bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Coding of reward uncertainty and probability by orexin neurons

Ed F. Bracey, Nikola Grujic, Daria Peleg-Raibstein, Denis Burdakov

ETH Zürich, Department of Health Sciences and Technology
Correspondence: denis.burdakov@hest.ethz.ch

ABSTRACT

Appropriate arousal in our brains critically relies on hypothalamic hypocretin/orexin neurons (HONs). HON activity regulation has been a subject of intense research, uncovering slow responses to nutrients and hormones, and fast responses to external sensory inputs. However, it remains unclear whether HONs also respond to statistical patterns in the environment. Among these, reward probability and uncertainty are fundamental measures of information that enable evaluation of predictions, and guide diverse aspects of cognition such as learning and motivation. Using different cues to denote reward likelihood across the full probability range (0-1) , we found that HON population responses to the reward-predicting cues co-vary with reward uncertainty. Single-cell analyses additionally revealed HON subsets specializing in monotonic representations of probability, as well as cells sharply tuned to uncertainty. Furthermore, simultaneous recordings of HONs and midbrain dopamine neurons (DANs) indicated that HONs and DANs diverge in their coding of reward probability, both in responses to reward-predicting cues and to rewards. These results demonstrate previously unseen coding of reward expectation in genetically-defined hypothalamic cells central to clinical conditions and neurobehavioral physiology, and suggest a possible role for HONs in converting reward statistics into arousal-based allocation of cognitive and autonomic resources.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

INTRODUCTION

The brain constantly creates expectations and compares them to inputs (Bar, 2007; Clark, 2013). This ability is considered fundamental for behavioral adaptations that increase survival (Engel et al., 2001; Keller and Mrsic-Flogel, 2018; Livneh and Andermann, 2021; Rao and Ballard, 1999; Volkow et al., 2010; Wolpert and Ghahramani, 2000; Zimmerman et al., 2016). In particular, neural representations of reward expectation are considered essential for interrelated brain functions such as arousal, learning, and motivation in both humans and animals (Dreher et al., 2006; Nassar et al., 2012; O'Doherty, 2004; Schultz, 2000; Schultz et al., 1997). Neural substrates of reward expectations have thus been a subject of intense research. These substrates are considered to be neurons whose activation varies monotonically with either probability or uncertainty of rewards (Dreher *et al.*, 2006; Fiorillo et al., 2003; Ma and Jazayeri, 2014). Probability (*P*) quantifies the degree to which a reward is expected, and uncertainty is related to *P* being maximal at $P = 0.5$ and decreasing at higher and lower *P*s (Dreher *et al.*, 2006; Fiorillo *et al.*, 2003).

Genetically-targeted lesions, opto/chemogenetic manipulations, and pharmacological effects indicate that hypothalamic hypocretin/orexin-producing neurons (HONs) (de Lecea et al., 1998; Sakurai et al., 1998) are essential for orchestrating adaptive arousal (Adamantidis et al., 2007; Chemelli et al., 1999; de Lecea et al., 2006; Hara et al., 2001; Harris and Aston-Jones, 2006; Lin et al., 1999; Sakurai, 2014; Tsunematsu et al., 2011). A small number of studies indicate that HONs are activated by reward-associated stimuli, suggesting that they may represent reward information (Harris et al., 2005; Hassani et al., 2016). However, HON responses to reward probability (i.e. to cues indicating the full probability range, 0-1) remain undefined, and thus whether HONs code reward probability or uncertainty is unknown. Here we designed and performed experiments that reveal this important missing information. We then asked whether HONs and other subcortical neurons implicated in reward

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

processing, specifically the dopamine neurons of the ventral tegmental area (DANs) (Fiorillo *et al.*, 2003; Schultz *et al.*, 1997), code redundant or distinct aspects of rewards.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

RESULTS

Experience alters the influence of reward probability on hypocretin/orexin neurons

To examine the influence of reward probability on HONs, mice were conditioned in a Pavlovian procedure with distinct auditory stimuli, each stimulus predicting a reward (milkshake, delivered at the end of the 2 s stimulus) with a different probability (P = 0, 0.5, 1) (Figure 1A). After conditioning, anticipatory movement evoked by the stimuli increased with the reward probability indicating that mice learnt to behaviourally discriminate the stimuli (Figure S1A). To determine how conditioning influenced the activity of HONs, we optically monitored their responses to the probability-indicating stimuli by imaging the dynamics of individual HON-GCaMP6s neurons using 2-photon microscopy (Figure 1B,C), or by tracking HON-GCaMP6s population responses using fiber photometry (Figure 1D,E; Figure S1B). We found that conditioning significantly changed the influence of probability on conditioned stimulus (CS) responses of HONs, indicating neural learning (2-photon data, Figure 2F: 2-way RM ANOVA, probability x experience $F(2, 946) = 31.988$, $P < 0.0001$, n = 475 cells from four mice; photometry data, Figure 2G: 2-way RM ANOVA probability x experience $F(2, 24) = 8.204$, $P < 0.01$, n = 13 mice).

Coding of reward probability and uncertainty in subsets of hypocretin/orexin neurons

We next examined whether, after conditioning, the probability representations during the CS-responses vary across HONs, and whether these representations code reward probability or uncertainty. To achieve this, we first classified the probability response profiles (i.e. responses to P = 0, 0.5, 1) of 278 individual HONs that responded to CSs into groups using unsupervised K-means clustering (Figure S2A-C, see Methods). Within these clusters, we then quantified the “classic” coding of probability (Spearman’s correlation test to determine if response magnitudes monotonically increase with probability) and uncertainty (Kruskal-Wallis test to determine if response magnitudes follow the ranking 0.5 > 0.75&0.25 > 1&0). We found significant coding of uncertainty, and no coding

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of probability, in a large subpopulation (n = 117/278, ~42%) of HONs (Figure S2A-C, clusters 1 and 6, statistical analyses are given in the figure legend).

A subpopulation of these "uncertainty cells" (Figure 2A, n = 37/278 cells, corresponding to cluster 1 in Figure S2A-C) selectively responded to CSs indicating uncertain rewards (0 > P > 1), and not to CSs indicating certain rewards (P = 1 or 0) (Figure 2B). The CS responses of these cells did not monotonically follow reward probability (Spearman's correlation: Rs =- 0.3000, P = 0.6833), but were significantly influenced by reward uncertainty (Kruskal-Wallis H(2) = 70.05, P < 0.0001) (Figure 2B). A non-overlapping subpopulation of "probability cells" (Figure 2C, n = 41/278 cells, corresponding to cluster 2 in Figure S2A-C) monotonically represented reward probability (Spearman's correlation: Rs = 1, P < 0.05), but not were not significantly influenced by reward uncertainty (Kruskal-Wallis H(2) = 5.625, P = 0.0601) (Figure 2D).

Other cell classes were also identified (Figure S2A-C), but the probability and uncertainty coding cells (clusters 1, 2, and 6 in Figure S2) accounted for majority of the population response, due to their larger response amplitudes (Figure S2B). An alternative cell classification method (manual threshold-based clustering, see Methods) produced broadly similar results (Figure S2D-F).

Reward coding in hypocretin/orexin vs dopamine neurons: responses to conditioned stimuli

We next asked whether reward probability and uncertainty exert a similar influence on LH HON and VTA DAN populations. To address this as directly as possible, we used dual fiber photometry to concurrently record LH HON and VTA DAN neural responses to reward-indicating conditioned stimuli in the same mince (n = 10 mice, Figure 3A,B). Statistical analysis revealed that the HON and DAN population responses to conditioned stimuli were differentially influenced by probability, suggesting distinct probability coding in the two neural systems (Figure 3D; 2-way RM ANOVA, effect of probability x neural type F(4, 72) = 4.001, P < 0.01, n = 10 mice).

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Within these datasets (Figure 3D), DAN population responses monotonically increased with reward probability (Spearman r = 1, P < 0.01), but were not significantly influenced by reward uncertainty (Kruskal-Wallis H(2) = 0.4410, P = 0.8021). In contrast, the HON population responses did not monotonically follow reward probability (Spearman r = 0.6000, P = 0.1750), but were significantly influenced by reward uncertainty (Kruskal-Wallis H(2) = 10.05, P = 0.0066). Furthermore, the HON responses showed no consistent relation to coincident DAN responses across mice (Figure 3E; monotonic analysis: Spearman r = 0.1862, P = 0.1955; linear analysis: Pearson r = 0.046, P = 0.7511; analyses within individual probabilities gave similar responses, Figure S3).

Collectively, these analyses indicate that phasic population responses of HONs and DANs to reward-predicting cues code distinct aspects of reward expectation.

Reward coding in hypocretin/orexin vs dopamine neurons: responses to rewards

Lastly, we sought to determine whether HON responses to rewards depend on reward probability (i.e. do they code reward prediction error), and whether they are similar to DAN responses within the same brain. To this end, we analysed dual HON-DAN photometry recordings in mice presented with further CS-reward pairings including delivery of unexpected rewards at a low rate (n = 7 mice, Figure 4A,B, see Methods). Statistical analysis revealed that HON and DAN population responses to reward were differently influenced by reward probability, suggesting that the two neural system perform distinct comparisons between expected and actual reward (Figure 4C; 2-way RM ANOVA, effect of probability x neural type F(4, 48) = 2.982, P < 0.05, n = 7 mice).

Within these datasets (Figure 4C), DAN responses monotonically decreased with reward probability, i.e. they coded reward prediction error as previously described (Spearman r = -0.5143, P < 0.001), but were not significantly influenced by reward uncertainty (Kruskal-Wallis H(2) = 0.9367, P = 0.6260). In

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

contrast, HON responses did not monotonically follow reward probability (Spearman r = 0.1266, P = 0.1286), nor were they influenced by reward uncertainty (Kruskal-Wallis H(2) = 0.9367, P = 0.6260). Furthermore, the HON reward responses showed no consistent relation to coincident DAN reward responses across mice (Figure 4D; monotonic analysis: Spearman r = 0.1664, P = 0.3394; linear analysis: Pearson r = 0.2302, P = 0.1833; analyses within individual probabilities gave similar responses, Figure S4).

Collectively, these results suggest, when measured within the same subject experiencing the same task, the reward responses of the DAN population code reward prediction error, but the HON population reward responses do not appear to be influenced by reward expectation.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

DISCUSSION

HONs are powerful controllers of brain state, and factors that govern their activity have been the subject of intense research. Our present findings reveal that HON activity is sensitive to statistical associations within the world, in particular to reward probability which is a factor not previously studied as a HON modulator. HON responses to reward-predicting cues transmitted a reward uncertainty signal at the population level (Figure 1, Figure 3E). At the single cell level, we discovered distinct HON subsets that specialized in coding reward uncertainty (Figure 2A,B), or reward probability (Figure 2C,D). Thus, distinct statistical features of reward expectation were represented as parallel streams of information within the HON system. However, the uncertainty representation predominated in the HON population (Figure 1, Figure 3E), in contrast to the DAN population whose phasic responses to reward-predicting cues coded for reward probability and not uncertainty (Figure 3E). In addition to these responses to reward-predicting cues, responses to rewards themselves revealed no clear relation to reward probability in HON reward responses, while DAN reward responses coded reward-prediction error (Figure 4) as previously described (Schultz *et al.*, 1997; Schultz and Dickinson, 2000). Because in these experiments HONs and DANs were recorded simultaneously using the same tools during the same experiences in the same subjects, these differences cannot be explained by variations in experimental conditions that typically confound direct comparisons of neural responses across different experiments, and suggest that HONs and DANs code different aspects of reward-related information.

HONs were initially conceptualized as energy sensors that control arousal and energy balance based on their slow (minute time scale) responses to glucose and hormones (Burdakov, 2005;

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Burdakov and Alexopoulos, 2005; Yamanaka et al., 2003a). Subsequent studies revealed that HONs also receive brain-wide neural inputs and rapidly (seconds) respond to multiple neurotransmitters (Gonzalez et al., 2016a; Li et al., 2002; Sakurai et al., 2005; Yamanaka et al., 2003b). These studies shifted attention to rapid modulation of HONs by external sensory streams (Bracey and Burdakov, 2020; Gonzalez *et al.*, 2016a; Hassani *et al.*, 2016; Karnani et al., 2020; Mileykovskiy et al., 2005), and raised the question of whether HONs simply relay sensory input or display "neural learning" allowing experience-dependent remodeling of their fast sensory responses. Our findings demonstrate that such neural learning does occur in HONs to a previously unsuspected degree, enabling HONs to communicate reward expectation in the form of probability and uncertainty in distinct cell subsets.

In the ventral tegmental area, DANs have been proposed to be a downstream excitatory target of HONs (Baimel et al., 2017; Thomas et al., 2022). Based on this, we thus expected VTA DAN activation to be correlated to some extent with HON activation in the same brain. Surprisingly, we did not detect such correlations, neither at the level of reward cue responses (Figure 3F, Supplementary Figure 3), nor of reward responses (Figure 4C, Supplementary Figure 4). This does not necessarily contradict the previous studies, since our results do not rule out that HON→DAN connectivity exists on a yet-to-be-defined subpopulation level, or that DANs may receive HON information and integrate it with other inputs. However, it does indicate that, at the population level, HONs and DANs display distinct coding of probability and uncertainty in their reward-cue and reward responses (Figure 3E, Figure 4B). Our findings thus reveal a reward expectation signal in HONs that is not redundant with the classic DAN coding of reward (Schultz, 2002).

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In summary, our findings reveal previously unobserved HON responses that co-vary with aspects of reward statistics which are fundamentally important as measures of information and for assessing the accuracy of predictions. In future work, it will be important to determine which of the numerous, brain-wide direct neural inputs and outputs of HONs (Gonzalez *et al.*, 2016a; Peyron et al., 1998) correspond to upstream and downstream circuits that read and write the HON representations of reward probability. Given that HONs are central to the control of arousal, movement, panic, and reward-seeking; and implicated in learning and risk-avoidance (Blomeley et al., 2018; Boutrel et al., 2005; Burdakov and Peleg-Raibstein, 2020; Hara *et al.*, 2001; Harris *et al.*, 2005; Harris et al., 2007; Johnson et al., 2010; Karnani *et al.*, 2020; Sakurai, 2014), deeper knowledge of their control by statistical patterns in the environment will illuminate multiple aspects of adaptive and maladaptive physiology and behavior.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

METHODS

Animals and gene targeting

All animal procedures were performed in accordance with the Animal Welfare Ordinance (TSchV 455.1) of the Swiss Federal Food Safety and Veterinary Office, and were approved by the Zurich Cantonal Veterinary Office; or approved by the UK Home Office. Mice were kept on a reversed 12-h/12-h light/dark cycle and water, and unless otherwise stated, chow ad libitum. Experiments were performed during the dark phase. All mice were over 8 weeks old. WT C57BL6N mice were used for 2-photon recordings of orexin neurons, and DAT-Cre mice (Jax: Slc6a3tm1(cre)Xz) were used for dual photometry recordings from orexin and dopamine neurons. For recording of LH orexin/hypocretin cell activity, we used AAV1-hORX.GCaMP6s (2.0 x 10^{13} vg/ml, Penn vector core). This virus targets orexin/hypocretin neurons with >96% specificity (Gonzalez et al., 2016b). For recording VTA dopamine cell activity, we used pAAV9.CAG.Flex.GCaMP6s.WPRE.SV40 (1.9 x 10^{13} vg/ml, Addgene) or pGP-AAV1-syn-FLEX-jGCaMP7f-WPRE (titre 2.1 x 10^{13} vg/ml, Addgene). 6s and 7f dopamine recordings were pooled in presented analysis, because their response amplitudes were statistically indistinguishable.

Surgery

Implanting GRIN lenses into the lateral hypothalamus

Mice were anesthetized with isoflurane, placed in a stereotaxic frame (Kopf instruments), and a small incision was made to expose the skull surface. A 0.2 mm craniotomy was drilled at AP: -1.35 mm, ML: +/- 0.9 mm and the skull was roughened using a 0.8 mm diameter dental drill to remove periosteum, and promote adhesion. A Nanoject III (Drummond Instruments) with a glass micropipette was used to inject three 50 nL injections of hORX.GCaMP6s at a rate of 1nL/s, with coordinates from bregma of AP: -1.35, ML: +/- 0.9, DV: -5.7, 5.4 and -5.1. After removal of the injection pipette, the GRIN lens (0.39 NA, 7.3mm long, 0.6mm diameter cylindrical lens, Inscopix), was lowered in the center of the craniotomy to a depth of -5.3 mm from the height at the nearby skull surface using a linear manipulator (Scientifica), at a rate of 2 um/s. The lens was secured using three-part dental cement (Superbond, Sun Medical), and a custom made stainless steel head plate (Protolabs) was attached to the skull and GRIN lens cap using histoacryl, then three-part cement. The cement was then coated in black nail varnish (Essie), and the skin glued using Histoacryl around the implant site.

Implanting dual fiberoptic sensors into the lateral hypothalamic and the ventral tegmental areas

Mice were anaesthetised, given analgesia and fixed in the stereo-taxic frame as described above, and two small craniotomies were drilled at AP: -1.35, ML: +/- 0.9 and AP: -3.32 (to target the LH) and ML: +/- 0.5 (to target the VTA). A Nanoject III (Drummond Scientific) was used to deliver three 50nL injections of AAV1-hORX.GCaMP6s in to the LH at a rate of 1nL/s, with coordinates from bregma of AP: -1.35, ML: +/- 0.9, DV: -5.7, 5.4 and -5.1, and two doses of 50nL flex-GCaMP6s or 7f in the VTA in the contralateral hemisphere, with coordinates from bregma of AP: -3.32, ML: +/- 0.5, DV: -4.4 and -4.2. The bone surface was then roughened with a larger bore dental drill, and the steel headplate was glued to the skull surface using clear Histoacryl (Braun). Optic cannulas with a 200 µm core fiber optic (ThorLabs CFML12U-20, NA 0.39) in custom-built holders were then advanced to a depth 50 -100 µm above the most superficial viral injection site in each brain region. One of the two fibers was angled at 10 degrees to make space for securing optical cables later. This was randomised between VTA and LH. optic-fiber cannulas and headplate were then secured using three-part dental cement (Superbond, Sun Medical), and the fiber holders retracted. The cement was coated in black nail varnish (Essie), and the skin glued using Histoacryl around the implant site.

Neural recording apparatus

2-photon imaging experiments

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Hypothalamic 2-photon imaging was performed similarly to our previous work (Karnani *et al.*, 2020). Briefly, the setup consisted of two head-fixation bolts made of customised steel-optical posts (Thorlabs), which were positioned to allow mice to run on a Perspex disk. A reward port, mounted on a lockable hydraulic arm (Noga Engineering) was placed around 2 mm from the animal's tongue. Rewards were delivered using a peristaltic pump (Peripump WPM1), driven by a valve driver interfaced with a multifunction I/O device (National instruments, USB-6343), running custom-built software. The mouse's head was positioned under the microscope objective. Changes in GCaMP fluorescence were then imaged with a custom-built electro-tunable lens equipped resonant/galvanometer scan head two-photon microscope (INSS) and a femtosecond-pulsed mode-locked Ti:sapphire laser (Spectra-physics Mai Tai HP Deepsee 2) at 950 nm through a 20 (0.45 NA, Olympus) air-IR objective at 31 frames/s with 512 x 512 pixel resolution using custom Labview software (Karnani *et al.*, 2020). Six planes were imaged using the electro-tunable lens, yielding a volume rate of 5.1 volumes/s. Images were obtained with a 510/80 nm band-pass emission filter. A National Instruments USB-6008 DAQ board was used to output TTL triggers and count imaging frames. Auditory tones and reward delivery were controlled using custom-built Labview software, and outputs from a lick sensor attached to the reward spout and rotary encoder were recorded in the 6008 DAQ board. Frame outputs from the DA board were used to trigger trial onsets in Labview software on a separate PC, to ensure stimuli delivered by the USB 6343 (NI) were always aligned to the same plane. Bottom planes were aligned as best as possible to the same depth each session.

Fiber photometry experiments

In a similar head-fixation setup as described above for 2-photon imaging, cells in the LH and VTA of fiberoptic-implanted mice were imaged using fiber photometry similar to our previous work (Kosse and Burdakov, 2019). Briefly, GCaMP was excited with alternating blue and purple LEDs (465 and 405 nm respectively, Thorlabs), and resulting fluorescence emission collected using a photoreceiver (Thorlabs) through a fiber-connectorized GFP filter cube (Doric, FMC_GFP_FC). Average excitation light power was ~100 µW at the fiber tip. Fluorescence produced by 405-nm excitation provides a real-time control for motion artifacts ((Kim et al., 2016), see Figure S1B). Auditory tones and reward delivery were controlled using custom-built Labview software, and outputs from a lick sensor attached to the reward spout, the rotary encoder and photometry signal were acquired at a rate of 400 Hz using a HEKA LH1+1 I/O box, and HEKA Patchmaster software.

Histology and Immunohistochemistry

Mice were terminally anaesthetized and perfused transcardially with 4% PFA. Brains were post-fixed for 24 h, placed in sucrose for a further 24 hours, then frozen in dry ice and stored at -80^{o}c. 50 um coronal brain slices were later sectioned using a cryostat. For immunohistochemistry, sections were blocked in PBS with 0.3 % Triton X-100 and 1% bovine serum albumin (blocking solution) for 1 h, incubated with goat anti-orexin (1:1000; Santa Cruz) over-night, washed, incubated with Alexa 647 conjugated donkey anti-goat (1:1000; Invitrogen) for 3.5 h, washed and mounted. Epifluorescence images were then acquired and stitched using an epifluorescence microscope (Nikon) and merged in ImageJ.

Associative conditioning

Reward probability conditioning

For classical conditioning in both photometry and GRIN lens experiments, after at least 14 days of recovery post surgery, mice were habituated to headfixation and running on the disc for 2 sessions of around 20 minutes each. In the next 2 sessions of headfixation, they were habituated to auditory stimuli with 15 - 30 trials of each frequency, presented in a random order. Stimuli were pure tones of 4, 6, 8, 10 and 12 Hz, 2s in duration, at an acoustic power of 60-70 dB, played using the PC soundcard,

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

through a speaker (Visaton Titankalotte DSM 25 FFL, 8 Ohm). Next, they were food restricted and habituated to reward by positioning the lick spout 2-3 mm from the tongue, and presenting milkshake rewards (Strawberry Energy Milk Drink, Emmi) of 14 ul at random intervals from 5 - 9 seconds, until the mouse stopped licking. Typically this took 10 - 40 trials. To ensure homogeneity in baseline state before stimulus onset, stimuli or rewards were only delivered after the rotary encoder detected no movement for 2 seconds. Food restriction was maintained from then on. They were given full access to food at weekends. In the next session, food restricted mice were presented with the 5 pure tones of 2 second duration in a random order, each tone paired with a given probability of reward, which was delivered at the offset of the auditory tone (Fiorillo *et al.*, 2003). Reward probabilities (P_{REW}) of 0, 0.25, 0.5, 0.75 and 1.0 were paired with tones of 4, 6, 8, 10, 12, or 12, 10, 8, 6, 4 Hz, respectively. Tones were only delivered after no changes were detected using the rotary encoder for 2 s. In a given session, trials were presented until mice stopped licking, typically after 15 - 30 trials per stimulus, 70-150 trials per session. In total, mice performed 150 - 200 trials per reward probability over the course of 7 - 8 days.

Unexpected reward experiments

In Figure 4, mice were presented with the same reward pairings as before, but unexpected rewards were delivered at a rate of 0.02% on some P_{REW} = 0 trials, and trials with no rewards were delivered at the same rate for P_{REW} = 1 trials (Fiorillo *et al.*, 2003). Conditioning was continued for at least 1000 trials per reward probability, yielding around 10 unexpected P_{REW} = 0 rewards, and 10 P_{REW} = 1 no rewarded trials per mouse.

Statistics and data analysis

Statistical assessment of neural coding

Coding of reward uncertainty was evaluated by quantifying the influence of reward uncertainty on neural activity. To achieve this, we followed the method of (Fiorillo *et al.*, 2003), where each dataset in which the five reward probabilities were tested, the increases in neural activity during the CS (Figures 1-3) or reward (Figure 4) was ranked across the five probabilities for each cell (in 2-photon experiments) or each mouse (in photometry experiments). A Kruskal-Wallis test was then performed on the ranked values, with three groups defined by the degree of uncertainty (i.e. p=0.0 and 1.0; p=0.25 and 0.75; p=0.5). The ranking of the data points accounted for the paired nature of the data from each mouse or neuron, while the Kruskal-Wallis test is appropriate for multiple comparisons of nonparametric data. Coding of reward probability was evaluated by quantifying monotonic relationships between reward probability on neural activity (Fiorillo *et al.*, 2003). The strength and direction of monotonic relationships between probability and neural activity was assessed by Spearman's correlation analysis. The statistical analyses were performed using Prism 8.4.3 (GraphPad). In Figure 1F, G, asterisks represent P values from Bonferroni post-hoc analysis following RM ANOVA (ns = P>0.05, *P<0.05, **P<0.01, ****P<0.0001); main ANOVA results are reported in corresponding Results text.

Two-photon GRIN lens imaging

Image processing, including correcting motion artifacts and drift in the imaging plane, and ROI detection were done using suite2p (Pachitariu et al., 2017). To ensure the same neuron was not resampled across imaging planes, 100 um diameter circular ROI masks of were added to images of each plane in Adobe illustrator, and ROIs that overlapped were compared for SNR, the ROI with the lower value deleted. Mean intensity within each region of interest (ROI) was used to generate F(raw).

Fiber photometry

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

465-nm excited signal was calculated by taking the mean of 6 samples centered on the mid-point of the period when triggers activated 465-light, and similarly for 405-nm.

Normalization

To produce Z-score values for both photometry and GRIN lens data, raw fluorescence values were cut into peri-event windows 8.7 s (87 samples) either side of the stimulus onset, and the mean and standard deviation (SD) of the 2s pre-stimulus baseline was calculated. Each sample was then Z-scored using the following calculation [100*((465 sample - mean baseline)/SD)] as in (Duffet et al., 2022). These traces were then filtered using a Savitsky-Golay filter and the median of trials were used to generate traces for each probability.

K-means clustering

The max of the stimulus window was taken for each cell, resulting in a matrix with one row for each cell, and one column for each probability. Each row was then Z-scored using its own mean and STD. This matrix, excluding values for P_{REW} of 0.25 and 0.75, was then passed through Matlab k-means clustering kmeans_opt program, with the percentage variance set to 0.95, maximum number of clusters 10, number of repeats to 500, which yielded 7 clusters.

Threshold-based clustering

In Figure S2, the max of the stimulus window was taken for each cell, resulting a matrix with one row for each cell, and one column for each probability. The range of each cell was then taken, and a threshold value set at 30% of this range. Different cell types were then determined as follows: Uncertainty cell: P_{REW} 0.5 value – threshold value > than P_{REW} 1 and 0 values. Probability cell: P_{REW} 1 value – threshold value > 0.5 values & P_{REW} 0.5 value > 0 value. Mixed cell: P_{REW} 1 & P_{REW} 0.5 value > p 0 + threshold value, & neither P_{REW} 0.5 of p 1 above each other, plus threshold.

Unexpected reward response values

In Figure 4, for each mouse, the mean of traces recorded during unexpected reward trials was subtracted from the mean traces of unrewarded trials (separate rewarded and unrewarded responses are shown in Figure S4). The mean value of the last 5 samples (500 ms) of the stimulus was then subtracted from these traces to ensure differences in reward peaks were not simply caused by starting from different baselines due to stimulus response. The max of these traces from reward onset were then taken, and the mean and standard deviation of the 2s pre-stimulus baseline was calculated to determine unexpected reward values.

ACKNOWLEDGEMENTS

This work was supported by ETH Zürich. We would like to thank Mathilde Guillaumin and Alexander Tesmer for help with statistical analysis, Paulius Viskaitis for help with initial setting up of experiments, and Cristina Concetti for help with some of the surgeries.

AUTHOR CONTRIBUTIONS

DB conceived the study, EB and DB conceptualized and designed the experiments. EB performed the experiments and analyzed data. NG contributed to experiments and analysis. All authors contributed to data interpretation and the writing of the manuscript.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

REFERENCES

Adamantidis, A.R., Zhang, F., Aravanis, A.M., Deisseroth, K., and de Lecea, L. (2007). Neural substrates of awakening probed with optogenetic control of hypocretin neurons. Nature *450*, 420-424.

Baimel, C., Lau, B.K., Qiao, M., and Borgland, S.L. (2017). Projection-Target-Defined Effects of Orexin and Dynorphin on VTA Dopamine Neurons. Cell reports *18*, 1346-1355. 10.1016/j.celrep.2017.01.030.

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. Trends Cogn Sci *11*, 280-289. 10.1016/j.tics.2007.05.005.

Blomeley, C., Garau, C., and Burdakov, D. (2018). Accumbal D2 cells orchestrate innate risk-avoidance according to orexin signals. Nature neuroscience *21*, 29-32. 10.1038/s41593-017-0023-y.

Boutrel, B., Kenny, P.J., Specio, S.E., Martin-Fardon, R., Markou, A., Koob, G.F., and de Lecea, L. (2005). Role for hypocretin in mediating stress-induced reinstatement of cocaine-seeking behavior. Proceedings of the National Academy of Sciences of the United States of America *102*, 19168-19173.

Bracey, E.F., and Burdakov, D. (2020). Fast sensory representations in the lateral hypothalamus and their roles in brain function. Physiol Behav *222*, 112952. 10.1016/j.physbeh.2020.112952.

Burdakov, D. (2005). Physiological Changes in Glucose Differentially Modulate the Excitability of Hypothalamic Melanin-Concentrating Hormone and Orexin Neurons In Situ. Journal of Neuroscience *25*, 2429-2433. 10.1523/JNEUROSCI.4925-04.2005.

Burdakov, D., and Alexopoulos, H. (2005). Metabolic state signalling through central hypocretin/orexin neurons. J Cell Mol Med *9*, 795-803.

Burdakov, D., and Peleg-Raibstein, D. (2020). The hypothalamus as a primary coordinator of memory updating. Physiol Behav *223*, 112988. 10.1016/j.physbeh.2020.112988.

Chemelli, R.M., Willie, J.T., Sinton, C.M., Elmquist, J.K., Scammell, T., Lee, C., Richardson, J.A., Williams, S.C., Xiong, Y., Kisanuki, Y., et al. (1999). Narcolepsy in orexin knockout mice: molecular genetics of sleep regulation. Cell *98*, 437-451.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav Brain Sci *36*, 181-204. 10.1017/S0140525X12000477.

de Lecea, L., Jones, B.E., Boutrel, B., Borgland, S.L., Nishino, S., Bubser, M., and DiLeone, R. (2006). Addiction and arousal: alternative roles of hypothalamic peptides. J Neurosci *26*, 10372-10375.

de Lecea, L., Kilduff, T.S., Peyron, C., Gao, X., Foye, P.E., Danielson, P.E., Fukuhara, C., Battenberg, E.L., Gautvik, V.T., Bartlett, F.S., 2nd, et al. (1998). The hypocretins: hypothalamus-specific peptides with neuroexcitatory activity. Proceedings of the National Academy of Sciences of the United States of America *95*, 322-327.

Dreher, J.C., Kohn, P., and Berman, K.F. (2006). Neural coding of distinct statistical properties of reward information in humans. Cereb Cortex *16*, 561-573. 10.1093/cercor/bhj004.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Duffet, L., Kosar, S., Panniello, M., Viberti, B., Bracey, E., Zych, A.D., Radoux-Mergault, A., Zhou, X., Dernic, J., Ravotto, L., et al. (2022). A genetically encoded sensor for in vivo imaging of orexin neuropeptides. Nat Methods *19*, 231-241. 10.1038/s41592-021-01390-2.

Engel, A.K., Fries, P., and Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. Nature reviews. Neuroscience *2*, 704-716. 10.1038/35094565.

Fiorillo, C.D., Tobler, P.N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. Science *299*, 1898-1902. 10.1126/science.1077349.

Gonzalez, J.A., Iordanidou, P., Strom, M., Adamantidis, A., and Burdakov, D. (2016a). Awake dynamics and brain-wide direct inputs of hypothalamic MCH and orexin networks. Nature communications *7*, 11395. 10.1038/ncomms11395.

Gonzalez, J.A., Jensen, L.T., Iordanidou, P., Strom, M., Fugger, L., and Burdakov, D. (2016b). Inhibitory Interplay between Orexin Neurons and Eating. Current Biology *26*, 2486-2491. 10.1016/j.cub.2016.07.013.

Hara, J., Beuckmann, C.T., Nambu, T., Willie, J.T., Chemelli, R.M., Sinton, C.M., Sugiyama, F., Yagami, K.I., Goto, K., Yanagisawa, M., and Sakurai, T. (2001). Genetic ablation of orexin neurons in mice results in narcolepsy, hypophagia, and obesity. Neuron *30*, 345-354. 10.1016/S0896-6273(01)00293-8.

Harris, G.C., and Aston-Jones, G. (2006). Arousal and reward: a dichotomy in orexin function. Trends in neurosciences *29*, 571-577.

Harris, G.C., Wimmer, M., and Aston-Jones, G. (2005). A role for lateral hypothalamic orexin neurons in reward seeking. Nature *437*, 556-559.

Harris, G.C., Wimmer, M., Randall-Thompson, J.F., and Aston-Jones, G. (2007). Lateral hypothalamic orexin neurons are critically involved in learning to associate an environment with morphine reward. Behavioural brain research *183*, 43-51.

Hassani, O.K., Krause, M.R., Mainville, L., Cordova, C.A., and Jones, B.E. (2016). Orexin Neurons Respond Differentially to Auditory Cues Associated with Appetitive versus Aversive Outcomes. J Neurosci *36*, 1747-1757. 10.1523/JNEUROSCI.3903-15.2016.

Johnson, P.L., Truitt, W., Fitz, S.D., Minick, P.E., Dietrich, A., Sanghani, S., Traskman-Bendz, L., Goddard, A.W., Brundin, L., and Shekhar, A. (2010). A key role for orexin in panic anxiety. Nature medicine *16*, 111-115. 10.1038/nm.2075.

Karnani, M.M., Schöne, C., Bracey, E.F., González, J.A., Viskaitis, P., Li, H.T., Adamantidis, A., and Burdakov, D. (2020). Role of spontaneous and sensory orexin network dynamics in rapid locomotion initiation. Progress in Neurobiology *187*, 101771. 10.1016/j.pneurobio.2020.101771.

Keller, G.B., and Mrsic-Flogel, T.D. (2018). Predictive Processing: A Canonical Cortical Computation. Neuron *100*, 424-435. 10.1016/j.neuron.2018.10.003.

Kim, C.K., Yang, S.J., Pichamoorthy, N., Young, N.P., Kauvar, I., Jennings, J.H., Lerner, T.N., Berndt, A., Lee, S.Y., Ramakrishnan, C., et al. (2016). Simultaneous fast measurement of circuit dynamics at multiple sites across the mammalian brain. Nat Methods *13*, 325-328. 10.1038/nmeth.3770.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Kosse, C., and Burdakov, D. (2019). Natural hypothalamic circuit dynamics underlying object memorization. Nature communications *10*, 2505. 10.1038/s41467-019-10484-7.

Li, Y., Gao, X.B., Sakurai, T., and van den Pol, A.N. (2002). Hypocretin/Orexin excites hypocretin neurons via a local glutamate neuron-A potential mechanism for orchestrating the hypothalamic arousal system. Neuron *36*, 1169-1181.

Lin, L., Faraco, J., Li, R., Kadotani, H., Rogers, W., Lin, X., Qiu, X., de Jong, P.J., Nishino, S., and Mignot, E. (1999). The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. Cell *98*, 365-376.

Livneh, Y., and Andermann, M.L. (2021). Cellular activity in insular cortex across seconds to hours: Sensations and predictions of bodily states. Neuron *109*, 3576-3593. 10.1016/j.neuron.2021.08.036.

Ma, W.J., and Jazayeri, M. (2014). Neural coding of uncertainty and probability. Annu Rev Neurosci *37*, 205-220. 10.1146/annurev-neuro-071013-014017.

Mileykovskiy, B.Y., Kiyashchenko, L.I., and Siegel, J.M. (2005). Behavioral correlates of activity in identified hypocretin/orexin neurons. Neuron *46*, 787-798. 10.1016/j.neuron.2005.04.035.

Nassar, M.R., Rumsey, K.M., Wilson, R.C., Parikh, K., Heasly, B., and Gold, J.I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. Nature neuroscience *15*, 1040-1046. 10.1038/nn.3130.

O'Doherty, J.P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. Current opinion in neurobiology *14*, 769-776. 10.1016/j.conb.2004.10.016.

Pachitariu, M., Stringer, C., Dipoppa, M., Schröder, S., Rossi, L., Dalgleish, H., Carandini, M., and Harris, K.D. (2017). Suite2p: beyond 10,000 neurons with standard two-photon microscopy. https://www.biorxiv.org/content/10.1101/061507v2.

Peyron, C., Tighe, D.K., van den Pol, A.N., de Lecea, L., Heller, H.C., Sutcliffe, J.G., and Kilduff, T.S. (1998). Neurons containing hypocretin (orexin) project to multiple neuronal systems. J Neurosci *18*, 9996-10015.

Rao, R.P., and Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nature neuroscience *2*, 79-87. 10.1038/4580.

Sakurai, T. (2014). The role of orexin in motivated behaviours. Nature reviews. Neuroscience *15*, 719-731. 10.1038/nrn3837.

Sakurai, T., Amemiya, A., Ishii, M., Matsuzaki, I., Chemelli, R.M., Tanaka, H., Williams, S.C., Richardson, J.A., Kozlowski, G.P., Wilson, S., et al. (1998). Orexins and orexin receptors: A family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. Cell *92*, 573-585. 10.1016/S0092-8674(00)80949-6.

Sakurai, T., Nagata, R., Yamanaka, A., Kawamura, H., Tsujino, N., Muraki, Y., Kageyama, H., Kunita, S., Takahashi, S., Goto, K., et al. (2005). Input of orexin/hypocretin neurons revealed by a genetically encoded tracer in mice. Neuron *46*, 297-308. 10.1016/j.neuron.2005.03.010.

Schultz, W. (2000). Multiple reward signals in the brain. Nature Reviews Neuroscience.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Schultz, W. (2002). Getting formal with dopamine and reward. Neuron *36*, 241-263. 10.1016/s0896-6273(02)00967-4.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. Science *275*, 1593-1599.

Schultz, W., and Dickinson, A. (2000). Neuronal coding of prediction errors. Annu Rev Neurosci *23*, 473-500. 10.1146/annurev.neuro.23.1.473.

Thomas, C.S., Mohammadkhani, A., Rana, M., Qiao, M., Baimel, C., and Borgland, S.L. (2022). Optogenetic stimulation of lateral hypothalamic orexin/dynorphin inputs in the ventral tegmental area potentiates mesolimbic dopamine neurotransmission and promotes reward-seeking behaviours. Neuropsychopharmacology *47*, 728-740. 10.1038/s41386-021-01196-y.

Tsunematsu, T., Kilduff, T.S., Boyden, E.S., Takahashi, S., Tominaga, M., and Yamanaka, A. (2011). Acute optogenetic silencing of orexin/hypocretin neurons induces slow-wave sleep in mice. J Neurosci *31*, 10529-10539. 10.1523/JNEUROSCI.0784-11.2011.

Volkow, N.D., Wang, G.J., Fowler, J.S., Tomasi, D., Telang, F., and Baler, R. (2010). Addiction: decreased reward sensitivity and increased expectation sensitivity conspire to overwhelm the brain's control circuit. Bioessays *32*, 748-755. 10.1002/bies.201000042.

Wolpert, D.M., and Ghahramani, Z. (2000). Computational principles of movement neuroscience. Nature neuroscience *3 Suppl*, 1212-1217. 10.1038/81497.

Yamanaka, A., Beuckmann, C.T., Willie, J.T., Hara, J., Tsujino, N., Mieda, M., Tominaga, M., ichi Yagami, K., Sugiyama, F., Goto, K., et al. (2003a). Hypothalamic orexin neurons regulate arousal according to energy balance in mice. Neuron *38*, 701-713.

Yamanaka, A., Muraki, Y., Tsujino, N., Goto, K., and Sakurai, T. (2003b). Regulation of orexin neurons by the monoaminergic and cholinergic systems. Biochem Biophys Res Commun *303*, 120-129.

Zimmerman, C.A., Lin, Y.C., Leib, D.E., Guo, L., Huey, E.L., Daly, G.E., Chen, Y., and Knight, Z.A. (2016). Thirst neurons anticipate the homeostatic consequences of eating and drinking. Nature *537*, 680-684. 10.1038/nature18950.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

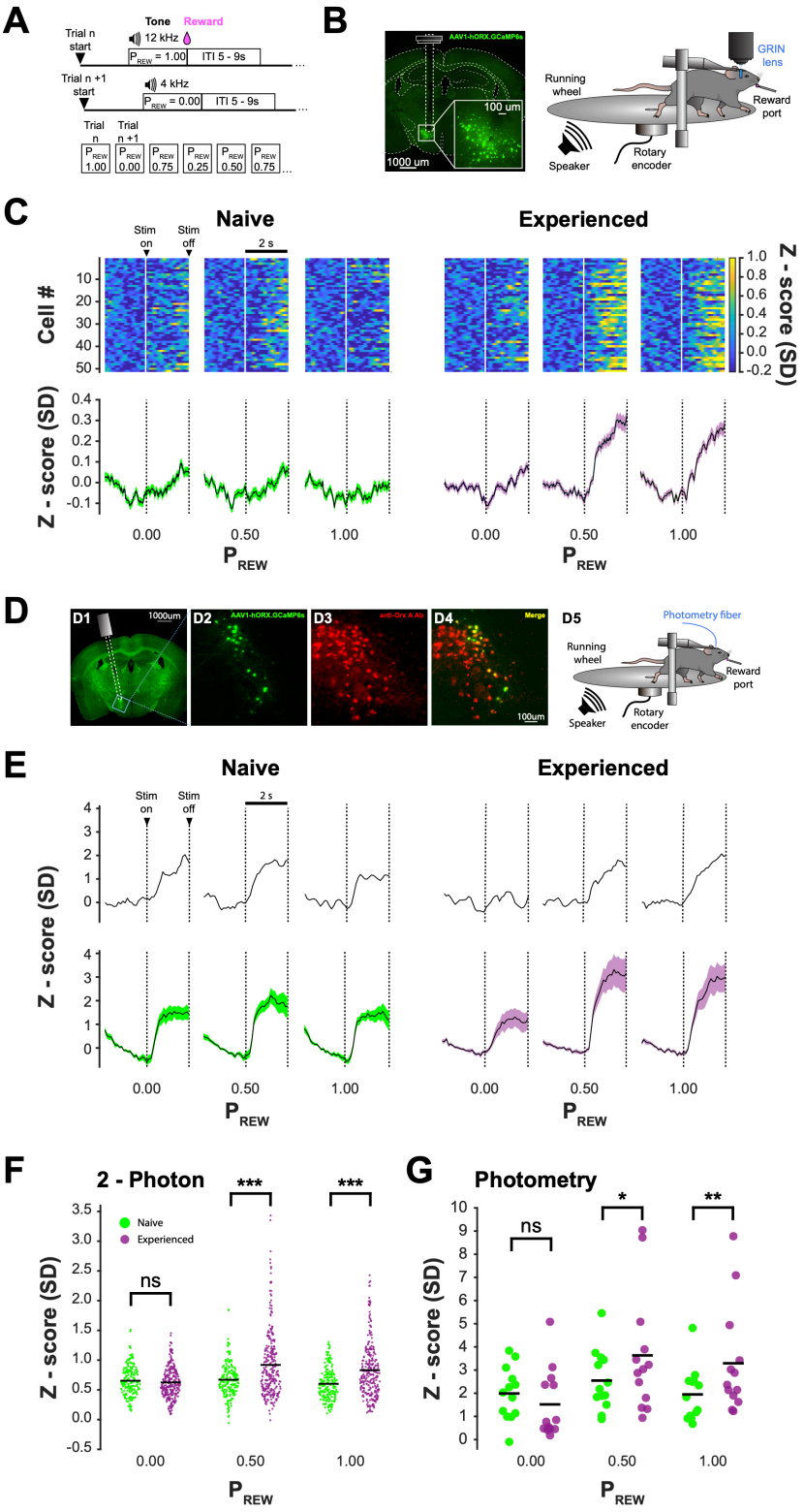


Figure 1. Experience alters the influence of reward probability on hypocretin/orexin neurons

A. Pavlovian associative-conditioning task trial-structure. Top, example probability of reward (P_{rew}) 1 and P_{rew} 0 trial structure. Tone offest occured simultaneously with reward onset. Bottom, trials of different reward probabilities were interleaved randomly.

B. Left, coronal section showing HON-GCaMP expression and GRIN lens position. Inset, magnified localisation of HON-GCaMP6s expression in LH. Right, behavioural and 2-photon imaging apparatus.

C. Top, example Z-scored fluorescence heatmaps of 50 cells recorded with 2-photon imaging from one animal, during stimulus presentation, across three reward probabilities. Left, early (“naïve”, median of trials 1 - 30, green) and right, late (“experienced”, median of trials 121 - 150, purple) associative conditioning trials.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Bottom, population Z-scored fluorescence traces for naïve (n = 197 cells, n = 4 mice), and experienced (n = 278 cell, n = 4 mice), expressed as mean +/- SEM.

D. D1, Coronal brain section showing position of photometric optical fiber and HON-GCaMP expression in LH. **D2 - 4** Histological confirmation of virus specificity showing h-orx GCaMP6s expression (**D2**), goat anti-orx A antibody with donkey anti-goat cy3-conjugated antibody (**D3**), and (**D4**), overlay of (D2) and (D3). **D5**, photometric and behavioural apparatus.

E. Top, example photometric HON-GCaMP Z-scored fluorescence trace from one animal during stimulus presentation across three reward probabilities, in naïve (left, median of trials 1 - 30) and experienced (right, median of trials 121 - 150) associative-conditioning trials. Bottom, population Z-scored fluorescence traces during stimulus presentation for 12 mice during naïve (left, green) and experienced (right, purple) trials, expressed as mean +/- SEM.

F. Population values of Z-scored fluorescence for cells recorded with 2-photon imaging. Values are max of stimulus window from cells recorded during naïve (green, n = 197 cells) and experienced trials (purple n = 278 cells), across three reward probabilities, n = 4 mice. Black horizontal bar indicates mean of cells.

G. As in (F), for max photometric Z-scored fluoresence during stimulus window, n = 12 mice.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

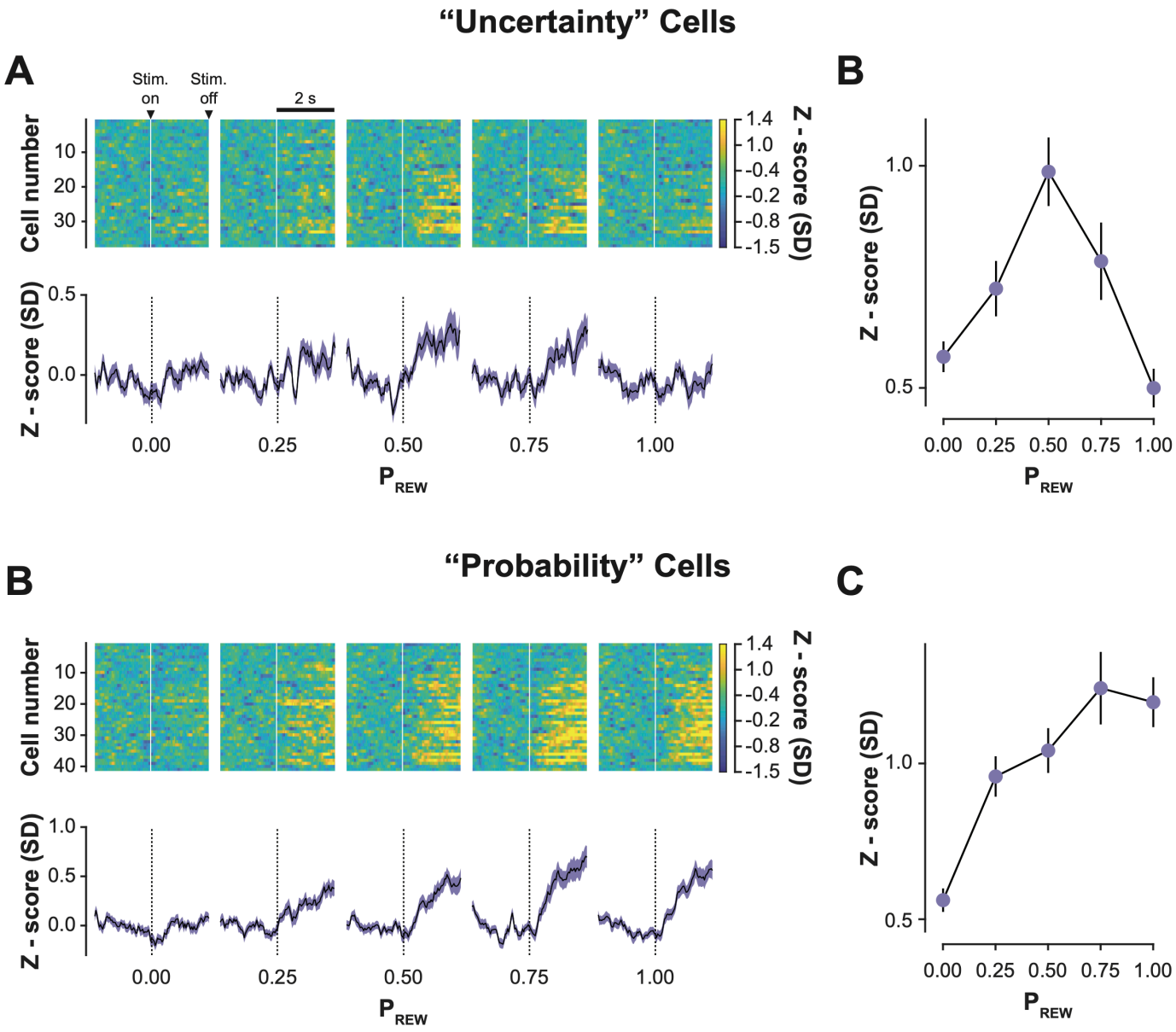


Figure 2. Coding of reward probability and uncertainty in subsets of hypocretin/orexin neurons.

A. Top, heatmaps of Z-scored fluorescence changes recorded with 2-photon imaging, during the stimulus period, across reward probabilities, for cells that displayed strongest responses to P_{rew} 0.5, categorised using unsupervised k-means clustering. Bottom, population traces of uncertainty cells, expressed as mean +/- SEM.

B. Population responses of max Z-scored fluorescence values of uncertainty cells during stimulus period, shown across probabilities, expressed as mean +/- SEM.

C. As in (A) for cells reward-probability tuning cluster.

D. As in (B), for cells with reward-probability tuning cluster.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

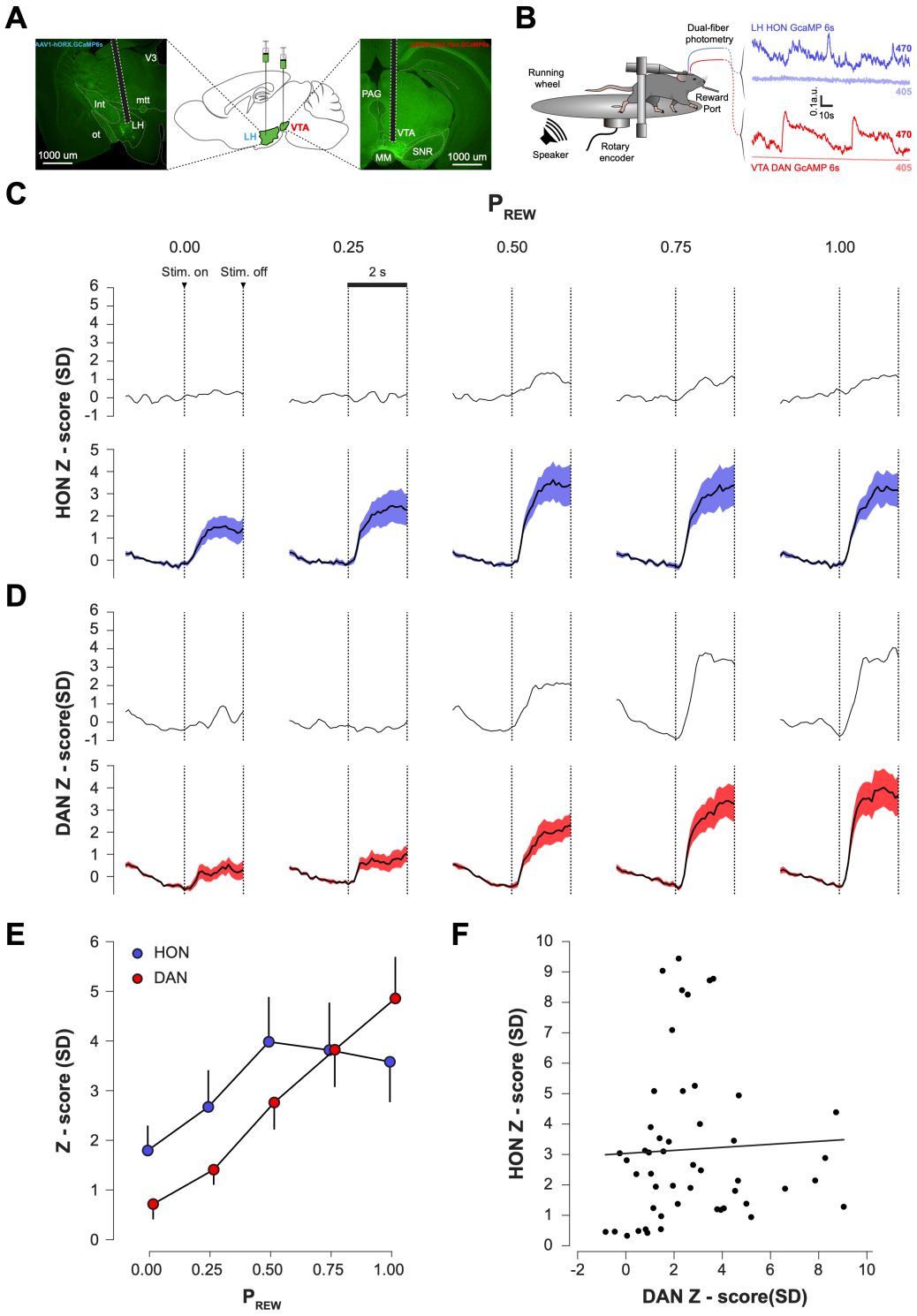


Figure 3. Reward coding in hypocretin/orexin vs dopamine neurons: responses to conditioned stimuli

A. Left, coronal brain section showing position of optical fiber and orx-GCAMP expression in the LH. Middle, sagittal section showing viral injection sites of orx-GCaMP in LH (blue), and flex-GCaMP (red) in the VTA. Right, coronal brain section showing position of optical fiber and flex-GCAMP expression in dopaminergic neurons (DANs) of the VTA. *Int: internal capsule, mtt: mammillothalamic tract, ot: optic tract, V3: third ventricle, MM: medial mamillary nucleus, PAG: peri-aqueductal gray, SNR: substantia nigra pars reticularis.*

B. Left, behavioural and photometric imaging apparatus. Right, LH HON and VTA DAN raw fluorescence traces recorded simultaneously in a mouse before it had been exposed to any tone–reward pairings.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

C. Top, example LH HON-GCaMP fluorescence trace from one animal during stimulus presentation of five reward probabilites, median of “experienced” trials 121-130. Bottom, population fluorescence traces, median of trials 121 - 130, n = 10 mice, expressed as mean +/- SEM.

D. As in (C), for VTA DAN-GCaMP fluorescence, recorded concurrently with LH HON-GCaMP.

E. Population fluorescence values of max of stimulus window across five reward probabilities, for HON-GCaMP (blue dots) and DAN (red dots), expressed as mean +/- SEM.

F. Spearman’s correlation of HON-GCaMP max Z-scored fluorescence stimulus values vs VTA DAN-GCaMP values pooled across all reward probabilities. $r^2 = 0.0021158$, $p = 0.7511$.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

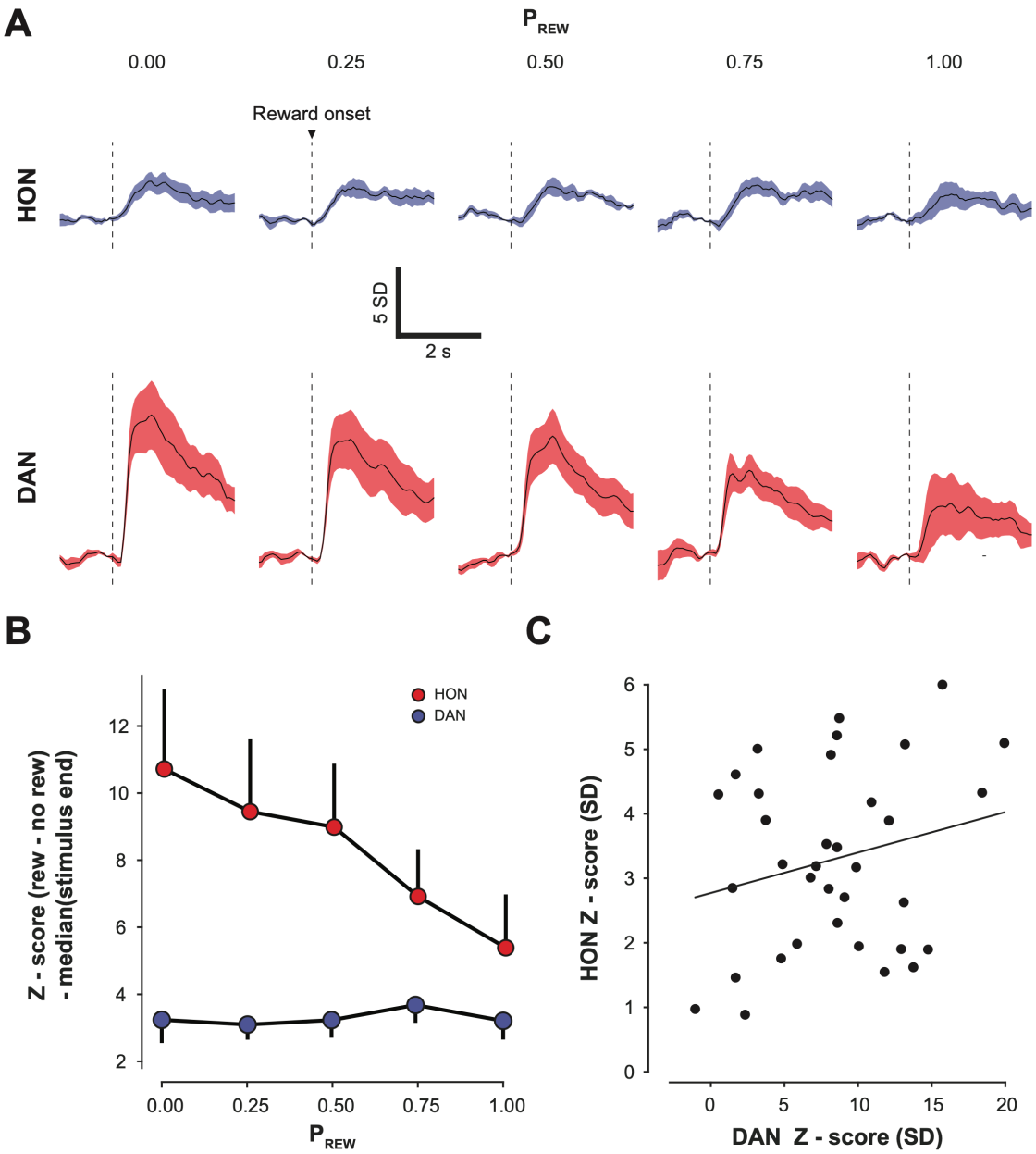


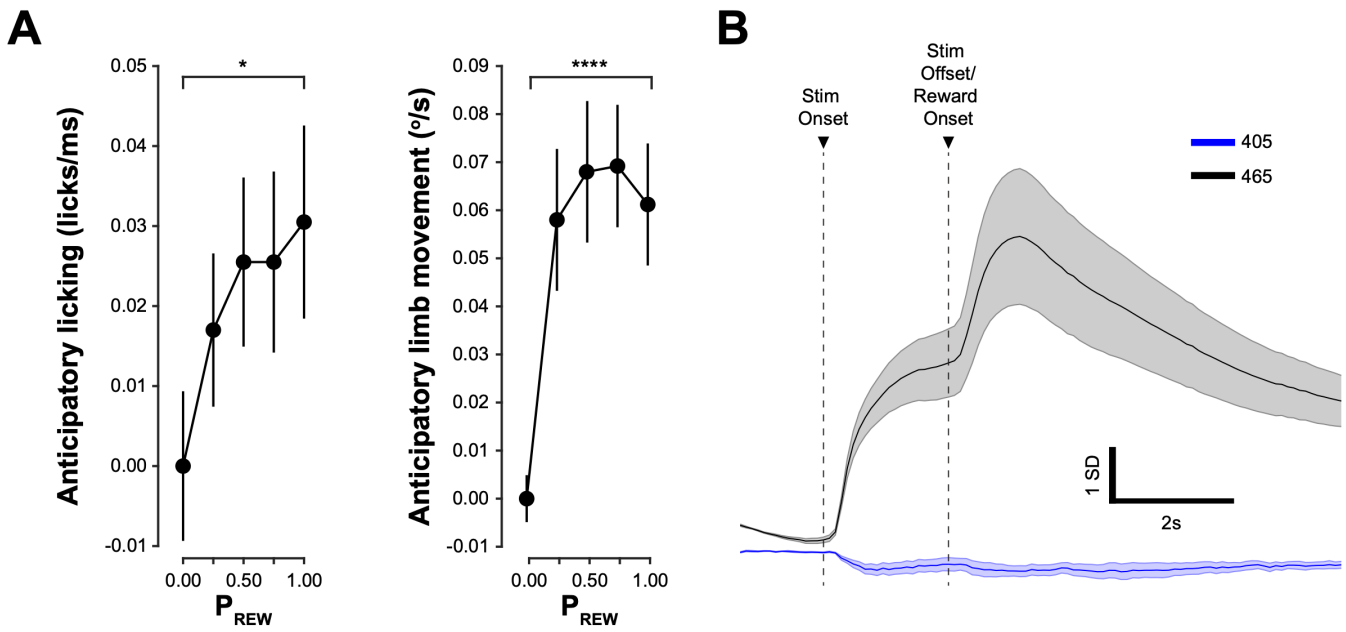
Figure 4. Reward coding in hypocretin/orexin vs dopamine neurons: population responses to rewards

A. Top, population Z-scored fluorescence traces of LH HON-GCaMP during reward presentations, across five reward probabilities. Non-rewarded traces were subtracted from rewarded traces, and then shifted down to normalize for differences created during stimulus response period, n = 10 trials, n = 7 mice, expressed as mean +/- SEM. Bottom, as in (A) for VTA DAN-GCaMP responses during reward presentations.

B. Max Z-scored fluorescence values during stimulus-normalized reward-period, across 5 reward probabilities.

C. Spearman's correlation of LH HON-GCaMP max fluorescence reward-response values vs VTA DAN-GCaMP values for all reward probabilities. $R^2 = 0.053024$, $p = 0.18326$.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

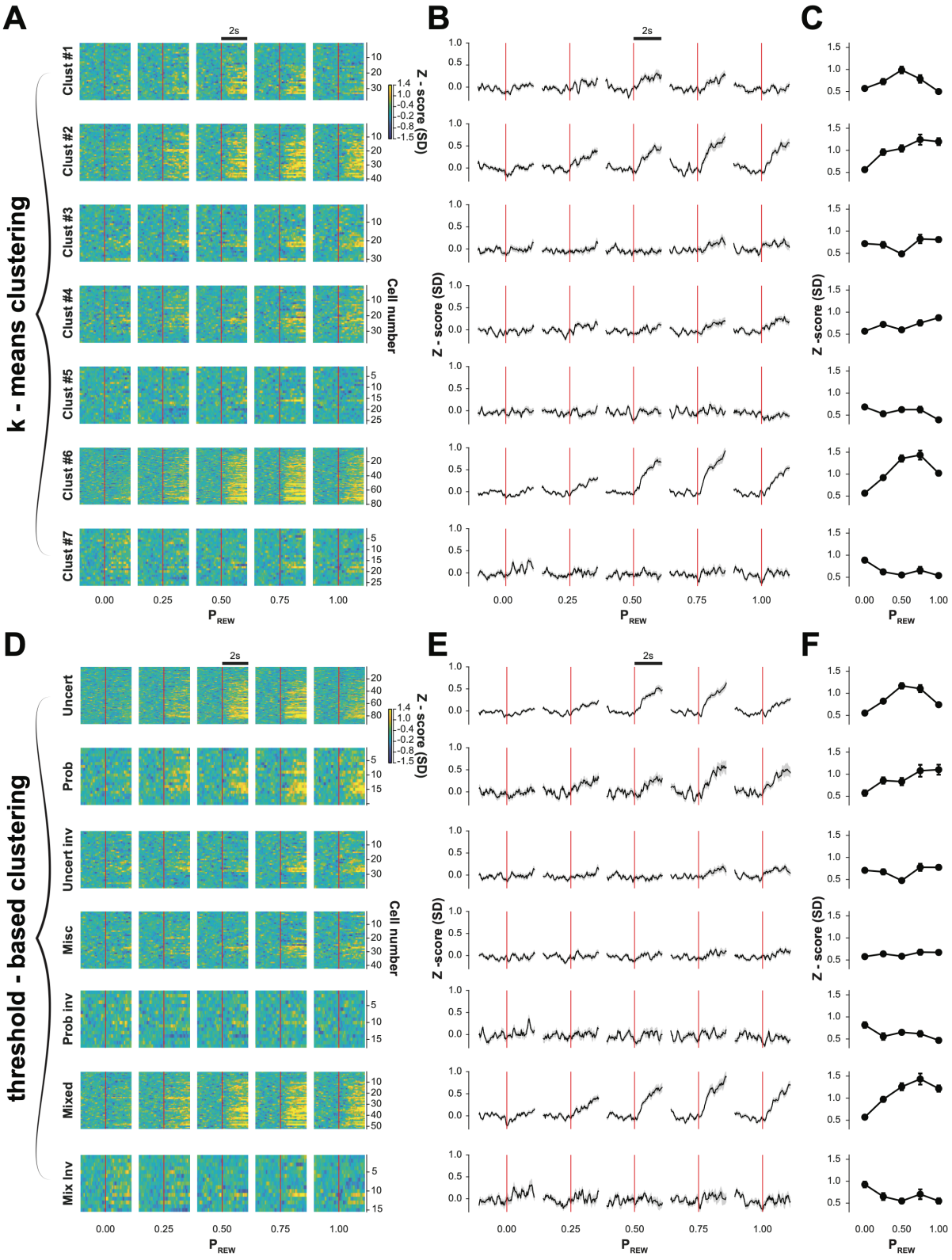


Supplementary Figure 1. Anticipatory movements during cue, and movement artefact controls in HON-GCaMP fluorescence.

A. Example anticipatory licking (left), and limb movements (right) averaged across 50 trials and five probabilities during stimulus presentation. Rank sum of P_{REW} 0 vs P_{REW} 1 is shown, *, $p < 0.02$, ****, $p < 0.0001$. All mice had to demonstrate a significant difference between P_{REW} 0 and P_{REW} 1 to be included in the study.

B. Mean Z-scored photometric fluorescence HON-GCaMP values during stimulus and reward presentation, averaged across all trials and all reward probabilities, n = 12 mice. Fluorescence recorded at 465nm (black) and 405nm (blue) shows calcium-dependent, and movement-related changes respectively.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Supplementary Figure 2. Reward probability prediction encoding in subsets of hypocretin/orexin neurons.

A. Heatmaps of fluorescence changes recorded using 2-photon imaging lens, during the stimulus period, across reward probabilities, for clusters of cells categorised using unsupervised k-means clustering.
Each row represents a different cluster. Vertical red lines show stimulus onset, duration of stimulus onset to end of traces is 2s.

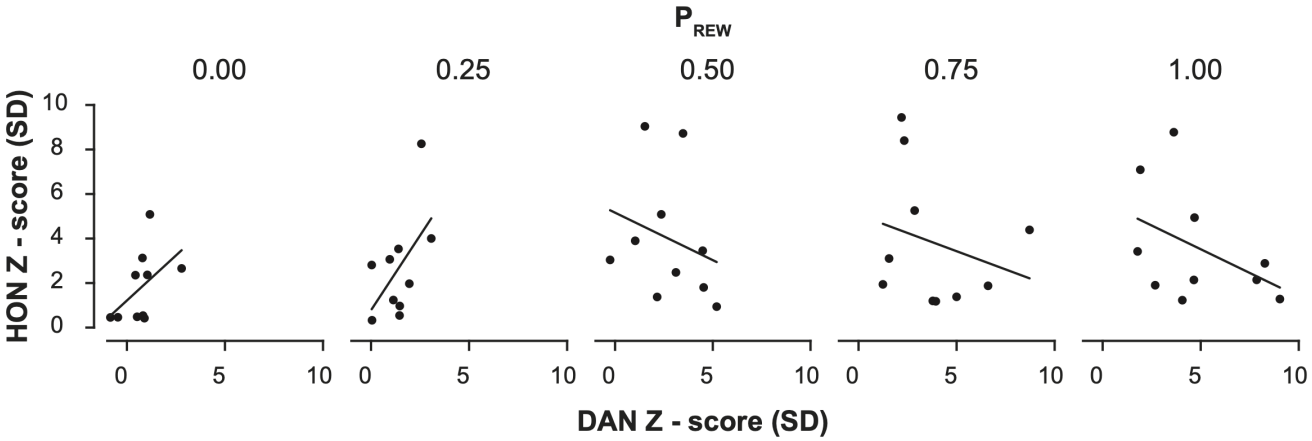
B. Population traces for results shown in A, expressed as mean +/- SEM.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

C. Population responses of max fluorescence values for different cell clusters during stimulus period, shown across probabilities, expressed as mean +/- SEM.
42% of neurons (117/278 cells, corresponding to clusters 1 and 6) displayed, as a population, significant coding of uncertainty (Kruskal-Wallis H(2)=228.3, P<0.0001) but did not monotonically follow probability (Spearman's correlation: Rs = 0.3000, P=0.6833). Proportion of cells: Cluster #1: 13.31%, 37/278 cells, cluster #2: 14.75%, 41/278 cells, cluster #3: 11.15%, 31/278 cells, cluster #4: 13.31%, 37/278 cells, cluster #5: 9.35%, 26/278 cells, cluster #6: 28.78%, 80/278 cells, cluster #7: 9.35%, 26/278 cells.

D, **E** and **F**, as in (A, B and C), for clusters determine using thresholding (see Methods). Proportion of cells: Uncert = uncertainty cells (33.45%, 93/278 cells), Prob = probability cells (6.12%, 17/278), Uncert inv = inverted uncertainty cells (14.03%, 39/278), Misc = miscellaneous cells (15.11%, 42/278), Prob inv = inverted probability cells, Mixed = mixed cells (18.71, 52/278), Mix inv = inverted mixed cells (5.40%, 15/278).

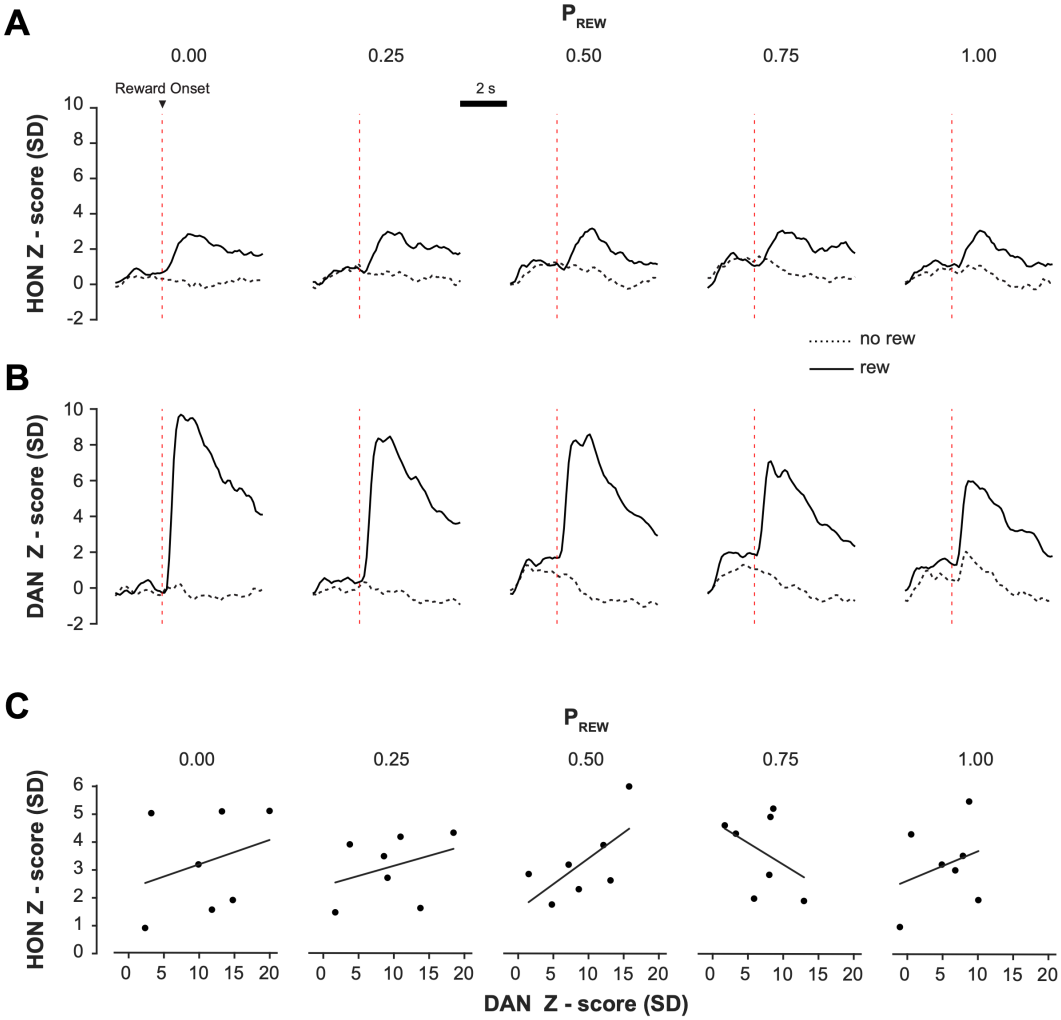
bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Supplementary Figure 3

Spearman's correlation of HON-GCaMP max fluorescence stimulus-response values vs DAN-GCaMP values, separated by reward probability. R2 and p values for each comparison indicated above each panel. Each point represents a mouse. P_{REW} 0; R^2 =0.24764, p = 0.14331, P_{REW} 0.25; R^2 =0.30783, p = 0.095979, P_{REW} 0.5; R^2 =0.066397, p = 0.47229, P_{REW} 0.75; R^2 =0.0656944, p = 0.47386, P_{REW} 1; R^2 =0.19296, p = 0.20403.

bioRxiv preprint doi: https://doi.org/10.1101/2022.04.13.488195; this version posted April 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Supplementary Figure 4.

A. Population Z-scored fluorescence traces of LH HON-GCaMP during reward presentations, across five reward probabilities, n = 10 trials, n = 7 mice, expressed as mean. Rewarded (filled lines), and unrewarded traces (dotted lines) are shown.

B. As in (A) for DAN-GCaMP fluorescence.

C. Spearman's correlation of LH HON-GCaMP max fluorescence stimulus-response values vs VTA DAN-GCaMP values separated by reward probability. Each point represents a mouse. P_{REW} 0; R^2 =0.09129, p = 0.51016, P_{REW} 0.25; R^2 =0.12141, p = 0.44372, P_{REW} 0.5; R^2 =0.44133, p = 0.10359, P_{REW} 0.75; R^2 =0.16861, p = 0.36016, P_{REW} 1; R^2 =0.09172, p = 0.50913