Cognitive, Affective, & Behavioral Neuroscience (2024) 24:292–301
https://doi.org/10.3758/s13415-023-01134-z

SPECIAL ISSUE/PRECLINICAL ASSAYS

Depth recordings of the mouse homologue of the Reward Positivity

Penelope Kehrer[1,2] · Jonathan L. Brigman[2] · James F. Cavanagh[1]

Accepted: 28 September 2023 / Published online: 18 October 2023

Abstract
We recently advanced a rodent homologue for the reward-specific, event-related potential component observed in humans known as the Reward Positivity. We sought to determine the cortical source of this signal in mice to further test the nature of this homology. While similar reward-related cortical signals have been identified in rats, these recordings were all performed in cingulate gyrus. Given the value-dependent nature of this event, we hypothesized that more ventral prelimbic and infralimbic areas also contribute important variance to this signal. Depth probes assessed local field activity in 29 mice (15 males) while they completed multiple sessions of a probabilistic reinforcement learning task. Using a priori regions of interest, we demonstrated that the depth of recording in the cortical midline significantly correlated with the size of reward-evoked delta band spectral activity as well as the single trial correlation between delta power and reward prediction error. These findings provide important verification of the validity of this translational biomarker of reward responsiveness, learning, and valuation.

Keywords Prediction error · Mouse · Reward Positivity · Infralimbic · Prelimbic

Functional comparisons between species are challenging. The *status quo* focus on similar behavioral tendencies between species is limited by the degenerative nature of action selection and the intrinsic weakness of applying psychological labels to complex cognitive functions (Keeler & Robbins, 2011; Sarter, 2004). Systems-level neural activities may offer a more sensitive measure of translational similarity beyond behavioral tendencies. Population-level electrophysiology is particularly well-suited for addressing such questions about translatable biomarkers. Task-evoked electrophysiological responses offer the opportunity to identify similar spectral fingerprints between species, which may reflect common latent computations. Probabilistic learning tasks are easily transferrable across vertebrates (Amitai et al., 2014; Amodeo et al., 2012; Bari et al., 2010; Hyman et al., 2017), and reinforcement learning theory allows a quantification of abstract latent processes (Sutton & Barto, 1998). Taken together, electrophysiological assessment during reinforcement learning offers an excellent opportunity for comparing neural activities between species.

✉ James F. Cavanagh
jcavanagh@unm.edu

[1] Psychology Department, University of New Mexico, Logan Hall, MSC03 2220, 87131 Albuquerque, NM, Mexico

[2] Department of Neurosciences, University of New Mexico School of Medicine, Albuquerque, NM, Mexico

We recently advanced a rodent homologue for a reward-specific, event-related potential (ERP) component observed in humans. The Reward Positivity (RewP) is a positive deflection in the human ERP that is most commonly quantified over frontocentral sites approximately 200–400 ms after reward presentation (Baker & Holroyd, 2011; Holroyd et al., 2008; Proudfit, 2015). Both the RewP and its delta-band spectral reflection scale with the degree of positive reward prediction error, whereby "better than expected" outcomes evoke increasingly larger RewP amplitudes (Baker & Holroyd, 2011; Cavanagh, 2015; Holroyd & Umemoto, 2016). The candidate rodent homologue shares a similar sensitivity to positive reward prediction errors (Cavanagh et al., 2021), and both species' signals are parametrically modulated by amphetamine (Cavanagh et al., 2022). This study was designed to determine the cortical source of this signal in mice to further test the nature of this homology.

Previous investigations of this reward-related local field feature only examined EEG from a frontal dura screw, leaving questions of the generative neural systems unaddressed. Other groups have advanced similar reward-related cortical signals in rats (Iturra-Mena et al., 2023; Warren et al., 2015), identifying cingulate gyrus (CG2) as a generative structure. However, other potential generative structures were not assessed. The source of the human RewP has not been rigorously tested either, but it also is often assumed to be generated in the dorsal cingulate (Walsh & Anderson,

2012). Importantly, we think that both of these assumptions are incomplete. As a value-related signal, we hypothesize that more ventral midfrontal regions may have a major generative role in this signal (Roy et al., 2012). Most recently, it has been suggested that perigenual and ventral midline cortices contribute meaningful variance to the RewP (Cavanagh et al., 2018; Crane et al., 2022; Whitton et al., 2023), although conclusive evidence is still lacking. There is considerable evidence for anatomical homology across midline structures between rodents and primates, yet little evidence for functional homologies (Bicks et al., 2015; Heilbronner et al., 2016; Laubach et al., 2018; Preuss & Wise, 2022; Schaeffer et al., 2020; van Heukelum et al., 2020). In line with these compelling cross-species similarities, we tested the hypothesis that more ventral midline regions in mice (e.g., prelimibic (PL) and infralimbic (IL) cortex) also are major generators of this signal.

The identification of the generative neural structures contributing to this signal could provide additional evidence on the nature of the computation reflected by the RewP. This inference requires a sophisticated assessment of the topography of function across the midline dorsoventral axis of the rodent. Functional segregations do not appear to confirm to standard dorsal-ventral subregion dissociations (e.g., CG1, CG2, PL, IL); rather a homologous comparison is maximally effective when considered as a caudal-rostral gradient around the genu of the callosum (Francis-Oliveira et al., 2022; Schaeffer et al., 2020; van Heukelum et al., 2020). This gradient may be considered a continuum from dorsal action selection to ventral affective representation (van Heukelum et al., 2020). In both humans and mice, this may be analogous to an actor-critic dissociation; dorsal premotor areas select actions based on context, but ventral areas represent goal value. We hypothesize that the RewP may be a marker of this critic-like goal value. Thus, we expect to find a strong role of ventral generators in mice (as well as in humans, someday). We investigated local field activity within two distinct groups of mice: CG1/CG2 vs. PL/IL. However, we also hypothesized that continuous measures of cortical depth across mice could provide a more sophisticated vectorized account of the topography of the generators of this candidate RewP homologue.

Materials and methods

Animal subjects

Female and male C57BL/6J mice were obtained from The Jackson Laboratory (Bar Harbor, ME), housed in same sex groupings of 2 per cage in a temperature- and humidity-controlled vivarium under a reverse 12-h light/dark cycle (lights off 0800 h) and tested during the dark phase. A total of 39 mice (23 males) were used. All experimental procedures were performed in accordance with the National Institutes of Health Guide for Care and Use of Laboratory Animals and were approved by the University of New Mexico Health Sciences Center Institutional Animal Care and Use Committee.

Behavioral chambers

All operant behavior was conducted in a custom acrylic chamber, measuring 21.6 × 17.8 × 12.7 cm, housed within a sound- and light-attenuating box (Med Associates, St. Albans, VT) as previously described (Marquardt, Sigdel, Caldwell, & Brigman, 2014). At one end of the chamber, a liquid dispenser delivered a strawberry milk liquid reward, which consisted of a mix of Nesquik (S.A., Vevey, Switzerland), Carnation powdered milk (Nestle Baking, a division of Nestle USA, Inc., Solon, OH), and water. There also was a house-light, a tone generator, and an ultrasensitive lever. On the other end of the chamber, there was a touch-sensitive screen (Conclusive Solutions, Sawbridgeworth, UK) covered by a black acrylic aperture plate, allowing two active touch areas measuring 7.5 x 7.5 cm separated by 0.6 cm and located at a height of 1.6 cm from the floor. Stimulus presentation in the response windows and touches were controlled and recorded by the K-Limbic Software Package (Conclusive Solutions, Sawbridgeworth, UK).

Pretraining

The weights of mice were first slowly reduced and then maintained at 85% free-feeding body weight. Before training, mice were acclimated to the reward by providing 3 mL of liquid per mouse on a weigh boat in the home cage for 2 days. After becoming acclimated to the reward, mice were habituated to the operant chamber with a 30-min session where they could retrieve 40 µL of liquid reward after each magazine head dip. Mice that retrieved 10 rewards within 10 min were then moved to bar training. During bar training, mice pressed a bar located to the side of the magazine to get a liquid reward, followed by a 5-second intertrial interval (ITI). The criterion for passing bar training was 30 bar presses in less than 30 min. Following completion of bar training, mice were moved to touch training. Touch training required mice to first initiate each trial with a bar press, then touch the screen in one of the two response windows. Upon touching either side of the screen, mice were rewarded with a 1-s tone and liquid reward. There was a 5-s ITI before they were able to initiate a new trial. The criterion for touch training was again 30 trials in less than 30 min. Upon completion of touch training, mice were moved to punish training. In punish training, the mouse initiated the task by pressing the bar and then touched one of the two response windows. Touching the window with a white stimulus (of various

shapes) resulted in a reward as described above. Touching the blank window resulted in punishment, where the white house light was turned on for 10 s. Mice were not allowed to progress to a new trial until they correctly touched the target screen. The criterion for passing punish training was again 30 trials in less than 30 min. The final stage of training involved simplified version of the Probabilistic Learning Task (PLT) called FAN100. Upon initiation of the task via bar press, a "fan" image was shown in one screen and a "marble" image in the other. In FAN100, the fan image was always the correct stimulus, and the marble image was always the incorrect stimulus. Touching the fan image would result in the above-described reward, whereas touching the marble image would result in a punishment and trigger a correction trial. The criteria for this task was >85% correct choices (26/30 trials) for two consecutive days.

Mice then progressed to the final version of the PLT task. This task is the same as FAN100, but the fan image was only rewarded 80% of the time, and the marble image was rewarded 20% of the time. There also were no correction trials for an incorrect choice at this stage. Four mice used a version with 100/50 probabilities instead of 80/20. The criterion for completion of the presurgical training portion of this task was 70% correct choices (21/30 trials) for two consecutive days.

Surgical implantation and testing

After completing PLT criteria, mice were put on a special diet of 3 g of food for 2 days leading up to implantation. They were then anesthetized with isoflurane and placed in a stereotaxic alignment system (Kopf Instruments, Tujunga, CA) for fitting with a 16-channel array of 35-µm diameter, 5-mm length tungsten electrodes (two lateral clusters separated by 500 µm with 150-µm spacing between electrodes in each cluster; Innovative Neurophysiology Inc.). A 0.6- x 0.6-mm window was drilled into the skull, and the array was lowered to the assigned depth. Two skull screws were secured to the cerebellum, and a ground wire from the array was wrapped around them. The entire area was then solidified with dental cement. The mouse was then given 0.15 ml of buprenorphine postoperatively. After 5 days of recovery, body weight reduction resumed, and mice were given a post-surgery reminder session that consisted of the last pretraining regimen to ensure retention of pretraining criterion.

During each recording session, electrophysiological activity was recorded via a multichannel acquisition processor (PlexControl, Plexon) at a sample rate of 10,000 Hz for spikes and 1,000 Hz for local field potentials (LFPs). Only LFPs are reported here. Relevant task events and behaviors were time-coded as event markers in the recording file via a TTL pulse from the behavioral software. Each recording session consisted of 60 trials. Only sessions with more than 30 completed trials were included in analyses. Mice completed between six and 11 sessions with this criterion.

Computational modeling

In all models, state-action values were estimated for each cue, and a softmax choice function was used to predict the most likely action on each trial. State-action values (Q values) were updated according to the delta learning rule with a learning rate (α) scaling the prediction error (δ):

$$Q_t = Q_{t-1} + \alpha(\delta), \tag{1}$$

where prediction errors were calculated as the difference between reinforcements (r) and Q values:

$$\delta = r - Q, \tag{2}$$

and reinforcements were from a set of 0,1:

$$\mathrm{r} \in (0, 1). \tag{3}$$

The probability of action selection was predicted by using a softmax logistic function with a free parameter for gain adjustment to select the highest value option (β, also termed behavioral consistency or inverse temperature):

$$p\left(Q_{selected}\right) = exp\left(\beta * Q_{selected}\right) / \sum\nolimits_{all} exp\left(\beta * Q_{all}\right). \tag{4}$$

The computationally derived prediction error (Eq. 2) from the best-fitting model was used as a single trial regressor in LFP analyses. However, the estimation of prediction errors may vary depending on dynamic task qualities captured by free parameters. A number of competing models were formally compared and prediction errors from the best-fitting model were used as regressors for EEG analyses. The probabilities of action selection (Eq. 4) were used to compute the log likelihood estimate (LLE) of the subject having chosen that set of responses for a given set of parameters. The parameters that produced the maximum LLE were found by using the Nelder-Mead simplex method, a standard hill-climbing search algorithm (implemented with Matlab function fmincon.m). All models used the best-fitting outcome of ten different starting points (using Matlab function rmsearch.m).

All models included a parameter for softmax weighting. The first model (M1: Vanilla) had one free parameter for learning rate. The second model (M2: WinLoss) included separate learning rates to gain and loss. The third model (M3: WL_Qbias) modulated M2 with a free "omega" parameter that weighted the initial Q values ($Q_a = 0.5 + \Omega$, $Q_b = 0.5 - \Omega$) to account for preserved learning between sessions.

Following convention (Daw, 2011), all learning rates were constrained to remain between 0 and 1, and all softmax gain parameters were constrained to be between 0 and 10.

M3 omega was constrained between −0.5 and 0.5. Characterization of model fits were computed as pseudo-R^2 statistics: (LLE-chance)/chance (Camerer & Ho, 1999). For model comparison, the Akaike information criterion (AIC) was used to penalize the LLE based on the number of parameters. Given the multiple sessions, two strategies for model fitting were compared: 1) fitting individual sessions before within-subject averaging of parameters versus 2) one model fit to all aggregate session data per mouse. Because AIC cannot be compared between these approaches due to different sizes of data, pseudo R^2 was used as an index of fit. Pseudo R^2 was larger for individual session fits instead of one single aggregate fit in nearly all mice and model types (observed in 100% of mice in the best fitting model M3; pseudo-R^2 improvement M = 5.5%, range 3.2–8.5%).

EEG processing

Data were epoched around the time of the screen touch (−2000 ms to 2000 ms), which was synonymous with the indicator for reinforcement (tone for reward, lights for punishment). As in our previous work, we used the cerebellar screw as a reference for the midfrontal LFPs, but we also examined three other reference schemes for posterity: 1) average reference over all 16 leads; 2) eight anterior-to-posterior bipolar pairs; and 3) lead 1 to lead 16. Because all outcomes were highly similar across all reference schemes, we only report on the cerebellar reference findings. All local field potential findings were averaged across all 16 leads within each mouse.

Time-frequency measures were computed by multiplying the fast Fourier transformed (FFT) power spectrum of single trial LFP data with the FFT power spectrum of a set of complex Morlet wavelets defined as a Gaussian-windowed complex sine wave: $e^{i2\pi tf}e^{-t^2/(2x\sigma^2)}$, where t is time, f is frequency (which increased from 1–50 Hz in 50 logarithmically spaced steps), and the width (or "cycles") of each frequency band were set to increase from 3/(2πf) to 10/(2πf) as frequency increased. Then, the time series was recovered by computing the inverse FFT. The end result of this process is identical to time-domain signal convolution. It resulted in estimates of instantaneous power taken from the magnitude of the analytic signal. Each epoch was then cut in length from −500 to +1000 ms peri-feedback.

Averaged power was normalized by conversion to a decibel (dB) scale (10*log10[power(t)/power(baseline)]), allowing a direct comparison of effects across frequency bands. The baseline consisted of averaged power −300 to −200 ms before all imperative cues. A 100-ms duration often is used as an effective baseline in spectral decomposition, because pixel-wise time-frequency data points have already been resolved over smoothed temporal and frequency dimensions with the wavelets. All ERPs were filtered with a 20-Hz low-pass filter. Model-derived reward prediction error (+PE) was regressed on all single-trial spectral decompositions, as well as all single-trial broadband EEG.

Statistical analysis

Mice were excluded from analyses if they had an aggregate accuracy <55% (3 mice: 2 males), if histology failed to verify that recordings were in medial cortex (4 mice with leads in septum: 3 males), or if the head stage became detached (1 mouse: male). Two mice had exploratory arrays in orbitofrontal cortex (2 males) and were not included in this analysis. This left 29 mice (15 males) for the planned analyses (Fig. 1A). Because four subjects were recorded on an analogue amplifier, we removed these subjects from ERP analyses. However, all other analyses used decibel or regression-based outcomes, which were not affected by this variable.

Rewarding feedbacks were used for all analyses. Rewarded responses were immediately indicated by a 1-s, pure noise tone, which ended with the illumination of the magazine light and delivery of liquid reward. We used a priori regions of interest (ROIs) but also followed up on the a posterori findings observed here. For the ERP, we used an a priori ROI of 400–600 ms (Cavanagh et al., 2022) but also examined an a posteriori time window of 600–800 ms. For time-frequency regions of interest (tf-ROI), we used 1–1.4 Hz from 250–550 ms (Cavanagh et al., 2021) but also examined an a posterori time window of 550–850 ms. The time-frequency PE regressions revealed a strong effect in the alpha band (9.5–12 Hz; see also Iturra-Mena et al., 2023); we examined this feature as an a posteriori outcome.

Results

Performance

Mice learned the task well (mean [M] = 72%, standard deviation [SD] = 8%) with aggregate win stay (M = 66%, SD = 9%) and lose-switch (M = 56%, SD = 8%) tendencies (Fig. 1C). Table 1 shows model fits. Each increasingly complex model fit better than the previous one in Pseudo R^2, although AIC was rather similar across model types. This suggests that each model added an improvement in the prediction of behavior, but this was minimal and about equivalent to the AIC penalty for complexity.

Model M3 was chosen as the best-fitting model given better pseudo R^2, ambivalent AIC, and a verified conceptualization of maintained learning (positive omega) instead of a necessary high learning rate per each session. Notably, the LFP analyses only leveraged the relative ranking of +PE (i.e., nonparametric correlation), which is largely unaffected by

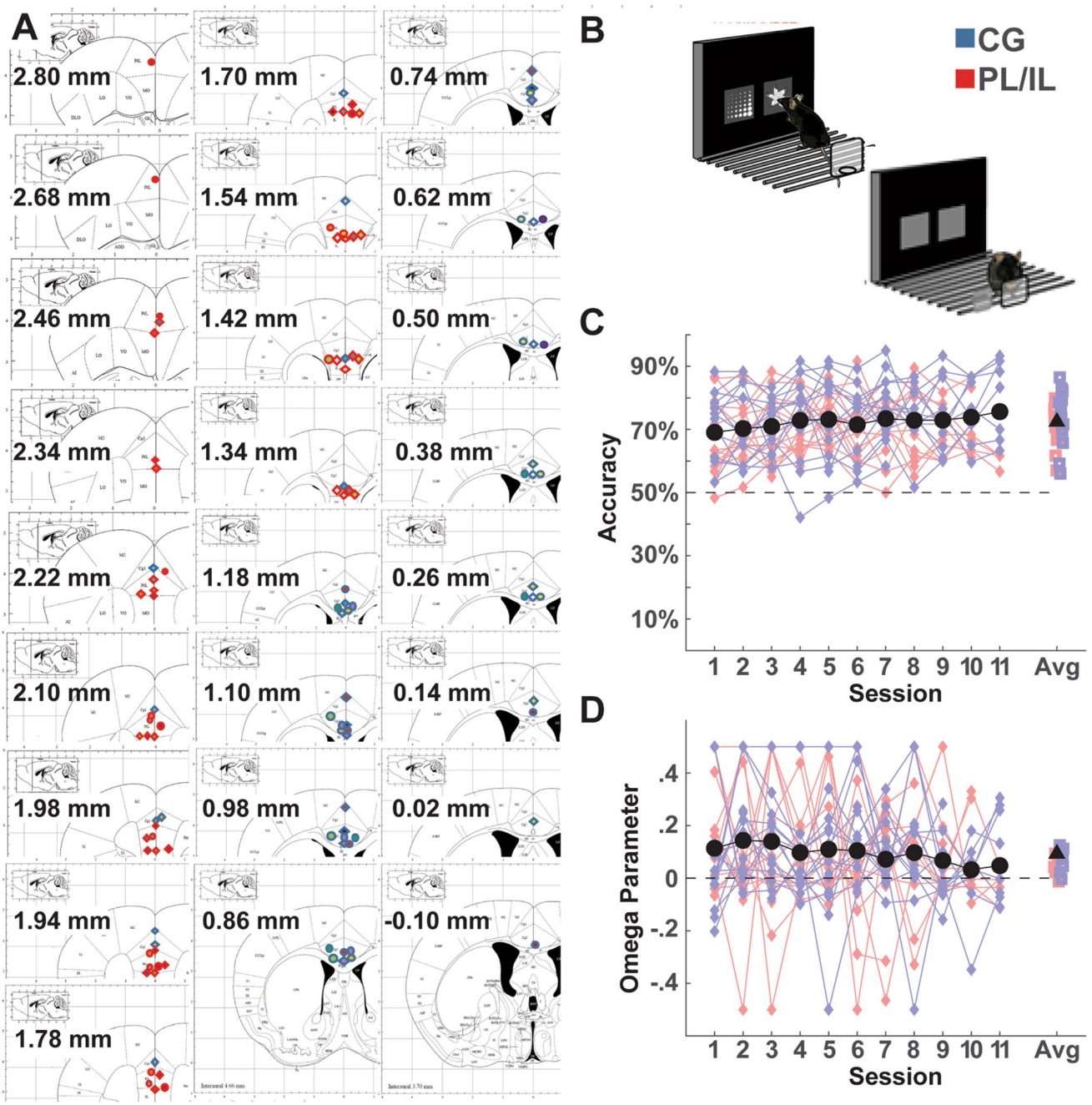

Fig. 1 Depth recordings on a touchscreen probabilistic learning task. Groups are identified by color: red = prelimbic or infralimbic (PL/IL); n = 14, blue = cingulate gyrus (CG), n = 15. **A** Electrolytic lesion identification of the placement of 16-channel leads; each mouse had four anterior-posterior slices with lead placements identified. Measurements are millimeters from bregma. **B** Mice touched one of two screens to identify the most rewarding stimulus. Immediately after touching the screen, a tone indicated the availability of liquid reward on reinforced trials. **C** Accuracy over sessions, as well as within-mouse average accuracy. Black icons show average across sessions. **D** Omega parameter from the best-fitting reinforcement learning model, indicating a maintained bias toward the action value for the optimal stimulus between sessions. Black icons show average across sessions

additional parameters. A generative model of model M3 had nearly identical aggregate accuracy (M = 72%, SD = 8%), win stay (M = 65%, SD = 8%), and lose-switch (M = 56%, SD = 8%) tendencies. M3 omega parameters did not reliably change over time within each mouse across sessions ($t<1$), (Fig. 1D). Overall, all behavioral tendencies and model fits indicated that all mice learned the task well with maintained action values between sessions. Singles trial prediction errors were taken from model M3 fits and used as regressors for neural recordings.

Table 1 Model parameters and fits. All values are mean (SD)

	Alpha	Alpha gain	Alpha loss	Beta	Omega	AIC	Pseudo R^2
M1: Vanilla	0.26 (0.11)			3.47 (1.24)		−70.00 (8.21)	17.81 (10.04)
M2: Win Loss		0.25 (0.14)	0.23 (0.12)	5.90 (1.49)		−70.05 (8.40)	20.27 (10.29)
M3: M2 + Qbias		0.10 (0.12)	0.29 (0.12)	7.68 (1.99)	0.11 (0.08)	−70.01 (8.95)	22.84 (11.05)

Local field potentials

Figure 2 shows the ERPs (Fig. 2A, B) and broadband regressions with +PE (Fig. 2C, D). As shown in Fig. 2B, the PL/IL and CG groups differed in RewP amplitude ($t(23) = 2.61$, $p = 0.02$) but not in the regression with +PE ($t(27) = -1.45$, $p = 0.16$; Fig. 2D). While there was an apparent representation of +PE-related variance in the late *a posteriori* window (Fig. 2C), this was not statistically significant, and none of the ERP or EEG +PE correlations scaled with the depth of the lead in either window.

Figure 3 shows the time-frequency outcomes for power (Fig. 3A) and power +PE correlation (Fig 3B). An unpredicted alpha-band response from 250–850 ms also was apparent in the power +PE correlation (Fig. 3B), particularly for the PL/IL group. Figure 3C-D shows how the a priori delta band tf-ROI scaled with the depth of the lead, both for raw power as well as for the tf-ROI +PE correlation. This depth-related variance was most strongly prevalent in CG subjects. These findings were similar for the a priori early (250–550 ms) and a posteriori late (550–850 ms) windows, as well as across all four reference schemes, suggesting a robust effect.

The a posteriori alpha band power tf-ROI from 250–850 ms was not different between groups ($t(27) = -0.54$, $p = 0.59$), although the alpha +PE correlation was significantly larger in the PL/IL group ($t(27) = 2.10$, $p = 0.04$). Power in the alpha band tf-ROI similar scaled with depth ($rho(27) = 0.49$, $p = 0.01$), but this was primarily observed in the PL/IL group ($rho(12) = 0.63$, $p = 0.01$) and not the CG group ($rho(13) = 0.41$, $p = 0.13$). The alpha +PE correlation also was scaled with depth ($rho(27) = 0.41$, $p = 0.03$), but this was not particularly prevalent within either subgroup (p's > 0.14).

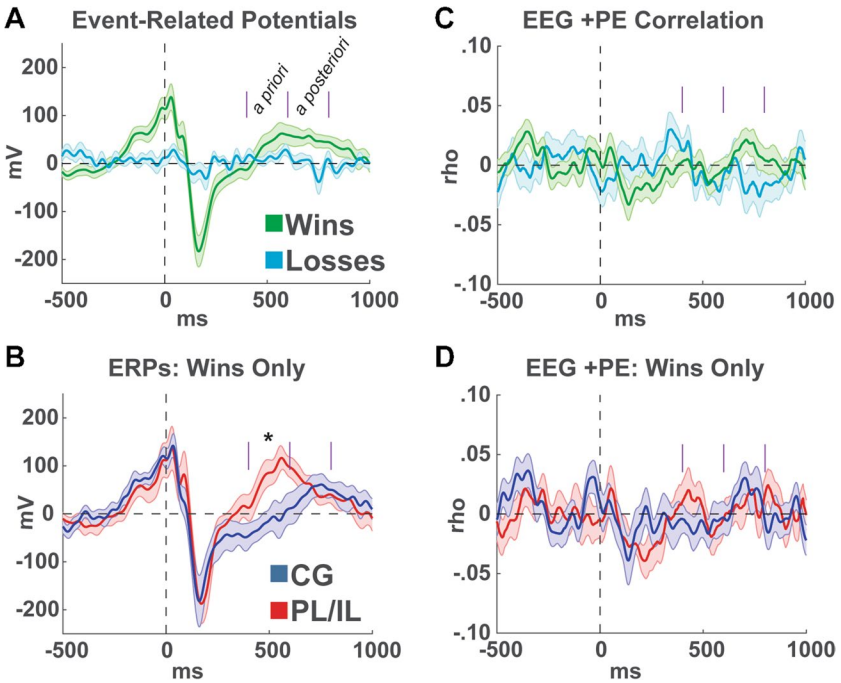

Fig. 2 Average event-related potentials from local field recordings. All leads and sessions were first averaged within each mouse; the ERP shown is the grand average across mice. **A** Win versus loss trials, with time windows of interested indicated in magenta. **B** The mouse homologue of the RewP was larger in the PL/IL group than the CG group. **C** Single trial correlation of +PE with broadband EEG activity. **D** This single trial correlation did not differ between groups in either time window of interest. *$p < 0.05$.

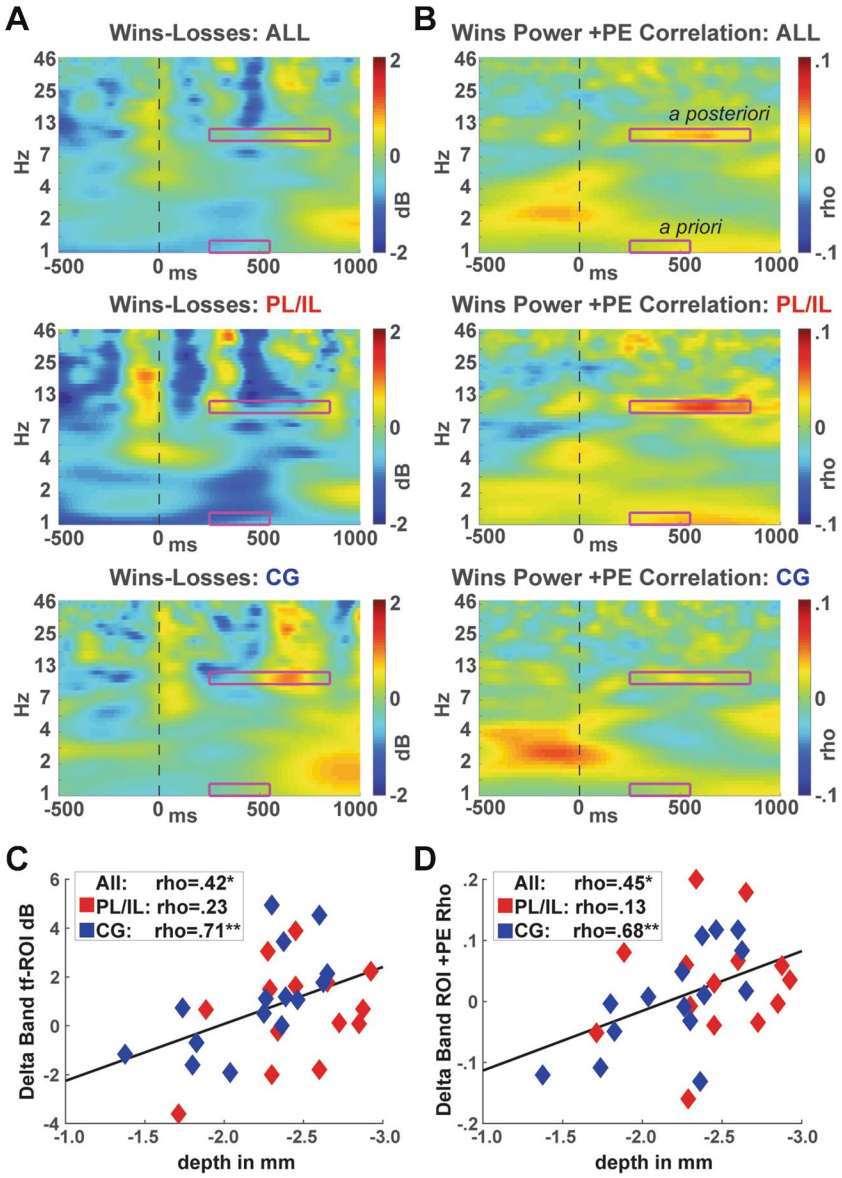

Fig. 3 Average time-frequency activity from local field recordings. **A** Win versus loss trials, with time windows of interested indicated in magenta. **B** Single trial correlation of +PE with spectral activity, highlighting a small low-frequency delta band effect which replicates prior findings, as well as a novel and unexpected alpha band finding particularly in the PL/IL group. **C** Recording depth predicted the size of the delta-band response to wins, particularly in the CG group. **D** Recording depth similarly scaled with the single trial +PE regressed delta-band activity, again primarily in the CG group. *$p < 0.05$; **$p < 0.01$

Discussion

We identified a robust correlation between reward-related local field signatures and cortical depth. The ERP feature previously identified as a RewP homologue was significantly larger in the PL/IL group than the CG group (Fig. 2B). The delta band spectral feature previously identified as a correlate of reward prediction error was larger in deeper probes (Fig. 3C), as was the single trial correlation between delta power and positive prediction error (Fig. 3D). These findings replicate and extend work from multiple different labs (Cavanagh et al., 2021; 2022; Iturra-Mena et al., 2023; Warren et al., 2015) and provide important verification of the validity of this translational biomarker.

These reward-specific delta-band responses appear to be unique in rodent frontal cortex. Other experiments have revealed cortical theta band activities during rewarding activities, but theta appears to be tightly linked to action-related processes, not feedback-related information like in this report. Theta is observed in general control of consummatory behavior (Amarante & Laubach, 2021), including motor actions associated with licking frequency (Horst & Laubach, 2013)

and reward-dependent vigor (Amarante, Caetano, & Laubach, 2017). Notably, this theta presence is not as simple as motor execution but can include control-related processes, such as motor inhibition (Müller Ewald et al., 2022) and post-error adjustment (Narayanan, Cavanagh, Frank, & Laubach, 2013; Olguin et al., 2023). Whereas these delta versus theta band features appear to align with evaluative versus executive processes, the current report also revealed prediction error sensitive alpha band activity, particularly in PL/IL areas. It is unknown what alternative or nested process this alpha activity may represent, but it appears similar to recent observations of reward-evoked alpha/beta power in rat CG (Iturra-Mena et al., 2023).

It remains unknown how this delta band signal reflects cortical interaction with dopaminergic midbrain and striatum, which are traditionally associated with the mechanistic instantiation of reinforcement learning. The RewP and rodent RewP homologue occur about the same time as phasic midbrain dopaminergic neuron responses to rewards in humans (200–400 ms; Zaghloul et al., 2009) and mice (200–700 ms, Dabney et al., 2020; Eshel et al., 2016). Notably, Iturra-Mena et al. (2023) demonstrated highly similar delta band activities in nucleus accumbens simultaneous with the CG delta band response. This suggests that these low-frequency activities may be used to communicate across cortical and subcortical areas. It is our working hypothesis that this cortical signal reflects cross-modal integration facilitating computationally complex assessments of the multidimensional value of a reward; we expect that this cortical activity may thus precede the striatal activity.

Limitations and future directions

There are many outstanding questions about the nature of the RewP homology between species. We have focused on one type of task, but we predict that the rodent RewP homologue will have the same cross-task and cross-modality domain generality as the human RewP. We further predict similar dorsal versus ventral specificities in the information content represented in this signal. We suggest that this might follow a rough actor-versus-critic dissociation, possibly preferentially signaled by theta versus delta-band activities. The unexpected alpha band representation of +PE in PL/IL suggests a possible mechanism for area-specific network interactions, possibly with striatum (Iturra-Mena et al., 2023).

A fully validated animal model of a human, domain-specific, neural response has high biomedical value. There is strong enthusiasm for animal models of psychiatric deep brain stimulation predicated on translatable functional biomarkers (Roberts & Clarke, 2019; Rudebeck et al., 2019). The similarities between rodent IL and human subgenual cingulate offer a compelling base for such a model (Balsters et al., 2020; Heilbronner et al., 2016), as well as for investigation of pharmacological antidepressants (Fullana et al., 2019). This rodent model also may inform somewhat untestable hypotheses in humans, such as mechanisms of cortico-striatal or cortico-midbrain network interaction.

Conclusions

The human RewP appears to be a transdiagnostic biomarker of reward responsiveness, learning, and valuation. An animal homologue of the RewP offers a chance to test mechanistically the network-level systems underlying these reward domains. We anticipate that the cross-species translatability of this bio-signal will further bolster our mechanistic understanding of reward-related disfunctions in major depression, schizophrenia, addiction, and Parkinson's disease.

Funding This project was funded by NIMH 1R01MH119382.

References

Amarante, L. M., Caetano, M. S., & Laubach, M. (2017). Medial frontal theta is entrained to rewarded actions. *The Journal of Neuroscience, 37*(44), 1965–17. https://doi.org/10.1523/JNEUROSCI.1965-17.2017

Amarante, L. M., & Laubach, M. (2021). Coherent theta activity in the medial and orbital frontal cortices encodes reward value. *ELife, 10*, 1–25. https://doi.org/10.7554/eLife.63372

Amitai, N., Young, J. W., Higa, K., Sharp, R. F., Geyer, M. A., & Powell, S. B. (2014). Isolation rearing effects on probabilistic learning and cognitive flexibility in rats. *Cognitive, Affective and Behavioral Neuroscience, 14*(1), 388–406. https://doi.org/10.3758/s13415-013-0204-4

Amodeo, D. A., Jones, J. H., Sweeney, J. A., & Ragozzino, M. E. (2012). Differences in BTBR T+ tf/J and C57BL/6J mice on probabilistic reversal learning and stereotyped behaviors. *Behavioural Brain Research, 227*(1), 64–72. https://doi.org/10.1016/j.bbr.2011.10.032

Baker, T. E., & Holroyd, C. B. (2011). Dissociated roles of the anterior cingulate cortex in reward and conflict processing as revealed by the feedback error-related negativity and N200. *Biological Psychology, 87*(1), 25–34. https://doi.org/10.1016/j.biopsycho.2011.01.010

Balsters, J. H., Zerbi, V., Sallet, J., Wenderoth, N., & Mars, R. B. (2020). Primate homologs of mouse cortico-striatal circuits. *ELife, 9*, 1–24. https://doi.org/10.7554/eLife.53680

Bari, A., Theobald, D. E., Caprioli, D., Mar, A. C., Aidoo-Micah, A., Dalley, J. W., & Robbins, T. W. (2010). Serotonin modulates sensitivity to reward and negative feedback in a probabilistic reversal learning task in rats. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology, 35*(6), 1290–1301. https://doi.org/10.1038/npp.2009.233

Bicks, L. K., Koike, H., Akbarian, S., & Morishita, H. (2015). Prefrontal cortex and social cognition in mouse and man. *Frontiers in Psychology*, *6*(NOV), 1–15. https://doi.org/10.3389/fpsyg.2015.01805

Camerer, C., & Ho, T. H. (1999). Experience-weighted attraction learning in normal form games. *Econometrica, 67*(4), 827–874.

Cavanagh, J. F. (2015). Cortical delta activity reflects reward prediction error and related behavioral adjustments, but at different times. *NeuroImage, 110*, 205–216. https://doi.org/10.1016/J.NEUROIMAGE.2015.02.007

Cavanagh, J.F., Gregg, D., Light, G. A., Olguin, S., Sharp, R. F., Bismark, A. W., … Young, J. W. (2021). Electrophysiological biomarkers of behavioral dimensions from cross-species paradigms. *Translational Psychiatry*, *11*(482), 1–11.

Cavanagh, James F., Olguin, S. L., Talledo, J. A., Kotz, J. E., Roberts, B. Z., Nungaray, J. A., … Brigman, J. L. (2022). Amphetamine alters an EEG marker of reward processing in humans and mice. *Psychopharmacology*, *239*(3), 923–933. https://doi.org/10.1007/s00213-022-06082-z

Cavanagh, James F., Bismark, A. W., Frank, M. J., & Allen, J. J. B. (2018). Multiple dissociations between comorbid depression and anxiety on reward and punishment processing: Evidence from computationally informed EEG. *Computational Psychiatry*. https://doi.org/10.1162/cpsy

Crane, N. A., Burkhouse, K. L., Gorka, S. M., Klumpp, H., & Phan, K. L. (2022). Electrocortical measures of win and loss processing are associated with mesocorticolimbic functional connectivity: A combined ERP and rs-fMRI study. *Psychophysiology, 59*(12), 1–17. https://doi.org/10.1111/psyp.14118

Dabney, W., Kurth-Nelson, Z., Uchida, N., Starkweather, C. K., Hassabis, D., Munos, R., & Botvinick, M. (2020). A distributional code for value in dopamine-based reinforcement learning. *Nature, 577*(7792), 671–675. https://doi.org/10.1038/s41586-019-1924-6

Daw, N. D. (2011). Trial by trial data analysis using computational models. In M. R. Delgado, E. A. Phelps, & T. W. Robbins (Eds.), *Decision Making, Affect, and Learning: Attention and Performance XXIII* (pp. 1–26). Oxford University Press.

Eshel, N., Tian, J., Bukwich, M., & Uchida, N. (2016). Dopamine neurons share common response function for reward prediction error. *Nature Neuroscience, 19*(3), 479–486. https://doi.org/10.1038/nn.4239

Francis-Oliveira, J., Leitzel, O., & Niwa, M. (2022). Are the anterior and mid-cingulate cortices distinct in rodents? *Frontiers in Neuroanatomy, 16*(June), 1–7. https://doi.org/10.3389/fnana.2022.914359

Fullana, M. N., Ruiz-Bronchal, E., Ferrés-Coy, A., Juárez-Escoto, E., Artigas, F., & Bortolozzi, A. (2019). Regionally selective knockdown of astroglial glutamate transporters in infralimbic cortex induces a depressive phenotype in mice. *Glia, 67*(6), 1122–1137. https://doi.org/10.1002/glia.23593

Heilbronner, S. R., Rodriguez-Romaguera, J., Quirk, G. J., Groenewegen, H. J., & Haber, S. N. (2016). Circuit-based corticostriatal homologies between rat and primate. *Biological Psychiatry, 80*(7), 509–521. https://doi.org/10.1016/j.biopsych.2016.05.012

Holroyd, C B, Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, *45*(5), 688–697. https://doi.org/10.1111/j.1469-8986.2008.00668.x

Holroyd, Clay B., & Umemoto, A. (2016). The research domain criteria framework: The case for anterior cingulate cortex. *Neuroscience and Biobehavioral Reviews, 71*, 418–443. https://doi.org/10.1016/j.neubiorev.2016.09.021

Horst, N. K., & Laubach, M. (2013). Reward-related activity in the medial prefrontal cortex is driven by consumption. *Frontiers in Neuroscience*, *7*(7 APR), 1–15. https://doi.org/10.3389/fnins.2013.00056

Hyman, J. M., Holroyd, C. B., & Seamans, J. K. (2017). A novel neural prediction error found in anterior cingulate cortex ensembles. *Neuron, 95*(2), 447-456.e3. https://doi.org/10.1016/j.neuron.2017.06.021

Iturra-Mena, A. M., Kangas, B. D., Luc, O. T., Potter, D., & Pizzagalli, D. A. (2023). Electrophysiological signatures of reward learning in the rodent touchscreen-based probabilistic reward task. *Neuropsychopharmacology, 48*(4), 700–709. https://doi.org/10.1038/s41386-023-01532-4

Keeler, J. F., & Robbins, T. W. (2011). Translating cognition from animals to humans. *Biochemical Pharmacology, 81*(12), 1356–1366. https://doi.org/10.1016/j.bcp.2010.12.028

Laubach, M., Amarante, L. M., Swanson, K., & White, S. R. (2018). What, If Anything, Is Rodent Prefrontal Cortex? *ENeuro*, *5*(5).

Marquardt, K., Sigdel, R., Caldwell, K., & Brigman, J. L. (2014). Prenatal ethanol exposure impairs executive function in mice into adulthood. *Alcoholism: Clinical and Experimental Research*, *38*(12), 2962–2968. https://doi.org/10.1111/acer.12577

Müller Ewald, V. A., Kim, J., Farley, S. J., Freeman, J. H., & LaLumiere, R. T. (2022). Theta oscillations in rat infralimbic cortex are associated with the inhibition of cocaine seeking during extinction. *Addiction Biology, 27*(1), 1–12. https://doi.org/10.1111/adb.13106

Narayanan, N. S., Cavanagh, J. F., Frank, M. J., & Laubach, M. (2013). Common medial frontal mechanisms of adaptive control in humans and rodents. *Nature Neuroscience, 16*(October), 1–10. https://doi.org/10.1038/nn.3549

Olguin, S. L., Cavanagh, J. F., Young, J. W., & Brigman, J. L. (2023). Impaired cognitive control after moderate prenatal alcohol exposure corresponds to altered EEG power during a rodent touchscreen continuous performance task. *Neuropharmacology, 236*(May), 109599. https://doi.org/10.1016/j.neuropharm.2023.109599

Preuss, T. M., & Wise, S. P. (2022). Evolution of prefrontal cortex. *Neuropsychopharmacology, 47*(1), 3–19. https://doi.org/10.1038/s41386-021-01076-5

Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology, 52*(4), 449–459. https://doi.org/10.1111/psyp.12370

Roberts, A. C., & Clarke, H. F. (2019). Why we need nonhuman primates to study the role of ventromedial prefrontal cortex in the regulation of threat- and reward-elicited responses. *Proceedings of the National Academy of Sciences of the United States of America, 116*(52), 26297–26304. https://doi.org/10.1073/pnas.1902288116

Roy, M., Shohamy, D., & Wager, T. D. (2012). Ventromedial prefrontal-subcortical systems and the generation of affective meaning. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2012.01.005

Rudebeck, P. H., Rich, E. L., & Mayberg, H. S. (2019). From bed to bench side: Reverse translation to optimize neuromodulation for mood disorders. *Proceedings of the National Academy of Sciences of the United States of America, 116*(52), 26288–26296. https://doi.org/10.1073/pnas.1902287116

Sarter, M. (2004). Animal cognition: Defining the issues. *Neuroscience and Biobehavioral Reviews, 28*(7), 645–650. https://doi.org/10.1016/j.neubiorev.2004.09.005

Schaeffer, D. J., Hori, Y., Gilbert, K. M., Gati, J. S., Menon, R. S., & Everling, S. (2020). Divergence of rodent and primate medial frontal cortex functional connectivity. *Proceedings of the National Academy of Sciences of the United States of America, 117*(35), 21681–21689. https://doi.org/10.1073/pnas.2003181117

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning : an introduction. In Adaptive computation and machine learning*. MIT Press.

van Heukelum, S., Mars, R. B., Guthrie, M., Buitelaar, J. K., Beckmann, C. F., Tiesinga, P. H. E., … Havenith, M. N. (2020). Where is cingulate cortex? A cross-species view. *Trends in Neurosciences*, *43*(5), 285–299. https://doi.org/10.1016/j.tins.2020.03.007

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: event-related potential correlates of reward processing, neural

adaptation, and behavioral choice. *Neuroscience and Biobehavioral Reviews, 36*(8), 1870–1884. https://doi.org/10.1016/j.neubiorev.2012.05.008

Warren, C. M., Hyman, J. M., Seamans, J. K., & Holroyd, C. B. (2015). Feedback-related negativity observed in rodent anterior cingulate cortex. *Journal of Physiology Paris, 109*(1–3), 87–94. https://doi.org/10.1016/j.jphysparis.2014.08.008

Whitton, A., Kumar, P., Treadway, M. T., Rutherford, A. V., Ironside, M. L., Foti, D., … Pizzagalli, D. (2023). Distinct profiles of anhedonia and reward processing and their prospective associations with quality of life among individuals with mood disorders. *Molecular Psychiatry*, (June), 1–19. https://doi.org/10.1038/s41380-023-02165-1

Zaghloul, K. A., Blanco, J. A., Weidemann, C. T., McGill, K., Jaggi, J. L., Baltuch, G. H., & Kahana, M. J. (2009). Human substantia nigra neurons encode unexpected financial rewards. *Science (New York, N.Y.)*, *323*(5920), 1496–1499. https://doi.org/10.1126/science.1167342

Open practices statement All data and Matlab code to recreate these analyses are available at OSF.io: HTTPS://DOI.ORG/10.17605/OSF.IO/T3SNW. This experiment was not preregistered.