firing rate to the CS and behaviors of the opposite valence.

To study valence coding among BLA neurons, we plotted the observed normalized change in firing rate of LA and BL cells to reward approach (Figure 6, y axes) as a function of that seen in relation to active avoidance (Figure 6, x axes) and color-coded the points corresponding to each cell based on the model-estimated modulation (blue, inhibition; red, excitation) by the various task variables (Figure 6: top CSs; middle, active CRs; bottom, passive CRs). First, we verified that the model-estimated modulation by reward approach and active avoidance matched the change in firing rate observed in relation to these behaviors. Generally, the two matched: cells with strong positive estimated modulations by reward approach (red dots in Figure 6A2) or active avoidance (red dots in Figure 6B2) had higher values on the y and x axes, respectively ($RA r = 0.56$, $p = 1.75 \times 10^{-17}$; $AA r = 0.66$, $p = 3.29 \times 10^{-25}$). Cells that deviated from this trend strongly encoded speed, as discussed above (not shown).

Figure 5. Multi-dimensional Coding by BLA Neurons, as Determined by the GLM

In (A)–(C), the left, middle, and right columns show data obtained in LA, BL, and striatum, respectively. (A1 and A2) Frequency distributions of the number of task variables encoded by presumed principal cells (A1) and interneurons (A2). (B1) Proportion of presumed principal cells (y axis) that exhibited excitatory (red) or inhibitory (blue) coding of different task variables (x axis). (B2) Absolute average ± SEM modulation of firing rates in relation to each variable. (C1) Proportion of presumed interneurons (y axis) that exhibited excitatory (red) or inhibitory (blue) coding of different task variables (x axis). (C2) Absolute average ± SEM modulation of firing rates in relation to each variable. Abbreviations: AA, active avoidance; CS-R, reward-predicting conditioned stimulus; CS-S, shock-predicting conditioned stimulus; Frz, freezing; ITNs, interneurons; PA, passive avoidance; PNs, principal neurons; RA, reward approach; RAnt, reward anticipation. Related to Figures S5–S7.
Consistent with the notion of valence coding in the BLA, a minority of principal BLA cells encoded valence along both stimulus and behavior dimensions. We first provide a qualitative description of these findings and then summarize the results of the statistical analyses. In Figure 6, cells encoding positive valence are huddled near and high along the y axis. They tend to show strong positive modulations by the CS-R (Figure 6A1), reward approach (Figure 6A2), and reward anticipation (Figure 6A3). Conversely, cells encoding negative valence hug the x axis. They tend to show strong positive modulations by the CS-S (Figure 6B1), active avoidance (Figure 6B2), and freezing (Figure 6B3). Additionally, a subset of these valence-coding cells flipped their responses from an excitation to an inhibition in relation to variables of the opposite valence.

Overall, a low proportion of principal cells encoded positive or negative valence in LA (3.4% and 9.5%) and BL (1.9% and 17.5%), respectively. To test whether this incidence was significantly higher than expected by chance given the proportions of cells whose activity was modulated by the various CSs and CRs, we permuted the normalized peak values of all task features (see STAR Methods) 10,000 times to break any relationships between stimuli and behaviors and tested whether the proportion of valence-coding cells fell outside the two-sided 95th percentile of the null distribution. The incidence of valence-coding cells did not exceed chance levels in LA or BL (33rd and 84th percentile, respectively). Of note, this negative result was not due to the fact that BLA neurons responded sparsely: most encoded multiple task features (Figure 5A), including CSs and CRs (Figure 5B). The low incidence of valence-coding we report is due to the fact that our definition of valence coding not only takes into account the CS specificity of the cells responses, as in prior studies, but also their behavioral correlates.

**Heterogeneous Coding in LA and BL**

Since explicit coding of valence is not prominent at the single cell level in the BLA, we tested whether a valence code is present in the population activity. To examine this, for each region and cell type, we correlated the cells’ peak modulations of all task variables (Figure 7, dots mark significant values at p < 0.001). Principal LA neurons exhibited a marked tendency to show correlated responses to task events of the same valence, and decreased or negative correlations between those of different valences. This pattern was less striking for BL, but still apparent. To test the possibility that the correlations might partly result from correlations between the features themselves, we re-assessed the significance of the correlation matrices by comparing them to shuffled neural activity. To this end, the peak values of each task feature were shuffled 10,000 times and the correlation...
matrices were recomputed for each shuffled distribution. The actual correlation coefficients were considered significant if they fell outside the two-sided 99.9th (p < 0.001) percentile of the null distribution. However, the two approaches yielded nearly identical results: there was a 98% agreement between the two methods for principal cells and 100% for interneurons.

Similar results were obtained using a d-prime score (d'; Keene et al., 2016; McKenzie et al., 2014) that measured the degree to which within-valence (CSs plus CRs) correlations exceeded those between-valence (bootstrap confidence intervals from zero, LA = 1.56, p < 10^{-4}; BL = 0.99, p = 0.0009). In LA, valence coding was also evident when restricting the d-prime analysis to the valence of behaviors (d' = 5.99, p < 10^{-4}) or CSs (d' = 1.26, p = 0.004), whereas in BL it was significant for behaviors (d' = 1.96, p = 0.0009) but not CSs (d' = 0.58, p = 0.1). The same analyses applied to ITNs revealed significant valence coding for behaviors (LA = 3.27, p = 0.014; BL = 6.57, p = 0.0028), but not CSs (LA = 2.48, p = 0.27; BL = 0.09, p = 0.4).

While the above analyses indicate that valence is coded along multiple dimensions in the BLA, it remains unclear how these codes are distributed at the population level and how they relate to each other. For instance, do cells similarly encode the valence of stimuli and behaviors? To address these questions, we applied multidimensional scaling (MDS; Figure 8A), a dimensionality reduction strategy that proved useful to study coding in the gustatory and olfactory systems, where sensory representations are obscure (Di Lorenzo et al., 2009; Youngentob et al., 2006).

In brief, the peak firing rate modulations related to the ten variables of interest (3 CS-Ss, 2 CS-Rs, 4 CRs, and speed) were used to plot the position of neurons in a ten-dimensional space (Figure 8A1). This map was then collapsed into two dimensions (Figures 8A2 and 8A3) based on the similarity of the cells’ peak responses, while minimizing errors in setting their relative distance (see STAR Methods). Once this map was generated for each region, we examined how it related to various coding dimensions (valence-behavior, valence-CS, active-passive, behavior-CS, speed) by fitting a plane in 3D space where the first two dimensions were provided by the MDS and the third was one of the coding dimensions (Figures 8A4 and 8A5; labels at top of plots in Figure 8B). In these maps, the steepness of the color gradient (vector length) is proportional to how strongly the coding dimension is represented in the population. The relative orientation of the color gradients (angle between them) indicates whether the different coding dimensions are related. That is, orthogonal orientations indicate independent codes, whereas parallel orientations indicate closely related codes. For simplicity, Figure 8C represents the steepness and orientation of the color gradients by vectors, one for each of the coding dimensions examined. For consistency, in each region, we set to vertical the vector associated with the coding dimension most strongly represented, while preserving the relative orientation of the other vectors.

One concern with MDS is the possibility that the algorithm distorts the relationship between coding dimensions, precluding the interpretations outlined above. To check whether this was...
Figure 8. Coding for Different Task Dimensions at the Population Level

(A1) Each neuron was described by a vector composed of their responses to CSs and behaviors. (A2 and A3) These were mapped to a two-dimensional space using MDS. (A4) For each neuron we also computed a code value that was derived by contrasting responses to different stimuli and behaviors. Specifically, we ranked each dimension relative to the other units in the dataset. Then, for each unit we constructed a feature vector (based on ranks) by taking the GLM-estimated peak modulations by the five stimuli (CS-R1, CS-R2, CS-S1, CS-S2, CS-S3) and five behaviors (RA, Rant, AA, Frz, speed). The coding dimensions were then calculated as specified in (A4). (A5) PNs from each region were placed in a space where the first two dimensions were their MDS values, and the third was the value for one of their codes. A three-dimensional plane was fit in this space that could capture a systematic mapping of the code under consideration in the low dimensional space.

(B) Low dimensional maps of coding for LA PNs. Each PN (filled circle) was placed into a two-dimensional space such that nearby PNs had similar response vectors. The code value for each neuron was then added (color of the filled circle). A code that is systematically represented in the low dimensional response space produces a strong gradient (e.g., ValBeh), and one that is not produces a weak gradient (e.g., BehCS).

(legend continued on next page)
the case, we examined other measures that did not depend on MDS. We found that vector length positively correlated with d-prime (Figure S8A, Spearman’s r = 0.86, p = 0.01, df = 6). Moreover, the relative orientation of the gradients corresponded with the correlation between coding dimensions (Figure S8B, Spearman’s r = –0.66, p = 0.02, df = 10). That is, coding dimensions with a small angle between them were positively correlated, while those orthogonal to each other tended toward zero correlation, and those with angles approaching 180 degrees were negatively correlated.

Consistent with the correlation matrices (Figure 7), valence behavior was the strongest coding dimension in LA and BL (Figures 8C1 and 8C2). Although less salient, the valence-CS dimension was also expressed in LA and BL and aligned to valence behavior. Coding for speed also fell roughly along the same axis as valence, except that its direction was opposite in LA and BL. Evidence of this can also be found in the correlation matrices where the speed-related activity of LA neurons tended to correlate positively with their modulation by aversive events (Figure 7, top left) and less so or inversely in BL (Figure 7, top middle). The active-passive coding dimension was also present in LA and BL (Figures 8C1 and 8C2), but it was orthogonal to the valence code, implying that the two populations could simultaneously represent both types of information with minimal interference. As for speed, the modulation along the active-passive coding dimension was opposite in LA versus BL. Finally, while the same coding dimensions were observed in the striatum (Figure 8C3), their relative importance and overall organization differed markedly from LA and BL. In particular, speed emerged as the dominant coding dimension followed by the active-passive, behavior-CS, valence-CS, and valence-behavior codes.

To assess the variability of vector length estimates, we used a 95% bootstrap confidence interval (Figure 8D). For each vector length, we also defined 95% confidence intervals returned by computing null gradients where code values were randomly permuted across the population (dashed lines in Figure 8D denote 97.5 and 2.5 percentile of the null distributions). Using this approach, it was determined that all vector lengths were significant with the exception of behavior-CS in LA (Figure 8D1). To assess whether vector lengths differed significantly from the others within a region, we used a permutation test with correction for multiple comparisons, confirming that valence behavior was the dominant coding dimension in LA and BL (p = 0.04; Figures 8D1 and 8D2), whereas speed dwarfed the others in the striatum (p = 0.04; Figure 8D3). Bootstrap resampling with threshold p of 0.05 (corrected for ten comparisons) was also used to assess the difference of angle estimates between codes (Figure 8C, concentric circles at tip of vectors), indicating that the active-passive coding dimensions differed significantly from the valence dimensions in LA and BL (valence-behavior versus active-passive BL, p = 0.04, LA p < 0.002; valence-CS versus active-passive BL, p = 0.04; LA p < 0.002), but not in the striatum (both p > 0.05). See Figures S8C and S8D for similar analyses in interneurons.

**DISCUSSION**

It was reported that appetitive or aversive conditioning leads to the potentiation of CS inputs onto valence-specific BLA neurons (McKernan and Shinnick-Gallagher, 1997; Rumpel et al., 2005; Tye et al., 2008). As a result, CSs would trigger approach or defensive CRs through the activation of valence-coding neurons with distinct outputs (Beyeler et al., 2016; Namburi et al., 2015). However, because of the close temporal relation between CSs and CRs, and the fact that in most prior studies, each CS could only trigger one CR, it remains unclear whether BLA neurons encode CS identity, the behavior it elicits, or if the two are one and the same. Our study aimed to shed light on this question using a task where rats could emit different CRs in response to the same CSs, allowing us to dissociate normally intertwined task features that BLA neurons potentially encode.

Using this approach, we found that in most LA and BL cells, the magnitude and time course of their activity during the CS depended on the type and timing of the CRs. Thus, the CS responses of BLA cells, that is, their sensory aspects, are separable from the activity that drives CRs; one does not necessarily cause the other. Second, while valence could be decoded at the population level, the incidence of valence-coding cells, as defined by their corresponding CS- and CR-related selectivity, did not exceed chance. Last, a majority of BLA neurons concurrently encode multiple task features and behaviors. Below, we consider the significance of these findings for the contribution of the amygdala to the acquisition and expression of valenced behaviors.

**Independent Encoding of CSs and CRs by BLA Neurons**

In the present study, two distinct approaches revealed that sensory responses and CRs are encoded concurrently, yet independently, by the same BLA neurons. First, separately plotting BLA firing rates when CRs were emitted at different latencies revealed an early CS-locked response and a later component whose latency increased with that of the CRs. Similarly, comparing BLA activity on correct and error trials provided evidence for two response components: an early sensory-related part that was present on correct and error trials and a late component, present on correct trials but nearly absent on error trials.

Second, we used a GLM to identify the response components associated with stimuli and behaviors. The challenge when isolating these factors is that they co-occur in time, and the performance of the CR is contingent on the presentation of the CS. Given their overlap, it is reasonable to ask to what degree can we separate their individual contributions to the cells’ activity. One way to address this is to operationally define these components by clarifying exactly what aspects of the spike trains were used by the GLM to extract CS and CR components. CS responses...
were those components of the spiking that were time locked to the presentation of the CS, while CR responses were time locked to the emission of the behavior. This approach hinges on the fact that a particular CR is only present on some trials, and when it does occur its latency and duration will not always be the same. The component of the spiking that covaries with this behavioral variability will be assigned to the CR, while the remaining component that was consistently related to the onset of the stimulus across trials will be assigned to the CS. This codifies the intuition that spiking related to a particular factor will systematically covary with it, and it is difficult to envision how this approach could be improved upon without making inferences about the cognitive state of the animal, such as perception of the stimulus, which we do not have direct access to.

The fluctuating character of the late, CR-related component despite the stability of the early CS-locked phase is inconsistent with the view that, due to the potentiation of CS inputs on BLA neurons, CRs are automatically triggered by CSs. Instead, these results suggest that the receipt of sensory information about CSs and the emission of CRs are distinct, yet related processes. Most likely, interactions between the BLA and other structures such as the medial prefrontal and temporal cortices (Pitkanen, 2000) regulate whether and when a given CS will elicit a CR of a particular type. Such interactions would account for the behavioral flexibility exhibited by animals depending on context (Perusini and Fanselow, 2015).

The realization that BLA neurons separately, yet concurrently encode CSs and CRs sheds new light on earlier findings. For instance, using Ca²⁺ imaging of large groups of BLA neurons in vivo, Grewe et al. (2017) reported that after fear conditioning, the representation of the CS by BLA neurons morphed into that of the US. Since our results indicate that individual BLA cells show similar activity in relation to behaviors of the same valence, and considering that after (but not before) training, both USs and CSs elicited negatively valenced behaviors, the change in representation Grewe et al. (2017) observed could have resulted from the shift in the behavioral outcome of the CS.

**At the Level of Individual Cells, Valence Coding Is Not Prevalent in the BLA**

Here, we define valence-coding neurons as cells that showed a firing rate increase in relation to at least one CS and CR of the same valence coupled to an inhibition or no change in activity in relation to all CSs and CRs of the opposite valence. Under this definition, few valence-coding cells were observed in the BLA (<20%). The proportions of cells whose activity was modulated by the various CSs and CRs, the incidence of valence-coding cells did not exceed chance. Yet, our GLM analyses indicated that valence information could be decoded from the collective activity of BLA neurons. To gain insights into this apparent contradiction, we correlated the estimated modulation of BLA activity by CSs and CRs. This revealed that in LA, the firing rate modulations by CSs of opposite valence were positively correlated, whereas those associated with CRs of opposite valence were not. Thus in our paradigm, valence information is mostly derived from the modulation of activity by CRs.

Of note, the correlated modulations of BLA activity by the CS-Rs and CS-Ss were expected in our paradigm given that rats had to rely on location to determine whether similar light stimuli were CS-Ss or CS-Rs. The fact that BLA neurons showed differentiated modulations to CRs of opposite valence despite the similar sensory properties of the CSs cast further doubt on the notion that potentiated CS inputs onto distinct valence-coding BLA neurons is the sole factor supporting the differentiated expression of conditioned emotional behaviors. Rather, these findings again suggest that in between the receipt of sensory information about CSs by BLA neurons and the emission of CRs are additional processing steps where other streams of information, in this case about place, determine the outcome of the CS.

Another unexpected finding in our study was the positive correlation, particularly marked in LA, between the activity modulations associated with freezing and active avoidance. Although these CRs constitute markedly different defense strategies, BLA cells showed similar activity modulations in relation to both behaviors. This observation suggests that while different subsets of BLA neurons may generate distinct CRs via their particular outputs, it is incumbent on target effector neurons to decode what CRs should be generated. Presumably, other inputs allow them to disambiguate the significance of BLA inputs at any given time.

**Multi-dimensional Coding by BLA Neurons**

While explicit coding of valence by individual neurons was not prominent in the BLA, valence information was present at the population level. In fact, our MDS analyses suggest that several coding dimensions are represented in the ensemble activity of BLA neurons, including valence, speed, and active versus passive behaviors. Unexpectedly, valence coding was primarily seen in the behavioral domain. The preferential association of valence coding with behavior rather than CSs is reminiscent of a trace conditioning study in monkeys where valence coding by BLA neurons was more prominent during the trace period, shortly before the CR (Paton et al., 2006). Furthermore, in the present study, the active-passive coding dimension was orthogonal to valence coding, which suggests that LA and BL neurons can concomitantly represent different types of information with minimal interference. Interestingly, the active-passive and speed dimensions had opposite directions in LA versus BL. Whereas in LA, a negative modulation by speed was associated with a positive modulation by valence behavior, the opposite relationship was found in BL. The origin of these differences remains unclear. A major challenge for future investigations will be to determine the identity of the extrinsic afferent(s) (or potential intrinsic interactions) contributing to each stream of information represented in the activity of BLA neurons.

Consistent with our findings, other behavioral paradigms have uncovered coexisting codes in BLA ensemble activity. For instance, Munuera et al. (2018) reported that in non-human primates, cells that encode the reward value of non-social stimuli also encode the hierarchical rank of conspecifics. Moreover, in a paradigm where the reward associated with different stimuli varied in a context-dependent fashion, BLA cells encoded context, CS identity, and reinforcement expectations (Saez et al., 2015).

Overall, these considerations suggest that conditioned emotional behaviors do not depend on the recruitment of
discrete, dedicated subsets of cells that explicitly encode a particular valence, but on multiplexed representations distributed across the ensemble activity of BLA neurons. These high-dimensional coding schemes confer major computational advantages (Fusi et al., 2016), notably a dramatic increase in the dimensional coding schemes confer major computational advantages. Even the control of spinal motoneurons by cortical and rubral cells depends on such distributed codes (Fetz, 1992).

The presence of distributed coding would seem to contradict the numerous demonstrations of emotional behaviors controlled by specific cell types in the amygdala. While our findings argue against a strong view where neurons are treated as labeled lines for specific emotional behaviors and states, they are consistent with the fact that those variables are only partially correlated with anatomical position (Beyeler et al., 2018), projection target (Beyeler et al., 2016), or genetic profile (Kim et al., 2016). Thus, it might be better to view coding in the amygdala as a spectrum, not a switchboard.

**STAR METHODS**

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **CONTACT FOR REAGENT AND RESOURCE SHARING**
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Overview of the experimental timeline
  - Risk Reward Interaction Task
  - Surgical procedures
  - Unit recordings and clustering
  - Histology
  - Behavioral Analyses
  - Classification of neurons
  - Generalized Linear Model (GLM)
  - Multidimensional scaling (MDS)
- **QUANTIFICATION AND STATISTICAL ANALYSES**
  - Observed behavior- or CS-related changes in firing rates
  - Normalized peak firing rate modulations
  - Incidence of valence Cells
  - Number of Features Encoded
  - Similarity Matrices
  - D-prime Analyses
  - Assessment of population response maps
- **DATA AND SOFTWARE AVAILABILITY**

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes eight figures and can be found with this article online at https://doi.org/10.1016/j.neuron.2018.07.036.

**ACKNOWLEDGMENTS**

This material is based upon work supported by NIMH R01 grants MH107239 and MH112505 to D.P. P.K. was supported by the Behavioral and Neural Sciences Graduate Program of Rutgers University.

**AUTHOR CONTRIBUTIONS**

D.B.H. came up with the initial concept of the RRI task. D.B.H. and P.K. built the behavioral apparatus. P.K., D.B.H., and D.P. contributed to designing the experiments as well as to the data and statistical analyses. P.K. conducted most of the experiments. P.K. and D.P. wrote most of the first draft of the manuscript and made most of the figures. All authors contributed to refining the manuscript.

**DECLARATION OF INTERESTS**

The authors declare that they have no competing financial interests.

Received: April 5, 2018
Revised: June 4, 2018
Accepted: July 20, 2018
Published: August 23, 2018

**REFERENCES**


STAR METHODS

KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deposited Data</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>The full dataset and custom MATLAB code will be made available upon request.</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Experimental Models: Organisms/Strains</td>
<td>Long Evans rat</td>
<td>Charles River Laboratories</td>
</tr>
<tr>
<td>Software and Algorithms</td>
<td>MATLAB</td>
<td>MathWorks</td>
</tr>
<tr>
<td>RStudio</td>
<td><a href="https://www.rstudio.com/">https://www.rstudio.com/</a></td>
<td>Version 1.0.136</td>
</tr>
<tr>
<td>KlustaKwik</td>
<td><a href="http://klustakwik.sourceforge.net/">http://klustakwik.sourceforge.net/</a></td>
<td>N/A</td>
</tr>
<tr>
<td>Klusters</td>
<td>Hazan et al., 2006</td>
<td>N/A</td>
</tr>
<tr>
<td>Other</td>
<td>Silicon probe</td>
<td>NeuroNexus</td>
</tr>
<tr>
<td>Intan amplifier</td>
<td>Intan technologies</td>
<td>N/A</td>
</tr>
</tbody>
</table>

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Denis Pare (pare@andromeda.rutgers.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Procedures were approved by the Institutional Animal Care and Use Committee of Rutgers University, in compliance with the Guide for the Care and Use of Laboratory Animals. We used adult male Long Evans rats (300 – 350 g; Charles River Laboratories).

METHOD DETAILS

Overview of the experimental timeline
Rats were housed individually with *ad libitum* access to food and water and maintained on a 12 hr light/dark cycle. Animals were habituated to the animal facility for one week and then to handling 15 min daily for 3-5 days. One week prior to training in the RRI task, animals were placed on water restriction while maintaining them at ≥ 85% of initial body weight. The water restriction protocol consisted of 6 consecutive days of restriction followed by one day of *ad libitum* access to water. All experiments were performed during the light cycle.

Risk Reward Interaction Task
The apparatus (Figure 1) consisted of a dimly lit rectangular arena (90 cm in length x 30 cm in width) with high walls (60 cm), a floor of metal bars (5 mm od, spaced 8 mm) and no ceiling. The floor was divided into 3 equal sectors (30 cm x 30 cm), each with an array of light-emitting diodes (LEDs) below it (Figure 1A). At each end of the arena, there was a water port with an LED behind it (Figure 1A). A programmable microcontroller (Arduino, SparkFun, Niwot, CO) controlled the activation of the LEDs, the delivery of water through the water ports (60 μL), and the delivery of footshocks in each sector (0.4 mA, 10 s).

Reward availability was signaled by activation of an LED (CS-R) behind the active water port for 20 s. The reward was delivered 10 s after CS-R onset through a dipper which retracted at the end of the CS-R. Similarly, shock delivery was signaled by an LED (CS-S) below one of the floor sectors for 20 s. A mild footshock was delivered to the corresponding sector 10 s after CS-S onset. The shock and CS-S co-terminated. If rats were located in the sector when the CS-S activated, they could actively avoid the shock by leaving that sector in the first 10 s following CS-S onset. If rats were not in the shock sector at CS-S onset, they could passively avoid the shock. An overhead video camera (29.97 frames/s) recorded the behavior. A white noise maker was used to mask any sounds from the room during the task. In the absence of CS, light levels were very low in the RRI task apparatus (0 lux). When a CS was activated, light levels increased considerably (≥27 lux) such that they could not go undetected, irrespective of the rats’ position.
The day before training began, rats (n = 8) were habituated to the arena for 1 hr. No light stimuli, water rewards, or shocks were delivered during habituation. The white noise maker was on and the lights were off for the duration of habituation. Behavioral training on the RRI task occurred over 7 days. During this period, rats gradually learned to retrieve water rewards at each of the two ports and avoid shocks from each of the three sectors when signaled by the corresponding LED. The first two days, rats were trained for two 15 min sessions (9 shock trials, 6 reward trials); one in the morning and one in the afternoon. Once animals acquired the active avoidance behavior, the sessions were extended to one hour (39 shock trials, 26 reward trials). Training continued on days 3 through 7 with 1 one-hour long session per day until animals reached at least 80% performance, on all trial types combined, usually within 7 days (Figure S1B).

Surgical procedures
Once they reached criterion, animals were implanted with silicon probes aimed at the BLA. Rats were anesthetized with a mixture of isoflurane and O2 and administered atropine sulfate (0.05 mg/kg, i.m.) to aid breathing. They were then placed in a stereotaxic apparatus with nonpuncture ear bars. A local anesthetic (bupivacaine, S.C.) was injected in the scalp. Fifteen minutes later, an incision was made to expose the skull and a craniotomy was performed over the amygdala. One of three types of silicon probes (see below) was attached to a 3D-printed microdrive, and aimed just dorsal to the BLA (coordinates in mm relative to bregma: AP −2.2 to −3.6, ML 5.2, DV 6.0). The microdrive allowed us to lower the probe gradually between recording sessions. The three types of silicon probes used were Buzsaki32L (n = 3 rats; Figure S2A), Buzsaki64L (n = 1 rat; Figure S2B), and a custom-designed probe (n = 4 rats; Figure S2C). The Buzsaki32L and 64L probes respectively consisted of four and eight shanks spaced 200 μm apart. The custom-designed probe had 4 shanks with 16 channels on each shank arranged in a tetrode formation. Each tetrode was 333 μm from the neighboring tetrode and the inter-shank distance was 250 μm. Rats were allowed 1-2 weeks to recover from surgery.

Unit recordings and clustering
Once animals recovered from surgery, unit recordings began. During the recording sessions, the various CS-Rs and CS-Ss were presented in random order, not in blocks. Red and green LEDs added to the rat’s headcap during the recording sessions allowed us to track the rat’s activity throughout the task. Red LEDs were always positioned on the rostral end of the headcap and green LEDs on the caudal end. This allowed us to track head position by taking the midpoint between the red and green LEDs, and head direction by calculating the angle between the midpoint and the red LEDs with respect to the apparatus. Individually and as a group, all rats spent more time in sector and 1 and 3 compared to sector 2. This difference in chamber occupancy was observed during the inter-trial intervals (% time ± SE in sector 1, 42% ± 5%; sector 2, 11% ± 2%; sector 3, 48% ± 5%) as well as during the trial periods (sector 1, 43% ± 3%; sector 2, 6% ± 1%; sector 3, 51% ± 3%).

The silicon probe was lowered ≥ 140 μm after every recording session to avoid recording the same cells across days. Data were sampled at 30 kHz and stored on a hard drive. A high pass filter was initially applied to the recording data, followed by a median filter (window size of 1.1 ms). A threshold was applied to extract spikes. Principal component analysis was used on the spikes and the first three components were clustered using KlustaKwik (http://klustakwik.sourceforge.net/). Finally, spike clusters were refined manually using Klusters (Hazan et al., 2006). In order to reliably separate clusters, we inspected autocorrelograms and cross-correlograms of clusters. Autocorrelograms had to show a refractory period of at least 2 ms. Cross-correlograms could not have a refractory period as this indicated that the same unit was shared between clusters. If a unit had unstable spike shape, it was excluded.

Spike duration was calculated for a given unit by selecting the channel with the largest action potentials and peak-to-trough amplitude. The time between spike trough and peak represented the spike’s duration (Bartho´ et al., 2004). BLA units were classified as presumptively projection cells or interneurons based on their baseline firing rates (6 Hz cutoff) and trough-to-peak durations. Striatal units were divided into fast-spiking interneurons and presumed medium spiny neurons based on interspike interval and trough-to-peak durations. Based on Berke (2008), highly active cells with less than 2% of their interspike intervals longer than 1 s were classified as interneurons. The trough-to-peak duration used for striatal units was the same as for BLA cells (0.55 ms). All cells that failed to meet both criteria were labeled unclassified and excluded from analyses.

In a prior study (Amir et al., 2018), we tested the reliability of these classification criteria on a different sample of rat BL neurons recorded with the same methods. Using cross-correlograms in ~17,500 pairs of BL neurons, we looked for evidence of monosynaptic inhibition from putative principal cells to other cells (a case where an interneuron would be misclassified as a projection cell) or excitation from putative interneurons to other cells (indicating a principal cell misclassified as an interneuron). Only 0.99% of principal cells and 3.6% of interneurons with putative connections were found to be misclassified.

With respect to the usefulness of spike duration to distinguish principal BL neurons, only one study (Bienvenu et al., 2012) directly compared the duration of extracellularly measured spike waveforms in principal cells (n = 23) and different types of interneurons (n = 51) of the basolateral amygdala. On average, all four classes of BLA interneurons examined in this study (pavulubin positive, calbindin positive, axo-axonic, amygdalo-striatal projecting) generated spikes of shorter duration than principal cells. Of the 51 interneurons described in this study, only two interneurons had a spike duration as high as the principal cell with the lowest spike duration. Together, these earlier findings support the reliability of the criteria we used to distinguish principal cells and interneurons.
Histology
At the end of experiments, while rats were anesthetized, electrolytic marking lesions were made on either the most dorsal or ventral electrodes, alternating on each shank (10 μA for 16 s), so that lesions marking different shanks would be distinct. They were then perfused-fixed through the heart, their brains extracted and cut on a freezing microtome. The sections (80 μm) were counterstained with 1% thionine solution. Only neurons histologically determined to have been recorded in the striatum or BLA were analyzed. Cells located within ± 150 μm of nuclear boundaries were excluded. See Figure S3.

Behavioral Analyses
All behaviors were analyzed using a custom made graphical user interface (GUI) in MATLAB (The MathWorks, Natick, Massachusetts, USA). Behavior start and end times were determined in the GUI by indicating the first and last frame of the video recording when the rat initiated and ended a behavior. The start of active avoidance was noted as the first frame when the rat began to move off the lit sector, and the end of the behavior was denoted as the first frame when the rat ended the continuous avoidance behavior on one of the unlit sectors. Passive avoidance start time was denoted as the first frame of a CS-S trial where the animal was away from the lit sector and ended either at the end of the trial, as long as the animal remained away from the lit sector for the entire trial duration, or until the animal approached or engaged with the lit sector. Freezing was denoted as any period of immobility except for respiration lasting at least one second during CS-S trials. For reward approach, the start time was the first frame the rat began movement toward the reward port and the end was the frame when his nose reached the port. Reward anticipation began when the rat placed both front paws on the reward port and waited there until reward delivery. We considered a total of five behaviors: active avoidance, passive avoidance, freezing, reward approach, and reward anticipation.

Classification of neurons
A cell was defined as a R-cell if, relative to the 5 s immediately preceding the CS-R, it had a significant increase in firing rate (based on rank-sum test p < 0.005) either during the first 1 s after CS-R onset or during the initial 10 s CS-R period and an inhibition or no change to the CS-Ss. S-cells were similarly defined using the opposite valence CSs. AA-cells and RA-cells were defined as cells with a significant increase in firing rate for 1 s after behavior onset compared to 2 s prior to behavior onset.

Generalized Linear Model (GLM)
A regularized regression, group Lasso, with Poisson distribution (grpreg R package; Brehey and Huang, 2015) and ten-fold cross-validation was used to fit the spiking of individual cells for the duration of the task. Spiking was binned (50 ms bins) across the entire recording session and stimuli and behaviors were indicated with ones when they occurred. After Park et al. (2014), the stimulus and behavior events were convolved with basis functions defined by log-time scaled raised cosine bumps separated by π/2 radians (50 ms). Each event kernel was represented as a linear combination of basis functions spanning a duration of time (see below).

The model was fit by minimizing the value β in

$$Q(\beta) = \frac{1}{2n} \| y - X\beta \|^2 + \lambda \sum_j \sqrt{K_j} |\beta_j|$$

where y represents the spike train across time, X is the design matrix with the basis functions for stimuli and behaviors, βs are the least-squares regression coefficients, λ is the regularization penalty for an $L_1$ norm, and $\sqrt{K_j}$ normalizes across groups of different sizes. The lasso penalty parameter (λ) to the Euclidean ($L_1$) norm is chosen based on the lowest cross-validation error and it is applied to each group, creating sparsity and variable selection at the group level. (Brehey and Huang, 2015; Tibshirani, 1996; Yuan and Lin, 2006). Cross-validation sets were assigned by dividing the recording session into ten equal segments. These approaches allowed us to be confident that our beta coefficients were significantly different from zero.

GLM Basis Functions and Kernels
The GLM kernels for each stimulus and behavior type that best fit spiking activity were created by combining a set of pre-determined basis functions. We used two different sets of basis functions to represent CSs and CRs. Stimulus basis functions covered their initial 10 s, had a sharp onset and narrow width which became smoother and wider across time (Figure S5A), reflecting stimulus responses that tend to have sharp onsets and decay quickly. Behavior basis functions extended before and during the behavior onset to capture spiking activity related to planning and executing the behavior (Figure S5B). However, they were bounded by the start of a trial. These functions were then scaled by the model’s β values and summed within time bins to create a single kernel for each stimulus and behavior (see Figures S5 and S5D for example stimulus and behavior kernels from a cell fit by the model). Most importantly, if a cell did not encode a parameter, the group Lasso GLM gave beta values of 0 for the corresponding basis functions. Thus, only parameters that best fit observed spiking of a given cell are selected by the model, allowing us to perform dimensionality reduction.

To compare the model fit to observed spiking we used the coefficient of determination ($R^2$). Kernels fit with all model parameters for appetitive and aversive trials during the 10 s CS period were adjusted for firing rate and concatenated. Similarly, observed PETHs for
the 10 s CS period during all appetitive and aversive trials were concatenated and smoothed with a Gaussian window of 550 ms. These model-estimated kernels and observed PETHs were used to calculate the coefficient of determination as follows:

\[ R^2 = 1 - \frac{\sum (y_i - f_i)^2}{\sum (y_i - \bar{y})^2} \]

where \( y_i \) represents the observed PETH, \( f_i \) is the model-estimated kernel, at different time points, \( i \), and \( \bar{y} \) is the mean of the observed PETH.

**Assessing the independence of GLM variables**

When GLM variables are correlated or non-independent, a phenomenon termed multi-collinearity, the variance of the regression parameters is inflated, leading to an incorrect selection of predictors. Recommendations for how to address the problem of collinearity include using penalized regression methods (such as Ridge and Lasso regression; Dormann et al., 2013), as we have done in this study. Penalized regression methods, such as the group Lasso, allow for the selection of variables through cross-validation and are equipped to deal with correlated variables (Tibshirani, 1996; Breheny and Huang, 2015). One of the methods for identifying collinearity in predictors is to assess the correlation between predictors: it should not exceed a threshold of 0.7 (Dormann et al., 2013). To verify whether this was the case in our dataset, we computed all pairwise correlations between the 18 grouped variables used in our regression, in each session and for all rats. This analysis revealed that most of the recording sessions (82%) did not have any pairwise correlations exceeding the threshold (mean ± SEM absolute pairwise correlations across all sessions: 0.1 ± 0.0017).

**Multidimensional scaling (MDS)**

To visualize and quantify the coding of task dimensions at the population level for each region, we mapped our codes (see Figure 8A4) into a low dimensional space created using multidimensional scaling (MDS). This technique projects a set of points in a high dimensional space into a low dimensional space, while trying to preserve their pairwise distances. For a given unit we constructed a feature vector by taking the GLM-estimated peak modulations by the five stimuli (CSR1, CSR2, CSS1, CSS2, CSS3) and five behaviors (RA, Rant, AA, Frz, speed) and ranked each dimension relative to the other units in the dataset. We then calculated the Euclidean distance between all pairs of feature vectors within a region. This set of distances was passed to the MDS algorithm (non-metric, squared stress criterion) to reduce the distance from ten to two dimensions, while still preserving the relative distances between the cells. Then, for each unit type and region we added a third dimension that corresponded to one of the codes suggested by the correlation matrices (behavior valence, stimulus valence, stimulus/behavior) or previous studies (speed, active versus passive behaviors; Amir et al., 2015). Since coding dimensions are linear combinations of features, they do not provide any additional coding information beyond the covariance between sets of features at the population level. This created a collection of points in a three dimensional space, and we fit a plane (code-value = \( \beta_1 \text{MDS-Dim1} + \beta_2 \text{MDS-Dim2} + \text{intercept} \)) to those points to determine if the feature of interest was systematically mapped across the population. The magnitude of the slope of the plane, \( \sqrt{\beta_1^2 + \beta_2^2} \), indicated the degree to which that feature was mapped across the population. By comparing the difference in the direction, \( \arg(\beta_1 + i\beta_2) \), between two feature planes, we could determine the degree of overlap in their representations at the population level.

**QUANTIFICATION AND STATISTICAL ANALYSES**

Group data are reported as average ± SEM. All neurons with stable firing rates and spike shapes were included. Firing rates were logarithmically transformed as natural logarithms. All statistical tests were two sided. Different procedures were used to assess statistical significance depending on the type of data, as specified below.

**Observed behavior- or CS-related changes in firing rates**

To determine if individual neurons showed significant behavior- or CS-related changes in firing rates, we used the Wilcoxon rank-sum test (\( p < 0.005 \)) on the observed binned spiking locked to either the behavior or CS onset. Spiking was binned in 50 ms windows. A baseline of 5 s before the event onset was used to compare binned spiking with either 1 s after CS/behavior onset or 10 s after CS onset.

**Normalized peak firing rate modulations**

The absolute peak of each CS and CR kernel was normalized by the baseline firing rate as follows: (Peak-Baseline)/Baseline. The sign of the modulation was preserved in the normalized peak in order to identify cells excited or inhibited in relation to each task feature. Normalized peak modulations ≤ 0.001 were considered non-significant and were set to zero. The average modulations to the task features were computed by taking the mean of the absolute value of all significant peak modulations for each parameter. For Figure 6, the observed normalized peak firing rate during active avoidance and reward approach are plotted on the x and y axes, respectively. These values were calculated by binning the observed spiking in 50ms bins and referencing PETHs to behavior onset. The mean firing rate during the behavior was normalized by the firing rate during the 5 s preceding behavior onset, similarly to the normalized peaks for the GLM.
Incidence of valence Cells
In order to determine if the proportion of valence cells was significantly higher than expected by chance given the proportions of cells whose activity was modulated the various CSs and CRs, we permuted the model-estimated peak firing rate modulations for all task variables across cells 10,000 times. This shuffling procedure randomizes any potential associations between encoded variables and yields a null distribution for the incidence of valence-coding cells. Using this null distribution, we then calculated the percentile of the observed proportions. If this value fell outside the two-sided 95th percentile of the null distribution it was considered significant.

Number of Features Encoded
To identify how many features cells encoded, we used the significant normalized peak firing rate modulations of each cell task features: CS-R, CS-S, reward approach, reward anticipation, freezing, active avoidance, and passive avoidance. For the CSs, we averaged modulations within CS types (CS-R1-2; CS-S1-3). The number of significant excitatory modulations was counted for each cell and that value represented the number of features encoded. To test if the distributions of features encoded differed between LA, BL, and the striatum, we used the Kullback-Leibler divergence as a test statistic. A null distribution was created by shuffling the location of each cell randomly 10,000 times and computing a divergence score after each permutation. This gave us a distribution of divergence scores which we then used to compare the observed scores. If the observed score was outside the two-sided 95th percentile of the null distribution, it was considered significant.

Similarity Matrices
Similarity matrices were computed using a Spearman correlation of the peak firing rate modulations by all stimuli and behaviors across cells. The p value was corrected for multiple (45) comparisons (p = 0.001).

D-prime Analyses
The d-prime (d') metric is used to measure the separation between different coding dimensions based on the distributions of the correlation coefficients in the similarity matrices (Figure 7; Keene et al., 2016; McKenzie et al., 2014). The coding dimensions tested were valence (CS-R1, CS-R2, RA, RAnt versus CS-S1, CS-S2, CS-S3, Freezing, AA), valence-CS (CS-R1, CS-R2 versus CS-S1, CS-S2, CS-S3), and valence-behavior (RA, RAnt versus Freezing, AA). The d’ metric was calculated for each coding dimension by comparing the degree to which within-valence correlations exceeded those between-valence, as follows:

\[
d' = \frac{\mu_{\text{within}} - \mu_{\text{between}}}{\sqrt{\frac{1}{2} (\sigma_{\text{within}}^2 + \sigma_{\text{between}}^2)}}
\]

Where \(\mu_{\text{within}}\) and \(\mu_{\text{between}}\) are the means of the correlation coefficients for the within and between coding dimension features, respectively, and their corresponding variances, \(\sigma_{\text{within}}^2\) and \(\sigma_{\text{between}}^2\). To identify d’ values significantly different from 0, we compared the observed d’ values to a bootstrap distribution in which normalized peak values were randomly sampled 10,000 times, and similarity matrices and d’ values were then re-computed for each bootstrap sample. The p value of the observed d’ was calculated based on the bootstrapped distribution and considered to be significant if p < 0.05. To compare if two d’ values were significantly different from each other, we computed a shuffled distribution of d’-difference values (10,000 permutations). If the observed d’-differences were >95% of the shuffled distribution then the d’-differences were considered significant. For example, when comparing if Valence d’ differed between PNs and ITNs, we permuted the cell identities 10,000 times and re-calculated the similarity matrices and d’ values each time. We then calculated the d’-difference, defined as \(d'_{\text{PNs}} - d'_{\text{ITNs}}\). This gave us a null distribution of d’-differences which could be used to compare the observed d’-difference value.

Assessment of population response maps
We estimated the distribution of the slopes using bootstrap resampling. The plane was refit 500 times using different collections of neurons that were resampled with replacement. Each of the coding planes was recalculated for the same collection of resampled data points, generating matched pairs. For each code these were used to generate 95% confidence intervals. To determine if a coding dimension was significantly mapped across the MDS space, we permuted code values across neurons and recalculated the planes (500 times), yielding the distribution of null planes. To determine if two codes were significantly different in their strength or direction, we took the difference in the magnitude or angle (using circular distance) of the slopes between the matched pairs and found the probability of a difference of zero. When comparing between codes, a Bonferroni correction for multiple comparisons (10 potential comparisons) was performed on all p values.

DATA AND SOFTWARE AVAILABILITY
The full dataset and custom MATLAB code will be made available upon request.