Title

Electrophysiological Signatures of Reward, Surprise, and Conflict in the Human Brain

Author

Hoy, Colin Weir

Publication Date

2021

Peer reviewed|Thesis/dissertation

Electrophysiological Signatures of Reward, Surprise, and Conflict in the Human Brain

By

Colin W Hoy

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Neuroscience

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Robert T Knight, Chair
Professor Joni Wallis
Professor Steven Piantadosi
Professor Rich Ivry

Summer 2021

Abstract

Electrophysiological Signatures of Reward, Surprise, and Conflict in the Human Brain

By

Colin W Hoy

Doctor of Philosophy in Neuroscience

University of California, Berkeley

Professor Robert T. Knight, Chair

Flexible, intelligent behavior requires choosing actions predicted to lead to optimal outcomes in an ever-changing environment and learning from feedback to improve future predictions. This action-feedback loop is governed by cognitive control, which allocates attention, working memory, and decision making resources according to changes in task demands. Predictive coding frameworks describe these dynamics in terms of prediction errors (PEs), or the difference between expected and actual outcomes. Despite decades of cognitive neuroscience research, the neural mechanisms of cognitive control remain elusive, in part because complex cognition depends on rapid interactions between widespread brain networks. Studying cognitive control in humans provides the opportunity to design complex tasks that dissociate the various computations underlying rich behaviors, but these advantages are typically offset by the limited spatial or temporal resolution of non-invasive measures of neuronal activity. This dissertation pairs behavioral modeling with the combined spatial and temporal resolution of direct brain recordings in humans to test foundational theories of cognitive control and predictive coding. Chapter 2 addresses a longstanding debate over whether scalp electroencephalography (EEG) signatures represent a valenced, quantitative reward prediction error (RPE) or the non-valenced magnitude of RPE. We use reinforcement learning principles to model individual participant behavior in an interval timing task and apply powerful single-trial regression analyses across time, space, and frequency dimensions to disentangle multiple overlapping components. Our results show valenced RPE effects are an artifact of component overlap between the early, frontal, theta frequency feedback-related negativity tracking non-valenced RPE magnitude for negative outcomes and the subsequent, more posterior reward positivity in delta frequencies tracking non-valenced RPE magnitude for positive outcomes. Our modeling approach also uncovered a novel, late frontal P300-like component elicited by low probability outcomes. Chapter 3 uses the same paradigm and behavioral model to investigate RPE value and magnitude coding at the local circuit level based on high frequency broadband (HFB) activity extracted from intracranial EEG (iEEG) recording in lateral prefrontal cortex (LPFC), medial prefrontal cortex (MPFC), and insular cortex (IC). We show that many sites in each of these three control regions represent either RPE value or magnitude, but also that

some electrodes show mixed selectivity for both RPE features. Interestingly, RPE value and magnitude representations were most common in IC, which also showed the greatest proportion of electrodes with mixed selectivity. Furthermore, onsets of RPE value effects were earlier in IC than MPFC, suggesting a potential leading role for IC in reward processing. Finally, Chapter 4 characterizes the spatiotemporal evolution of conflict signals using HFB activity in a color-word Stroop task. In order to distinguish detection, resolution, and monitoring phases of conflict processing, we use a dynamic sliding window analysis strategy to segregate effects into stimulus, decision, and responses stages of the trial. We observed widespread conflict effects in LPFC, MPFC, IC, orbitofrontal, sensorimotor, and temporal cortices that formed partially overlapping but largely distinct networks for stimulus, decision, and response stages. Contrary to serial processing hypotheses proposed by classic conflict monitoring theory, we found the onsets of these conflict signals to be heterogeneous within and across regions. These results indicate conflict processing unfolds across distributed networks that work in parallel throughout the trial. Future studies can extend these findings using connectivity network analyses to determine information flow between these networks, which will help constrain the functional roles of each region. In particular, analyzing the relationship between HFB activity and rhythmic low frequency responses will help identify circuit mechanisms that bridge local HFB activity and non-invasive biomarkers from scalp EEG. Collectively, the experiments in this thesis provide novel insights into how reward, surprise, and conflict signals are processed in parallel across distributed networks and emphasize the importance of combining experimental design, behavioral modeling, and advanced signal processing to understand the neural computations underlying cognitive control in the human brain.

Table of Contents

Dedication ..... *iii*

Acknowledgements ..... *iv*

Chapter 1: Introduction and Background *1*

Chapter 2: Single-trial modeling separates multiple overlapping prediction errors during reward processing in human EEG *6*

ABSTRACT **6**

INTRODUCTION **6**

RESULTS **10**

Behavioral Modeling 10

Single-trial regression reveals a spatio-temporal cascade of PE components 12

Scalp topography dynamics delineate PE components 15

Single-trial regression of time-frequency power dissociates PE effects in theta and delta bands 16

Correlations with Oddball task ERPs quantify contributions of overlapping components 17

Separating outcomes by valence dissociates FRN and RewP RPE magnitude effects 18

ERP PE sequence results are robust to biases in subjective reward expectations 21

DISCUSSION **21**

Methodological implications 22

Theoretical implications: Predictive coding 23

Theoretical implications: ERP components 24

METHODS **28**

Experimental model and participant details 28

Method details 28

Statistics and reproducibility 32

Data availability 34

Software availability 35

Acknowledgements 36

Author contributions 36

Competing interests 36

SUPPLEMENTARY INFORMATION **37**

SUPPLEMENTARY NOTE 1: FRN peak latency shifts with RPE valence 38

SUPPLEMENTARY NOTE 2: ERP PE sequence results are robust to biases in subjective reward expectations 38

SUPPLEMENTARY FIGURES 40

Chapter 3: Human intracranial recordings reveal mixed coding of reward and salience prediction errors in fronto-insular cortices *49*

ABSTRACT **49**

INTRODUCTION **49**

METHODS **51**

Participants 51

Target Time behavioral task 52

Behavioral modeling 53

iEEG data collection, localization, and preprocessing 54
High frequency broadband power extraction and modeling 54

RESULTS 55
Behavioral modeling 56
High frequency broadband power reveals the spatiotemporal distribution of RL variables 57
Separate and mixed selectivity for RPE value and magnitude 60
Temporal dynamics of RPE features in HFB power 61

DISCUSSION 62

Chapter 4: Intracranial Stroop recordings reveal parallel conflict processing in networks distributed across human cortex 66

ABSTRACT 66

INTRODUCTION 66

METHODS 69
Participants 69
Stroop behavioral task 69
iEEG data collection, localization, and preprocessing 70
High frequency broadband power extraction and modeling 72

RESULTS 73
iEEG HFB power separates stimulus, decision, and responses stages of conflict processing 74
Different stages of conflict processing recruit distributed, partially overlapping networks 75
Conflict onset latencies support parallel processing 76

DISCUSSION 77

Chapter 5: Conclusions 81

References 84

Dedication

To my parents, for teaching me compassion and kindness, integrity and respect, dedication and hard work, and how to do what must be done.

Acknowledgements

This dissertation describes much but not all of seven years of intense work, none of which would have been possible without help from a wide range of colleagues, friends, and family members. It would be impossible to give each of them the credit they deserve, but here I'll mention a few key people.

First and foremost, thank you to my advisor Bob Knight. I will be forever grateful for the numerous opportunities you've given me, for surrounding me with the tools, resources, and people necessary to take advantage of them, and for your continuing support across scientific, professional, and personal domains. I am a better scientist and person because of your mentorship, and I hope to use the platform built on our work together to launch a career that helps as many people as you have.

Thank you to my many other mentors throughout my PhD, including Rich Ivry, Joni Wallis, Steve Piantadosi, and Mark D'Esposito for helping form and guide my research as members of my thesis and qualifying exam committees. I also learned so much from my colleagues in the extended Knight lab community. I'd especially like to thank Matar Haller and Avgusta Shestyuk for helping me get started in electrophysiology as a rotation student, and Randolph Helfrich, Arjen Stolk, and Nick Myers for bolstering my knowledge and code base for signal processing and statistics. I'm also grateful to Vinitha Rangarajan, Lisa Johnson, Julia Kam, Anais Llorens, and so many other wonderful lab mates for providing a rich and stimulating work environment filled with interesting ideas and friendly faces.

I also want to thank the many collaborators that I was fortunate to work with over the years. The most important of these are the clinicians, EEG technicians, nurses, and especially the patients contributing to our intracranial research program, as well as the entire Knight lab for their collective efforts in intracranial data collection and electrode reconstructions. Thanks to my undergraduate research assistants Rana Eser, Aditya Goel, Ian Griffith, Sheila Steiner, Aastha Shah, and Jason Abbas for their contributions to our research, and for helping me develop the skills to become a better mentor. Thanks to Kris Anderson and Vitoria Piai for setting up the Stroop project; to Michael Pereira for his helpful perspective and hard work on errors in continuous motor control; to Regina Lapate, Alan Cowen, and Kuan Chen for helping me dip my toes into the world of emotion; to Jose del R. Millan, Nikolai Axmacher, Leila Reddy, and the other scholars that came to visit our lab and share their insights.

I also want to thank the members of the UC Berkeley and Helen Wills Neuroscience Institute community for providing a warm, fun, and inclusive environment that helped me learn and grow throughout these formative seven years. Thanks to my classmates entering the year of 2014 and my fellow graduate students for giving me a vibrant social community. Thanks to HWNI staff Candace Groskreutz, Natalie Terranova, Tony Leonard, and Kati Markowitz, HWNI graduate program directors Dan Feldman and Michael Silver, and Kirsten Upson and the many administrative staff that shielded me from bureaucracy to let me focus on research.

I'm also grateful to be surrounded by my roommates, my Quaranteam, my Very Best Friends, and my Ursa Major and Firefly ultimate frisbee communities, which helped me stay healthy and sane through the tough times of graduate school. I want to thank Erik Urban specifically for reintroducing me to ultimate frisbee, which has provided me with an outlet for physical activity and many friends and fond memories.

Thanks to my best friend Omar Mahmoud for getting me away from my desk on his annual adventures out West, for keeping me grounded, and for talking me through the biggest challenges of life and free agency.

Thanks to my partner Rose Larios for her love and support, especially during difficult pandemic and thesis writing times. You've broadened my horizons and helped me grow into a stronger, more compassionate, and more understanding person, and I can't wait to see where we go next!

Lastly, I want to thank my immediate family and extended Weir and Hoy families for their love and support. In particular, thanks to my parents and all those who worked so hard to give me these incredible opportunities.

I am grateful to the National Science Foundation and the Greater Good Science Center for fellowships that helped fund the work in this thesis.

Chapter 1: Introduction and Background

"Engage people with what they expect; it is what they are able to discern and confirms their projections. It settles them into predictable patterns of response, occupying their minds while you wait for the extraordinary moment—that which they cannot anticipate." – Sun Tzu, The Art of War

Behavior is based on a continuous cycle of perception and action in an ever-changing environment, and surviving and thriving depend on choosing actions that maximize reward and minimize punishment. Intelligent organisms can learn from previous actions and outcomes to improve future choices by exerting effortful, top-down control. Cognitive control is the process of deploying, adjusting, and coordinating cognitive processes such as attention, working memory, decision making, planning, and memory to optimize this action-feedback loop in a given environmental context. Cognitive control is especially important for overcoming distractions or overriding automatic, habitual responses, and it is also referred to as executive function due to its similarities to a CEO managing the many components of a large organization. Practically speaking, cognitive control can be thought of as the collection of brain mechanisms that help you navigate to a restaurant, compare and choose between options on the menu, maintain attention on a friend's voice amid background noise, and inhibit the urge to check your cell phone during the meal.

Many wise souls have agreed that everything in life is judged relative to expectations, and this principle also applies to brain function. One of the most successful frameworks for understanding perception, action, and control is called predictive coding. Predictive coding argues that the brain is constantly predicting the future, whether it's the face of the person who's voice you hear, the taste of your favorite dish, or the feeling of anticipation before seeing a loved one after a long time**.** The second component of predictive coding is feedback that can match the prediction or result in a different, surprising outcome. Surprise is quantified by prediction errors (PEs), or the difference between the outcome and the prediction, which then updates predictions in the future.

The most well-known PEs in neuroscience are in the field of reward processing, where they signal the positive or negative valence (better or worse than expected?) and magnitude of surprise (how different than expected?). Seminal animal research during the late 1990's showed that firing in midbrain dopamine neurons was sensitive to rewards but tracked reward prediction errors (RPEs) rather than the value of the reward itself. This translates to firing increases only when rewards exceeded expectations and decreases when rewards were worse than expected (Bayer & Glimcher, 2005; Schultz et al., 1997; Zaghloul et al., 2009). In the following decades, neuroscientists have found neural activity encoding other types of PEs in visual (Rao & Ballard, 1999), auditory (Schneider et al., 2014, 2018), olfactory (Zelano et al., 2011), and other sensory input systems (Schneider, 2020), motor outputs circuits (Adams et al., 2013; Shipp et al., 2013), cognitive (Pezzulo et al., 2015; Pezzulo & Cisek, 2016), emotion (Hoemann et al., 2017), and social networks (Koster-Hale & Saxe, 2013), and many other domains (Clark, 2013). Importantly, these predictions and PEs about perception or action do

not have inherent valence like reward and punishment but reflect non-valenced surprise or salience. Neurons encoding these kinds of salience PEs for specific features in the environment like color or shape can help organisms identify what caused a certain outcome, solving what is known as the credit assignment problem (Akaishi et al., 2016; Asaad et al., 2017; Oemisch et al., 2019). The ubiquitous nature of PEs has led some neuroscientists to propose that predictive coding is a general foundation of all brain function (Bastos et al., 2012; Clark, 2013; Friston, 2010).

In contrast, early cognitive control research interpreted errors in the traditional sense of mistakes. To understand why participants made mistakes and how they learned to improve their performance, researchers contrasted measurements of behavior and neuronal activity between incorrect and correct responses and/or feedback. Mapping error responses in humans using non-invasive functional magnetic resonance imaging (fMRI) measurements of blood oxygenation identified distributed cognitive control networks spanning regions in frontal, parietal, and insular lobes (Gratton et al., 2018). Prominent error signals in medial prefrontal cortex (MPFC) led to proposals that this region plays a key role in performance monitoring and is sensitive to events like errors, difficult choices, and uncertainty that indicate control should be adjusted (Alexander & Brown, 2011; Kolling et al., 2016; Shenhav et al., 2013). One important control signal is conflict between mutually exclusive options, which is especially important to detect and resolve if the appropriate response is in competition with default, automatic responses that would otherwise result in an error. **Chapter 4** addresses a prominent conflict monitoring theory that posits MPFC monitors for conflict and errors to relay adjustments in control demands to lateral prefrontal cortex (LPFC) (M M Botvinick et al., 2001; Tang et al., 2016; Yeung et al., 2004). LPFC in turn has been proposed to be responsible for implementing top-down regulation of stimulus-response mappings, i.e., attending to relevant sensory cues and selecting and executing appropriate motor responses (Knight, 1984; Miller & Cohen, 2001; Shenhav et al., 2013).

In terms of electrical brain activity, non-invasive electroencephalography (EEG) recordings of electrical potentials from the human scalp revealed several event-related potentials (ERPs) sensitive to errors (Falkenstein et al., 1991; Gehring et al., 1993; Nieuwenhuis et al., 2004). Specifically, the error-related negativity (ERN) and the feedback-related negativity (FRN) are negative EEG deflections appearing within a few hundred milliseconds of erroneous responses and negative feedback, respectively (Gehring et al., 2012; Miltner et al., 1997), and source localization techniques localized these signatures to sources in MPFC (Gehring & Willoughby, 2002; Hauser et al., 2014; E. H. Smith et al., 2015). However, these results were largely derived from correct-incorrect contrasts in easy tasks in which errors are unexpected, meaning the negative valence of errors was confounded with surprise.

In the mid-2000's, several studies challenged error-based interpretations by dissociating errors and surprise and observing ERN and FRN responses following surprising correct responses and unexpected positive feedback (Ferdinand et al., 2012; Oliveira et al., 2007; Talmi et al., 2013; Vidal et al., 2003). These contradictory findings lead to new theories of cognitive control based on predictive coding such as the Prediction-Response Outcome (PRO) model, which argues the primary function of MPFC in control is to detect unexpected action-outcome prediction errors (Alexander & Brown, 2011, 2014, 2017). **Chapter 2** directly addresses the longstanding debate over whether the FRN represents a valenced RPE signal or a non-valenced

action-outcome prediction error, and **Chapter 3** examines the representations of valenced and non-valenced RPEs in local population activity in the putative sources of these scalp EEG ERPs.

The action-outcome prediction errors underlying the PRO model explain a wide variety of control signals in MPFC related to goal-directed cognition, but they are not the only types of control captured by predictive coding. The dynamics of sensorimotor learning are well described by control loops between visual and tactile PEs that refine reaching movements to a target (Morehead & Xivry, 2021; Sohn et al., 2020), which have applications in improving robotics and brain-machine interfaces (Wolpert et al., 2011). Similar predictive coding principles govern circuit-level interactions between motor and auditory cortices that facilitate learning to speak, walk, or play instruments (Schneider & Mooney, 2018). Collateral fibers from motor cortex send efference copies of motor commands to the specific local inhibitory neurons in auditory cortex needed to suppress expected sounds associated with the movement, leaving excitatory neurons to focus on auditory PEs that indicate motor errors or environmental effects (Schneider et al., 2014, 2018).

Predictive coding has also been proposed as a mechanism for regulating internal perception of the organs and bodily sensations (i.e., interoception) and maintaining energy use in the body (i.e., allostasis) (Barrett & Simmons, 2015; Livneh et al., 2020). The insula is a cortical region that receives primary interoceptive inputs such as pain, itch, and temperature, as well as multimodal sensory inputs (A. D. Craig, 2002; A. D. B. Craig, 2009; Nieuwenhuys, 2012). The insula is heavily connected with MPFC (Evrard, 2019), and this network is proposed to coordinate with visceromotor functions in MPFC to control autonomic and vascular activity (Kleckner et al., 2017). Predictive coding in the insula has also been shown to play a causal role in hunger and thirst (Huang et al., 2021; Kusumoto-Yoshida et al., 2015). Furthermore, modern systems neuroscience techniques such as optogenetics have revealed dopaminergic midbrain neurons encode a much greater diversity of PEs than the classic RPE signal, including non-valenced physical salience of stimuli (Matsumoto et al., 2016), surprising aversive PEs (i.e., how much worse than expected) (Fiorillo, 2013; Jong et al., 2019), and even coding of specific sensory and cognitive variables (Engelhard et al., 2019). Collectively, these multiple control circuits and their feature-specific PEs reflect the rich sources of information in structured environments that can be leveraged for complex behavior. Notably, the logic and principles of these neurobiological computations have greatly informed the development of advanced machine learning algorithms in artificial intelligence (Neftci & Averbeck, 2019).

Predictive coding frameworks also capture hierarchical interactions between these brain networks. Hierarchy is common organizational principle throughout the brain that describes cortical areas such as LPFC (Badre & D'Esposito, 2007; Badre & Nee, 2017) and MPFC (Alexander & Brown, 2018; Holroyd & Verguts, 2021), as well as subcortical areas like the basal ganglia (Haber et al., 2000; Haynes & Haber, 2013) and thalamus (Sherman, 2007). In the context of predictive coding, higher order regions pass predictions down the hierarchy while lower order regions pass prediction errors back up to refine future predictions (Alexander & Brown, 2015, 2018; Bastos et al., 2012). Models based on these hierarchical interactions capture many aspects of motivation and control in neural activity and behavior (Pezzulo et al., 2015, 2018).

Studying cognitive control requires a combination of theoretical models of the potential algorithms used to generate behavior, experimental paradigms designed to dissociate the

relevant variables and computations in those models, and measurements of neuronal activity to test whether those hypothesized models accurately predict how the brain solves the problem (Alexander et al., 2017; Brown, 2014; Collins, 2019; Frank & Badre, 2015; Watabe-Uchida et al., 2017). Systems neuroscience tools in animals allow the greatest access to and control over the control circuits distributed across multiple brain regions, particularly bridging sub-cellular, single neuron, and population levels of analysis. These techniques are the most advanced in rodent models, largely due to the widespread availability of genetic manipulations and other systems neuroscience tools, but the homology between rodent and human PFC is intensely debated (Laubach et al., 2018; Merre et al., 2021; Wise, 2008), which limits cross-species inferences in higher-order functions like cognitive control. Non-human primates have both LPFC and MPFC regions with substantial homology to humans and therefore provide excellent animal models of human brain function, but they face a variety of challenges ranging from technical difficulties recording from multiple regions in large brains to important ethical considerations.

Ultimately, the only way to understand the human brain is to study it directly. In the vast majority of cases, human neuroscience is limited to non-invasive techniques with relatively impoverished resolution. EEG provides millisecond level temporal resolution but poor spatial resolution on the scale of a centimeter, and fMRI provides excellent spatial resolution at the millimeter scale but poor temporal resolution (Nikos K Logothetis, 2008). In rare cases, neurosurgical procedures in human patients provide opportunities for direct electrophysiological recordings from the human brain. For some of the experiments presented in this thesis, intracranial EEG (iEEG) data is recorded from patients with epilepsy that are implanted with stereo-electroencephalography (SEEG) or electrocorticography (ECoG) electrodes to monitor epileptic activity and identify the source of their seizures for potential neurosurgical treatment. This iEEG data provides excellent spatiotemporal resolution on the scale of millimeters and milliseconds, enabling researchers to test hypotheses about both "where" and "when" brain activity occurs.

Each of these techniques carries advantages and disadvantages. One important advantage of high temporal resolution in electrophysiology is the ability to measure activity across different timescales. Microcircuits in the cortex with particular connectivity patterns between excitatory and inhibitory cell types can generate rhythmic activity at specific frequencies (Womelsdorf et al., 2014; Womelsdorf & Everling, 2015), and this coordinated activity across many neurons generates signals strong enough to be detected at the scalp with EEG. Rhythmic neural activity has traditionally been divided into bands as follows: delta (1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-30 Hz), and gamma (40-60 Hz). The amplitude and phase of these different rhythms can provide important information about cognitive functions. For example, **Chapter 2** uses frequency decomposition analyses of scalp EEG data to separate distinct PE signals into delta and theta frequencies.

One important advantage of iEEG relative to non-invasive methods is access to high frequency broadband (HFB) activity. HFB activity is an aperiodic (non-rhythmic) signal present across all frequencies, but it is typically measured from 70-150 Hz to avoid contamination from strong contributions of lower frequency rhythms. HFB power is known to correlate with fMRI and neuronal firing activity (Leszczyński et al., 2020; N K Logothetis et al., 2001; Manning et al., 2009; Rich & Wallis, 2017) on a local scale within a radius of ~1.5 mm (Dubey & Ray, 2019). This

spatial specificity makes it a powerful tool providing access to a critical meso-scale level of population coding in the brain. Another important advantage of HFB activity is its high signal-to-noise ratio (Lachaux et al., 2012), which is characteristic of invasive recordings and enables single-trial analyses that track the minutiae of behavior within subjects instead of averaging across trials to test effects at the group level. These properties are leveraged in **Chapter 3** to investigate the putative sources of reward and surprise PEs observed using EEG in **Chapter 2**, and more generally, they situate iEEG data as a powerful bridge between systems and cognitive neuroscience and between animal and human studies.

One important limitation of iEEG data is that neurosurgical treatment of epilepsy is rare but underutilized given the efficacy in treating phamacoresistant epilepsy (Bjellvi et al., 2019; Dewar & Pieters, 2015). Recording locations are determined solely by clinical needs of each patient, making comparing results at the group level difficult due to small sample sizes and inconsistent coverage of brain regions or subregions. For example, estimates of the proportion of electrodes in a region that represent a given variable may be unreliable in patients with very few electrodes in that region, and patients with limited sampling may skew statistical distributions and obscure true effects.

Overall, this thesis investigates multiple control signals including reward and surprise PEs at the level of both scalp and intracranial EEG, as well as conflict in an iEEG Stroop dataset. **Chapter 2** addresses the longstanding debate over whether the FRN represents a valenced RPE signal or a non-valenced action-outcome prediction error using behavioral modeling and EEG in an interval timing task. These scalp PE signatures have been linked to depression and anxiety symptoms (Cavanagh et al., 2019), highlighting the potential utility of non-invasive biomarkers for diagnosis and treatment of neuropsychiatric conditions. However, the overlap of these components in time, space, and frequency dimensions and the limited spatial resolution of EEG source localization techniques limits understanding of the underlying neural mechanisms. **Chapter 3** addresses these limitations by using the same paradigm and modeling approach but measuring HFB activity from iEEG recordings to map RPE value and magnitude representations in the putative sources of the EEG ERPs. Similarly, **Chapter 4** leverages the spatiotemporal resolution of HFB activity to delineate conflict processing into multiple stages throughout trials in the Stroop task. These iEEG data address classical conflict monitoring theories regarding interactions between MPFC and LPFC and characterize the evolution of conflict processing across a broader network of regions. Collectively, these findings bridge multiple levels of neuroscience to characterize core signatures of reward, surprise, and conflict that facilitate cognitive control computations underlying human behavior.

Chapter 2: Single-trial modeling separates multiple overlapping prediction errors during reward processing in human EEG

Citation:
Hoy, C. W., Steiner, S. C., & Knight, R. T. (2021). Single-trial modeling separates multiple overlapping prediction errors during reward processing in human EEG. *Communications Biology*, *4*(1), 1-17.

Authors:
Colin W. Hoy[1]*, Sheila C. Steiner[1], Robert T. Knight[1,2]

[1]Helen Wills Neuroscience Institute, University of California Berkeley
[2]Department of Psychology, University of California Berkeley

Address for Correspondence:
*Colin Hoy
210-A1 Barker Hall
University of California, Berkeley
Berkeley, CA 94720
hoycw@berkeley.edu

ABSTRACT:

Learning signals during reinforcement learning and cognitive control rely on valenced reward prediction errors (RPEs) and non-valenced salience prediction errors (PEs) driven by surprise magnitude. A core debate in reward learning focuses on whether valenced and non-valenced PEs can be isolated in the human electroencephalogram (EEG). We combine behavioral modeling and single-trial EEG regression to disentangle sequential PEs in an interval timing task dissociating outcome valence, magnitude, and probability. Multiple regression across temporal, spatial, and frequency dimensions characterized a spatio-tempo-spectral cascade from early valenced RPE value to non-valenced RPE magnitude, followed by outcome probability indexed by a late frontal positivity. Separating negative and positive outcomes revealed the valenced RPE value effect is an artifact of overlap between two non-valenced RPE magnitude responses: frontal theta feedback-related negativity on losses and posterior delta reward positivity on wins. These results reconcile longstanding debates on the sequence of components representing reward and salience PEs in the human EEG.

INTRODUCTION:

Adaptive behavior requires predicting relationships between stimuli, actions, and outcomes to decide which choices maximize reward. Predictive coding is a general computational framework that learns these mappings based on surprise as measured by

Reward prediction errors (RPEs) convey the valence (better or worse?) and magnitude (how surprising?) of a reward relative to expectations. RPEs are encoded by transient firing rate modulations of midbrain dopamine (DA) neurons (Eshel et al., 2016; Schultz et al., 1997; Zaghloul et al., 2009) , and early work in reinforcement learning (RL) established that simple and efficient model-free algorithms such as temporal difference learning can account for basic reward learning phenomena by using RPEs to update the expected value associated with a stimulus without the need for additional modeling of the influence of actions (R. S. Sutton & Barto, 1998).

In contrast, sensorimotor and abstract cognitive control loops depend on non-valenced PEs to learn action-outcome contingencies (Wessel, 2018). For example, the predicted response-outcome (PRO) model asserts that the medial prefrontal cortex (MPFC) controls action selection by tracking salient, unexpected outcomes independent of valence (Alexander & Brown, 2011). Valenced and non-valenced PEs have complimentary contributions to learning (Rouhani & Niv, 2021), and current model-based reinforcement learning algorithms combine both valenced reward and non-valenced state, action, and outcome predictions to account for more complex behaviors by modeling relationships between an agent's actions and the environment (Gershman & Uchida, 2019; Neftci & Averbeck, 2019; Silver et al., 2016; Wang et al., 2018). Recent animal investigations of reward and control emphasize the multitude of learning signals represented by subpopulations of midbrain DA neurons (Bromberg-Martin et al., 2010), including aversive outcomes and stimulus salience independent of valence (Fiorillo, 2013; Jong et al., 2019; Matsumoto et al., 2016), as well as sensory, cognitive, and reward variables (Engelhard et al., 2019).

Non-invasive human electroencephalography (EEG) findings have identified a variety of learning-related event-related potentials (ERPs) and time-frequency signatures (Glazer et al., 2018), but their specific relationships to reward and control PEs are still debated. Seminal early studies identified a posterior scalp positivity generated ~300 ms after detection of an infrequent stimulus called the P3 (Donchin et al., 1978; S. Sutton et al., 1965) and a fronto-central negativity elicited ~80 ms after incorrect compared to correct responses called the error-related negativity (ERN) (Falkenstein et al., 1991; Gehring et al., 1993). Subsequent extensive literature has revealed these ERPs to be part of large families of similar components. P3-style ERPs are characterized by slow ramping dynamics peaking from ~300-600 ms and in delta frequencies (~1-4 Hz) (O'Connell et al., 2012) with at least two different scalp topographies. The P3b has a posterior maximal topography and is elicited by detected events conveying various forms of salient information leading to working memory and model updates, while the P3a has an earlier latency, fronto-central topography and is generated by attention and orienting to novel, task-irrelevant stimuli (Knight, 1984; Mars et al., 2008; Seer et al., 2016) (see (Polich, 2007) for review). Note that both the P3a and P3b are not unitary physiological events but rather reflect the summed activity of multiple intracranial sources (Kam et al., 2016). ERNs are related to a family of faster latency (~200-300 ms) N2 negativities over fronto-central sensors generated in part by phase-locking in theta frequency (~4-8 Hz) MPFC activity triggered by unexpected events requiring behavioral adjustment (Cavanagh & Frank, 2014; Hauser et al., 2014; E. H. Smith et al., 2015) (for reviews, see (Folstein & Petten, 2008) for stimulus-locked N2s and (Gehring et al., 2012) for response-locked ERNs).

Reward feedback conveys multiple informative variables with varying salience and elicits both N2- and P3-like ERPs that overlap in time and space, leading to longstanding debates over which components track different aspects of feedback. Foundational studies focused on the N2-like feedback-related negativity (FRN) occurring ~200 ms at frontal sites after loss feedback compared to wins (Gehring & Willoughby, 2002; Miltner et al., 1997). An early, influential RL theory (originally called the RL-ERN theory) proposed the FRN represents valenced, quantitative RPE value driven by midbrain DA projections to MPFC (Holroyd & Coles, 2002). This hypothesis predicts FRN sensitivity to all the feedback properties that determine RPEs: outcome valence, magnitude, and probability (Nieuwenhuis et al., 2004). However, two recent meta-analyses found mixed evidence for magnitude and probability effects (Sambrook & Goslin, 2015; Walsh & Anderson, 2012). Reports of larger FRNs to unexpected positive outcomes (Ferdinand et al., 2012; Hauser et al., 2014; Oliveira et al., 2007; Soder & Potts, 2018; Talmi et al., 2013) led to an alternative account called the salience theory, which proposes the FRN represents the degree of surprise of an outcome regardless of valence, similar to non-valenced action-outcome PEs driving cognitive control in the PRO model.

A third prominent proposal called the independent coding hypothesis posits the FRN represents binary reward valence instead of scalar RPE value, while the subsequent P3 tracks non-valenced RPE magnitude (Hajcak et al., 2006; Philiastides et al., 2010; Yeung & Sanfey, 2004). This interpretation is complicated by more recent observations of a P3-like positivity called the Reward Positivity (RewP) that tracks RPE magnitude specifically on positive outcomes (see (Proudfit, 2015) for review) (Baker & Holroyd, 2011; Becker et al., 2014; Foti et al., 2011, 2015; Holroyd et al., 2008; Meadows et al., 2016) that overlaps in space and delta frequencies with other non-valenced P3 components (Bernat et al., 2015; Williams et al., 2021). Importantly, this suggests losses and wins generate distinct FRN and RewP ERPs with opposite polarities that interact to some degree depending on their overlap in time and space (Baker & Holroyd, 2011; Holroyd et al., 2011). As a result, it remains unclear after decades of research whether FRN and/or RewP ERPs are driven by valenced RPEs, non-valenced salience PEs, or one of the valenced/non-valenced input variables contributing to these PEs (e.g., outcome valence, magnitude, or probability).

An important challenge in resolving this debate is disentangling overlapping ERP components. For example, the FRN is commonly measured by averaging ERP amplitude across time, but the epochs used in mean window analyses cover both classic N2 and P3 windows (Sambrook & Goslin, 2015). The FRN is also often measured by the peak-to-peak amplitude difference between the N2 and the preceding P2 positivity to account for influences of early P3 ramping. However, individual ERP peaks are variable and may not correspond to unique neural sources (Gruber et al., 2005; Luck, 2014; Sauseng et al., 2007), rendering their use as reference measures questionable. For example, the P2 shows confounding effects of surprising positive reward (Potts et al., 2006; Williams et al., 2021). Difference waves are commonly used to isolate target variables such as valence by subtracting ERPs across conditions matched for confounding variables (e.g., magnitude or probability). Indeed, win-loss difference waves are commonly used as the operational definition of the FRN and RewP (Krigolson, 2018; Sambrook & Goslin, 2015; Williams et al., 2017, 2021). However, this subtraction logic cannot determine which ERP nor which condition were modulated and thus is not well suited to unraveling the dual multiplicity of ERPs and learning signals, particularly since the FRN and RewP may have

distinct neural mechanisms supporting different computational roles (Bai et al., 2015; Cavanagh et al., 2019; Meadows et al., 2016). Here, we use the term FRN to refer to the early, feedback-locked frontal N2-like ERP that is most prominent on but not specific to losses and the term RewP to refer to the subsequent feedback-locked P3-like ERP specifically following wins. We return to the definitions of FRN and RewP components and their relationship to other ERPs and difference wave contrasts in the discussion.

The overlap of ERPs in time-domain analyses has made time-frequency decompositions an important tool for separating the FRN and RewP into theta and delta frequencies, respectively. Several studies have shown theta is sensitive to negative RPEs (Bernat et al., 2011, 2015; Cavanagh et al., 2019; Foti et al., 2015), but it has also been reported to track non-valenced probability and magnitude (Cavanagh et al., 2010, 2012; Hajihosseini & Holroyd, 2013). Likewise, delta activity is linked to both non-valenced surprise (Mars et al., 2008; Seer et al., 2016; Wessel & Huber, 2019) and positive RPEs (Bernat et al., 2015; Cavanagh, 2015; Cavanagh et al., 2019; Foti et al., 2015; Proudfit, 2015). These mixed results highlight how individual measures of neural activity may be insufficient to distinguish between ERP components such as the RewP, P3a, and P3b that overlap in one or more dimensions (e.g., frequency) but correspond to distinct cognitive variables.

Single-trial modeling methods have provided clarity into the theoretical debates in reward processing EEG signals. RL models estimate latent cognitive variables such as reward expectations, which can change the subjective meaning of and ERP response to reward feedback (Hajcak et al., 2007; Ichikawa et al., 2010; Nieuwenhuis et al., 2004). Importantly, model-based single-trial correlation or regression analyses provide enhanced statistical power than traditional categorical statistics applied to condition-averaged data (Aarts et al., 2014; Frömer et al., 2018), which can be leveraged to map the evolution of cognitive variables in high resolution across multiple time points, channels, and frequencies. This framework has been used to separate overlapping components (Cavanagh, 2015; Cavanagh et al., 2012, 2019; Fischer & Ullsperger, 2013; Foti et al., 2015). In particular, this approach enables data-driven discovery of RL variable representations in EEG data that aren't time-locked to ERP peaks (Collins & Frank, 2018; Collins & Frank, 2016), and allows formal model comparisons between competing hypotheses (Mars et al., 2008; Seer et al., 2016; Wessel & Huber, 2019). Here we use these methods to compare the predictions of the main competing hypotheses across the different measurements of the FRN and RewP.

Our goal was to combine these modeling and signal analysis tools to provide a comprehensive assessment of the core theoretical and measurement issues underlying the FRN and RewP debate. We start by estimating PEs from individual participant behavior in an interval timing task designed to dissociate the valence, magnitude, and probability of outcomes. We then use formal model comparisons to test the predictive power of outcome and PE features central to RPE, salience, and independent coding hypotheses using mixed-effects multiple regression analyses applied across temporal, spatial, and spectral dimensions of wins and losses to separate overlapping components in initial and replication cohorts. To relate our results to previous EEG literature, we perform analogous modeling of mean window and peak-to-peak FRN and RewP metrics, in addition to quantifying the overlap in our reward feedback ERPs based on correlations with canonical N2 and P3 benchmark ERPs measured in a three-tone oddball task collected from a subset of the same participants.

We found that when modeling wins and losses together as done in standard analyses, early, frontal theta activity underlying the FRN is best described by valenced RPE value, while non-valenced RPE magnitude and probability effects drive later delta activity consistent with P3 ERPs. Model comparisons show these data are better explained by RL-based PEs than outcome features and confirm the importance of predictive coding principles. Mean window and peak-to-peak FRN analyses replicated these RPE value effects but also showed non-valenced effects and were unstable across the two cohorts. These results suggest the FRN represents a scalar, quantitative RPE while two P3 ERPs encode non-valenced RPE magnitude and outcome probability. However, these conventional analyses combining wins and losses confound the FRN on negative trials and the RewP on positive trials, and comparisons to oddball task ERPs suggest a mixture of N2 and P3 contributions to the RPE value window. Indeed, modeling EEG amplitude, topographies, and time-frequency power after separating wins and losses reveals that the valenced RPE value effect is an artifact of non-valenced RPE magnitude driving two overlapping FRN and RewP components in theta for losses and delta for wins respectively, while the late frontal probability effect in delta is stable across outcomes. Finally, we use subjective ratings obtained in a follow up behavioral experiment to confirm these EEG results cannot be explained by subjective biases in reward contingencies. Collectively, these results provide strong evidence that human EEG following reward feedback is composed of a sequence of multiple overlapping neurophysiological signatures best accounted for by specific PEs in a predictive coding framework.

RESULTS:

We collected and analyzed EEG from 32 cognitively normal young adults split into initial (n = 15) and replication (n = 17) cohorts. Participants performed an interval timing task designed to dissociate the key variables underlying the PEs central to the debate between RL, salience, and independent coding accounts: outcome valence, magnitude, and probability. At the beginning of each trial, participants saw a target zone cue whose size indicated the temporal range of responses tolerated as correct (Figure 1a). Participants then estimated the temporal interval by means of extrapolation from visual motion, and received audiovisual feedback indicating their reaction time (RT) and whether it was within or outside the tolerance (i.e. a win or loss). After each trial, the error tolerance was titrated by two staircase algorithms (Fig. 1b) to clamp accuracy at $82.7 \pm 1.7\%$ and $18.1 \pm 2.5\%$ (mean ± SD) in easy and hard blocks, respectively (Fig. 1c). This design dissociates outcome valence and probability to separate valenced and non-valenced PE features by comparing surprising wins and losses. Neutral outcomes with no RT feedback were also delivered on a random subset of 12% of trials to manipulate outcome magnitude as another source of surprise.

Behavioral Modeling:

To directly compare the predictive power of RL, salience, and independent coding theories, we used computational modeling of individual participant behavior to derive single-trial estimates of valenced RPE value, as well as two sources of salience: non-valenced RPE magnitude and outcome probability. For each participant, we used logistic regression to fit the

relationship between the interval tolerance and binary win/loss outcomes across the entire session (Fig. 1d; see inset for group model fits). The resulting model yields the probability of that participant winning for any given tolerance, which was then linearly scaled to the range of rewards (1, 0, and -1 for winning, neutral, and losing outcomes) to quantify expected value for every trial. We then contrasted expected value with actual outcomes to obtain single-trial RPE values and derived the absolute value of RPEs to obtain RPE magnitudes. Outcome probability was determined by the frequency of each outcome in each condition. Notably, RPE values for neutral outcomes were non-zero and switched valence across blocks (negative for easy and positive for hard blocks; see model predictions in Supplementary Figure 1a), suggesting they could be interpreted as omissions of the expected outcome.

To compare this RL model to simple win/loss contrasts standard in the FRN and RewP literature that don't account for predictive coding, we also computed an outcome-based model composed of reward value (1, 0, and -1), reward magnitude (absolute value of reward value), and outcome probability. Finally, to test the hypothesis that the FRN tracks binary valence but not scalar value or magnitude, we added a modified outcome model that replaced the outcome value on neutral trials with valence based on reward omission relative to expected value (see Sup. Fig. 1b for outcome-based model predictions).

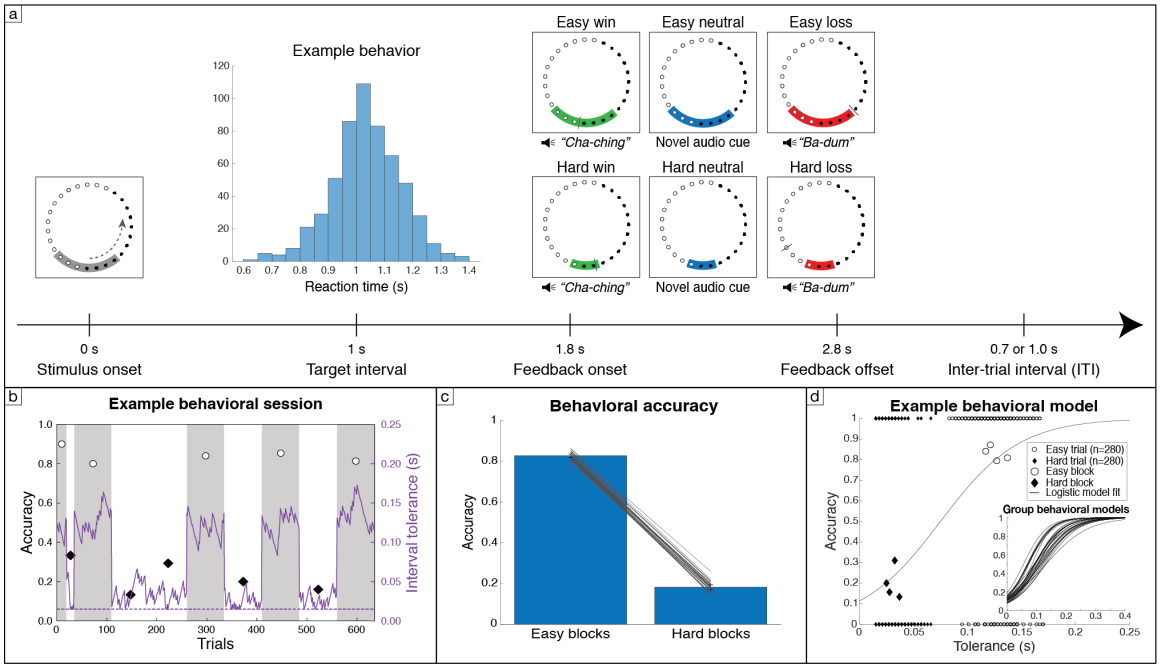


Figure 1: Task design, performance, and behavioral modeling of prediction errors. (a) Participants pressed a button timed to the estimated completion of lights moving around a circle. The gray target zone cue displayed error tolerance around the 1 s target interval. An example participant RT distribution is centered at the target interval. Audiovisual feedback is indicated by the tolerance cue turning green for wins and red for losses. A black tick mark displayed RT feedback. On 12% of randomly selected trials, blue neutral feedback was given with no RT marker. (b) Example recording session for one participant for training (first 35 trials) and experimental blocks. Staircase adjustments of tolerance are plotted in solid purple, and the dotted purple line indicates the minimum bound on tolerance at ±15 ms. Accuracy for easy and hard blocks is plotted as white circles on gray backgrounds and black diamonds on white backgrounds, respectively. (c) Separate staircase procedures resulted in group accuracies of (mean ± SD) 82.7 ± 1.7% for easy and 18.1 ± 2.5% for hard blocks. Error bars indicate standard deviation across participants, with individual participant accuracy overlaid in gray. (d) Tolerance and outcome data for the same

example participant. Larger markers show block level accuracy; smaller markers show binary single trial outcomes. Model fit using logistic regression provides single trial estimates of win probability, which can be converted to expected value. Inset shows win probability curves across all participants.

Single-trial regression reveals a spatio-temporal cascade of PE components

According to the RL theory, the valenced RPE values derived from our behavioral model should predict FRN amplitude (Holroyd & Coles, 2002; Nieuwenhuis et al., 2004), while the independent coding hypothesis suggests the FRN is sensitive only to binary reward valence and not scalar value (i.e., combined valence and magnitude)(Hajcak et al., 2006; Philiastides et al., 2010; Yeung & Sanfey, 2004). In contrast, the salience theory predicts that FRN amplitude should scale with how surprising each outcome is (Alexander & Brown, 2011; Cavanagh et al., 2012; Oliveira et al., 2007), which increases with non-valenced RPE magnitude and decreases with outcome probability. Importantly, disentangling these outcome and PE features and resolving this debate depends on addressing the overlapping ERP component problem that confounds traditional mean window and peak-to-peak amplitude measurements (Krigolson, 2018). Here, we leverage known timing differences between early FRN and later P3 activity by predicting single-trial evoked amplitude at every time point from 50 to 500 ms post-feedback using mixed-effects multiple regression analyses to adjudicate between different PEs. Within this framework, we use formal model comparisons to test whether RL models combining expected value, RPE value, RPE magnitude, and outcome probability predict ERP amplitudes better than standard models composed of outcome magnitude, probability, and either scalar value or binary valence. These analyses were conducted separately at frontal Fz and posterior Pz electrodes to assess the FRN and longer latency positivities.

Grand-average ERPs show FRN peaks ~200-250 ms post-feedback at frontal electrode Fz and P3 peaks ~300-350 ms at posterior Pz (Fig. 2a and 2b; see also Sup. Fig. 9a). Model coefficients are plotted below in Figures 2c and 2d for the best performing model, which includes RL features of expected value, RPE value, RPE magnitude, and probability (see Sup. Fig. 2 for difference wave contrasts for each of these variables). The most prominent result is a large effect of RPE value peaking in the FRN window at 216 ms in electrode Fz (β_{max} = 4.572, $q_{FDR} < 10^{-10}$; Fig. 2c). In accordance with the RL theory, this positive model coefficient indicates that more negative RPE values are associated with more negative amplitudes. In other words, larger FRNs with more negative amplitude are associated with worse-than-expected outcomes, and better-than-expected outcomes drive more positive amplitude. The RPE value effect decreases as the FRN subsides, and a significant positive RPE magnitude effect emerges. This RPE magnitude effect is maximal in electrode Pz at 308 ms (β_{max} = 1.703, $q_{FDR} < 10^{-10}$; Fig. 2d), indicating larger non-valenced RPE magnitudes are associated with larger P3 amplitudes. This result replicates the non-valenced effect of magnitude matching a posterior P3b predicted by the independent coding hypothesis (Yeung & Sanfey, 2004), but this analysis included both positive and negative outcomes and cannot disambiguate potential contributions of a RewP specific to positive RPEs.

The temporal coincidence of significant model coefficients for RPE value and magnitude in the epoch between FRN and P3 peaks suggests that previous findings supporting the salience theory could be explained by component overlap confounds, particularly since FRN and RewP amplitude is commonly quantified as the mean amplitude from approximately 228-344 ms

(Sambrook & Goslin, 2015), an epoch that encompasses both FRN and P3 activity. To compare our time-resolved single-trial regression analyses to metrics more commonly used in the field, we computed mean window and peak-to-peak estimates of the FRN, as well as mean window measures of the P3 (Sup. Fig. 3a). As predicted, RL model coefficient results using traditional mean window and peak-to-peak estimates of the FRN confirmed the strong RPE value effect. However, they also show significant but divergent non-valenced effects, with mean window predicted by probability and peak-to-peak predicted by RPE magnitude (Sup. Fig. 3b and 3c). Importantly, these conflicting non-valenced FRN effects using traditional methods were unreliable across replication cohorts (see Supplementary Table 1 for RL model results across cohorts for FRN and P3 mean window and peak-to-peak metrics). Similarly, RPE magnitude was significant in the RL model regression for the P3 mean window analysis at Pz in both cohorts (Sup. Fig. 3d), but RPE value was significant in only one cohort, confirming that mean window and peak-to-peak metrics are less reliable.

Lastly, the probability predictor reveals a significant relationship to ERP amplitude that peaks later at 380 ms in Fz (β_{max} = -2.842, $q_{FDR} < 10^{-10}$; Fig. 2c). Observing this late, frontal positivity in response to unlikely outcomes was possible because we dissociated outcome probability and RPE magnitude as two distinct sources of salience.

Time-resolved model comparisons for Fz and Pz (Fig. 2e and 2f, respectively) are plotted as Akaike Information Criterion (AIC) values relative to a baseline model containing only random intercepts capturing each participant's mean amplitude across conditions. Lower AIC values mean better model performance. Figure 2e shows that RL-based models including RPE value capture more variance in EEG amplitude during the FRN window at Fz than outcome value or valence models (see Sup. Fig. 4c-f for outcome-based model coefficients; see Sup. Table 2 for AIC model comparison value averaged within peak model coefficient windows). The model with binary outcome valence performs better than the outcome-based model with scalar value, and these model results hold for both mean window and peak-to-peak estimates of FRN amplitude (Sup. Fig. 3e and 3f). The only difference between outcome value and outcome valence models is whether neutral trials in easy and hard blocks are treated as outcomes with identical values (zero) or as omissions of expected rewards with opposite valence (1 or -1). Similarly, Figure 2f shows that the RL-based models outperform the outcome value and valence models at Pz throughout the FRN and P3 epochs, which is confirmed by model comparisons using the mean window estimates of P3 amplitude in Supplementary Figure 3g. These results confirm that FRN and P3 ERPs are best viewed through the predictive coding lens of PEs.

Since RPE magnitude and probability are correlated (see Methods), we also used model comparisons as a control to examine whether the variance explained by these two non-valenced salience PEs dissociated in time and space. When the RPE magnitude predictor is excluded, RL model performance drops at Pz near the peak of the P3 (Fig. 2f). When the probability predictor is excluded, performance drops later during the downslope of the P3 at Pz and from approximately 350-450 ms at Fz (Fig. 2e and 2f; model coefficients plotted in Sup. Fig. 4g-j), confirming that these two sources of salience correspond to separable EEG components (see also dissociations with Oddball ERP correlations in Fig. 5).

Collectively, these results characterize a cascade of multiple PEs unfolding during the FRN and RewP epochs in reward processing EEG, starting with an early, frontal, valenced RPE value signal in the FRN time window, followed by a later, posterior, non-valenced RPE

magnitude effect in the P3 window, and finally a later, fronto-central probability effect. These findings replicate evidence supporting the RL account of the FRN as a scalar, quantitative RPE instead of binary valence (Holroyd & Coles, 2002; Nieuwenhuis et al., 2004; Walsh & Anderson, 2012) and the independent coding proposal's separation of early valenced effects in the FRN window from two later non-valenced P3 effects(Philiastides et al., 2010; Yeung & Sanfey, 2004). We re-assess the interpretations of these effects after characterizing their spatial and frequency distributions, their correspondence with benchmark N2 and P3 ERPs from the oddball task, and most importantly when separating losses and wins to avoid confounding overlapping FRN and RewP ERPs.

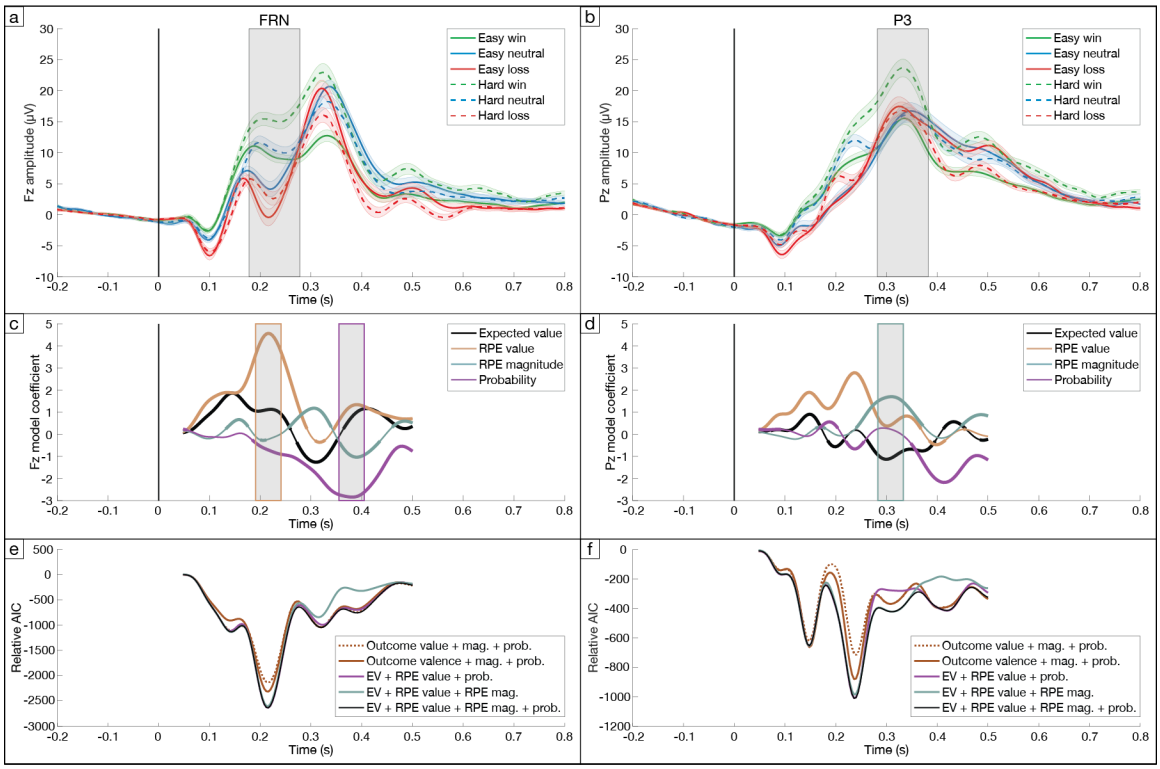


Figure 2: Single-trial modeling of ERP amplitude reveals a sequence of prediction errors. (a) Feedback-locked grand-average ERPs at Fz plotted for each condition, with shaded error bars indicating standard error of the mean across participants. FRN is evident as prominent negative deflections at Fz ~200 ms post-feedback. Gray shading shows 100 ms window used to average FRN amplitude. (b) Same for posterior electrode Pz, which shows large P3 positivities at ~300 ms. Gray shading shows 100 ms window used to average P3 amplitude. (c) Model coefficients from single-trial multiple regression at each time point from 50-500 ms at Fz show a strong early peak of valenced RPE in the FRN window, followed by later probability effect in the P3 time range. Bolding indicates significant time points ($q_{FDR} < 0.05$). Gray shading shows 50 ms windows used to average ERP amplitude at maximal RPE value and probability effects. (d) Same for electrode Pz. Note the increased non-valenced RPE magnitude coefficient in the P3 window. Gray shading shows 50 ms window used to average ERP amplitude at the maximal RPE magnitude effect. (e) Comparison of model performance at Fz over time via Akaike Information Criteria (AIC; more negative indicates higher performance) relative to baseline model. RL model with RPE value beats outcome-based models in the FRN window, and model performance drops in the P3 window when probability is excluded. (f) Same for Pz. RL model performance drops during P3 peak window when RPE magnitude is excluded, while RL model performance drops during late positivity window when probability is excluded. See Sup. Fig. 5 for RL model performance (R^2) across electrodes, which peaks in the FRN time window for Fz and the P3 time window for Pz.

Scalp topography dynamics delineate PE components

To further disentangle the spatial topographies of sequential PE effects and relate them to known N2 and P3 scalp distributions, we applied single-trial multiple regression analysis across all electrodes in three 50 ms windows centered on the peaks of each model coefficient from the time-resolved analysis (see highlighted windows in Fig. 2c and 2D). Model coefficients for valenced RPE value, non-valenced RPE magnitude, and outcome probability are plotted as scalp topographies in Figure 3 (see Sup. Fig. 6 for evoked amplitude topographies by condition). This analysis confirmed that the largest effect was the valenced RPE value in the early window at anterior frontal sites (β_{max} = 3.887 in 216 ms window at electrode F1, $q_{FDR} < 10^{-10}$), which then dropped off in magnitude in the middle window before a smaller resurgence in the late window. The non-valenced RPE magnitude effect was maximal in the middle 308 ms window at posterior parietal electrodes (β_{max} = 1.859 at electrode PO3, $q_{FDR} < 10^{-10}$). Finally, the probability effect was focused in fronto-central electrodes in the later 380 ms window (β_{max} = -2.523 at electrode C4, $q_{FDR} < 10^{-10}$). The spatio-temporal distributions of these effects confirms the association between valenced RPE value and the early FRN epoch in frontal electrodes (Fischer & Ullsperger, 2013; Gehring & Willoughby, 2002; Holroyd & Coles, 2002), while non-valenced RPE magnitude shows a posterior, parietal distribution matching the P3b (Mars et al., 2008; Polich, 2007).

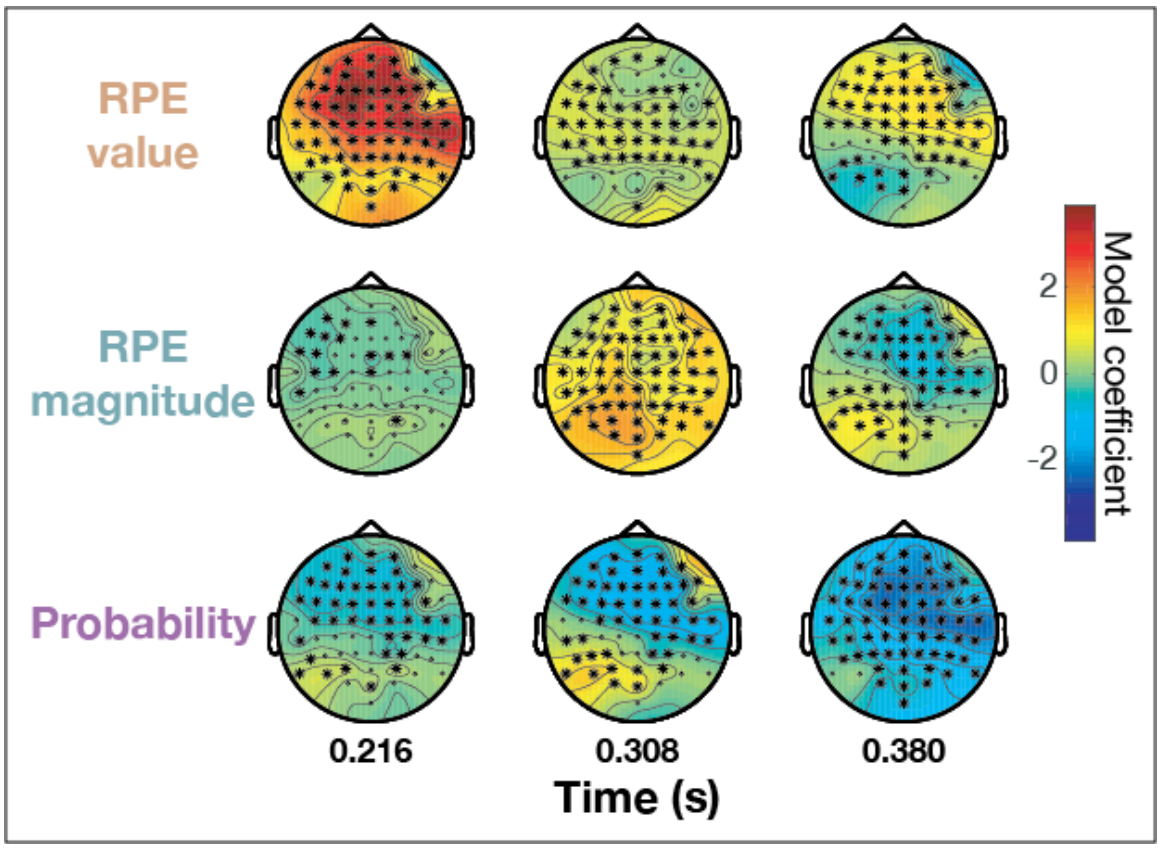


Figure 3: Spatio-temporal dynamics of prediction errors across ERP scalp topographies. Single-trial regression over all 64 electrodes is computed for three 50 ms windows centered on the largest peak in model coefficients for RPE value (216 ms at Fz), RPE magnitude (308 ms at Pz), and probability (380 ms at Fz) from Fig. 2. Stars indicate significant electrodes ($q_{FDR} < 0.05$). Valenced RPE value shows a frontal distribution in the early window (top left).

Non-valenced RPE magnitude is maximal at posterior electrodes in the middle window (center). Probability is maximal at fronto-central sensors in the late window (bottom right). See Sup. Fig. 6 for evoked potential voltage topographies.

Single-trial regression of time-frequency power dissociates PE effects in theta and delta bands

Although the N2/FRN and P3/RewP are defined as ERP phenomena, their waveform characteristics are associated with theta (4-8 Hz) and delta (1-4 Hz) frequencies, respectively (Bernat et al., 2007; Cavanagh et al., 2010; Foti et al., 2015; Polich, 2007; E. H. Smith et al., 2015). To further dissociate contributions of RPE value, RPE magnitude, and probability to these overlapping components, we extracted feedback-locked time-frequency representations (TFRs) of evoked power at Fz (Sup. Fig. 7) and Pz (Sup. Fig. 8). Single-trial multiple regression with our RL model across wins and losses revealed a strong negative relationship between RPE value and theta power (β_{max} = -1.405 at [292 ms, 6 Hz] in electrode Fz, $q_{FDR} < 10^{-10}$; Fig. 4a). The delayed peak of this theta effect relative to the FRN latency highlights the spread of this RPE value effect spanning across several cycles of theta (see also theta frequency fluctuations of RPE value coefficients in Fig. 2c). In contrast, RPE magnitude significantly predicted posterior delta power (β_{max} = -0.447 at [260 ms, 3 Hz] in electrode Pz, $q_{FDR} = 5.05 * 10^{-10}$; Fig. 4b), consistent with the upward ramp of the P3. Probability best predicts 4 Hz power at 392 ms post-feedback (β_{max} = -0.871 in electrode Fz, $q_{FDR} < 10^{-10}$; Fig. 4a). Overall, more negative RPEs predicted stronger theta power in accordance with the RL theory of the FRN (Holroyd & Coles, 2002), while delta power associated with P3 ERPs increased with larger RPE magnitudes and more unlikely events as predicted by the independent coding hypothesis (Yeung & Sanfey, 2004).

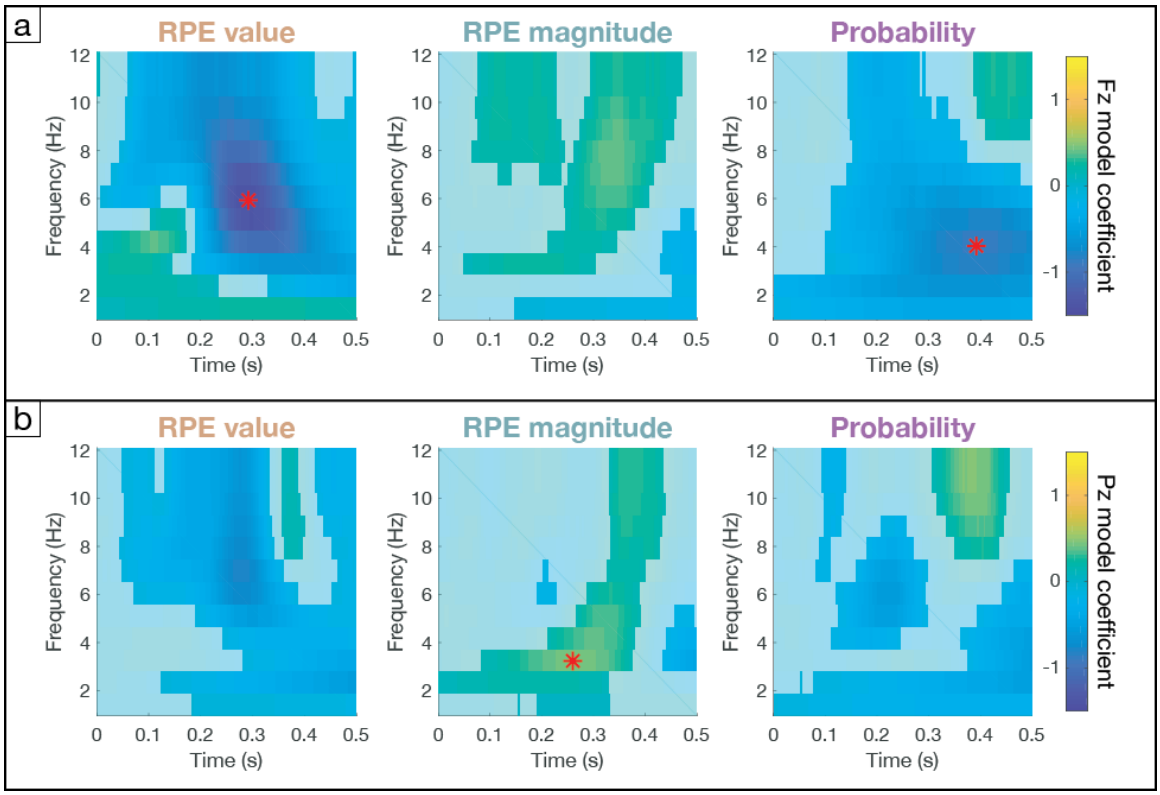


Figure 4: Time-frequency power signatures of prediction errors. (a) Model coefficients fit to evoked power at each time-frequency point for frontal electrode Fz, with non-significant points ($q_{FDR} > 0.05$) plotted opaquely. Red stars indicate maximal coefficients for each model predictor across both electrodes. Valenced RPE value coefficients peak in frontal theta power, and non-valenced probability coefficients peak later in frontal delta power. (b) Same for Pz. Non-valenced RPE magnitude coefficients peak in posterior delta power.

Correlations with Oddball task ERPs quantify contributions of overlapping components

To further disambiguate the contributions of overlapping FRN and P3 ERPs to RPE value, RPE magnitude, and probability effects, we compared feedback-locked ERPs in the Target Time task to reference N2 and P3 ERPs in a canonical Oddball task. A subset of the Target Time participants (n = 22) performed a three-tone Oddball task in which they attended to a stream of audiovisual stimuli and were instructed to press a button after rare targets (12.25% of trials) among common standard (75.5% of trials) and rare, task-irrelevant novel stimuli (12.25% of trials; see Methods for details). Figure 5a shows frontal Oddball ERPs at Fz with a prominent N2c in the target condition and a smaller N2b in the novel condition, both associated with control allocation in response to unexpected rare stimuli but differing in task contingencies (Cavanagh et al., 2011; Folstein & Petten, 2008). Posterior Oddball ERPs at Pz in Figure 5b and spatial topographies in Figure 5c show novel stimuli elicit a central P3a associated with bottom-up orienting of attention, and target stimuli elicit a classic posterior P3b related to top-down model updating (Polich, 2007). Since the timing, topographies, frequency characteristics, and potential intracranial sources of the FRN and RewP are shared with ERPs in the larger families of N2 and P3 components, respectively (Cavanagh et al., 2011; Donkers et al., 2005; Folstein & Petten, 2008; Hajihosseini & Holroyd, 2013; Holroyd et al., 2008; Wessel et al., 2012; Wessel & Huber, 2019), we used individual participant ERPs averaged in a 50 ms window centered on the grand-average N2 and P3 peaks in target and novel Oddball conditions as benchmarks to determine the relative contributions of N2-like and P3-like activity to ERPs underlying the PE effects observed in our multiple regression analyses (as highlighted by colored window overlays in Fig. 2c and 2d).

Group-level correlations between Oddball ERPs and Target Time condition ERPs averaged at Fz in the RPE value peak window showed significant relationships between both Novel N2b and Target N2c ERPs and negative RPE conditions, with weaker but significant relationships between the stronger target N2c and positive valence RPE conditions (Fig. 5d). These relationships suggest strong contributions of N2-like activity to the RPE value epoch in the Target Time task, particularly for conditions requiring control allocation such as button presses to targets in the Oddball task and adjustments in RTs following losses in the Target Time task (see elevated FRN mean window and peak-to-peak estimates in Sup. Fig. 3a). However, significant correlations between the fronto-central P3a from the novel Oddball condition and four out of six Target Time conditions indicate additional influences of P3-like activity, reinforcing the risks of interpreting reward processing in the FRN/RewP epoch as a unitary phenomenon.

Comparisons between Oddball and Target Time ERPs at Pz in the RPE magnitude window show significant correlations specifically with Novelty P3a and Target P3b ERPs (Fig. 5e). Interestingly, the fronto-central Novelty P3a correlated with all Target Time conditions

except hard losses, while the posterior Target P3b correlations were weaker and only significant for wins and easy neutral outcomes, despite these analyses being conducted at posterior Pz where the RPE magnitude effect was maximal. No Oddball ERPs correlated significantly with Target Time ERP amplitudes in the later probability window at Fz (Fig. 5f). These results suggest the RPE magnitude effect is mainly driven by P3-like activity, while the later frontal probability effect has no clearly analogous ERP in the Oddball task.

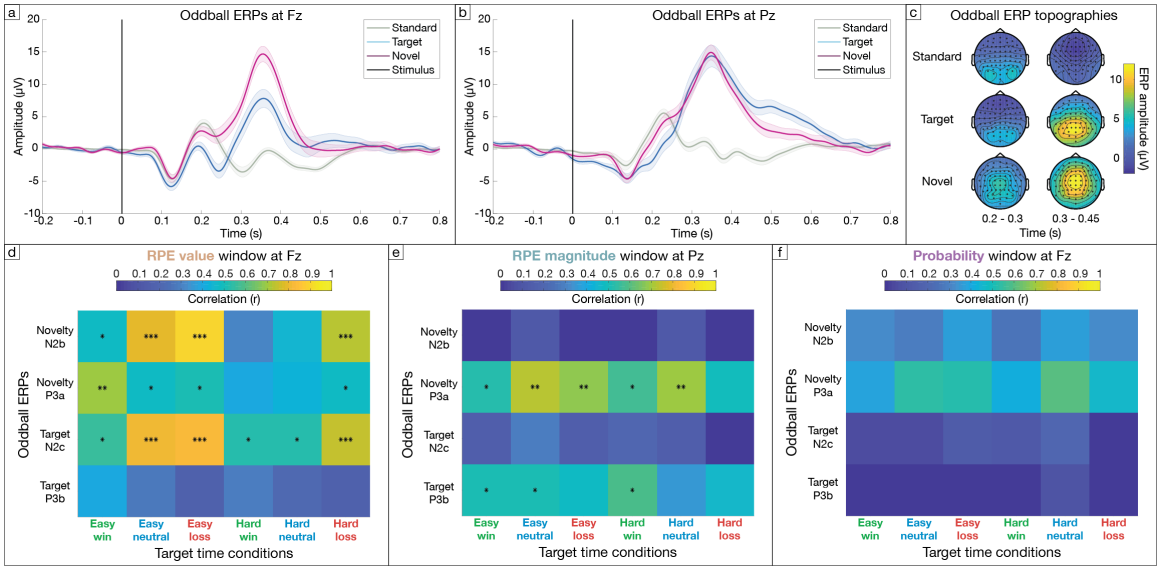


Figure 5: ERP comparison between Target Time feedback and 3-tone Oddball tasks. (a) Stimulus-locked grand-average Oddball ERPs at Fz plotted for each condition, with shaded error bars indicating standard error of the mean across participants. N2 is evident ~250 ms in Target and to a lesser extent in Novel conditions. (b) Same for Pz. P3 is evident ~350 ms in Target and Novel conditions. (c) Spatial topographies of grand-average ERPs averaged during N2 (200-300 ms) and P3 (300-450 ms) windows. Note the posterior distribution of the Target P3b and central distribution of the Novel P3a. (d) Correlation matrix between N2 and P3 ERPs in Oddball task and mean amplitude of 50 ms window centered on maximal RPE value effect at Fz in Target Time conditions. Asterisks indicate significance at $q_{FDR} < 0.05$ (*), $q_{FDR} < 0.01$ (**), and $q_{FDR} < 0.001$ (***). Note strong correlation between Oddball N2s and Target Time conditions with negative valence, especially for the Oddball Target N2c. (e) Same but correlating N2 and P3 amplitudes from Oddball task with mean amplitude of 50 ms window centered on maximal RPE magnitude effect at Pz in Target Time conditions. (f) Same but correlating N2 and P3 amplitudes from Oddball task with mean amplitude of 50 ms window centered on maximal probability effect at Fz in Target Time conditions.

Separating outcomes by valence dissociates FRN and RewP RPE magnitude effects

The mixture of N2 and P3 contributions to the RPE value window identified by Oddball ERP correlations revives concerns that the sequence of PE effects described above may be confounded by component overlap. In particular, our multiple regression results thus far account for competition between valenced and non-valenced PE predictors, but like difference waves, they still rely on contrasts between negative and positive outcomes that cannot distinguish between overlapping ERPs. To disentangle the roles of FRN and RewP ERPs in the sequence of PEs described above, we repeated our ERP and TFR multiple regression analyses separately for negative (easy loss, easy neutral, and hard loss) and positive (easy win, hard win,

Model coefficients at Fz confirm that RPE magnitude exerts opposite effects across negative and positive valenced outcomes in the FRN window (Fig. 6a and 6b). On negative outcomes, the RPE magnitude predictor shows a significant negative effect in the FRN window only at frontal electrode Fz (β_{max} = -1.833 at 216 ms, $q_{FDR} < 10^{-10}$; Fig. 6a) but also shows notable rhythmic fluctuations such that the maximal effect is at a second peak at 384 ms (β_{max} = -1.953 at Fz, $q_{FDR} < 10^{-10}$). On positive outcomes, RPE magnitude shows a sustained positive ramp that builds up to a maximum in the P3 window at posterior site Pz (β_{max} = 2.743 at 312 ms, $q_{FDR} < 10^{-10}$; Fig. 6b), with a similar but smaller significant effect at frontal Fz (β_{max} = 2.065 at 272 ms, $q_{FDR} < 10^{-10}$). In contrast, the late probability effect has a consistent sign and frontal distribution across positive and negative domains (positive: β_{max} = -2.646 at 336 ms in electrode Fz, $q_{FDR} < 10^{-10}$; negative: β_{max} = -3.742 at 392 ms in electrode Fz, $q_{FDR} < 10^{-10}$).

Modeling ERP amplitude topographies averaged in 50 ms windows centered on the peaks of the original RPE value, RPE magnitude, and probability effects in Figure 2 confirms the distinct spatio-temporal dynamics of RPE magnitude effects for negative and positive feedback. For negative outcomes, RPE magnitude model coefficients show significant negative effects in fronto-central sensors in the early FRN window at 216 ms (β_{max} = -1.556 at electrode F2, $q_{FDR} < 10^{-10}$) which dissipate in the middle P3 window and return to their strongest levels at the late 380 ms window (β_{max} = -1.684 at electrode C4, $q_{FDR} < 10^{-10}$; Fig. 6c). For positive outcomes, RPE magnitude shows significant positive model coefficients that are maximal at central and posterior sites in the middle P3 window (β_{max} = 2.683 at electrode CP1, $q_{FDR} < 10^{-10}$; Fig. 6d). As in the ERP time-domain analyses, the probability predictor shows a significant negative effect strongest in the late window at fronto-central sensors regardless of valence (negative: β_{max} = -3.093 at 380 ms in electrode F1, $q_{FDR} < 10^{-10}$, see Fig. 6c; positive: β_{max} = -2.099 at 380 ms in electrode C4, $q_{FDR} < 10^{-10}$, see Fig. 6d).

These ERP results suggest qualitative differences in responses to positive and negative feedback, and separate regressions of TFR power for positive and negative outcomes reveal a dissociation of theta and delta power underlying the FRN and RewP components driving these RPE effects. Specifically, RPE magnitude shows a strong positive effect in frontal theta frequencies on negative trials (β_{max} = 0.823 at [292 ms, 7 Hz] in electrode Fz, $q_{FDR} < 10^{-10}$), but this effect shifts to posterior delta power on positive outcomes (β_{max} = 0.871 at [260 ms, 3 Hz] at Pz, $q_{FDR} < 10^{-10}$). Again, the late probability effect shows a consistent negative effect strongest in delta frequencies at Fz that is stable regardless of valence (negative outcome: β_{max} = -0.734 at [416 ms, 4 Hz] at Pz, $q_{FDR} < 10^{-10}$; positive outcomes: β_{max} = -1.057 at [388 ms, 4 Hz] at Pz, $q_{FDR} < 10^{-10}$). In sum, the large valenced RPE value effect seen when contrasting wins and losses in the FRN/RewP window is composed of the superposition of two separate non-valenced RPE magnitude effects: early frontal theta activity drives negative FRN amplitudes on negative outcomes, and prolonged, more posterior delta activity increases positive RewP amplitudes.

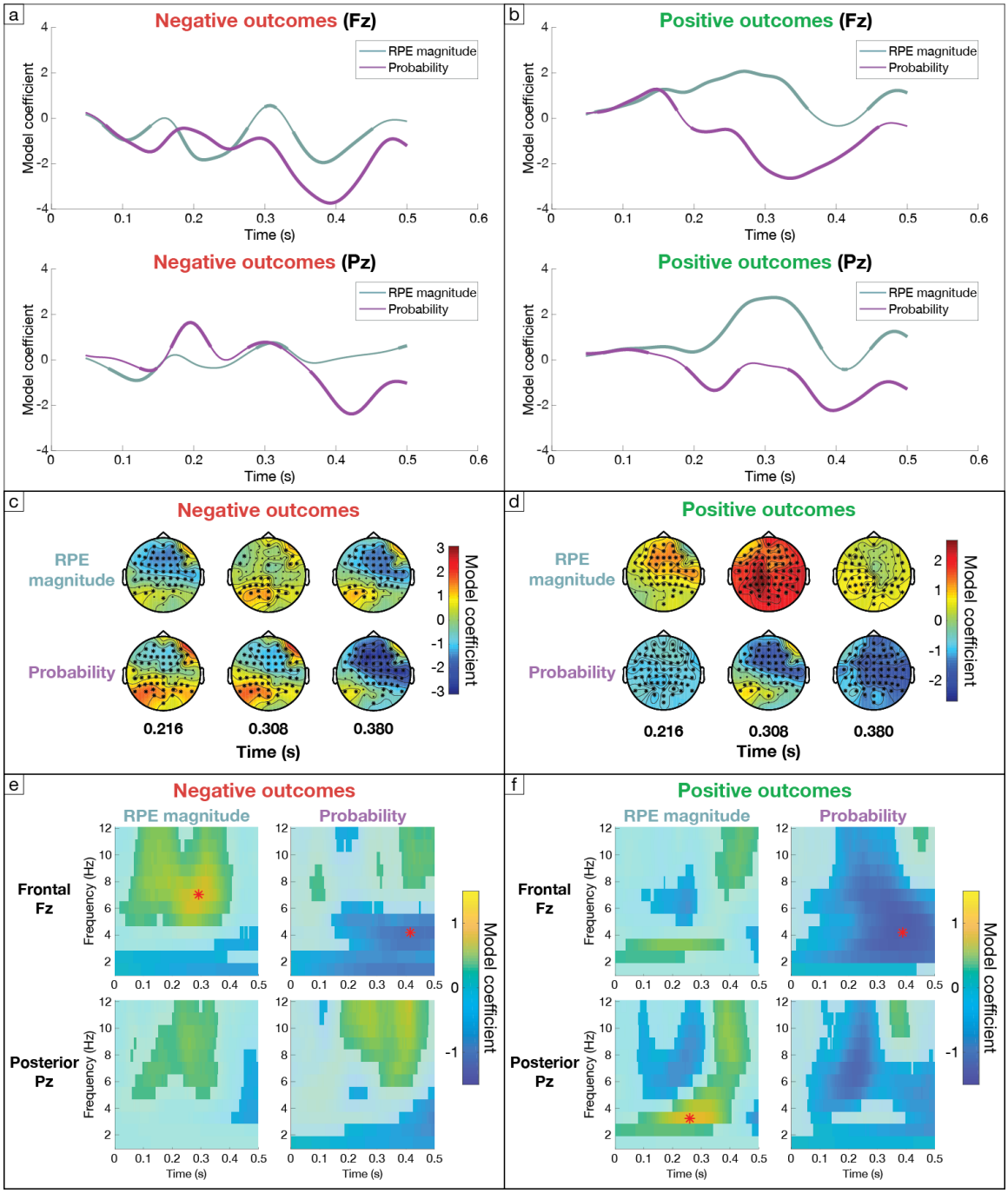


Figure 6: Separating outcomes by valence disentangles overlapping ERP and time-frequency power signatures of prediction errors. Left column is for negative outcomes (easy loss, easy neutral, and hard loss), while right column shows results for positive outcomes (easy win, hard win, and hard neutral). (a) Model coefficients for effects of RPE magnitude and probability on ERP amplitude in only negative outcomes at frontal site Fz (top) and posterior site Pz (bottom). Bolding indicates significant time points ($q_{FDR} < 0.05$). ERP amplitude significantly decreases with RPE magnitude in the early FRN window only at Fz, and this effect fluctuates rhythmically to rebound and peak at a later epoch. (b) Same for only positive outcomes. RPE magnitude shows a different pattern than on negative outcomes, with a significant positive relationship with ERP amplitude peaking in the RewP/P3 window and at posterior Pz. In contrast, the late, frontal probability effect shows a consistent late, negative relationship with amplitude across positive and negative valence. (c) Single-trial regression over all 64 electrodes is computed for only negative outcomes in the same three 50 ms windows from Fig. 2 and 3. Stars indicate significant electrodes

(q_{FDR} < 0.05). RPE magnitude shows significant negative effects in frontal sensors during the early and late windows, while probability shows a strong negative effect in fronto-central sites in the late window. (d) Same for only positive outcomes. The late, fronto-central negative effect of probability matches negative outcomes, but RPE magnitude effects are strongest in central and posterior sensors in the middle window. (e) Model coefficients fit to evoked power on negative outcomes for frontal electrode Fz and posterior electrode Pz, with non-significant points (q_{FDR} > 0.05) plotted opaquely. Red stars indicate maximal coefficients for each model predictor across both electrodes. RPE magnitude coefficients peak in frontal theta power, and probability coefficients peak later in frontal delta power. (f) Same for positive outcomes. The maximal RPE magnitude effect shifts to delta frequencies in posterior channels, indicating distinct mechanisms for RPE encoding on wins and losses. In contrast, probability effects maintain their late frontal delta distribution.

ERP PE sequence results are robust to biases in subjective reward expectations

Reward expectations are as critical to RPEs as the outcome, and failure to account for differences in these predictions across paradigms contributes to the disagreements in the reward EEG literature. For example, comparing ERPs following easy and hard neutral trials with identical feedback but opposite reward expectations showed FRN latency shifts according to RPE valence that matched those observed in losses and wins, respectively (see Supplementary Note 1). Furthermore, previous work has indicated participants' subjective reward expectations may deviate from objective probabilities established by the experimental design or modeled from behavior (Hajcak et al., 2007; Oliveira et al., 2007). To assess whether this issue impacted our results, we collected an additional behavioral dataset to measure subjective ratings of reward expectations before feedback (see Supplementary Methods). Ratings tracked difficulty (easy/hard) and trial outcomes (win/loss) and revealed subjective biases such that participants underestimated their probability of winning in easy conditions and overestimated it in hard conditions (Supplementary Note 2 and Sup. Fig. 10). The EEG results in Fig. 2 were reproduced when incorporating similar biases into our RL model, providing evidence that our conclusions are robust to differences between subjective and model-based reward expectations.

DISCUSSION:

We tested competing valenced and non-valenced explanations of reward feedback-locked EEG signatures by separating overlapping ERP components and cognitive variables using single-trial behavioral modeling and multiple regression across temporal, spatial, and spectral dimensions. Analyses using the standard approach of combining wins and losses implied early frontal theta activity in the FRN epoch represented valenced, scalar RPE value, and subsequent posterior delta band activity at the P3 peak indexed non-valenced RPE magnitude, seemingly supporting a combination of classical RL and independent coding theories (Holroyd & Coles, 2002; Yeung & Sanfey, 2004). However, repeating the regression analyses for wins and losses separately revealed the early valenced RPE value effect was an artifact of overlap between two distinct non-valenced RPE magnitude effects, providing support for both the salience account (Alexander & Brown, 2011; Cavanagh & Frank, 2014) and a revised version of the RL theory (Holroyd et al., 2008; Proudfit, 2015). Specifically, our data confirm recent studies arguing negative RPEs elicit frontal negativity and theta power consistent with the FRN (Bai et al., 2015; Cavanagh et al., 2019; Fischer & Ullsperger, 2013; Holroyd & Krigolson, 2007; Wu & Zhou, 2009), while positive RPEs elicit a slower, ramping positivity in delta frequencies consistent with

the RewP (Becker et al., 2014; Bernat et al., 2015; Cavanagh, 2015; Cavanagh et al., 2019; Foti et al., 2015; Proudfit, 2015; Williams et al., 2017). Correlations with benchmark ERPs from an Oddball task in the same participants highlighted mixed contributions of both N2-like and P3-like components in the window of the original valenced RPE value effect, consistent with FRN and RewP overlap. In contrast, we also observed a novel later fronto-central positivity on the downslope of the P3 that tracks outcome probability, was stable across wins and losses, and had no corresponding Oddball ERP. Notably, traditional mean window and peak-to-peak ERP metrics were less reliable than single-trial modeling and failed to disentangle these overlapping components and their relationships to PEs. Finally, model comparisons showed PEs captured EEG features better than outcome properties, and reward expectations modulated FRN latency on neutral trials, emphasizing that EEG signatures of reward processing are best viewed through the lens of predictive coding. Below, we discuss how our analysis strategy addresses core theoretical and measurement issues in the literature and the implications of our findings on the nature of reward processing EEG components and their proposed relationships to learning signals.

Methodological implications

Methodologically, our behavioral modeling and multi-dimensional regression approach improves on traditional ERP analyses in several ways. In terms of theory, many reward EEG studies use categorical ANOVA statistics (Donkers et al., 2005; Ferdinand et al., 2012; Wu & Zhou, 2009) and test ERP sensitivity to experimental manipulations of outcome valence, magnitude, and probability that indirectly reflect hypothesized PE computations (Foti et al., 2011, 2015; Meadows et al., 2016). Instead, we directly test the central tenets of the main proposals in the field by combining single-trial estimates of RPEs estimated from individual participant behavior into mixed-effects multiple regression analyses providing several advantages. Regression analyses incorporate directional hypotheses with continuous instead of categorical variables, produce signed model coefficients that obviate the need for post-hoc tests, and provide flexibility to analyze negative and positive outcomes separately and avoid confounding the overlapping FRN and RewP in a win-loss difference wave contrast. In contrast to single variable correlations, multiple regression partitions variance appropriately between the model's competing valenced and non-valenced predictors in a single analysis. Including random intercepts for each subject in a mixed-effects model also enhances statistical power to allow high resolution analyses (Aarts et al., 2014; Frömer et al., 2018). Importantly, regression frameworks provide formal model comparisons that quantitatively determined that EEG signatures of reward feedback are better described by modeling latent cognitive variables in a predictive coding framework (Cavanagh et al., 2010; Collins & Frank, 2018; Collins & Frank, 2016) than by the experimentally manipulated outcome features used in standard analyses (Foti et al., 2015; Meadows et al., 2016; Yeung & Sanfey, 2004).

Regarding ERP measurement, single-trial regression at each time, electrode, and time-frequency point provides the high spatio-tempo-spectral resolution needed to disentangle multiple overlapping components (Bridwell et al., 2018; Glazer et al., 2018; N. J. Smith & Kutas, 2015). In contrast, traditional mean window and peak-to-peak metrics provide a single measurement influenced by mixtures of components, usually averaged across trials and tested

at the group level (Sambrook & Goslin, 2015). For the FRN window, these traditional methods replicated the valenced RPE value regression result, but also found significant non-valenced effects of RPE magnitude or probability depending on the metric, suggesting potential confounds from overlapping components. Notably, these non-valenced effects using traditional metrics were the only results that did not replicate across both cohorts, potentially due to the lower statistical power relative to the single-trial mixed-effects modeling. Also, mean window and peak-to-peak measurements are usually aligned to observable peaks, a strategy that would have failed to detect the late probability effect. In sum, these results highlight the benefit of unbiased, high-resolution, multi-dimensional regression analyses for separating and interpreting overlapping components.

Theoretical implications: Predictive coding

Our results emphasize the importance of accounting for reward expectations when interpreting reward processing EEG. Quantitative model comparisons showed a model comprised of outcome value, magnitude, and probability—features commonly associated with the FRN and RewP but that do not account for expectations—performed worse than the RL-based model across all ERP features. The outcome-based model was improved when outcome value was replaced with outcome valence to account for different reward expectations on neutral trials in easy and hard conditions, though this outcome valence model was still worse than the RL model based on PEs. The importance of predictive coding was especially apparent on neutral trials with identical outcomes but different feedback valence based on expectations. FRN peak latencies were modulated by RPE value, and similar to losses and wins, FRN peak latency was earlier for neutral feedback with negative than positive RPEs. However, Williams et al. found an opposite FRN latency shift with wins occurring earlier than losses in a large (n=500) gambling dataset using only visual feedback (Williams et al., 2021). The latency of FRN and RewP ERPs in our data are also earlier than reported in recent meta-analyses (Sambrook & Goslin, 2015), likely because feedback in our task includes auditory components which generate faster FRN latencies (Miltner et al., 1997). These discrepancies and the presence of multiple overlapping components indicate that although our latency results provide evidence of the influence of reward predictions on neutral trials, they should be interpreted with caution and may not generalize to paradigms with different feedback stimuli or task demands (Krigolson, 2018; Luck, 2014; Picton et al., 2000).

In contrast to these different brain responses to neutral outcomes, post-experiment survey data indicated participants' explicit interpretations of neutral feedback did not differ across easy and hard conditions. This finding supports assertions that the FRN is generated by a habitual, model-free reward learning circuit that bypasses conscious representations of goal-directed task structure and instead relies on implicit associative learning mechanisms (Ichikawa et al., 2010; Walsh & Anderson, 2012). Nonetheless, previous work indicates that failure to account for subjective expectations can obscure reward EEG effects and confound their interpretation (Hajcak et al., 2007). To address this issue, we collected subjective ratings of win probabilities in an additional behavioral experiment that showed participants' subjective expectations tracked our difficulty manipulation and their likelihood of winning or losing. We employed a control RL model incorporating subjective biases representative of those ratings to

eliminate concerns that differences between subjective and model-based reward expectations influenced our EEG conclusions. However, these behavioral data and modeling analyses do not account for differences in subjective biases across participants or predictions on individual trials (e.g., accurately identifying errors on easy trials before feedback). Future studies may define how these effects modulate EEG signatures of PEs.

Theoretical implications: ERP components

Taken together, these observations clarify core issues in human reward processing EEG, confirming and extending recent proposals by disentangling the sequence of scalp ERP and TFR signatures unfolding after feedback and assessing their relationships to various cognitive PEs and canonical N2 and P3 ERPs. We employed mean window, peak-to-peak, and high-resolution regression analyses across wins and losses aiming to replicate commonly used win-loss difference waves (Krigolson, 2018; Proudfit, 2015; Sambrook & Goslin, 2015). Our results support a combination of the early, classical RL and independent coding theories, mainly that the FRN window represents a scalar, quantitative RPE instead of binary valence and the P3 represents non-valenced RPE magnitude. The overlapping epochs of significance for valenced RPE value and non-valenced RPE magnitude and probability predictors may also explain why some studies find no difference of valence and instead support the salience account, particularly since mean window and peak-to-peak metrics vulnerable to P3 confounds introduce significant non-valenced effects in our data. However, these approaches failed to capture the true nature of these components.

As suggested by the original authors of the RL theory, interpretations of any analysis that fails to separate wins and losses are flawed because they confound the FRN and RewP contributions on negative and positive outcomes, respectively (Cavanagh et al., 2019; Holroyd et al., 2008, 2011). Instead, temporal and spatial overlap of two distinct quantitative but non-valenced RPE magnitude effects in frontal theta on losses and posterior delta on wins creates the appearance of a valenced RPE value effect. Note that symmetric deflections in opposite directions caused by the FRN and RewP could potentially cancel out the quantitative tracking of RPE magnitude and explain why some authors observed only binary valence in this time window (Hajcak et al., 2006; Nieuwenhuis et al., 2004; Philiastides et al., 2010; Yeung & Sanfey, 2004).

Cavanagh et al. and others have argued the frontal theta response underlying the FRN is an instance of a general MPFC control mechanism elicited by feedback that requires adaptation (Cavanagh et al., 2011; Cavanagh & Frank, 2014; Hajihosseini & Holroyd, 2013; Hauser et al., 2014). Time estimation tasks like ours employ an implicit win-stay lose-switch strategy in which control adjustments are needed after losses but not wins. Consequently, the FRN mainly appears following negative outcomes, though we do observe weak FRN-like deflections after positive feedback, potentially because vectoral RT feedback allows adjustments of motor timing even when previous responses were correct (e.g., when the target zone is large on easy trials). Importantly, this argument can also explain evidence supporting the salience theory based on paradigm differences in which conditions elicit non-valenced control PEs that trigger MPFC theta responses (Alexander & Brown, 2011; Cavanagh & Frank, 2014). For example, FRN and theta responses have been observed following surprising positive outcomes in probabilistic

learning tasks when unexpected rewards may indicate shifting reward contingencies that require modifying decision policies, thereby dissociating control PEs from negative valence (Cavanagh et al., 2010, 2012; Hauser et al., 2014; Soder & Potts, 2018; Talmi et al., 2013). These results suggest the FRN is not specific to negative RPEs as proposed by the original RL theory but is instead better described by a version of the salience theory in line with action-outcome PEs from the PRO model (Alexander & Brown, 2011).

An outstanding question in the reward processing EEG literature is the nature of the RewP and other P3-like components. Our data substantiate claims that RPE magnitude on positive outcomes modulates a delta frequency P3-like component matching the RewP (Baker & Holroyd, 2011; Cavanagh, 2015; Cavanagh et al., 2019; Holroyd et al., 2008; Proudfit, 2015). This effect appeared strongest in posterior electrodes similar to the topography of a P3b when analyzing wins and losses together. However, when examining only positive outcomes, the RewP had a more central distribution closer to a P3a, matching previous reports (Cavanagh, 2015; Cavanagh et al., 2019; Foti et al., 2015; Proudfit, 2015). Further, Oddball ERP comparisons showed activity in this epoch correlated more with the P3a than with the P3b.

Whether the RewP is specific to positive RPEs is complicated by the fronto-central evoked positivity following the FRN on negative outcomes that also increases with RPE magnitude (see Fig. 2 and Sup. Fig. 9). Previous studies analyzing wins and losses together have interpreted these two positive peaks as the same P3 component representing non-valenced RPE magnitude in both conditions along the lines of the independent coding hypothesis (Fischer & Ullsperger, 2013; Ullsperger, Fischer, et al., 2014; Yeung & Sanfey, 2004). However, our data show stronger theta than delta power representations of RPE magnitude at that time point for negative outcomes (see Fig. 6e), as well as rhythmic fluctuations at theta frequencies in frontal RPE magnitude coefficients (Fig. 6a). These observations fit with an alternative interpretation that this positive peak on negative trials is not a P3 but instead due to phase reversal of the theta rhythm underlying the FRN (Bernat et al., 2011, 2015). This later interpretation would also explain why the RPE magnitude effect is stronger for losses in frontal sensors where theta is maximal but stronger for wins at posterior sensors where delta is maximal. Importantly, if the FRN and the following positivity are in fact generated by the same phase reset in MPFC theta activity—which is a more parsimonious explanation—they should not be interpreted as unique components representing different aspects of PEs (Gruber et al., 2005; Luck, 2014; Sauseng et al., 2007). These observations also imply the theta response on negative trials may mask the RewP in commonly used difference wave contrasts. Notably, when the RewP is defined as the peak of the win-loss difference wave, it is maximal on the upslope of the P3 and at Cz (Williams et al., 2021), which are the precise points of maximal spatial and temporal overlap between FRN and P3 components. Taken together, these data suggest the RPE magnitude effect in our data is not a single non-valenced P3 as posited by the independent coding hypothesis but is due to spatiotemporal overlap between the P3-like RewP specific to positive RPEs and the positivity generated by phase reversal of FRN-linked theta tracking negative RPEs (Bai et al., 2015; Bernat et al., 2015; Williams et al., 2021).

Our experimental design and modeling approach dissociated outcome probability as a second source of non-valenced salience which has recently been shown to modulate DA coding of RPEs in monkeys (Rothenhoefer et al., 2021). Our multiple regression analyses revealed this predictor captured a novel late frontal positivity in delta frequencies with no observable peak in

grand-averaged ERPs and no direct correlation with any Oddball ERPs. The latency of this late frontal positivity matches the timing of a late probability effect identified in the meta-analysis reported by Sambrook & Goslin (Sambrook & Goslin, 2015), but it's relationship to previously reported ERPs is unclear. Its anterior topography and sensitivity to relatively novel feedback fit the description of the classic P3a (Polich, 2007), but previous studies have ascribed the P3a label to the positivity following the FRN (Fischer & Ullsperger, 2013; Ullsperger, Fischer, et al., 2014). Also, frontal P3a ERPs typically occur before posterior P3b ERPs, which contradicts the order of posterior RPE magnitude and anterior probability positivities in our results. Furthermore, this effect did not correlate with the novelty P3a from the Oddball task, although it is possible other components such as the negativity driven by the second cycle of frontal theta may have obscured this relationship. Alternatively, our late probability effect could be related the late positive potential (LPP), a positive ramping ERP that starts ~400-600 ms after feedback and is enhanced by motivational significance, but this seems unlikely given the LPP's posterior distribution and sustained time course of up to several seconds (Hajcak & Foti, 2020). It is also possible that this effect is an artifact of correlations between RPE magnitude and probability in our model, but it's replication across cohorts and in wins and losses separately, as well as the dissociable spatiotemporal patterns of model performance when excluding either source of salience argue against this possibility. Ultimately, this late fronto-central probability effect may correspond to a P3a-like ERP with altered timing due to the specific design of our task, but this finding should be replicated using different paradigms.

Understanding the nature of these different components is important for maximizing their potential in clinical applications. For example, in a population with comorbid anxiety and depression symptoms performing a probabilistic learning task, theta power on losses correlated with anxiety subscores while delta power on wins predicted depression, highlighting the dissociable relationships between these EEG signatures and correlated dimensions of psychopathology (Cavanagh et al., 2019). Strong links between DA and reward circuits and the RewP make it a promising biomarker for mood disorders and addiction (Holroyd & Umemoto, 2016; Nusslock & Alloy, 2017; Proudfit, 2015). DA markers predict personality traits like extraversion and sensation seeking (Fischer et al., 2018; Gjedde et al., 2010), as well as psychiatric risk for schizophrenia (Maia & Frank, 2017), mood disorders (Lammel et al., 2014), and addiction (Nutt et al., 2015). RewP amplitude also predicts extraversion (Smillie et al., 2019), depression symptoms (Brush et al., 2018; Foti & Hajcak, 2009) and onset (Bress et al., 2013), substance misuse (Joyner et al., 2019), and may mediate the relationship between DA and aberrant reward sensitivity in these disorders (Baker et al., 2016). In contrast, components in the family of ERPs driven by mid-frontal theta power (e.g., N2, ERN, FRN) predict anxiety symptoms across multiple psychiatric conditions (Weinberg et al., 2015). These theta signals are hypothesized to reflect enlarged control PEs due to abnormal sensitivity to threat, reward, and punishment and hyperactive performance monitoring, resulting in more cautious and avoidant behavior (Moser et al., 2013; Riesel et al., 2017). Finally, P3 abnormalities are also observed in a host of disabling neuropsychiatric disorders including substance use disorders (Euser et al., 2012), bipolar disorder (Wada et al., 2019), and schizophrenia (Jeon & Polich, 2003). Given the complex relationships between these overlapping components and dissociable dimensions of clinical symptoms, refining specific mappings between reward and control PEs and EEG features

across multiple dimensions could enhance the power and reliability of these biomarkers and improve their diagnostic and therapeutic potential.

In conclusion, our experimental design, computational modeling, and signal analysis approach provide a comprehensive assessment of the sequence of EEG components elicited by reward feedback and their relationships to control and reward PEs. Multiple regression analyses across temporal, spatial, and frequency dimensions of the data and correlations with canonical N2 and P3 ERPs from the Oddball task elucidate a succession of overlapping components, each corresponding to distinct PEs. We demonstrate the pitfalls of using standard mean window, peak-to-peak, and win-loss difference wave techniques that confound the early, frontal, theta frequency FRN tracking non-valenced RPE magnitude on negative trials and the concurrent ramping of more posterior delta frequency RewP responses driven by RPE magnitude on positive outcomes. Separating positive and negative outcomes and distinguishing temporal, spatial, and frequency dimensions confirmed an updated version of the salience account of the FRN for negative RPEs that concurs with the PRO model, provided evidence that positive RPEs elicit a P3-like RewP, and identified a novel late frontal P3 tracking low probability outcomes. In summary, we used traditional analyses contrasting wins and losses to reproduce classical evidence of valenced RPE value effects in the FRN window and non-valence RPE magnitude effects in the P3 window that formed the foundations of early RL and independent coding theories. However, follow up analyses separating wins and losses revealed those interpretations were confounded by overlap of three distinct non-valenced salience components. Instead, our data corroborate and extend modern accounts of the FRN as an instance of control PEs generated by MPFC theta and of the RewP as a P3-like component tracking positive RPEs. Our findings demonstrate the power of behavioral modeling and single-trial EEG regression to separate overlapping components, adjudicate longstanding theoretical debates, and improve the utility of potential ERP biomarkers for diagnosis and treatment of neuropsychiatric disorders.

METHODS:

Experimental model and participant details

Target Time EEG data were collected from 41 adult healthy participants (mean ± SD [range]: 20.5 ± 1.4 [18-25] years old; 28 women; 37 right-handed) at the University of California, Berkeley. Oddball EEG data were collected during the same session from a subset of 30 of these participants, 19 of which were in the replication cohort. A separate cohort of 24 healthy adults (mean ± SD [range]: 30.6 ± 5.3 [21-44] years old; 12 women; 22 right-handed) completed a follow up remote behavioral version of the Target Time task to obtain subjective ratings of win probability. All participants reported no history of psychiatric or neurological disorders and had normal, or corrected-to-normal, vision. All participants were either financially compensated or given course credit and gave written informed consent to experimental protocols approved by the University of California, Berkeley Committees on Human Research.

Method details

Target Time behavioral task:

The Target Time interval timing task was written in PsychoPy (Peirce, 2008) (v1.85.3) and consisted of eight blocks (four easy and four hard in randomized order) of 75 trials. Following central fixation for an inter-trial interval randomly chosen as 700 or 1000 ms (an earlier task design also included 200 and 400 ms ITIs for four participants in the initial cohort), trials began with presentation of counter-clockwise visual motion from the bottom of a ring of dots at a constant speed to complete the circle at the one-second temporal interval. Participants estimated the interval via button press using an RTBox (v5/6) response device (Li et al., 2010). The width of a gray target zone indicated the tolerance for successful responses. Veridical win/loss feedback was presented from 1800-2800 ms and composed of (1) the tolerance cue turning green/red, (2) cash register/descending tones auditory cues, and (3) a black tick mark denoting the response time (RT) on the ring. Participants received ±100 points for wins/losses. Tolerance was bounded at ± 15-400 ms, and separate staircase algorithms for easy and hard blocks adjusted tolerance by -3/+12 and -12/+3 ms following wins/losses, respectively. Participants learned the interval in five initial training trials in which visual motion completed the full circle. For all subsequent trials, dot motion halted after 400 ms to prevent visuo-motor integration, forcing participants to rely on external feedback. Training concluded with 15 easy and 15 hard trials to initialize both staircase algorithms to individual performance levels. Main task blocks introduced neutral outcomes on a random 12% of trials that consisted of blue target zone feedback, a novel oddball auditory stimulus, no RT marker, and no score change.

Oddball behavioral task and performance:

The three-tone Oddball target detection task was written in PsychoPy (Peirce, 2008) (v1.85.3) and consisted of 10 training trials followed by three blocks of 130 trials each. Following central fixation for an inter-trial interval randomly chosen as 1.3 or 1.5 s, participants were presented with either a standard (75.5% of trials), target (12.25% of trials), or novel (12.25% of trials) audiovisual stimulus for 0.2 seconds. Participants were required to press a button using an RTBox (v5/6) response device (Li et al., 2010) when they detected a target stimulus. The visual stimulus was the same as the feedback stimulus in the Target Time task—a ring of dots with a

colored bar for the target zone—except without the black tick mark indicating response time and the green, red, and blue colors of the target zone were randomized to the standard, target, and novel conditions across participants. The accompanying auditory stimulus was a 440 Hz tone for standards, 1760 Hz tone for targets, and a novel, randomly selected oddball stimulus for the novel condition (different from Target Time neutral stimuli). Group-level accuracy was 99.68% ± 0.0061% (mean ± SD), and reaction times were 0.364 ± 0.063 s (mean ± SD).

Post-experiment behavioral survey:
Immediately following the EEG experiment, 33 participants (n = 24 remain after all exclusion criteria, see Methods) were given a six-question survey to assess their interpretation of example pictures of winning, neutral, or losing outcomes with small or large target zones to indicate easy or hard contexts. This data is only available in a subset of participants because data collection began 3 participants before implementation of the Oddball task. In response to the question "How would you feel about this feedback?", participants rated each outcome on a 9-point Likert scale, where 1 indicated "Terrible!", 5 indicated "I don't care…", and 9 indicated "Great!". Answers are reported after centering the ratings at the indifference point of 5.

Subjective ratings Target Time task:
In a follow up remote behavioral experiment, participants downloaded and completed a version of the Target Time task optimized to assess subjective ratings of win probability. In this version of the task, participants completed six blocks of 75 trials after the same 35 training trials and responded using the mouse. Importantly, after responding but before feedback, participants were asked on every third trial to rate "How likely is it that you won on this trial by responding in the target zone?" responded by clicking on a slider bar rating scale between two "0" and "100" tick marks at the far left and far right labeled "Definitely Lost" and "Definitely Won", respectively. Participants were given 20 seconds to respond before the rating timed out and feedback was presented. Participants with prior knowledge of outcome probabilities in the task (n = 2) were excluded after preliminary analyses revealed qualitatively different (less variable) ratings.

Behavioral modeling:
Target Time EEG participants were excluded because of technical recording errors (n = 4 datasets with missing EEG or metadata necessary for analysis), excessively noisy data (n = 2 datasets with >3 standard deviation outliers in number of epochs or time points rejected based on visual identification of large, global artifacts), or poor behavioral performance (n = 3 datasets where RT outlier exclusion criteria resulted in <20 trials in any condition), leaving 32 participants for analysis. All Target Time analyses were piloted on an initial cohort of 15 participants to finalize model parameters and statistical tests before results were replicated in a second cohort of 17 participants. All findings except point estimate results in Supplementary Figures 3b-d successfully replicated across cohorts (see Sup. Table 1 for differences in these results across cohorts), so all other results presented in the text and figures reflect all 32 datasets combined.

The relationship between the tolerance around the target interval and expected value was fit to individual participant behavior using logistic regression. Specifically, tolerance was used to predict binary win/loss outcomes across trials using the MATLAB function *glmfit* with a binomial distribution and logit linking function. Trials with neutral outcomes were excluded because they were delivered randomly and thus not reflective of performance. The probability of winning (p_{win}) for each participant was computed as:

$$p_{win} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 t)}}$$

where β_0 is the intercept and β_1 is the slope from the logistic regression, and *t* is the tolerance on a given trial. Expected value was derived by linearly scaling the probability of winning to the reward function ranging from -1 to 1. RPE value was then computed by subtracting expected value from the actual reward value, and RPE magnitude was computed as the absolute value of RPE value. Outcome probability was simply the proportion of each outcome across easy and hard blocks separately. See Supplementary Figure 1 for model predictions by condition.

Although RPE magnitude and probability predictors were correlated (r = -0.71), variance inflation factors, which measures the degree of collinearity, were VIF_{RPEmag} = 2.0 and VIF_{Prob} = 2.0, which is below even the most stringent recommended thresholds of 2.5 for excluding them from the same model (Johnston et al., 2018). Nonetheless, the separate contributions of these two predictors were assessed using versions of the RL model excluding each of these two predictors, and time-resolved model comparison and coefficient results are plotted in Figures 2e and 2f and in Supplementary Figures 4g-j respectively (see below for details).

Notably, this model was fit across all blocks after training under the assumption that participants learned the task during the 35 training trials, and that the staircase algorithm was appropriately initialized to the participant's skill level in the training. Since our model is fit using behavior over the entire session, it is possible that it would not describe early trials well, especially if learning occurs over the course of the session. As control analyses, we computed expected value after replacing single-trial win probabilities with block-level accuracy, as well as a rolling average of accuracy on the last 5 or 10 trials. Our single-trial logistic regression model outperformed all of these control models (higher R^2 and lower AIC) for mean window, peak-to-peak, and single-trial amplitude regression analyses.

To compare our main model based on RL principles to models similar to those commonly used in the literature, we computed a similar model using only outcome features that did not account for reward expectations. The Outcome Value model included the value (-1, 0, and 1 for losses, neutral, and wins), magnitude (0 for neutral, 1 for wins and losses), and probability (same as above) for each outcome. The Outcome Valence model was identical to the Outcome Value model except the value predictor was replaced by a valence predictor that treats neutral trials as valenced reward omissions, meaning losses and easy neutral outcomes were coded as -1 and wins and hard neutral outcomes were coded as 1.

Electrophysiology recording:

EEG data were recorded using a BioSemi ActiveTwo amplifier with a 64-channel active electrode system arranged according to the extended 10-20 system at a sampling rate of 512

Hz. Horizontal electrooculogram (EOG) were recorded from electrodes placed at both outer canthi, and vertical EOG were recorded from an electrode placed below the right eye and right frontopolar electrode FP2. Additionally, two external electrodes were placed on each ear lobe for use in offline re-referencing.

Electrophysiology and behavior preprocessing:
Preprocessing and analysis used the Fieldtrip toolbox (Oostenveld et al., 2011) and custom code in MATLAB. EEG data were bandpass filtered from 0.1-30 Hz, de-meaned, re-referenced to the average of both ear lobe channels, and then downsampled to 250 Hz. Excessively noisy epochs and channels were removed by visual inspection. Independent component analysis (ICA) was used to remove artifacts due to channel noise, muscle activity, heartbeat, and EOG (i.e., components correlated with bipolar derivations of horizontal or vertical EOG signals bandpass filtered from 1-15 Hz). Trials were segmented from -0.15 to 2.8 s relative to trial onset, and missing channels were interpolated from neighboring channels via Fieldtrip function *ft_channelrepair*. Final quality checks rejected trials for behavioral outliers (RTs missing, < 0.6 s, or > 1.4 s) or EEG artifacts including muscle activity, large voltage shifts, and amplifier saturation identified via visual inspection and using the Fieldtrip function *ft_reject_visual*, resulting in trial counts ranging from 448-524 (mean ± SD: 498.4 ± 20.1). In the remote behavioral subjective rating task, outliers were rejected for the same interval timing response RT criteria and for slow subjective ratings with RTs greater than three standard deviations from the mean, resulting in 382-450 trials (mean ± SD: 437.1 ± 17.6) and 125-150 ratings per participant (mean ± SD: 143.5 ± 6.2). Oddball EEG preprocessing was identical except trials were initially segmented from -0.2 to 1.3 s, and RTs were rejected as outliers if less than 0.1 s or greater than 1.3 s, resulting in trial counts ranging from 334-389 (mean ± SD: 379.5 ± 10.9).

Event-related potentials and difference waves:
EEG data were re-aligned to feedback onset and cut to -0.2 to 1 s. ERPs were calculated for each participant by bandpass filtering from 0.5-20 Hz, baseline corrected by subtracting the mean of 200 ms immediately preceding feedback, and averaging across trials. Oddball ERPs were identical except aligned to stimulus onset.

Difference waves were computed to facilitate visual comparisons to previous FRN/RewP studies and to valence, magnitude, and probability effects from our model-based multiple regression results. All difference waves were computed at the individual level for both Fz and Pz and then plotted as grand-average waveforms with standard error of the mean across participants. The simplest RewP contrast was computed by subtracting the ERP averaged over all negative valence conditions (easy neutral, easy loss, and hard loss) from the ERP averaged over all positive valence conditions (easy win, hard neutral, and hard win). Outcome valence difference waves were also computed between condition pairs matched for outcome magnitude and probability: hard win minus easy loss (large magnitude, low probability), easy win minus hard loss (large magnitude, high probability), and hard neutral minus easy neutral (small magnitude, low probability). Outcome magnitude difference waves were computed by subtracting small from large magnitude outcomes matched for valence and probability: easy loss minus easy neutral (negative valence, low probability) and hard win minus hard neutral (positive valence, low probability). Outcome probability difference waves were computed by

subtracting likely from unlikely outcomes matched for valence and magnitude: easy neutral minus hard loss (negative valence, low magnitude) and hard neutral minus easy win (positive valence, low magnitude).

ERP point estimates and latencies:
To facilitate comparisons with prior FRN/RewP studies, we computed traditional mean window and peak-to-peak point estimates of FRN amplitude at electrode Fz. The mean window metric was calculated as the mean amplitude of each participants' condition-averaged ERPs in a 100 ms window centered on that participant's FRN peak latency computed across all conditions. The peak-to-peak metric was calculated by subtracting the FRN peak amplitude from the amplitude of the preceding positivity (P2) for each condition and participant. To account for variability in ERP waveshapes at the single participant level, peak-to-peak amplitude was only computed if a positive peak was found in the interval 100-260 ms post-feedback that preceded a negative peak in the interval 180-300 ms. According to these criteria, peak-to-peak amplitude could not be reliably computed on 11/192 ERPs. Additionally, the latency of the negative peak in this analysis was used as the FRN peak latency, which was then normalized within participant by subtracting the mean latency across all conditions.

To aid interpretation of ERP features underlying model-based results, we computed reference ERPs in the Oddball EEG data quantified using the mean across 50 ms windows of condition-averaged participant ERPs. For the Novel N2b and Target N2c, mean windows were centered on the peak negativity from the grand-average ERP at Fz between 0.2-0.3 s in their respective conditions. For the Novel P3a and Target P3b, mean windows were centered on the peak positivity from the grand-average ERP between 0.3-0.45 s at Cz and Pz, respectively.

Time-frequency representations:
EEG data were re-aligned and segmented from -0.2 to 1 s around feedback onset. Spectral decompositions were estimated at each time point by convolving the signal with a set of complex Morlet wavelets, defined as complex sine waves tapered by a Gaussian. The frequencies of the wavelets ranged from 1-12 Hz in 1Hz linear steps. The full-width at half-maximum (FWHM) ranged from 1.184-0.096 s with increasing wavelet peak frequency, which corresponds to 3 cycles per frequency. Task-evoked power was computed as the square of the magnitude of complex Fourier-spectra and baseline corrected by decibel conversion relative to a 200 ms baseline immediately preceding feedback.

Statistics and reproducibility:

Time-resolved modeling:
We adopted a multiple linear regression framework to directly compare the predictive power of valenced and non-valenced PEs derived from our RL-based behavioral model and simple outcome features that don't account for expectations. Our full RL model combines our single-trial model estimates of expected value, valenced RPE value, non-valenced RPE magnitude, and outcome probability in a linear mixed-effects model with random intercepts for each participant to maximize statistical power by accounting for within participant variance. This model was used to predict the temporal evolution of EEG amplitude at each time point from 50

to 500 ms post-feedback using the MATLAB function *fitlme*, which tests significance of model coefficients using two-sided *t*-tests under the null hypothesis the coefficient is equal to zero. Resulting *p* values were corrected for multiple comparisons using false discovery rate (Benjamini & Hochberg, 1995) across time points and model predictors. For clarity, any *p* values corrected for multiple comparisons are reported as q_{FDR} throughout the manuscript. These analyses were run separately for electrodes Fz and Pz to assess frontal FRN and posterior P3 ERPs.

To compare the performance of our RL-based and outcome-based models, we ran the same time-resolved regression analyses using Outcome Value and Outcome Valence models. Model performance was quantified using the Akaike Information Criterion (AIC), which scores model performance based on variance explained while penalizing models with extra parameters. Lower AIC values indicate better model performance. To emphasize differences between models, AIC is reported relative to a baseline model containing only random intercepts for each participant, which is equivalent to the mean ERP across all conditions. Lastly, the relative contributions RPE magnitude or probability to predicting ERP amplitude in the P3 window are assessed by reporting AIC of the RL model when leaving out either of these two correlated predictors.

FRN and P3 point estimate modeling:

This modeling procedure was also used to predict mean window and peak-to-peak measurements of FRN amplitude at electrode Fz and mean window P3 amplitude at Pz. Since these metrics yield one value per condition per participant, each model predictor was averaged within condition for each participant. FDR corrections were applied across model predictors. AIC is reported for this procedure to compare the RL model with Outcome Value and Outcome Valence models in the FRN point estimates and the RL model with and without RPE magnitude and probability predictors in the P3 mean window analysis. A nearly identical multiple regression analysis was used to predict FRN peak latency, except the MATLAB function *fitglm* Was used without the random intercept because peak latencies were already normalized within participant. A two-sided paired samples *t* test was used to test whether FRN peak latencies were different between neutral feedback in easy and hard conditions.

Topography modeling:

To examine the spatial distribution of PE effects on evoked potentials, ERP amplitudes were averaged for all electrodes in three 50 ms windows centered on the largest coefficient from the time-resolved regression for RPE value (216 ms at Fz), RPE magnitude (308 ms at Pz), and outcome probability (380 ms at Fz). The multiple regression model was then used to predict amplitude at each channel in each window, and FDR corrections were applied across all channels, model predictors, and windows.

Time-frequency power modeling:

Time-frequency representations were analyzed using the same mixed-effects multiple linear regression model to predict evoked power at each time-frequency point from 0 to 500 ms and 1-12 Hz. FDR multiple comparison corrections were applied across time points, frequencies, and

model predictors, again separately for Fz and Pz to assess frontal FRN and posterior effects. These analyses were repeated for only negative and only positive outcomes to test whether these results were driven by feedback of one particular valence.

Oddball-Target Time ERP correlations:
Post-hoc exploratory analyses compared Target Time ERP amplitudes in epochs showing maximal RL model-based effects to canonical Oddball ERPs via inter-participant correlations. These analyses were conducted in the subset of participants with both Target Time and Oddball EEG data (n = 22) after excluding one participant for excessive number of trials rejected due to noise in the Oddball EEG data and another participant for outlier Oddball behavioral accuracy (both outliers > 3 standard deviations from the group mean). For each Target Time condition, individual participant amplitudes of Novel N2b, Novel P3a, Target N2c, and Target P3b ERPs were used as benchmarks of participant's N2/P3 amplitudes and correlated with the 50 ms mean window amplitudes in three epochs used for topography modeling: one at Fz centered on the peak RPE value effect (216 ms), one at Pz centered on the peak RPE magnitude effect (308 ms), and one at Fz centered on the peak Probability effect (380 ms). Correlation *p* values were FDR corrected for the number of Oddball ERPs and the number of Target Time conditions.

Post-experiment behavioral survey ratings:
Subjective ratings for neutral trials were tested for significant differences from the indifference point on the 9-point Likert scale after subtracting 5 to center ratings, separately for easy and hard trials. Rating data were tested using two-sided independent samples *t* tests under the null hypothesis that ratings were from a normal distribution with mean equal to zero.

Remote behavioral task subjective ratings:
We evaluated whether objective win probabilities derived from behavior deviated from participants' subjective experiences of reward probabilities. To accomplish this, we compared our measure of expected value computed via logistic regression of observed wins and losses to subjective ratings of win probabilities measured in the remote behavioral Target Time task. We correlated single-trial model-based win probability with subjective ratings across all participants and conditions, as well as independently for easy and hard conditions. For each participant, subjective bias was quantified separately for easy and hard conditions by the mean difference between subjective ratings and model-based win probabilities. To test whether subjective ratings were sensitive to the probability of winning or losing, ratings were z-scored within each participant and condition. Independent samples *t* tests were used to compare normalized ratings before wins and before losses using group aggregated data for only easy, only hard, and both easy and hard conditions combined.

Data availability:

The datasets generated and/or analyzed during the current study will be made available in the CRCNS (https://crcns.org/) public repository prior to publication.

Software availability:

Custom Python and MATLAB code used for preprocessing and analysis is available as a GitHub repository (https://github.com/hoycw/PRJ_Error_eeg), which includes system requirements and dependencies.

Acknowledgements:

We thank I. Griffith for help piloting the paradigm and A. Shah and J. Abbas for help collecting the data. This work was supported by NINDS R37NS21135 (RTK), CONTE Center PO MH109429 (RTK), and NSF GRFP (CWH).

Author contributions:

C.W.H. and R.T.K. designed the experiment. C.W.H. and S.C.S. collected the data. C.W.H. and S.C.S. analyzed the data. C.W.H. and R.T.K. wrote the paper.

Competing interests:

The authors declare no competing interests.

SUPPLEMENTARY INFORMATION:

Single-trial modeling separates multiple overlapping prediction errors during reward processing in human EEG

Colin W. Hoy, Sheila C. Steiner, Robert T. Knight

Supplementary Note 1: FRN peak latency shifts with RPE valence
Supplementary Note 2: ERP PE sequence results are robust to biases in subjective reward expectations
Supplementary Figures

SUPPLEMENTARY NOTE 1: FRN peak latency shifts with RPE valence

Examination of the grand-average FRN waveshapes in Supplementary Figure 9a reveals shifts in the latency of the negativity, with losses peaking earlier than wins. Interestingly, neutral outcomes with identical reward values have earlier peak latencies in easy than hard blocks (see inset in Sup. Fig. 9a for direct comparison), despite explicit instructions stating neutral feedback does not reflect performance. Moreover, subjective ratings from post-experiment surveys confirmed EEG participants had explicit neutral feelings towards these outcomes in both easy (mean ± SD on zero-centered 9-point Likert scale: -0.2 ± 1.3; $t(23) = -0.53$, $p = 0.60$) and hard (-0.4 ± 1.6; $t(23) = -1.00$, $p = 0.33$) conditions. However, in our behavioral RL model, neutral outcomes had opposite valence based on contextual expectations, with negative RPE valence for omission of expected wins in easy blocks and positive RPE valence for omission of expected losses in hard blocks.

To determine whether FRN timing shifted systematically according to our RL model predictors, we applied the same multiple regression framework to predict FRN peak latencies. In a general linear model, the only model features predictive of FRN peak latency were RPE value ($\beta = 0.008$, $q_{FDR} = 2.80 * 10^{-8}$; Sup. Fig. 9b) and expected value ($\beta = 0.004$, $q_{FDR} = 0.047$), which are the two predictors encoding valence. A paired samples *t* test confirmed that FRN peak latencies following neutral feedback were significantly different between easy (mean ± SD: 228.3 ± 0.02 ms) and hard conditions (247.0 ± 0.02 ms; $t(27) = -3.70$, $p = 0.001$), consistent with different reward expectations for easy and hard conditions leading to negative and positive RPEs, respectively, in these otherwise identical outcomes. This dissociation between subjective ratings and neural signatures of valence suggests our approach can index implicit brain states not revealed by explicit participant report.

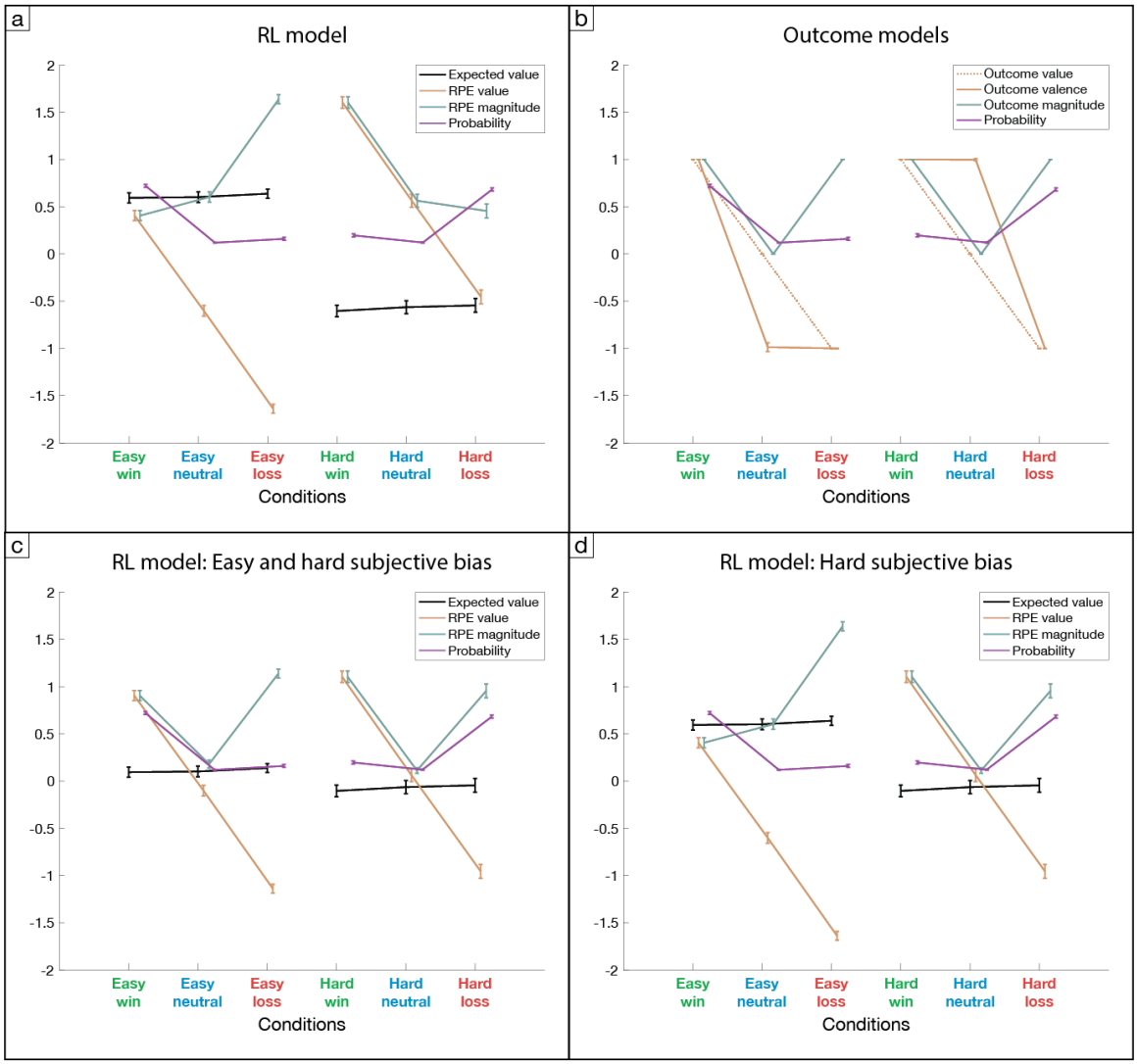
SUPPLEMENTARY NOTE 2: ERP PE sequence results are robust to biases in subjective reward expectations

To determine whether subjective reward expectations deviate from those derived from our RL model or impact our EEG results, we conducted an additional behavioral Target Time experiment in which we asked 22 participants to rate their subjective probability of winning using a slider bar after responding but before feedback (see Sup. Fig. 10a for example participant behavior). Comparing these subjective ratings to those derived from our behavioral modeling revealed strong correspondence overall ($r = 0.705$, $p < 10^{-10}$; Sup. Fig. 10b), but this relationship was likely driven by the large difference between easy and hard conditions, as the correlation was significant only within hard conditions ($r = 0.123$, $p < 10^{-6}$) and not within easy conditions ($r = 0.007$, $p = 0.79$). Nonetheless, participants rated their probability of winning significantly higher before wins than losses overall ($t(3156) = -4.27$, $p < 10^{-4}$), as well as within only easy ($t(1575) = -3.52$, $p < 10^{-3}$) and only hard conditions ($t(1579) = -4.29$, $p < 10^{-4}$). These data suggest that despite variance in subjective ratings, participants are sensitive to both the experimental manipulation of difficulty and their own behavior which combine to form reward expectations in our task.

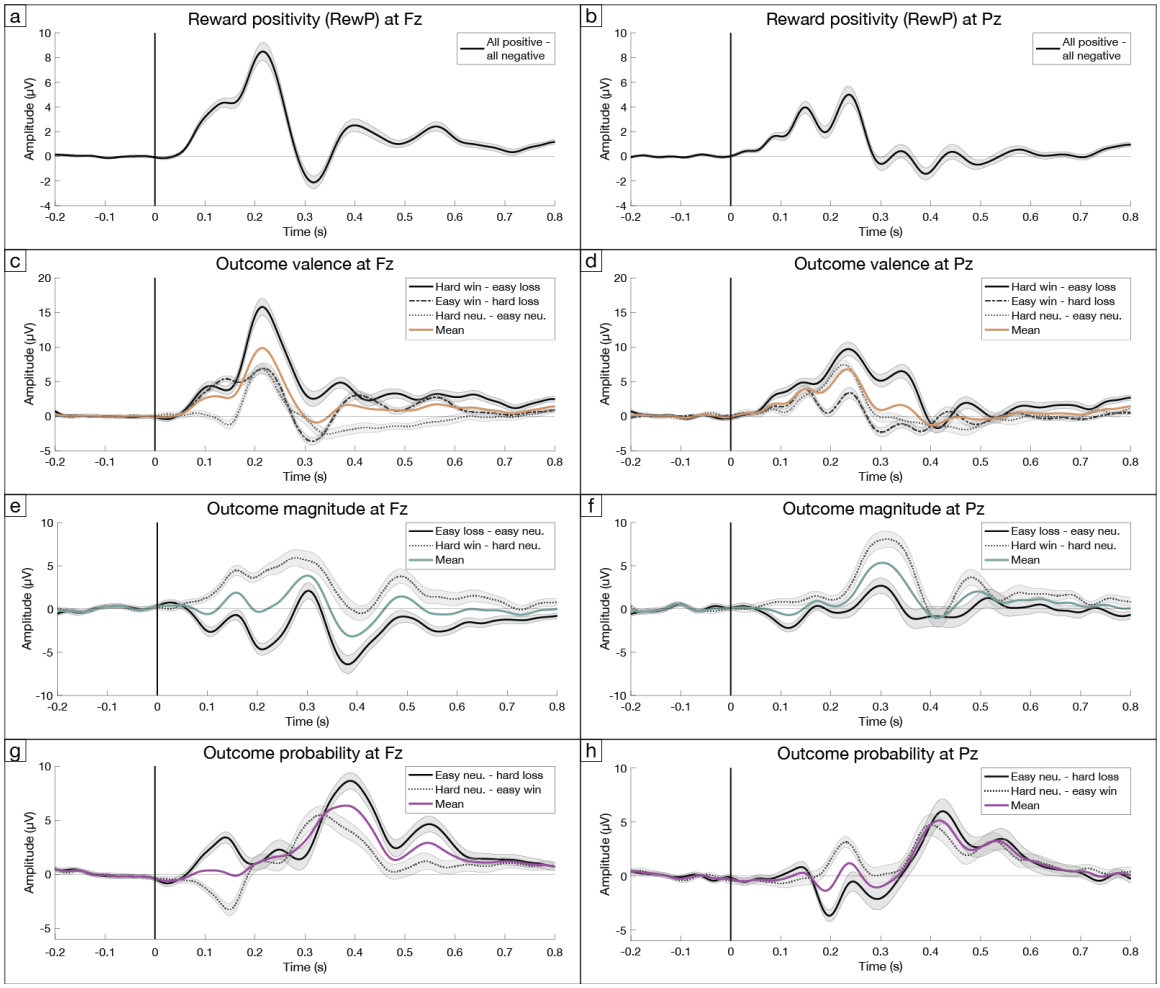
The superior performance of our RL model over outcome-based models and the outcome valence model over the outcome value model demonstrate the importance of incorporating reward expectations when evaluating EEG signatures of reward feedback.

However, the rating data from our behavioral experiment indicate participants' subjective expectations were biased relative to our behavioral model such that they overestimated their likelihood of winning in hard conditions and underestimated it in easy conditions (Sup. Fig. 10c). The biases measured in our behavioral experiment are small on average (mean ± SD in hard: -8.1 ± 10.2%; mean ± SD in easy: 6.3 ± 11.5%), but the results of our regression analyses could be affected by subjective biases in our sample of EEG participants. To address this potential confound, we simulated the effect of adding similar biases to the expected value predictor in our model. Due to the opposite directionality of bias in easy and hard conditions, adding the most extreme bias observed in our behavioral cohort (+/-25% shift in win probability) equalized reward expectations across easy and hard conditions and effectively reproduced the outcome value model tested above (compare model predictions in Sup. Fig. 1c to those in Sup. Fig. 1b). Since the optimistic bias in the hard condition was slightly larger and matches an overoptimistic bias reported in previous studies (Hajcak et al., 2007), we also simulated a model with +25% shift in win probability for only hard conditions (Sup. Fig. 1d). Repeating our multiple regression analyses at Fz and Pz with the RL model including subjective bias in hard conditions reproduced all of the main RPE value, RPE magnitude, and probability results from Fig. 2 (Sup. Fig. 4k and 4l). Overall, these analyses show that the level of subjective bias in reward expectations in our task do not affect our results and conclusions.

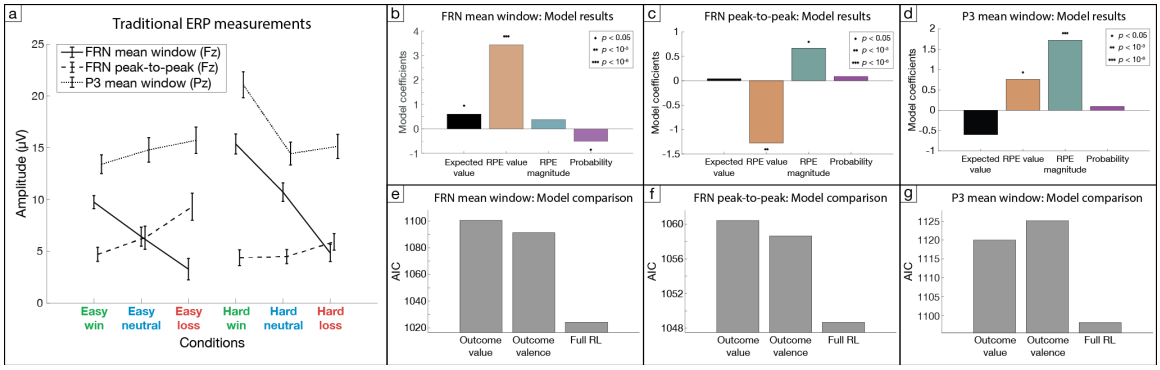
SUPPLEMENTARY FIGURES:



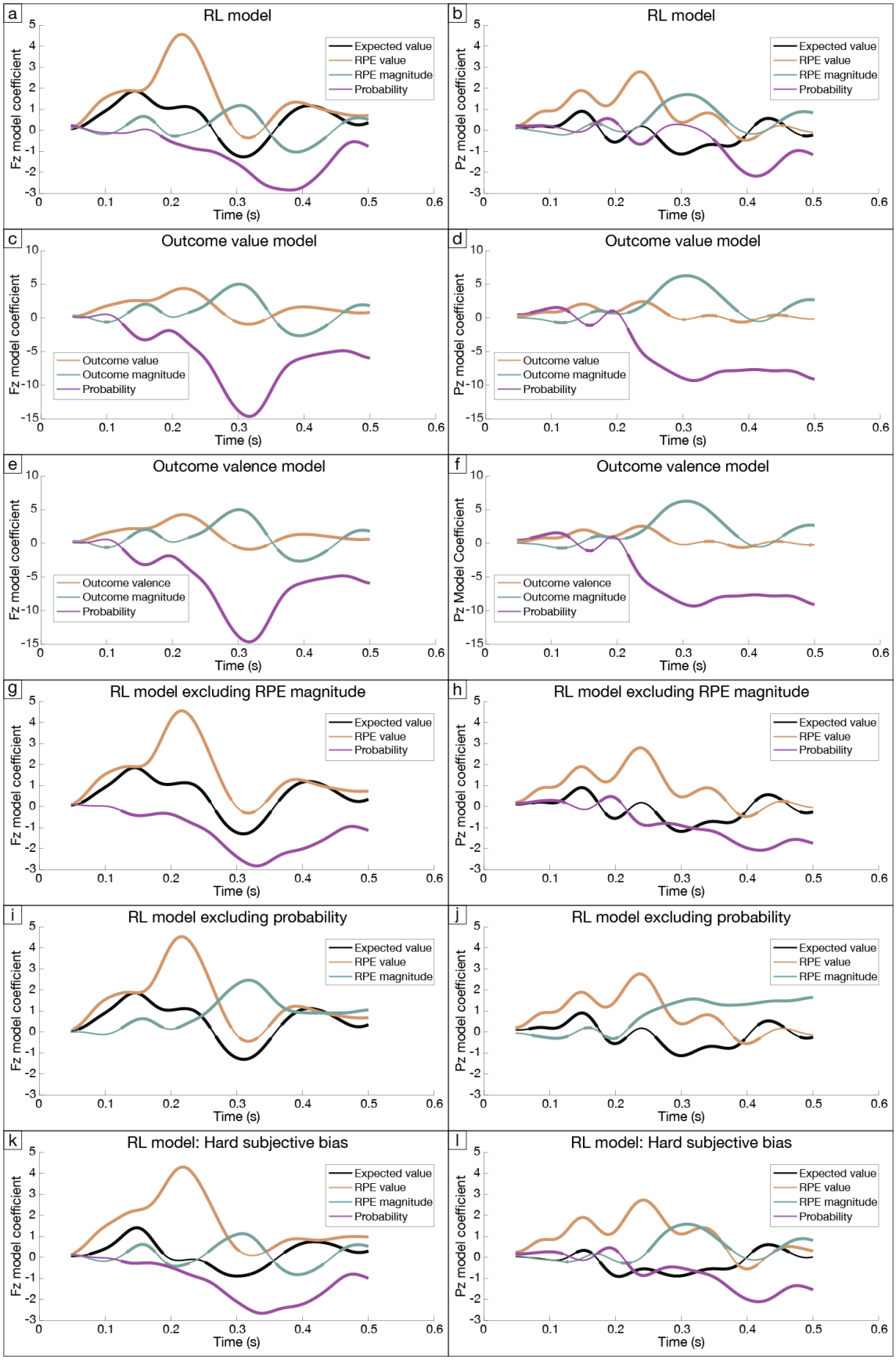
Supplementary Figure 1: Predictions for RL-based and outcome-based models. Error bars indicate standard deviation between participants. (a) Full RL model predictions by condition. (b) Outcome-based model predictors. Outcome value model includes value, magnitude, and probability, while outcome valence model is the same but replaces value with valence. (c) RL model with bias added to expected value in hard (+25% win probability) and easy (-25% win probability) conditions to match most extreme bias found in subjective rating data. Note that this effectively eliminates the difference in expected value across easy and hard conditions, resulting in predictors very similar to the outcome value model. (d) RL model with positive bias (+25% win probability) added to expected value in hard conditions to match most extreme bias found in subjective rating data.



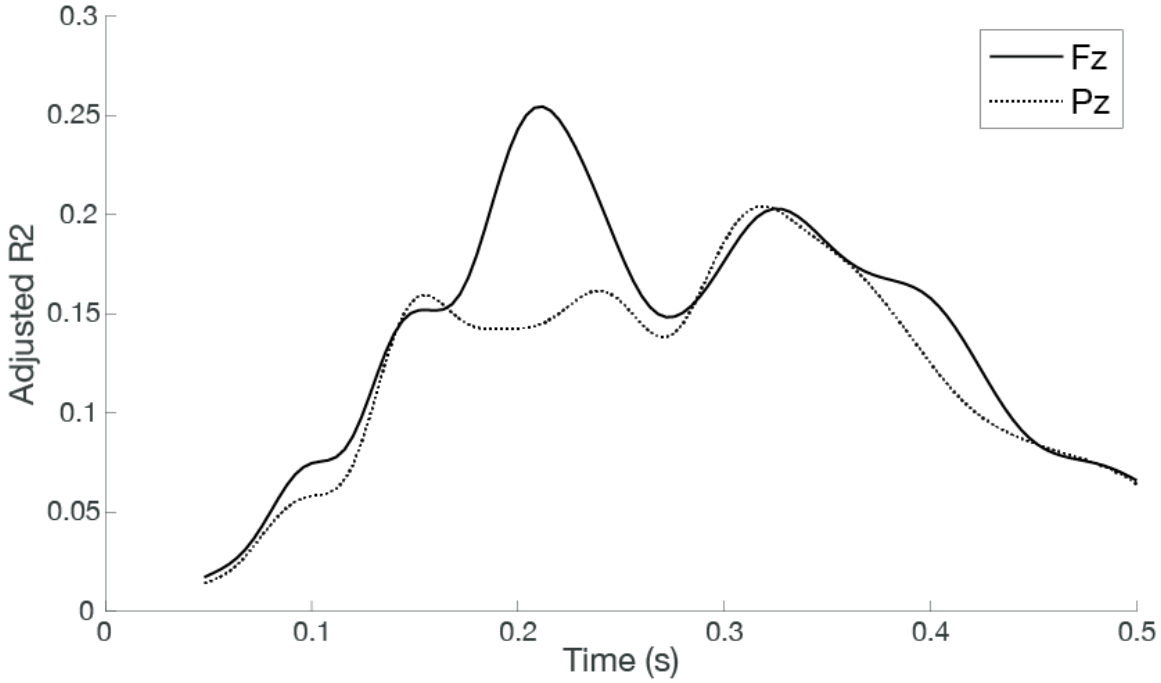
Supplementary Figure 2: Grand-average ERP difference waves to assess outcome valence, magnitude, and probability. Shaded error bars indicating standard error of the mean across participants. (a) Grand-average ERP difference wave at Fz between all conditions with positive RPE valence and all conditions with negative RPE valence, a contrast commonly used to derive the RewP. (b) Same at Pz. (c) Grand-average ERP difference wave at Fz to assess outcome valence by contrasting pairs of conditions matched for RPE magnitude and outcome probability. Tan line shows mean of these difference waves for comparison with RPE value model coefficients in Fig. 2. (d) Same at Pz. (e) Grand-average ERP difference wave at Fz to assess outcome magnitude by contrasting pairs of conditions matched for RPE valence and outcome probability. Teal line shows mean of these difference waves for comparison with RPE magnitude model coefficients in Fig. 2. (f) Same at Pz. (g) Grand-average ERP difference wave at Fz to assess outcome probability by contrasting pairs of conditions matched for RPE valence and magnitude. Purple line shows mean of these difference waves for comparison with probability model coefficients in Fig. 2. (h) Same at Pz.



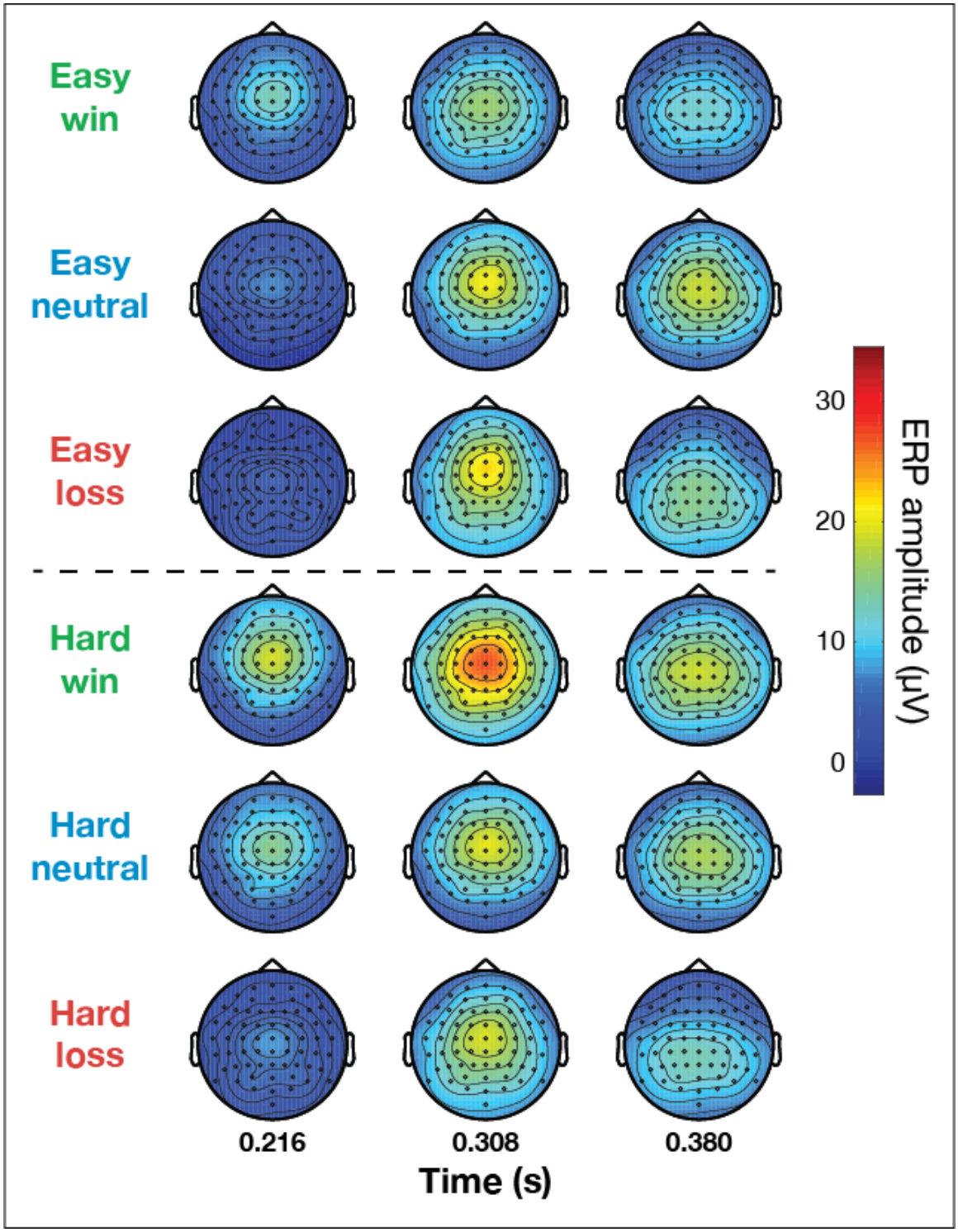
Supplementary Figure 3: Traditional mean window and peak-to-peak metrics support single-trial modeling results but cannot disambiguate individual predictors. (a) Traditional ERP measurements per condition. Error bars indicate standard error of the mean. FRN and P3 mean window estimates are averaged over 100 ms windows centered on the peak latency of the FRN at Fz and P3 at Pz in the grand-average ERP across all conditions as illustrated in Fig. 2a and Fig. 2b, respectively. FRN peak-to-peak is calculated as differences between amplitudes at FRN and preceding P2 positivity peaks for each participant. (b) RL model coefficients from multiple regression of FRN mean window estimates show significance for valenced expected value and RPE value, and also non-valenced probability ($q_{FDR} < 0.05$). (c) RL model coefficients from multiple regression of FRN peak-to-peak estimates are significant ($q_{FDR} < 0.05$) for valenced RPE value and non-valenced RPE magnitude. (d) RL model coefficients from multiple regression of P3 mean window estimates show significance for valenced RPE value and non-valenced RPE magnitude ($q_{FDR} < 0.05$). (e) Model comparison for FRN mean window estimates shows RL model outperforms outcome-based models that fail to account for expectations. (f) Same for FRN peak-to-peak. (g) Same for P3 mean window analysis.


a
RL model
Fz model coefficient
Time (s)
Expected value
RPE value
RPE magnitude
Probability
b
RL model
Pz model coefficient
Time (s)
Expected value
RPE value
RPE magnitude
Probability
c
Outcome value model
Fz model coefficient
Time (s)
Outcome value
Outcome magnitude
Probability
d
Outcome value model
Pz model coefficient
Time (s)
Outcome value
Outcome magnitude
Probability
e
Outcome valence model
Fz model coefficient
Time (s)
Outcome valence
Outcome magnitude
Probability
f
Outcome valence model
Pz Model Coefficient
Time (s)
Outcome valence
Outcome magnitude
Probability
g
RL model excluding RPE magnitude
Fz model coefficient
Time (s)
Expected value
RPE value
Probability
h
RL model excluding RPE magnitude
Pz model coefficient
Time (s)
Expected value
RPE value
Probability
i
RL model excluding probability
Fz model coefficient
Time (s)
Expected value
RPE value
RPE magnitude
j
RL model excluding probability
Pz model coefficient
Time (s)
Expected value
RPE value
RPE magnitude
k
RL model: Hard subjective bias
Fz model coefficient
Time (s)
Expected value
RPE value
RPE magnitude
Probability
l
RL model: Hard subjective bias
Pz model coefficient
Time (s)
Expected value
RPE value
RPE magnitude
Probability

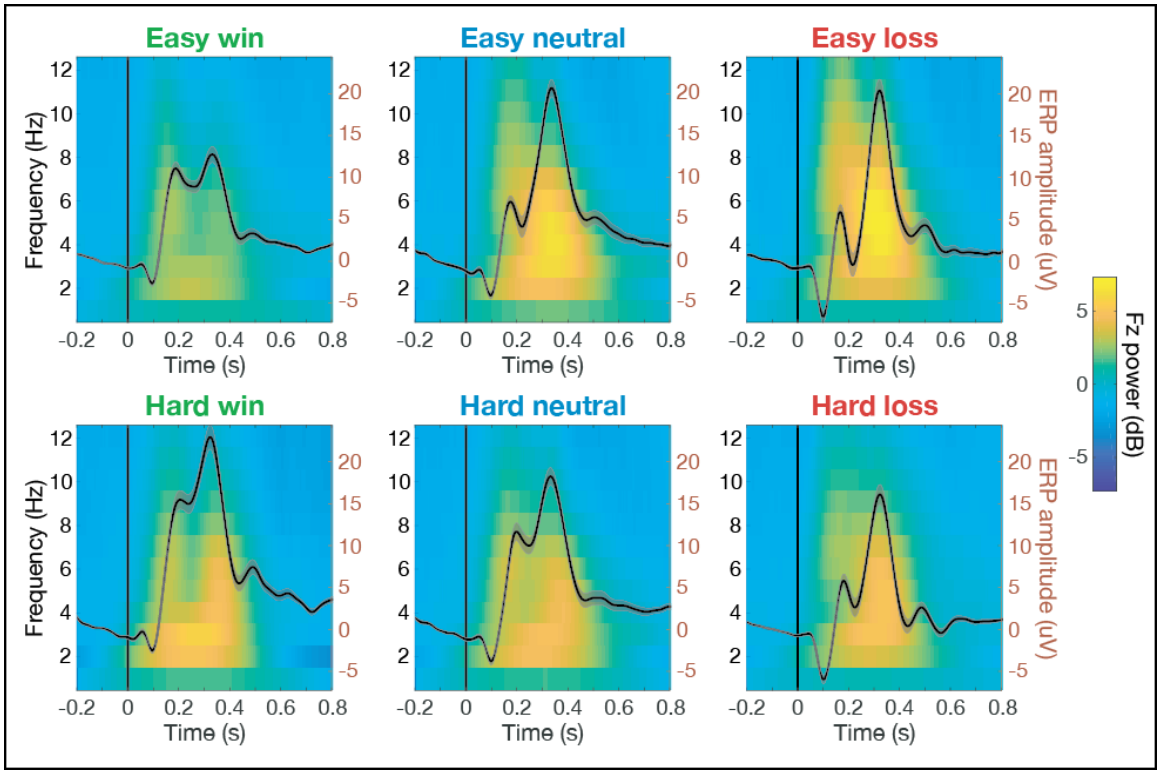
Supplementary Figure 4: Coefficients for alternative and control models. Left column shows results for Fz, and right column shows results for Pz. (a) Reference results from best performing RL model at Fz as displayed in Fig. 2c. (b) Same for Pz. (c) Coefficients at Fz for outcome value model show similar results as the RL model but with weaker model performance, especially during the FRN window. (d) Same for Pz. (e) Coefficients at Fz for outcome valence model show similar results as the RL model but with weaker model performance, especially during the FRN window. (f) Same at Pz. (g) Coefficients at Fz for RL model when excluding RPE magnitude, which shows RPE value and probability effects hold but RPE magnitude effect for the P3 is missing. (h) Same for Pz. (i) Coefficients at Fz for RL model when excluding probability, which shows RPE value and RPE magnitude effects hold but late frontal probability effect is missed. (j) Same for Pz. (k) Coefficients at Fz for the RL model after introducing a positive subjective bias on hard trials to match behavioral rating data, which shows the original RL model results are robust against simulated subjective biases. (l) Same for Pz.



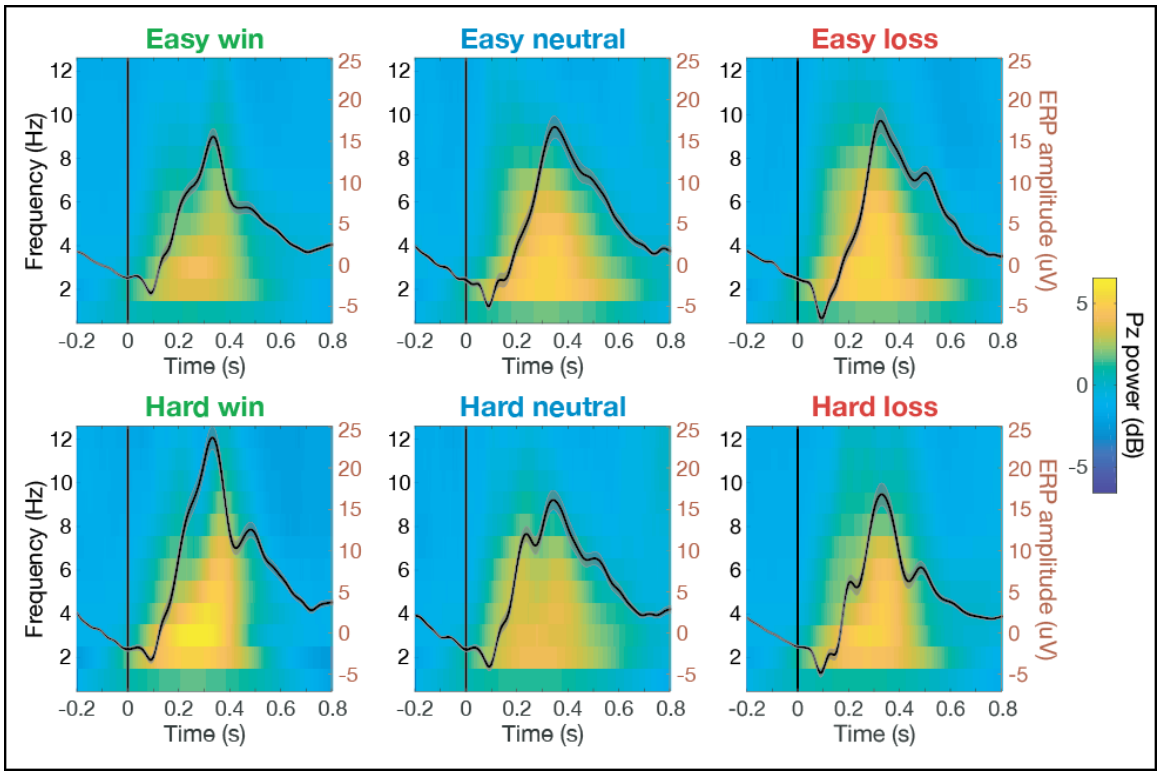
Supplementary Figure 5: ERP amplitude model fits. RL model fits at each time point for frontal electrode Fz and posterior electrode Pz plotted as adjusted R^2. Note that performance peaks in the FRN time window for Fz (R^2 = 0.254 at 212 ms) and in the P3 time window for Pz (R^2 = 0.204 at 320 ms).



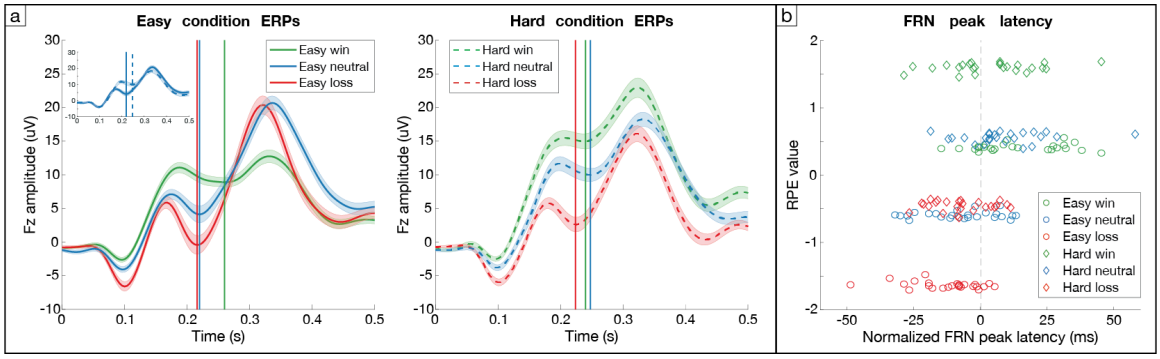
Supplementary Figure 6: ERP amplitude topography dynamics. ERP amplitudes are averaged within condition and participant in 50 ms windows centered on the peak model coefficients for RPE value (0.216 s), RPE magnitude (0.308 s), and probability (0.380 s) from Fig. 2. Topographies show group averaged amplitude for all electrodes in each window.



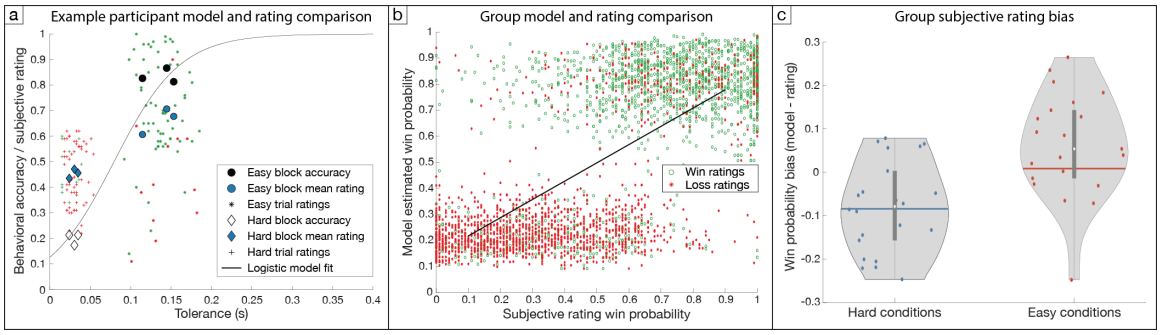
Supplementary Figure 7: Time-frequency evoked power at Fz for each condition. Grand-average ERPs are overlaid on the right y-axis to show the waveshape features driving evoked power.



Supplementary Figure 8: Time-frequency evoked power at Pz for each condition. Grand-average ERPs are overlaid on the right y-axis to show the waveshape features driving evoked power.



Supplementary Figure 9: FRN peak latency tracks valenced RPE value. (a) Grand-average ERPs at Fz for easy and hard conditions separately with the FRN peak latency marked by vertical lines. Shaded error bars indicate standard error of the mean across participants. Note that peak latencies for identical neutral outcomes shift early for negative RPEs in easy blocks versus late for positive RPEs in hard blocks (see top left inset for direct comparison). (b) Multiple regression revealed RPE value significantly predicted FRN peak latency ($q_{FDR} = 2.80 * 10^{-8}$). FRN peak latencies for all participants and conditions are plotted after normalizing for mean FRN latency within participant.



Supplementary Figure 10: Subjective ratings of reward expectations correspond with behavioral modeling but with slight bias. (a) Comparison of win probabilities derived from subjective ratings and logistic model fit to behavioral outcomes in an example participant. Smaller single trial rating markers colored green for correct and red for incorrect responses show overall distinction between easy and hard conditions. Larger markers of block-averaged data show subjective ratings are positively biased in hard conditions and negatively biased in easy conditions relative to objective behavioral accuracy. (b) Group level comparison of single-trial win probability as estimated by subjective ratings and behavioral modeling. Linear fit overlaid as black line for visualization of broad agreement. (c) Group level bias computed as the average difference between model-based and rating win probabilities confirm overall positive bias in hard conditions and negative bias in easy conditions.

Supplementary Table 1: Reproducibility of model coefficients for traditional ERP metrics across cohorts. Model coefficients from RL model for linear mixed-effects multiple regression of traditional FRN mean window, FRN peak-to-peak, and P3 mean window ERP metrics are reported for two cohorts separately and combined. RPE value and RPE magnitude results are reproducible for FRN and P3 metrics, respectively, but significance of other predictors varies across cohorts. Stars indicate significance (: $q_{FDR} < 0.05$; **: $q_{FDR} < 10^{-3}$; ***: $q_{FDR} < 10^{-6}$).*

ERP Metric	**Cohort**	**Model Coefficients**			
		Expected Value	**RPE Value**	**RPE Magnitude**	**Probability**
FRN Mean Window	1	0.843*	4.237***	0.960*	-0.053
	2	0.454	2.755***	-0.161	-0.859**

	Both	0.602*	3.443***	0.382	-0.508*
FRN Peak-to-Peak	1	-0.371	-1.686*	0.990	0.224
	2	0.413	-0.894*	0.391	-0.041
	Both	0.037	-1.277**	0.666*	0.083
P3 Mean Window	1	-0.822	0.567	1.887**	-0.028
	2	-0.422	0.862*	1.644**	0.208
	Both	-0.599	0.759*	1.721***	0.093

Supplementary Table 2: ERP amplitude model performance comparisons.
Model performance is reported as AIC averaged within three windows at Fz and Pz for all outcome-based and RL-based models relative to baseline AIC from a null model including only random effects for participants. Mean AIC for the null model is reported at the bottom. Lower AIC indicates better model performance. Stars indicate the best performing model for a given epoch and electrode. Relative likelihoods are reported for competing models with probabilities exceeding 1%.

Model	**Mean AIC at Frontal Site Fz**			**Mean AIC at Posterior Site Pz**		
	191-241 ms	**283-333 ms**	**355-405 ms**	**191-241 ms**	**283-333 ms**	**355-405 ms**
Outcome Value	-1906	-899	-705	-374	-347	-314
Outcome Valence	-2045	-901	-656	-509	-347	-316
Full RL	-2325*	-919*	-721*	-673 (RL = 0.54)	-409*	-336*
RL without RPE Magnitude	-2325 (RL = 0.78)	-862	-680	-674*	-276	-330 (RL = 0.04)
RL without Probability	-2293	-746	-294	-662	-409 (RL = 0.83)	-220
Null Model	142,897	145,788	145,056	141,780	146,384	147,407

Chapter 3: Human intracranial recordings reveal mixed coding of reward and salience prediction errors in fronto-insular cortices

Authors:
Colin W. Hoy[1], David King-Stephens[2], Kenneth D. Laxer[2], Peter Weber[2], Jack J. Lin[3], Robert T. Knight[1,4]

1. Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, CA 94720, USA
2. Departments of Neurology and Neurosurgery, California Pacific Medical Center, San Francisco, California
3. Departments of Neurology and Biomedical Engineering, University of California Irvine, Irvine, CA 92603, USA
4. Department of Psychology, University of California, Berkeley, Berkeley, CA 94720, USA

ABSTRACT:

Prediction errors are critical computations underlying reinforcement learning (RL) and cognitive control. The valenced scalar value and non-valenced magnitude of reward prediction errors (RPEs) play complimentary roles in learning and behavioral adaptation, but their corresponding neural circuitry is not clear. Medial prefrontal cortex (MPFC) and insular cortex (IC) are key regions for performance monitoring that encode both RPE value and magnitude. These signals are hypothesized to modulate top-down control in lateral prefrontal cortex (LPFC), but it is unknown how these reward and surprise circuits interact. Here, we use behavioral modeling and local high frequency broadband (HFB) activity in intracranial EEG during an interval timing task to track RPE value and magnitude representations across MPFC, LPFC, and IC. We show that all three regions show large proportions of electrodes (~30-40%) coding for valenced and non-valenced RPEs. These representations were separate in some electrodes, but we also observed electrodes showing mixed selectivity to both RPE effects in all regions. Surprisingly, IC showed the largest proportion of sites coding for RPE value and magnitude separately, as well as the greatest proportion of sites coding for a mixture of both RPE features. Finally, onsets of RPE value effects were earlier in IC than MPFC, suggesting a leading role for the insula in RPE coding. Collectively, these results clarify the interactions between neuronal circuits representing reward and surprise.

INTRODUCTION:

Adaptive behavior requires predicting which stimuli or actions are associated with the most valuable outcomes, and learning is dependent on monitoring feedback for surprising

Predictive coding frameworks formalize these computations in terms of prediction errors (PEs), or the difference between actual and expected outcomes. Reward prediction errors (RPEs) convey both valence (better or worse than expected?) and magnitude (how surprising?) and are important for optimizing future decisions to maximize reward. In particular, RPE valence determines whether actions should be repeated or extinguished, and RPE magnitude modulates the strength of adaptation or change in learning rate (Collins & Frank, 2015; Glimcher, 2011; Sutton & Barto, 1998). Salient outcomes that are unexpected or novel can also generate large PEs regardless of valence that trigger orienting processes such as reallocation of attention, cognitive control, and memory resources (Ullsperger, Danielmeier, et al., 2014; Wessel & Aron, 2017).

Valenced and non-valenced RPEs play complimentary roles in learning (Rouhani & Niv, 2021), but the nature of their neural representations is unclear. Foundational studies of reinforcement learning in systems neuroscience observed dopaminergic midbrain neurons that track both the valence and magnitude of RPEs (Schultz et al., 1997; Watabe-Uchida et al., 2017). Several decades of work have identified a core dopaminergic network for behavioral choices during value-based decision making including ventral striatum, orbitofrontal cortex, and anterior cingulate cortex (Litt et al., 2011; O'Doherty et al., 2007; Rangel et al., 2008). Neuroimaging research in humans has also found RPE coding in a variety of other areas, including insular, mid-cingulate, posterior cingulate, lateral prefrontal, parietal, and even sensory and motor cortices (Fouragnan et al., 2018; Vickery et al., 2011). These observations have led to the proposal that reward variables such as expected value and RPEs are not implemented in specialized regions but are instead computed in parallel across distributed circuits (Hunt & Hayden, 2017; Rushworth et al., 2012). However, distributed representations do not imply RPE signals are used for the same functions in all of these regions, so characterizing differences in coding schemes can help determine the specific roles of each node in the network.

Interpreting these widespread RPE signals is complicated in regions that also respond to non-valenced salience factors such as RPE magnitude. A recent fMRI meta-analysis identified medial prefrontal cortex (MPFC) and insular cortex (IC) as key regions of overlap between networks for PE valence and surprise (Fouragnan et al., 2018; McGuire et al., 2014). These regions are strongly interconnected and together form salience and cingulo-opercular control networks, both of which are involved in performance monitoring and adjusting control demands in response to both task-relevant and novel, unexpected stimuli (Gratton et al., 2018; Uddin, 2015). MPFC and IC are known to relay feedback to lateral prefrontal cortex (LPFC), which utilizes these updates to adjust task execution (M M Botvinick et al., 2001; Oemisch et al., 2019; Shenhav et al., 2013). In fact, LPFC activity has been observed to encode predictions (Boroujeni et al., 2021; Dürschmid et al., 2018), valenced RPEs (Asaad et al., 2017; Asaad & Eskandar, 2011; Bartolo & Averbeck, 2020), and non-valenced surprise effects (Alexander &

Eskandar, 2011; Bartolo & Averbeck, 2020), and non-valenced surprise effects (Alexander & Brown, 2018; Grohn et al., 2020; Oemisch et al., 2019). Progress in reinforcement learning and cognitive controls requires precise mappings of valenced RPE value and non-valenced RPE magnitude in time and space.

To resolve the distinct roles of these regions in reward and surprise processing, we characterized RPE value and magnitude coding across MPFC, IC, and LPFC using intracranial recordings in 10 epilepsy patients with implanted stereoelectroencephalography (SEEG) or electrocorticography electrodes (ECoG) during pre-surgical monitoring for treatment of medication refractory seizures. Patients performed an interval timing task with a difficulty manipulation designed to dissociate RPE valence and magnitude, and individual patient behavior was modeled to obtain single-trial estimates of expected value, which was then contrasted with the actual outcomes to obtain scalar RPEs. We then extracted high frequency broadband (HFB) power known to correlate with local population activity (Dubey & Ray, 2019; Lachaux et al., 2012; Leszczyński et al., 2020; Manning et al., 2009) and predicted these data using single-trial RL features in sliding window multiple regression analyses.

Our results show large proportions of LPFC, MPFC, and INS track expected value, RPE value, and RPE magnitude, confirming their multiple roles in reward and surprise processing. Notably, these signals were more common in the insula, which also showed the largest proportion of sites coding for both features together. Finally, onset latencies showed RPE value appears in IC before MPFC, with LPFC falling in between. These results suggest a potential leading role for the insula in surprise and valence processing.

METHODS:

Participants:

Data was collected from ten patients (mean ± SD [range]: 35.2 ± 13.1 [21-57] years old; 1 woman; see Table 1 for patient demographics and electrode coverage) undergoing neurosurgical treatment for medically refractory epilepsy. Patients were implanted with stereotactic (SEEG) or subdural grid or strip (ECoG) electrodes, and electrode placement and medical decisions were determined solely by the clinical needs of the patient. Patients were observed in the hospital for approximately a week, and those willing to participate performed the Target Time behavioral task during breaks in their clinical treatment. Informed consent was obtained according to experimental protocols approved by the University of California, Berkeley, University of California, Irvine, and California Pacific Medical Center Committees on Human Research. Patients had normal IQ (>85) and spoke fluent English.

Table 1: Patient demographics, electrode coverage, and behavior.

*For Button colum, "Kb" indicates responses were collected using the space bar on the built-in laptop keyboard, while "RTBox" indicates a USB button box was used. *For IR87, three runs used the RTbox device, while the keyboard was used to capture responses on the fourth run.*

SBJ	Age (years)	Sex	Task Version	Button	Number of Electrodes			Number of Trials		Accuracy (%)	
					LPFC	MPFC	Insula	Easy	Hard	Easy	Hard
S01	24	M	1.8.7	Kb	14	16	0	299	297	81.6	23.9
S02	23	M	1.8.2	Kb	43	16	13	140	138	67.9	18.8
S03	27	M	1.8.7	Kb	16	18	3	132	149	75.0	20.1
S04	28	M	1.8.7	Kb	25	17	6	145	149	62.8	21.5
S05	57	M	1.8.7	Kb	24	5	5	147	147	70.1	19.7
S06	47	M	1.8.8	Kb	15	8	6	141	133	77.3	21.8
S07	21	F	2.4.5	RTBox	13	9	8	144	143	68.3	20.0
S08	41	M	2.4.5	Kb	16	8	10	145	143	75.6	15.9
S09	52	M	2.4.7	RTBox*	19	8	8	444	446	83.9	15.6
S10	32	M	2.4.8	RTBox	3	1	5	286	282	81.2	17.4

Target Time behavioral task:

The Target Time interval timing task was written in PsychoPy (Peirce, 2008) (v1.85.3) and consisted of four blocks (two easy and two hard) of 75 trials each (see Fig 1 A for task schematic). Two patients completed the task twice, and one patient completed the task three times. The order of block difficulty was fixed as either two easy followed by two hard or alternating from easy to hard (Table 2). Following central fixation and a randomly chosen inter-trial interval ranging from 0.2 to 1.2 s (see Table 2), trials began with presentation of a visual motion cue at a constant speed to arrive at a target at the one-second temporal interval. Participants estimated the interval via button press using the space bar on a keyboard or an RTBox (v5/6) response device (Li et al., 2010). In the first version of the task (n = 6), the motion cue was upwards in a straight line towards a bullseye target, and in a second version (n = 4), the motion cue was counter-clockwise starting and ending at the bottom of a ring of dots on which a gray target zone was centered. The size of the bullseye and the width of a gray target zone indicated the tolerance for successful responses. Veridical win/loss feedback was presented from 1.8 s to either 2.6 or 2.8 s (Table 2) and composed of (1) the tolerance cue turning green/red, (2) cash register/descending tones auditory cues, and (3) a black tick mark denoting the response time (RT) on the ring. Participants received ±100 points for wins/losses. Tolerance was bounded at ± 15-200 or 15-400 ms (Table 2), and separate staircase algorithms for easy and hard blocks adjusted tolerance by -3/+12 and -12/+3 ms following wins/losses, respectively. Participants learned the interval in five initial training trials in which visual motion completed the full linear track or circle. For all subsequent trials, dot motion halted after 400 ms to

prevent visuo-motor integration, forcing participants to rely on external feedback. Training concluded with 15 easy and 15 hard trials to initialize both staircase algorithms to individual performance levels. For the second task version, main task blocks introduced neutral outcomes on a random 12% of trials that consisted of blue target zone feedback, a novel oddball auditory stimulus, no RT marker, and no score change.

Table 2: Target Time paradigm parameters.
For Block Order, E refers to easy blocks and H refers to hard blocks.

Task Version	Motion Cue	Inter-Trial Intervals (s)	Block Order	Error Tolerance Limits (s)	Neutral Outcomes
1.8.2	Linear	0.5, 0.85, 1.2	EEHH	0.2, 0.015	No
1.8.7	Linear	0.2, 0.4, 0.7, 1.0	EEHH	0.2, 0.015	No
1.8.8	Linear	0.2, 0.4, 0.7, 1.0	EEHH	0.2, 0.015	No
2.4.5	Circular	0.2, 0.4, 0.7, 1.0	EEHH	0.2, 0.015	Yes
2.4.7	Circular	0.7, 1.0	EHEH	0.4, 0.015	Yes
2.4.8	Circular	0.7, 1.0	EHEH	0.4, 0.015	Yes

Behavioral modeling:

The relationship between the tolerance around the target interval and expected value was fit to individual participant behavior using logistic regression. Specifically, tolerance was used to predict binary win/loss outcomes across trials using the MATLAB function *glmfit* with a binomial distribution and logit linking function. Trials with neutral outcomes were excluded because they were delivered randomly and thus not reflective of performance. The probability of winning (p_{win}) for each participant was computed as:

$$p_{win} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 t)}}$$

where β_0 is the intercept and β_1 is the slope from the logistic regression, and *t* is the tolerance on a given trial. Expected value was derived by linearly scaling the probability of winning to the reward function ranging from -1 to 1. RPE value was then computed by subtracting expected value from the actual reward value, and RPE magnitude was computed as the absolute value of RPE value. See Figure 1c for model predictions by condition.

iEEG data collection, localization, and preprocessing:

The data were recorded at either the University of California Irvine Medical Center (n = 9), USA or California Pacific Medical Center (n = 1), USA. Patients at Irvine were implanted with stereo-EEG (SEEG) electrodes with 5 mm spacing, and the patient at CPMC was implanted with strips of electrocorticography (ECoG) electrodes with 1 cm spacing. At both sites, electrophysiology and analog photodiode event channels were recorded using a 256-channel Nihon Kohden Neurofax EEG-1200 recording system and sampled at 500 (n = 3), 1000 (n = 3), or 5000 Hz (n = 4). For five patients, analog photodiode channels and a subset of iEEG channels were recorded in a separate Neuralynx ATLAS recording system at Irvine at 4000 (n = 1) or 8000 Hz (n = 4). Photodiode events were then aligned to the iEEG data acquired in parallel via the Nihon Kohden clinical amplifier via cross-correlation of shared iEEG channels.

Pre-operative T1 MRI and post-implantation CT scans were collected as part of standard clinical care, and recording sites were reconstructed in native patient space by aligning these scans via rigid-body co-registration according to the procedure described in Stolk et al. (Stolk et al., 2018). Anatomical locations of electrodes were determined by manual inspection in native patient space under supervision of a neurologist. Electrode positions were then normalized to group space by warping the patient MRI to a standard MNI 152 template brain using volume-based registration in SPM 12 as implemented in Fieldtrip (Stolk et al., 2018). Group-level electrode positions are plotted in MNI coordinates relative to the cortical surface of the fsaverage brain template from FreeSurfer (Dale et al., 1999).

Data cleaning, preprocessing, and analyses were conducted using the Fieldtrip toolbox (Oostenveld et al., 2011) and custom Python and MATLAB code. Raw iEEG traces were manually inspected by a neurologist for epileptic spiking and spread, as well as artifacts (e.g., machine noise, signal drift, amplifier saturation, etc.). Data in regions or epochs with epileptiform or artifactual activity were excluded from further analyses. Preprocessing included resampling data to 1000 Hz (for datasets recorded at sampling frequencies > 1000 Hz), bandpass filtered using a Butterworth filter from 0.5-300 Hz, re-referenced (bipolar to adjacent electrodes for SEEG data; common average reference across all channels for ECoG data), and bandstop filtered at 60, 120, 180, 240, and 300 Hz (Butterworth filter with 2 Hz bandwidth) to remove line noise and harmonics. Continuous data were then visually inspected to ensure all epochs with artifacts or spread from epileptic activity were removed. Finally, trials were rejected for task interruptions and behavioral outliers (RTs missing, < 0.5 s, > 1.5 s, or > 3 standard deviations from that patient's mean), resulting in 274-890 trials per patient (mean ± S.D.: 405.0 ± 210.6).

High frequency broadband power extraction and modeling:

Time series data were filtered to high frequency broadband (HFB) ranges known to correlate with local multi-unit activity (Leszczyński et al., 2020; Ray et al., 2008; Rich & Wallis,

2017). Specifically, data were segmented from -0.25 to 1.2 s relative to feedback onset, and multitaper time-frequency transformations with 50 ms windows were used to extract power from sub-bands ranging from 70 to 150 Hz in 10 Hz steps. These HFB power values were then log transformed to account for their log-normal distribution (Buzsáki & Mizuseki, 2014) in preparation for linear modeling. To normalize these power values against baseline activity, permutation distributions were created for each channel by taking the mean and standard deviation of baseline power values from -0.25 to -0.05 s relative to stimulus onset from 500 iterations of sampling trials with replacement. Feedback-lock power values were then z-scored using the average mean and standard deviation values from those permutation distributions of pre-stimulus baseline power values. This process avoids normalizing HFB power to pre-feedback data which may contain post-response activity and is robust to noisy outlier trials that can skew the baseline data. Finally, sub-bands were averaged together to create a single HFB power time series.

A sliding window approach was then used to average normalized single-trial HFB power values in 50 ms windows stepping by 25 ms from 0 to 0.6 s post-feedback. Multiple linear regression was then used to predict these single-trial HFB power data with the RL model containing expected value, RPE value, and RPE magnitude for each window. Statistical significance of model coefficients was assessed using a non-parametric bootstrap procedure by which the linear regression was repeated with shuffled model regressors across 1000 iterations, and two-sided *p* values were obtained by comparing the true model coefficients to this null distribution. These *p* values were corrected for multiple comparisons using the false discovery rate (FDR) methods of Benjamini & Hochberg (Benjamini & Hochberg, 1995) across regressors and time points for each channel. Corrected *p* values are referred to as q_{FDR} throughout the manuscript. Channels were considered to be significantly predicted by a model regressor if any HFB power window had a model coefficient with $q_{FDR} < 0.05$.

We compared the proportion of electrodes showing a significant effect across regions at the group level using Wilcoxon signed-rank tests to account for the lack of normal distributions in proportions, and these *p* values were FDR corrected for the number of region comparisons. Effect onsets are reported as the center of the first window showing significance for that channel and regressor, and these onsets are aggregated across subjects separately for RPE value and RPE magnitude and compared across regions using two-sided t-tests, which were FDR corrected for the number of regions compared.

RESULTS:

We collected behavioral data from 10 patients while recording from implanted SEEG and ECoG electrodes in LPFC, MPFC (primarily mid-cingulate cortex with some dorsomedial and anterior cingulate sites), and IC (Fig. 1d; see Table 1 for patient demographics, electrode coverage, and behavior). These patients performed a Target Time task that dissociates valenced

RPE value and non-valenced RPE magnitude by modulating reward expectations via task difficulty manipulations in an interval timing paradigm. At the beginning of each trial, patients saw a target zone indicating the temporal range of responses tolerated as correct (Fig. 1a). They then estimated the temporal interval via button press by means of extrapolation from linear or circular visual motion and received audiovisual feedback indicating their reaction time (RT) and reward (win or loss). Error tolerance was adjusted after each trial by two staircase algorithms to clamp accuracy at 74.4 ± 6.9% and 19.5 ± 2.6% (mean ± SD) in easy and hard blocks, respectively. This design dissociates outcome valence and probability by manipulating whether wins or losses are surprising, thereby separating valenced and non-valenced RPEs. Four patients performed a version of the task that delivered neutral outcomes with no RT feedback on 12% of trials as an additional source of surprise.

Behavioral modeling:

In order to separate valenced and non-valenced RPE features, we used computational modeling of individual patient behavior to derive single-trial estimates of expected value, RPE value, and RPE magnitude. For each patient, we used logistic regression to predict binary win/loss outcomes across the entire session using error tolerance (Fig. 1b; see inset for group model fits). This model yields patient-specific win probabilities for a given tolerance, which was linearly scaled to the reward function (1, 0, or -1 for winning, neutral, or losing outcomes) to quantify expected value for every trial. Single-trial RPE values were computed by subtracting the expected value from the outcome value, and RPE magnitudes were defined as the absolute value of RPE. Notably, different reward expectations across easy and hard conditions shift the RPE valence of neutral outcomes to negative in easy blocks and positive in hard blocks (see model predictions in Fig. 1c).

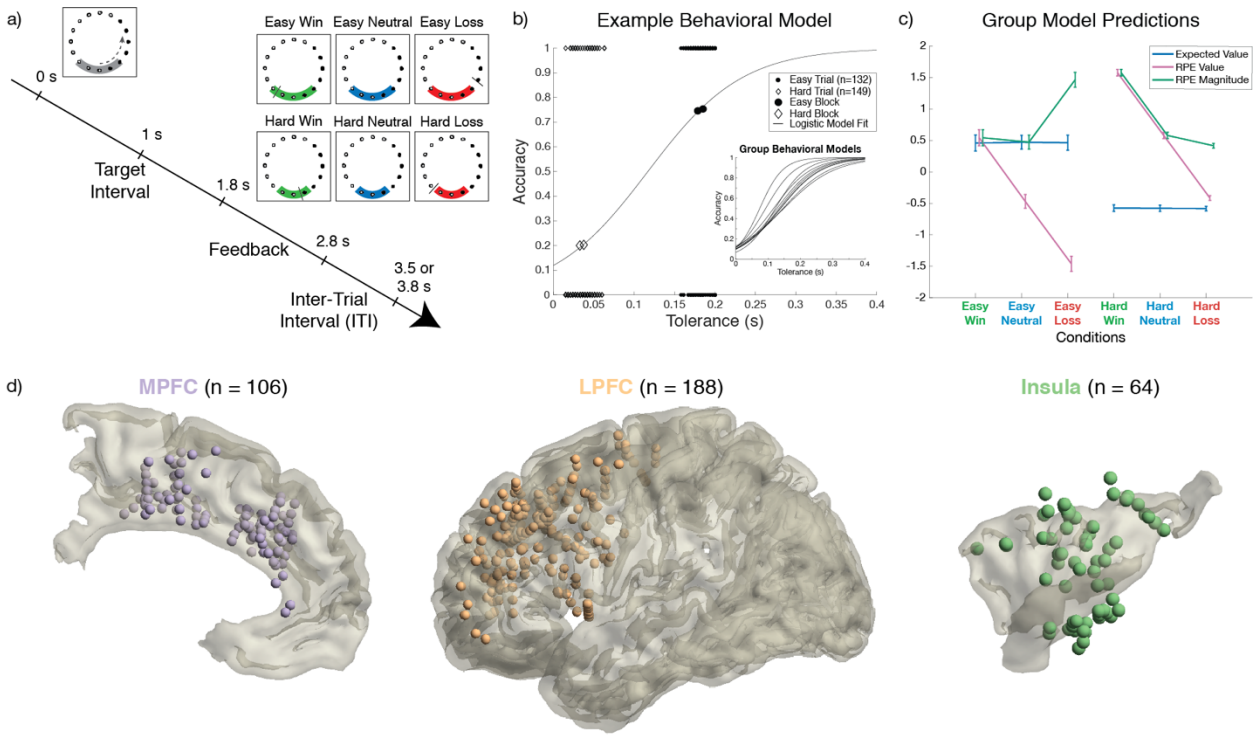


Figure 1: Task design, behavioral modeling, and iEEG recording sites. (a) Participants pressed a button timed to the estimated completion of lights moving around a circle. The gray target zone cue displayed error tolerance around the 1 s target interval. Audiovisual feedback is indicated by the tolerance cue turning green for wins and red for losses. A black tick mark displayed RT feedback. For 4 patients, blue neutral feedback was given with no RT marker on 12% of randomly selected trials. (b) Tolerance and outcome data for an example participant. Larger markers show block level accuracy; smaller markers show binary single trial outcomes. Model fit using logistic regression provides single trial estimates of win probability, which is converted to expected value. Inset shows win probability curves across all participants. (c) Predictions for RL model predictors. Error bars indicate standard deviation between participants. (d) Reconstruction of iEEG recording sites across all participants plotted by ROI on a standardized group brain after mirroring all electrodes to the left hemisphere.

High frequency broadband power reveals the spatiotemporal distribution of RL variables

To map the spatial and temporal distribution of RPE value and magnitude coding, we extracted and normalized high frequency broadband (HFB) power from 75-150 Hz at each electrode in LPFC, MPFC, and IC as a proxy for local population activity (Fig. 2a and 2b) (Lachaux et al., 2012; Manning et al., 2009). Single-trial HFB power was averaged in 50 ms windows sliding by 25 ms from 0 to 600 ms after feedback onset, and these averaged HFB power values were predicted by the RL model using multiple linear regression. The resulting model coefficients for each window provide time series depicting the evolution of expected value, RPE value, and RPE magnitude for each electrode (Fig. 2c).

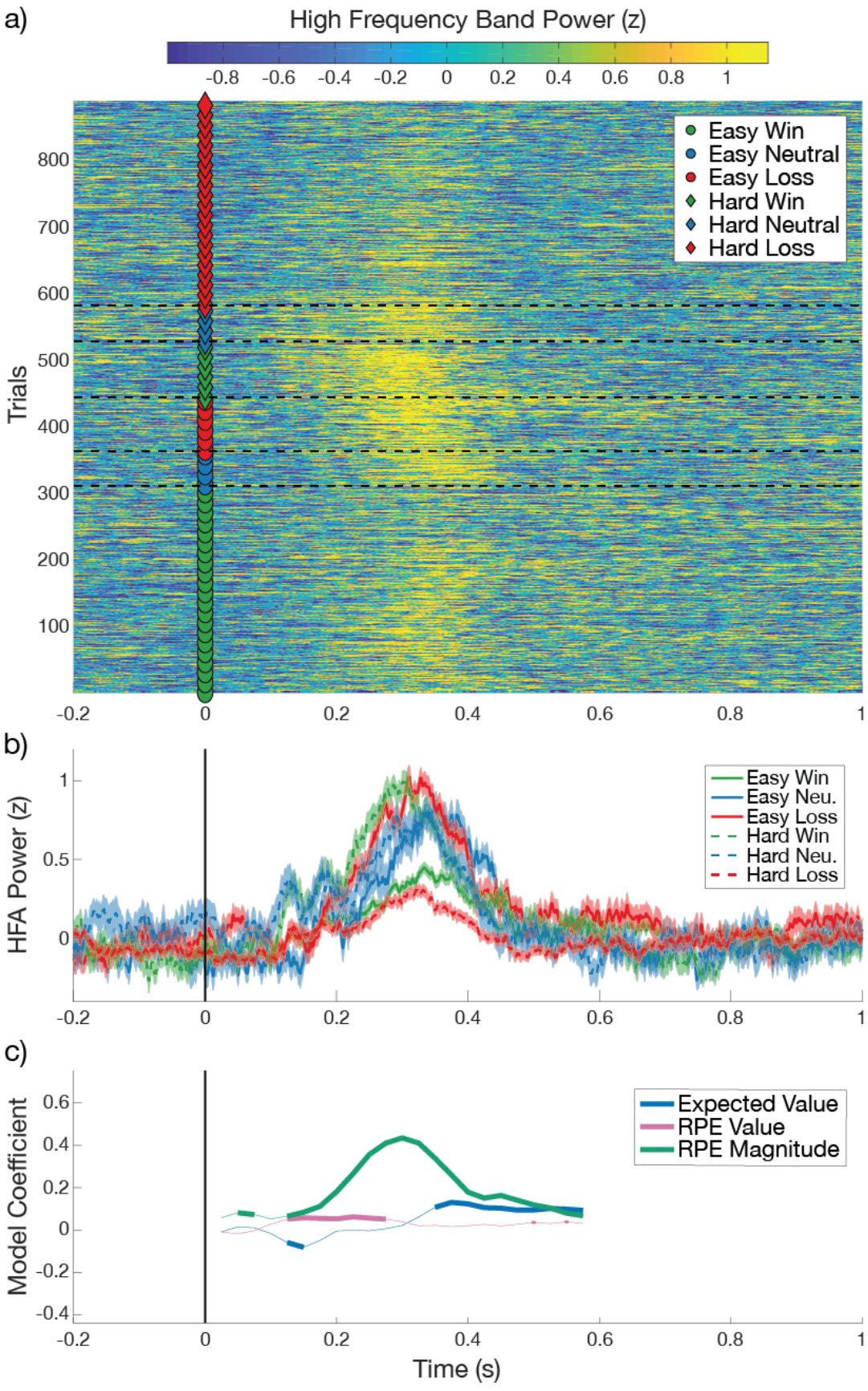

a)
High Frequency Band Power (z)
-0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1
Easy Win
Easy Neutral
Easy Loss
Hard Win
Hard Neutral
Hard Loss
Trials
800 700 600 500 400 300 200 100
-0.2 0 0.2 0.4 0.6 0.8 1
b)
HFA Power (z)
1 0.5 0
Easy Win
Easy Neu.
Easy Loss
Hard Win
Hard Neu.
Hard Loss
-0.2 0 0.2 0.4 0.6 0.8 1
c)
Model Coefficient
0.6 0.4 0.2 0 -0.2 -0.4
Expected Value
RPE Value
RPE Magnitude
-0.2 0 0.2 0.4 0.6 0.8 1
Time (s)

Figure 2: Example high frequency broadband (HFB) power modeling. (a) Single-trial HFB power at an example electrode in the insula is plotted time-locked to feedback (markers at feedback indicate condition). (b) Condition averaged HFB power (error bars represent standard error of the mean). (c) Model coefficients from regression analysis predicting single-trial HFB power with RL model variables (expected value, signed RPE, and unsigned RPE) averaged in 50 ms sliding windows (step size of 25 ms). Significant model coefficients ($q_{FDR} < 0.05$) are plotted in bold.

Aggregating across patients and time windows, substantial proportions of electrodes in each region encoded expected value, RPE value, and RPE magnitude (Fig. 3). In MPFC, 39.6 ± 7.5% (mean ± S.E.M.) of electrodes had HFB power significantly predicted by expected value, 41.5 ± 8.0% predicted by RPE value, and 35.8 ± 7.7% predicted by RPE magnitude. Similarly, in LPFC 32.4 ± 6.0% of electrodes showed HFB power tracking expected value, 38.8 ± 7.4% tracking RPE value, and 39.4 ± 8.7% tracking RPE magnitude. Surprisingly, IC showed the greatest proportion of sites coding for all RL variables, with 48.4 ± 9.1% (mean ± S.E.M.) of electrodes significantly predicted by expected value, 57.8 ± 8.7% predicted by RPE value, and 46.9 ± 7.8% predicted by RPE magnitude. Differences across regions in proportions of electrodes showing any of these effects were not significant after correcting for multiple comparisons (all $q_{FDR} > 0.07$).

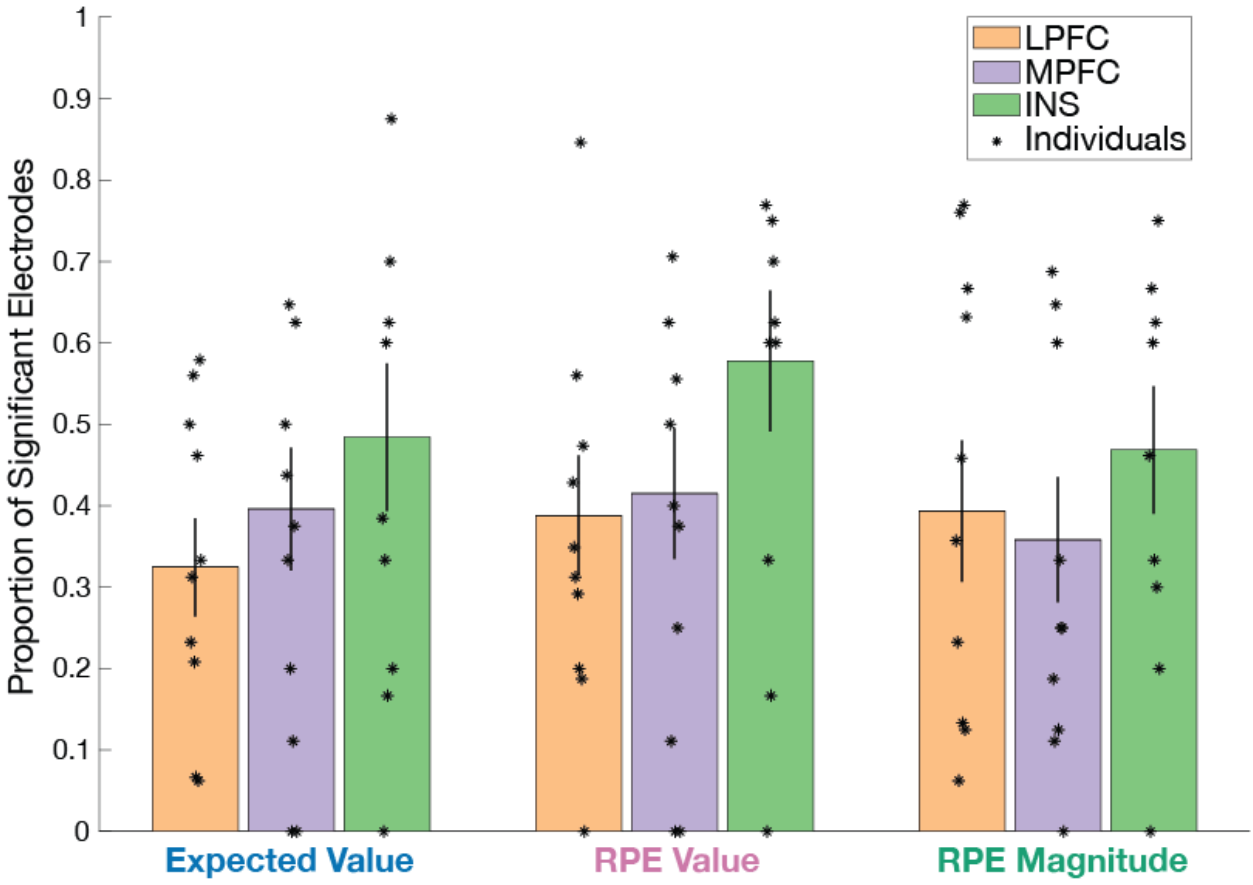


Figure 3: Proportion of sites showing significant effects by ROI. Bars show group averaged proportions of recording sites with at least one time bin where HFB power was significantly predicted by each of three RL model variables. Error bars represent standard error of the mean, and asterisks indicate data for individual patients.

Separate and mixed selectivity for RPE value and magnitude

In contrast to neurons in primary sensory and motor cortices, neurons in higher-order association regions often show "mixed selectivity" and respond to combinations of task-relevant features (Rigotti et al., 2014) . Figure 4 shows the anatomical distribution of RPE value and magnitude coding electrodes aggregated across all patients and colored by whether HFB power is selective to one or both RPE variables. Consistent with prior reports, MPFC and LPFC electrodes selectively encoded RPE value (mean ± SD for MPFC: 34.0 ± 21.7%; LPFC: 30.3 ± 26.9%) or RPE magnitude (MPFC: 35.2 ± 28.7%; LPFC: 32.0 ± 25.3%), with a minority of sites showing a mixture of both (MPFC: 30.8 ± 25.3%; LPFC: 37.7 ± 25.0%). In contrast, IC showed a small set of selective RPE value (25.8 ± 26.7%) or RPE magnitude (17.9 ± 22.3%) responses, but the majority of significant sites respond to both RPE features (56.3 ± 20.0%). However, among sites showing coding for either or both RPE effects, there were no significant differences between regions in the proportions of sites coding for RPE value, RPE magnitude, or mixed RPE value and magnitude (all $q_{FDR} > 0.14$).

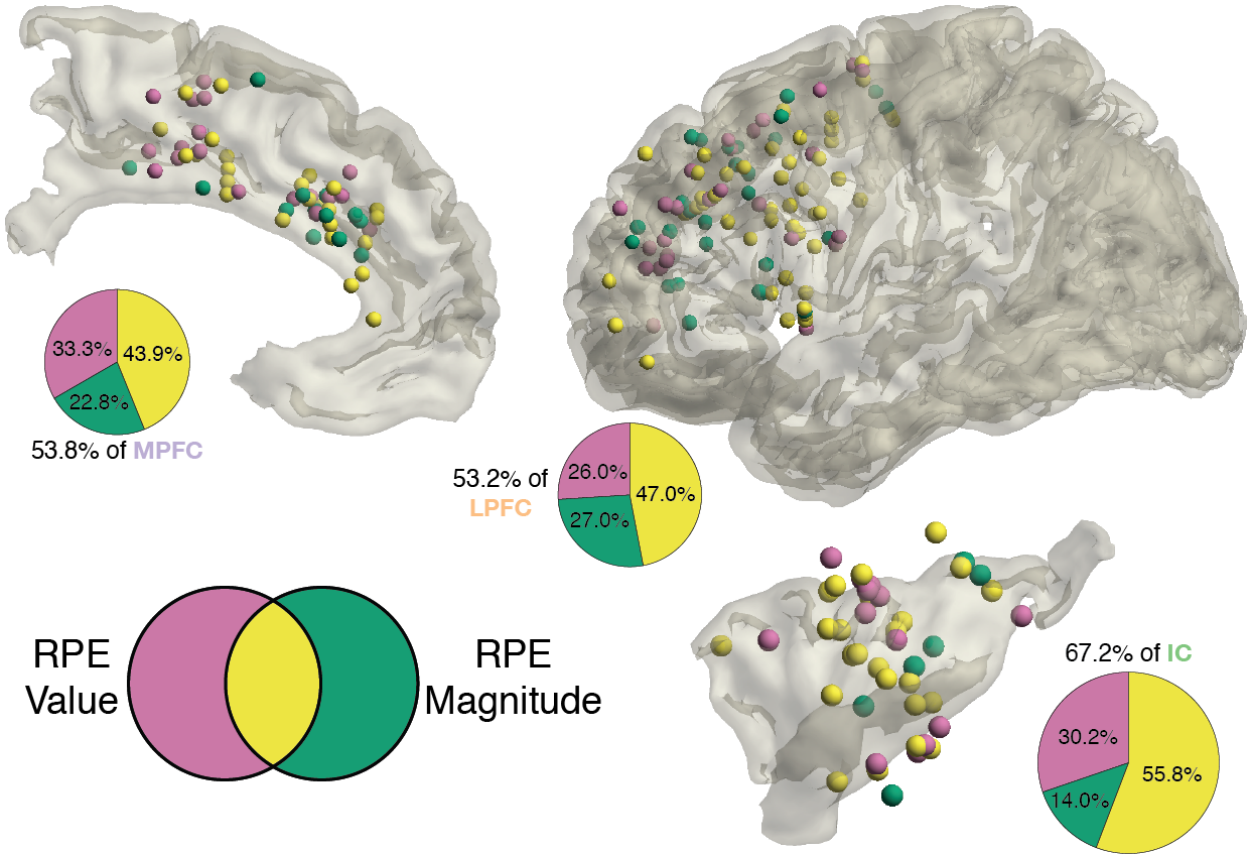


Figure 4: Anatomical distribution of separate and overlapping RPE value and magnitude effects. Electrodes with HFB power significantly predicted by RPE value and/or magnitude are plotted for MPFC (top left), LPFC (top right), and insular cortex (bottom right). As indicated by the Venn diagram color legend (bottom left), pink electrodes are predicted by RPE value only, green electrodes are predicted by RPE magnitude only, and yellow electrodes are predicted by both RPE effects. Pie charts display the percentage of electrodes aggregated across patients showing either or both effects out of all electrodes showing any significant RPE value or magnitude effects in that ROI, and the size of the pie chart and text label indicate the total percentage of sites in that region showing either or both RPE effects.

Temporal dynamics of RPE features in HFB power

We next examined the timing of RPE value and magnitude coding. Figure 5a and 5b shows the number of electrodes in each time window and region coding for RPE value and RPE magnitude, respectively. In both cases, RPE feature coding appears in IC and LPFC before MPFC. Onsets of RPE effects, which are defined as the middle of the first window to show a significant effect for each electrode, are displayed as violin plots for RPE value in Figure 5c and RPE magnitude in Figure 5d. RPE value onsets are significantly earlier in insula (mean ± SD: 0.212 ± 0.114 s) than MPFC (0.289 ± 0.127 s; $t(79) = 2.82$, $q_{FDR} = 0.018$), but not compared to LPFC (0.244 ± 0.112 s; $t(108) = 1.40$, $q_{FDR} = 0.25$). Mean RPE value onsets were also slightly earlier in LPFC than in MPFC, but the effect was not significant ($t(115) = -1.97$, $q_{FDR} = 0.15$). Onsets for RPE magnitude were also earliest in IC (0.228 ± 0.126 s), which was followed by MPFC (0.245 ± 0.081 s) then LPFC (0.262 ± 0.116 s), but these differences were not significant (all $q_{FDR} > 0.5$).

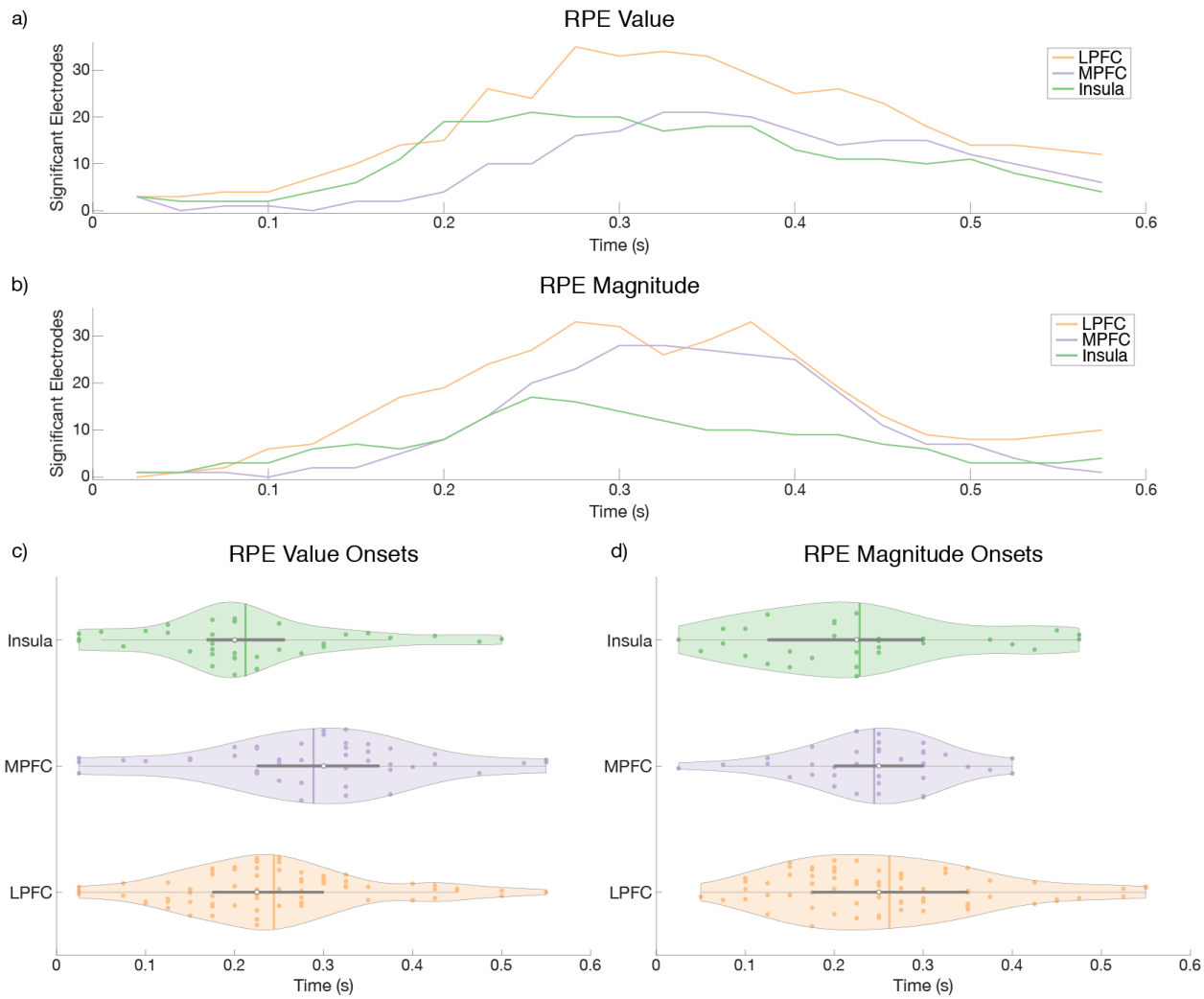


Figure 5: Temporal evolution and onset latencies of RPE value and magnitude effects by region. (a) Sum of electrodes significantly predicted by RPE value at each time window plotted for LPFC (orange line), MPF C (purple line), and insula (green line). (b) Same for RPE magnitude effects. Note the delayed onset of RPE value and magnitude effects in MPFC across both RPE features. (c) Violin plots show the distribution of the onset of RPE value effects for each electrode divided into ROIs. Onsets are defined as the center of the first window showing a significant effect. Mean onsets are plotted as vertical lines. Horizontal gray bars show inter-quartile ranges, and white dots show median onsets. (d) Same for onsets of RPE magnitude effects.

DISCUSSION:

Using HFB power as a proxy for local population activity, we observed that LPFC, MPFC, and IC all encode both valenced RPE value and non-valenced RPE magnitude, including mixed coding of both variables in individual electrodes. These signals were the most common in the insula, which also showed the greatest proportion of sites coding for overlapping RPE value and magnitude representations. Furthermore, RPE value onsets were significantly earlier in IC than MPFC, suggesting a potential leading role for the insula in reward coding.

The high spatiotemporal resolution of HFB power in iEEG data allowed us to resolve distinct RPE value and magnitude representations in neighboring MPFC sites, as well as electrodes that respond to both RPE features. This mixed coding observed with intracranial electrodes explains observations of scalp event-related potentials tracking reward and surprise that have both been localized to MPFC sources (Becker et al., 2014; Carlson et al., 2011; Cavanagh et al., 2019; Hauser et al., 2014). The close proximity of RPE value and magnitude signals in neighboring electrodes highlights the importance of intracranial data to segregate signals that cannot be distinguished in scalp recordings.

LPFC also showed diverse coding of RPE value and magnitude, which fits with its role in adjusting top-down control based on different types of feedback. These multiple PE types allow LPFC to update predictions for specific task features on future trials (Asaad et al., 2017; Bartolo & Averbeck, 2020; Boroujeni et al., 2021; Dürschmid et al., 2018; Oemisch et al., 2019). Interestingly, RPE feature onsets in MPFC were slower or not different than in LPFC, indicating the mechanisms by which MPFC relays feedback information to LPFC is not serial but likely occurs in parallel (Cisek & Kalaska, 2010; Shenhav et al., 2013).

Both IC and MPFC are part of the salience network, which is associated with bottom-up detection of novel or behaviorally relevant stimuli, but the earlier onset of RPE value in insula than MPFC suggests an asymmetric relationship between these regions. Indeed, fMRI data during resting state suggest insula is a key node for switching between states (Sridharan et al., 2008) and is also a causal outflow hub for controlling downstream regions (Cai et al., 2021). iEEG data also reveal effective connectivity flow from insula to MPFC signaling prediction errors (Bastin et al., 2017; Billeke et al., 2020). Overall, the latency differences we observed are consistent with a leading role for the insula in monitoring for reward and surprise. Future studies are needed to determine whether reward and surprise information are encoded in distinct sub-circuits with different directional influences between IC and MPFC.

One of the most striking features in our result is the mixture of RPE value and magnitude within individual electrodes. This pattern was most common in IC but occurred in all three regions. However, it is possible that separate neural populations are coding for reward and salience, but they are intermingled at a finer spatial scale than the resolution of HFB activity in iEEG, which is estimated at ~1.5 mm radius (Dubey & Ray, 2019). Alternatively, single unit recordings in animals have reported individual neurons in MPFC and LPFC with mixed selectivity for multiple reward-related decision variables (Hayden & Platt, 2010; Kennerley & Wallis, 2009). Future studies using single unit recordings in humans may determine whether valenced and non-valenced feedback processing is segregated, and whether these properties change across brain regions.

These coding schemes may also depend on functional and anatomical differences between subregions within the LPFC, MPFC, and IC regions of interest in this study. The insula has been parcellated into posterior and middle sensory, dorsal anterior cognitive, and ventral

anterior affective subregions (Evrard, 2019; Kurth et al., 2010; Nieuwenhuys, 2012) . Several studies have localized a variety of abstract reward and surprise signals to anterior insula (Loued-Khenissi et al., 2020; Vestergaard & Schultz, 2020; Yang et al., 2020), but valence processing related to feeding behavior has also been observed in posterior insula regions predominantly associated with sensory processing (Gehrlach et al., 2019). Similarly, MPFC shows variable sensitivity to reward and surprise depending on anterior posterior extent (Cole et al., 2009; Heukelum et al., 2020) and individual-specific sulcal morphometry (Amiez et al., 2013, 2016; Amiez & Petrides, 2014). One limitation of iEEG recordings in human patients is that recording locations are determined solely based on clinical rationale, so in our dataset, anterior IC and MPFC regions commonly studied in animal models of reward processing are underrepresented. Future studies may more densely sample these subregions and examine whether meso-scale anatomical differences affect RPE feature coding schemes, and whether those patterns reflect differences in function.

An important caveat to observing all modeled RL variables in each region analyzed is that this does not imply these regions use that information to perform the same function. For example, the strong connection to interoceptive circuits might enable MPFC and IC to utilize RPE magnitude to adjust pupil dilation, heart rate, and blood pressure (Gogolla, 2017; Kleckner et al., 2017; Kucyi & Parvizi, 2020) , while the same RPE magnitude signal in LPFC is likely more related to updating stimulus-response mappings in sensorimotor cortices (Bartolo & Averbeck, 2020; Norman et al., 2021). Causal interventions such as stimulation studies will be needed to better understand the computational role these signals play in each region (Nácher et al., 2019).

In summary, we find intermingled representations of valenced RPE value and non-valenced RPE magnitude in LPFC, MPFC, and IC, demonstrating the robust and redundant RPE coding in the brain. Although all regions showed a combination of separate and mixed coding of RPE value and magnitude, mixed selectivity electrodes for both effects were more common in the insula, consistent with its role in salience detection and valence processing. Additionally, RPE value appeared earlier in IC than MPFC, supporting a leading role for the insula in reward processing.

Acknowledgements:
We thank the patients for their time and selfless dedication.

Chapter 4: Intracranial Stroop recordings reveal parallel conflict processing in networks distributed across human cortex

Authors:
Colin W. Hoy[1], Kris L. Anderson[1], David King-Stephens[2], Kenneth D. Laxer[2], Peter Weber[2], Jack J. Lin[3], Robert T. Knight[1,4]

1. Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, CA 94720, USA
2. Departments of Neurology and Neurosurgery, California Pacific Medical Center, San Francisco, California
3. Departments of Neurology and Biomedical Engineering, University of California Irvine, Irvine, CA 92603, USA
4. Department of Psychology, University of California, Berkeley, Berkeley, CA 94720, USA

ABSTRACT:

Classical theories of conflict processing posit serial computations in which medial prefrontal cortex (MPFC) detects conflict before recruiting lateral prefrontal cortex (LPFC) to resolve the conflict and execute the correct response. However, the majority of studies testing this hypothesis have used EEG data unable to resolve distinct roles for these frontal regions or fMRI data unable to resolve within-trial conflict temporal dynamics. Consequently, the specific roles of MPFC, LPFC, and other regions such as orbitofrontal, insular, and temporal cortices during conflict processing remain poorly defined. Here, we utilized the high resolution of intracranial recordings in epilepsy patients performing a verbal color-word Stroop task to characterize the spatiotemporal dynamics of distributed brain networks across stimulus, decision, and response phases of conflict processing. High frequency broadband (HFB) activity detected substantial proportions of conflict sensitive sites spread across a distributed network of regions in frontal, insular, and temporal cortices. Segregating conflict effects by their trial dynamics revealed largely distinct sub-circuits within these networks for stimulus, decision, and response phases of conflict processing. Moreover, onsets and time courses of conflict effects provided evidence for parallel conflict processing across widely distributed functional networks, rather than a serial processing hierarchy composed of homogeneous anatomical regions. In sum, these findings emphasize heterogeneous functions within classical prefrontal control regions, contributions of non-classical regions such as orbitofrontal, insular, and temporal cortices, and rapid coordination of distributed networks in conflict processing.

INTRODUCTION:

Real life scenarios often involve a choice between multiple responses, and cognitive control networks in the brain must mediate between these options to implement adaptive

resolved to select the appropriate option. Default responses are fast, automatic, and prepotent, so implementing a different, less practiced response generates high conflict that requires increased control. For example, the classic color-word Stroop task presents the names of colors printed in different ink colors that are either congruent (ink and color name match) or incongruent (ink and color name differ), and participants must resolve conflict on incongruent trials by inhibiting the default, automatic color name reading response and instead report the ink color.

Previous literature has identified medial prefrontal cortex (MPFC) and lateral prefrontal cortex (LPFC) as key brain regions involved in conflict processing (Ridderinkhof et al., 2004; Ullsperger, Danielmeier, et al., 2014), but conflict effects have also been observed in a wider range of areas including orbitofrontal cortex (OFC), insula (INS), sensorimotor cortex (SM), parietal, and temporal cortices (TMP) (Freund, Bugg, et al., 2021; Haupt et al., 2009; Mansouri et al., 2017; Niendam et al., 2012). These results raise the question of whether these conflict effects reflect the same general neural computation in each region or the experimental contrasts used to identify them are non-specific and subsume multiple conflict-related functions. However, this is difficult to answer because conflict has been defined in a wide variety of conceptual and computational terms. Some neurobiological models of conflict specifically refer to co-activation of competing neural populations (Cisek & Kalaska, 2005; Pastor-Bernier & Cisek, 2011), but others broaden the term to refer to choice difficulty (Shenhav et al., 2013), error likelihood (Alexander & Brown, 2011; Brown & Braver, 2005), or time on task (Grinband et al., 2011). In this chapter, we adopt an inclusive operational definition of conflict as condition differences between high and low conflict trials and constrain the possible computations contributing to various conflict effects by characterizing the temporal dynamics of conflict effects in different brain regions.

Conflict monitoring theory has been one of the dominant theories of cognitive control to date and posits that MPFC detects conflict and recruits LPFC and attention networks to exert additional control and resolve the conflict in SM cortex (Botvinick et al., 2001; Botvinick et al., 2004; MacDonald et al., 2000). This implies a serial processing stream between conflict monitoring in MPFC and executive control functions in LPFC (Kerns, 2006; Kerns et al., 2004; Yeung et al., 2004). A large body of literature provides support for this account in the context of adaptation of behavior and neural activity to congruence sequence effects across trials, i.e. by comparing trials preceded by a congruent or incongruent stimulus (Braem et al., 2014; Duthoo et al., 2014). Specifically, error rates are reduced and the influence of conflict on reaction times (RTs) is weaker on post-conflict trials, presumably because more control resources are allocated to LPFC after performance monitoring regions like MPFC detected the difficult choice (Egner, 2007). However, it remains unclear whether similar mechanisms operate to detect and resolve conflict within a trial, primarily because standard non-invasive techniques like functional MRI and scalp EEG lack the combined spatiotemporal resolution to differentiate neural activity between control regions on the time scale of action.

In particular, within-trial conflict processing elicits distinct cognitive computations that unfold across different stages of the trial. Initial sensory processing creates a representation of the stimulus that must be transformed to a motor response, and in cases with multiple, mutually exclusive sensorimotor mappings, a decision must be made to select the appropriate transformation. Control systems must then inhibit the incorrect and execute the correct

response. Conflict effects occurring early in the trial before even the fastest responses likely reflect detection of the competition between responses. After this detection but before the response, control resources must be exerted to resolve the conflict in decision making circuits. Finally, performance monitoring circuits can detect the outcome of these difficult choices once the response begins, and those monitoring signals drive the between-trial adaptation underlying congruence sequence effects. Thus, conflict effects observed during stimulus, decision, and response phases of a trial can be interpreted as reflecting detection, resolution, and monitoring functions.

Intracranial EEG (iEEG) is ideally suited to address these issues due to its resolution on the scale of millimeters and milliseconds, and several iEEG studies have examined conflict-inducing paradigms such as the Stroop task (Bartoli et al., 2017; Cohen et al., 2008; Fu et al., 2018; Koga et al., 2010; Oehrn et al., 2014; Sheth et al., 2012; E. H. Smith et al., 2019). In particular, Tang et al. reported a cascade of control processing based on observations of earlier onsets of conflict effects in MPFC compared to LPFC and OFC (Tang et al., 2016), suggesting conflict monitoring theory may capture within-trial control dynamics and highlighting contributions outside MPFC and LPFC. However, this and other prior iEEG studies used relatively small sample sizes, focused exclusively on frontal cortex, and did not distinguish between processing stages. Furthermore, the most common behavioral measure of conflict is prolonged RTs on high conflict trials, and previous work did not fully account for confounds introduced by stimulus-locked averaging when comparing conditions with different RTs. Consequently, the full spatiotemporal evolution of within-trial conflict processing remains undefined.

In this study, we addressed conflict monitoring theory claims regarding the serial information flow between MPFC and LPFC for conflict detection, resolution, and monitoring functions by characterizing the evolution of conflict across stimulus, decision, and response phases of the trial. We utilized intracranial recordings in 17 patients with epilepsy performing a color-word Stroop task to examine conflict effects across a distributed set of frontal, insular, and temporal regions. High frequency broadband (HFB) power known to index local population activity (Leszczyński et al., 2020; Ray et al., 2008; Rich & Wallis, 2017) was used to track the spatial extent and temporal dynamics of conflict. We showed substantial proportions of sites in a distributed network of regions are involved in conflict processing. Furthermore, conflict effects in post-stimulus, pre-response, or post-response stages of the trial occurred in largely non-overlapping sub-circuits of these networks, suggesting stages of conflict processing evolve across independent but spatially intermingled networks. Finally, analysis of conflict onset latencies revealed heterogeneous timing within regions and parallel onsets across regions, providing evidence against the serial processing model implied by conflict monitoring theory. Collectively, these results emphasize the parallel and distributed nature of conflict processing in the human brain.

METHODS:

Participants:

Data was collected from 17 patients (mean ± SD [range]: 37.2 ± 13.4 [21-58] years old; 6 women; see Table 1 for patient demographics and behavior) undergoing neurosurgical treatment for medically refractory epilepsy. Patients were implanted with stereo-electroencephalography (SEEG; n = 15) or subdural electrocorticography (ECoG; n = 2) grid or strip electrodes. Electrode placement and medical decisions were determined solely by the clinical needs of the patient. Patients were observed in the hospital for approximately a week, and those willing to participate performed a color-word Stroop behavioral task during breaks in their clinical treatment and after informed consent was obtained according to experimental protocols approved by the University of California, Berkeley, University of California, Irvine, and California Pacific Medical Center Committees on Human Research. All patients had normal IQ and spoke English as a primary language except S08 and S12, which were both fluent in English with IQ above 85.

Table 1: Patient demographics and behavior.

Patient ID	Age (years)	Sex	Number of Trials	Mean RT (s)	Number of Errors
S01	25	M	605	0.702	3
S02	50	M	271	0.622	2
S03	34	M	290	0.697	0
S04	40	F	266	0.843	9
S05	21	F	245	0.778	0
S06	58	F	307	0.640	1
S07	34	M	301	0.582	6
S08	22	F	238	1.097	3
S09	33	M	128	0.831	3
S10	55	F	305	0.841	1
S11	23	M	281	0.557	5
S12	53	M	256	0.899	31
S13	28	M	293	0.764	9
S14	28	M	289	0.641	4
S15	57	M	269	1.153	6
S16	25	F	302	0.577	1
S17	47	M	307	0.760	3

Stroop behavioral task:

Patients performed a color-word Stroop task written in PsychoPy (Peirce, 2008) (v1.82.01) with congruent, incongruent, and neutral stimuli composed of the words “BLUE”,

"RED", "GREEN", and "XXXX" (neutral) printed in blue, red, and green ink on a gray computer screen (Fig. 1a). Color-word stimuli were presented for 1.5 s following inter-trial intervals randomly sampled from a uniform distribution between 1.05 and 1.8s s, except for patient S02 which had inter-trial intervals ranging from 2 to 2.3 s. They were asked to name the ink color, and verbal reaction times (RTs) were obtained from microphone recordings. RTs were manually marked as the first deflection in the audio waveform above baseline noise corresponding to color naming (i.e., pre-response vocalizations such as "uhh" were excluded). Error and partial error trials (e.g., "gree- blue!") were rare (see Table 1) and excluded from all analyses.

The task was organized in 9 blocks of 36 trials, except one patient (S02) with 12 blocks of 24 trials. One patient (S01) performed the task twice. The proportion of congruent trials was manipulated across blocks such that they were either mostly congruent (50% congruent, 33.3% neutral, and 16.7% incongruent), equal proportions (33.3% congruent, 33.3% neutral, and 33.3% incongruent), or mostly incongruent (16.7% congruent, 33.3% neutral, and 50% incongruent). Block types were randomized and occurred in equal numbers. RTs were z-scored for each patient across all conditions and tested at the group level using an ANOVA with fixed factors for trial type and block type and a random factors for patient.

iEEG data collection, localization, and preprocessing:

The data were recorded at either the University of California Irvine Medical Center (n = 16), USA or California Pacific Medical Center (n = 1), USA. Patients at Irvine were implanted with SEEG electrodes with 5 mm spacing or ECoG grids with 1 cm spacing. The patient at CPMC was implanted with ECoG strips with 1 cm spacing. At both sites, electrophysiology and analog photodiode event channels were recorded using a 256-channel Nihon Kohden Neurofax EEG-1200 recording system and sampled at 500 (n = 1), 1000 (n = 1), 5000 Hz (n = 14), or 10,000 Hz (n = 1). For one patient, analog photodiode channels and a subset of iEEG channels were recorded in a separate Neuralynx ATLAS recording system at Irvine at 4000 Hz. Photodiode events were then aligned to the iEEG data acquired in parallel via the Nihon Kohden clinical amplifier via cross-correlation of shared iEEG channels.

Pre-operative T1 MRI and post-implantation CT scans were collected as part of standard clinical care, and recording sites were reconstructed in native patient space by aligning these scans via rigid-body co-registration according to the procedure described in Stolk et al. (Stolk et al., 2018). Anatomical locations of electrodes were determined by manual inspection in native patient space under supervision of a neurologist. Electrode positions were then normalized to group space by warping the patient MRI to a standard MNI 152 template brain using volume-based registration in SPM 12 as implemented in Fieldtrip (Stolk et al., 2018). Group-level electrode positions are plotted in MNI coordinates relative to the cortical surface of the fsaverage brain template from FreeSurfer (Dale et al., 1999).

Data cleaning, preprocessing, and analyses were conducted using the Fieldtrip toolbox (Oostenveld et al., 2011) and custom MATLAB code. Raw iEEG traces were manually inspected by a neurologist for epileptic spiking and spread, as well as artifacts (e.g., machine noise, signal drift, amplifier saturation, etc.). Data in regions or epochs with epileptiform or artifactual activity were excluded from further analyses. Preprocessing included resampling data to 1000 Hz (for datasets recorded at sampling frequencies > 1000 Hz), bandpass filtered using a

Butterworth filter from 0.5-300 Hz, re-referenced (SEEG: bipolar to adjacent electrodes; ECoG: common average reference across all channels for strip cases or across all channels in each grid), and bandstop filtered at 60, 120, 180, 240, and 300 Hz (Butterworth filter with 2 Hz bandwidth) to remove line noise and harmonics. Continuous data were then visually inspected to ensure all epochs with artifacts or spread from epileptic activity were removed. Trials were rejected for task interruptions and behavioral outliers (errors and partial errors; RTs missing, <0.3 s, >2.0 s, or >3 standard deviations from the patient mean). Finally, trials were segmented from -0.25 to 2.5 seconds relative to stimulus onset and rejected for excessive variance in the preprocessed time series or the differentiated preprocessed time series. Exclusion criteria for trial variance were based on patient-specific thresholds of trial-level standard deviations ranging from 5 to 10 standard deviations. Between 0 and 13 trials per patient were rejected for excessive variance (mean ± S.D.: 3.5 ± 3.6 trials). In total, this process resulted in 128-605 trials per patient (mean ± S.D.: 291.4 ± 91.5) for further analyses (Table 1).

Table 2: Patient electrode coverage

Patient ID	Type	Lateral Prefrontal Cortex (LPFC)	Medial Prefrontal Cortex (MPFC)	Insula (INS)	Sensorimotor Cortex (SM)	Orbitofrontal Cortex (OFC)	Temporal Cortex (TMP)
S01	ECoG	14	16	0	6	5	15
S02	SEEG	19	5	5	0	6	4
S03	SEEG	18	1	0	0	1	2
S04	SEEG	11	3	1	5	4	12
S05	ECoG	38	17	0	21	1	0
S06	SEEG	24	3	6	3	5	14
S07	SEEG	17	3	3	1	3	12
S08	SEEG	39	15	5	9	7	7
S09	SEEG	22	17	8	5	2	20
S10	SEEG	10	10	3	0	0	24
S11	SEEG	38	14	10	7	5	3
S12	SEEG	14	8	0	0	6	1
S13	SEEG	11	11	0	0	2	18
S14	SEEG	25	17	6	0	3	32
S15	SEEG	21	5	5	7	3	6
S16	SEEG	31	40	2	0	6	40
S17	SEEG	15	14	8	14	9	19
Total	-	367	199	62	78	68	229

High frequency broadband power extraction and modeling:

Time series data were filtered to high frequency broadband (HFB) ranges known to correlate with local multi-unit activity (Leszczyński et al., 2020; Ray et al., 2008; Rich & Wallis, 2017) . Specifically, data were segmented from -0.25 to 2.5 s relative to stimulus onset, and multi-taper time-frequency transformations with 100 ms windows were used to extract power from sub-bands ranging from 70 to 150 Hz in 10 Hz steps. To normalize these power values, permutation distributions were created for each channel by taking the mean and standard deviation of power values across the entire trial from 500 iterations of sampling trials with replacement. Stimulus-locked power values were then z-scored using the average mean and standard deviation values from those permutation distributions of whole-trial power values, which allows analyses of baseline data that may be of interest for adaptation effects. Finally, sub-bands were averaged together to create a single HFB power time series. This procedure normalizes signal properties across channels, accounts for decreasing power in higher frequencies due to the 1/*f* property of electrophysiology data, and avoids the influence of outliers on amplitude normalization.

To interpret conflict effects according to their temporal context relative to the trial structure, we segregated single-trial HFB activity into stimulus, decision, and response windows. For stimulus windows, we averaged normalized single-trial HFB power values in 100 ms windows stepping by 25 ms from stimulus onset to the minimum RT for that patient. Decision making starts at an unknown time after stimulus information enters the system and before the response is prepared and executed. To account for this variability, a single decision window is computed by averaging normalized HFB power values on each trial from 100 ms before the patient's minimum RT through that trial's RT. This window captures the critical period in which control systems must resolve conflict to inhibit the incorrect response and prepare and execute the correct response. Finally, response windows are time-locked to RTs and are computed using the same 100 ms windows stepping by 25 ms from 50 ms pre-RT to 1 s post-RT.

HFB power in each window was then predicted using an ANOVA model with factors for current trial type (congruent, neutral, or incongruent), previous trial type, and block type (mostly congruent, equal proportions, mostly incongruent). Previous trial conflict and proportion congruent factors were included as controls for within trial dynamics. Statistical significance of main effects was assessed using a non-parametric bootstrap procedure by which the ANOVA analysis was repeated with shuffled design matrices across 1000 iterations, and two-sided *p* values were obtained by comparing the true *p* values to those obtained from the null distribution. These *p* values were corrected for multiple comparisons using the false discovery rate (FDR) methods of Benjamini & Hochberg (Benjamini & Hochberg, 1995) across factors and time points for each channel. Corrected *p* values are referred to as q_{FDR} throughout the chapter. Channels were considered to be significantly predicted by conflict if any HFB power window had a main effect of trial type with $q_{FDR} < 0.05$. Conflict effect onsets are defined using stimulus and decision windows and are reported as the first time point of the first window showing significance for that channel. To compare conflict onsets across patients with variable RTs, stimulus-locked onset latencies are normalized to mean RT across all conditions. To compare conflict onsets across regions, onset latency differences are computed within-subject between every pair of electrodes with significant conflict effects in different regions. Within-

subject onset latency difference analyses are only computed with stimulus-locked windows, i.e., stimulus and decision windows. A subset of these within-subject latency differences are plotted in Figure 4 to test the conflict monitoring theory hypothesis that conflict onsets should be earlier in MPFC than LPFC.

RESULTS:

We collected behavioral data from 17 patients implanted with iEEG electrodes performing a color-word Stroop task with congruent, neutral, and incongruent trials (Fig. 1a). The effect of conflict on RT data was tested using a three-way ANOVA with fixed effects for trial type (congruent, neutral, and incongruent) and block type (mostly congruent, equal proportions, and mostly incongruent) and random effects for patient, which revealed a significant main effect of trial type ($F = 156.4$, $p < 10^{-17}$; Fig. 1b). Errors were rare (see Table 1) and excluded from further analyses. To examine the neural activity underlying this conflict effect, we recorded from 1186 SEEG and ECoG electrodes (mean ± S.D. = 69.8 ± 28.9 electrodes per patient) across LPFC, MPFC, OFC, INS, SM, and TMP brain areas (Fig. 1c).

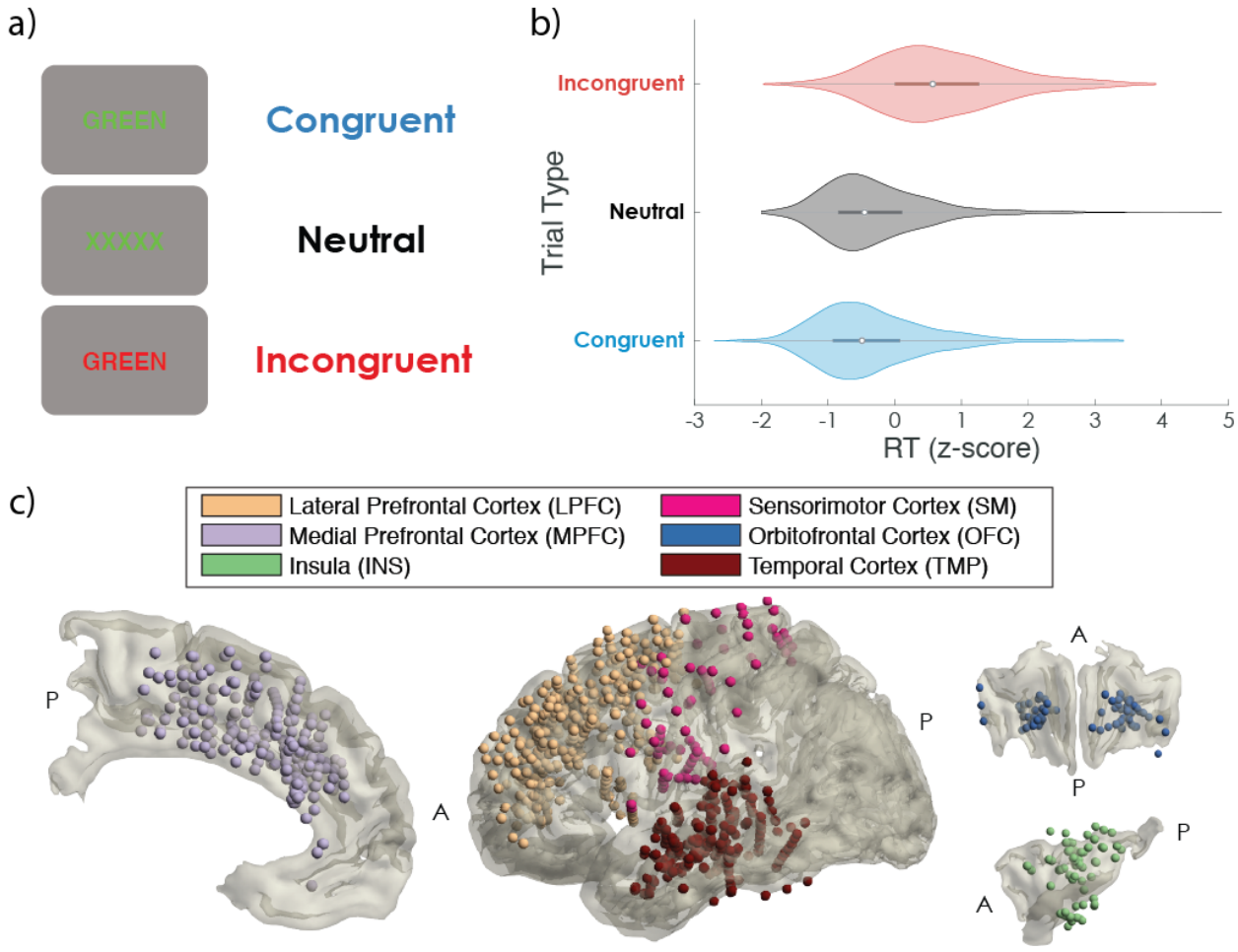


Figure 1: Stroop task design, behavior, and iEEG recording sites. (a) Participants spoke the ink color of visual text stimuli in three conditions. Congruent stimuli were color names matching the ink color, neutral stimuli were "XXXXX", and incongruent stimuli were color names that did not match the ink color. (b) Reaction times (RTs) were

normalized within patient via z-score and plotted at the group level. Violins show the full RT distribution for each condition, bars show the interquartile range, and white dots show the median. Note the prolonged RTs for incongruent trials due to conflict. (c) Reconstruction of iEEG recording sites across all participants (n = 17) plotted by region on a standardized group brain after mirroring all electrodes to the left hemisphere. A indicates anterior, and P indicates posterior.

iEEG HFB power separates stimulus, decision, and responses stages of conflict processing

In order to track the dynamics of within-trial conflict processing, we extracted HFB power from 75-150 Hz as a proxy for local population activity at each electrode in LPFC, MPFC, INS, SM, OFC, and TMP regions (Lachaux et al., 2012; Manning et al., 2009). Figure 2a demonstrates the temporal precision and high signal-to-noise ratio of single-trial HFB power for an electrode in anterior mid-cingulate cortex with pre-response activity specifically on incongruent trials, which is reminiscent of classic conflict detection and/or resolution (Kerns et al., 2004; MacDonald et al., 2000). Figure 2b shows single-trial response-locked motor activity in an electrode in the insular region responsible for speech articulation (Dronkers, 1996; Tomaiuolo et al., 2021; Woolnough et al., 2019), and Figures 2c and 2d show condition-averaged HFB power for these electrodes locked to stimulus and response onsets.

We leveraged the high spatiotemporal precision of the HFB signal to separate conflict processing into stimulus, decision, and response stages associated with conflict detection, resolution, and monitoring, respectively. Based on the rationale that conflict detection and resolution must precede successful responses while outcome monitoring is necessarily subsequent to responses, we adopted a dynamic sliding window approach that uses RTs to segment trials into processing stages. We defined stimulus stages as a series of 100 ms windows stepping by 25 ms from stimulus onset to the shortest RT. The decision stage was defined as a single window of variable length from 50 ms before the shortest RT through the RT on each trial. Lastly, the response stage was defined as a series of 100 ms windows stepping by 25 ms from 50 ms pre-response through 1000 ms post-response. HFB power values in each window were tested with ANOVA statistics (see Methods), and conflict effects were identified by significant effects of congruence and plotted as time series of variance explained (ω^2) for each electrode (Fig. 2e and 2f) (Olejnik & Algina, 2003).

Tailoring these analysis windows according to the epoch boundaries defined by RTs is necessary to avoid confounding different phases of conflict processing. This problem is especially important for decision and response stages that are temporally bounded by RTs, which are variable across trials and patients and also differ across conditions. Two examples of artifactual condition differences due to mis-aligned averaging are demonstrated by comparing condition-averaged HFB power traces during the peri-response epoch in Figures 2c and 2d. For the cingulate electrode in Fig. 2a and 2c, the prolonged RTs in the high conflict incongruent condition extend the temporal profile of the condition-specific activity increase, which overlaps with the general post-response decrease in activity on congruent and neutral trials when time-locked to stimulus onset. We account for these dynamics by locking the end of our decision windows and start of our response windows to single-trial RTs, thereby avoiding inaccurate comparisons across processing stages and revealing significant pre-response conflict resolution and post-response conflict monitoring effects (Fig. 2e). Similarly, our windowing strategy avoids

the artifactual condition differences created by averaging response-locked HFB power traces in the insula when mis-aligned to stimulus onset (Fig. 2d). Instead, the decision and response windows are not significant across conditions, indicating the large increase in activity is likely driven by a general motor response rather than conflict (Fig. 2f).

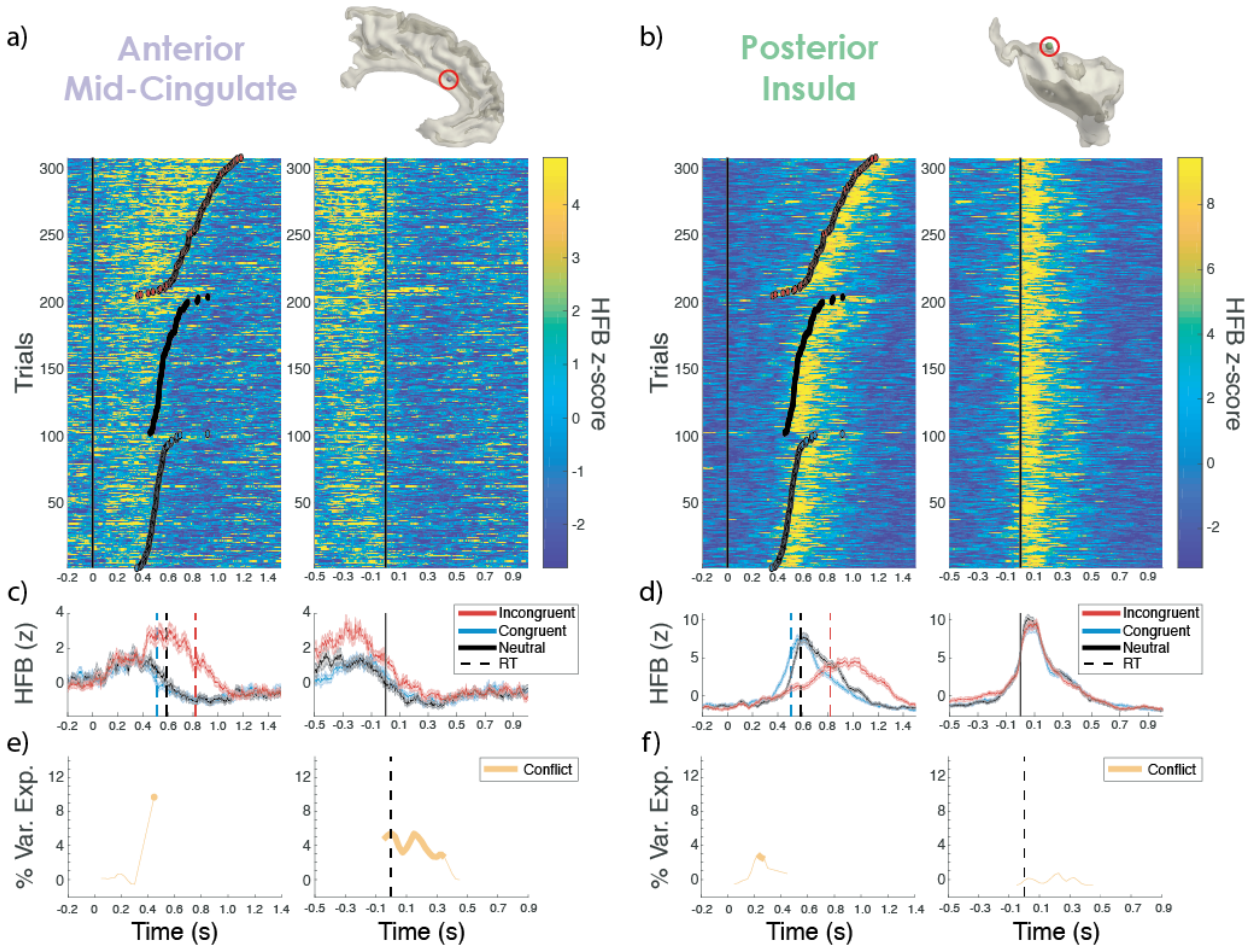


Figure 2: Example high frequency broadband (HFB) power analysis. (a) Normalized HFB from an example electrode in the anterior mid-cingulate cortex (recording site highlighted with red circle) is plotted time-locked to stimulus onset (left column; markers indicate RTs colored by condition; blue for congruent, black for neutral, and red for incongruent) and RT (right column) for each trial. Note the elevated pre-response HFB activity in incongruent trials with conflict. (c) Normalized HFB power averaged for each condition (error bars represent standard error of the mean). Dotted vertical lines in stimulus-locked left plot show mean RT per condition. (e) Percent variance explained by conflict (ω^2 from condition factor in ANOVA analysis) for single-trial HFB power averaged in 100 ms sliding windows (step size of 25 ms). Significant windows ($q_{FDR} < 0.05$) are plotted in bold. Dotted black line indicates RT in right plot. (b, d, f) Same for posterior insula electrode. Note how windowing strategy avoids analyzing epochs with artificial condition differences created by RT differences across conditions (e.g., comparing post-response activity when mis-aligned to stimulus onset).

Different stages of conflict processing recruit distributed, partially overlapping networks

Figure 3a shows the proportion of sites showing significant within-trial conflict effects for each region and processing stage. Conflict effects in each of the three stages recruited a

network of sites distributed across all regions analyzed. The proportion of conflict-sensitive sites was smallest in early stimulus processing windows and increased for decision windows when conflict must be resolved and for response windows requiring implementation of motor control and performance monitoring. Conflict effects were most common in MPFC, LPFC, and SM regions canonically associated with control and response execution, but conflict was also prominent in OFC and TMP regions and present in INS across all processing stages. The conflict processing networks involved in stimulus, decision, and response phases overlapped more with their immediately preceding or following stages (i.e., stimulus and response networks had the least overlap), but more than half of the conflict sensitive electrodes in each network were specific to that processing stage (Fig. 3a inset). Finally, every region analyzed showed involvement in more than one processing stage, providing evidence against specialized processing roles for individual regions (Fig. 3b).

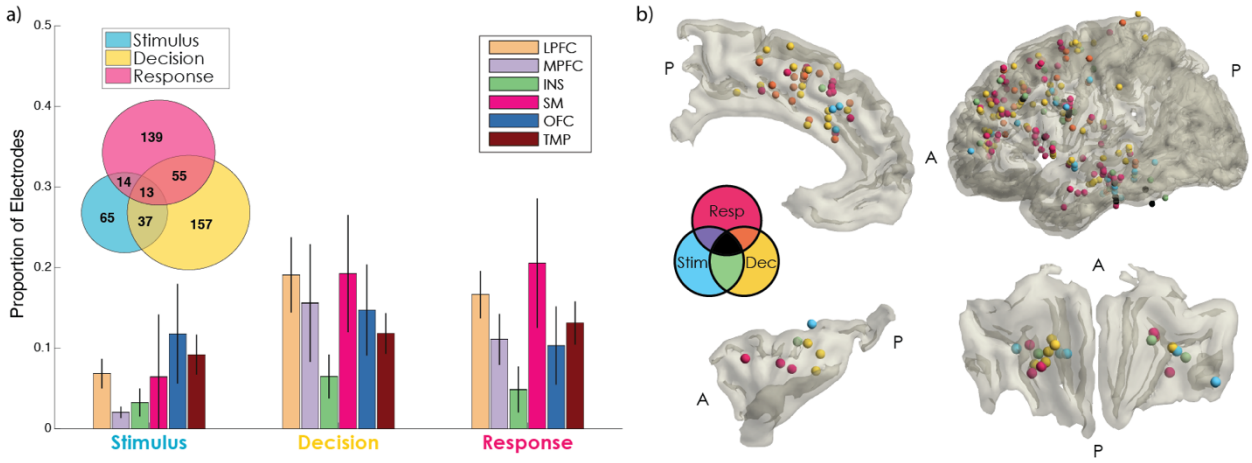


Figure 3: Spatial distribution and overlap of networks processing different stages of conflict. (a) Proportion of sites in each region showing significant conflict effects for stimulus, decision, and response processing stages. Bars show group averaged proportions of recording sites with at least one time bin with significant condition differences in HFB power. Error bars represent standard error of the mean across patients (n = 17). Inset shows Venn diagram overlap of electrodes participating in networks for stimulus, decision, and response stages of conflict processing. (b) Anatomical distribution of networks processing different stages of conflict. Electrodes with significant condition differences in HFB power are plotted for OFC (top left), MPFC (top right), insular cortex (bottom left), and LPFC, sensorimotor, and temporal cortices (bottom right). Venn diagram color legend (middle) indicates participation in one or more of stimulus, decision, and/or response stages of conflict processing. A indicates anterior, and P indicates posterior.

Conflict onset latencies support parallel processing

To test the conflict monitoring hypothesis that conflict effects should occur earlier in MPFC than LPFC (Botvinick et al., 2004; Shenhav et al., 2013; Yeung et al., 2004), as well as whether latency differences across regions could elucidate their specific roles, we measured the onset of conflict effects as the first time point of the first window showing a significant condition difference. Regions showed heterogeneous onset latencies, with no clear differentiation of timing across areas indicative of specialized regional functions (Fig. 4a). This pattern held across stimulus- and response-locked onsets, suggesting it is a general property of conflict processing across stimulus, decision, and response phases.

To compare onsets across regions while accounting for individual differences in RTs, we computed within-subject latency differences between pairs of significant electrodes across regions. Contrary to conflict monitoring theory, conflict onsets were relatively late in MPFC compared to other regions and no different from LPFC (mean ± S.D. = 9 ± 69 ms with LPFC leading). Onset latencies were earliest in TMP cortex, potentially due to its role in visual and linguistic stimulus processing, but in general the variability in latency differences was larger than any mean effects. These findings argue against the serial processing model posited by conflict monitoring theory and instead favor parallel and distributed processing theories of brain function.

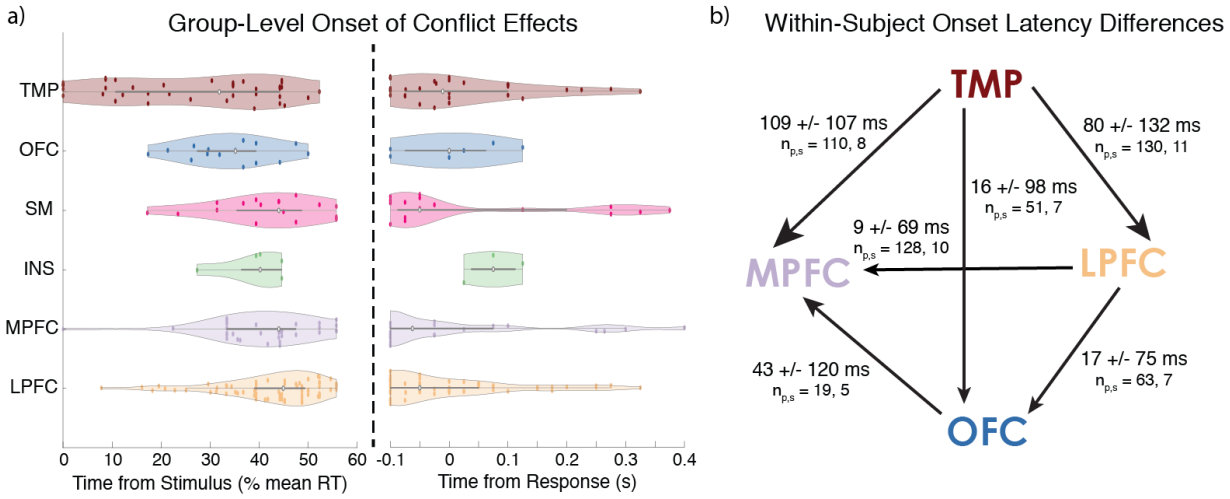


Figure 4: Parallel onsets of conflict effects across regions. (a) Violin plots show the group-level distribution of conflict effect onsets for each electrode divided by region. Left column shows stimulus-locked analyses, with onset times normalized within patient by mean RT across all conditions to account for individual differences in processing speed. Right column shows response-locked analyses. Onsets are defined as the first time point of the first window showing a significant effect. Horizontal gray bars show inter-quartile ranges, and white dots show median onsets. (b) Schematic diagram of within patient onset latencies differences across regions computed between pairs of significant electrodes in each region. Text shows mean ± standard deviation of latency differences between regions, with n_p indicating number of electrode pairs and n_s indicating number of subjects contributing to the analysis. Arrows show potential flow of conflict processing between regions based on onset latency differences, but note the large variability relative to mean effect, suggesting parallel processing across regions rather than serial flow between them.

DISCUSSION:

We leveraged the high spatiotemporal resolution of HFB power in iEEG data to characterize the evolution of conflict effects across different processing stages. We observed widespread conflict signals in classic MPFC, LPFC, and SM regions associated with cognitive and motor control, as well as in non-classical INS, OFC, and TMP regions. After dividing trial epochs into stimulus, decision, and response windows, we found separate but spatially intermingled networks embedded in these regions for each stage of conflict processing, with relatively small numbers of sites overlapping between stages. Lastly, within-trial conflict onsets were variable within all regions, with no clear within-subject latency differences between regions. These

results indicate conflict processing unfolds in parallel across a distributed network of regions extending beyond classic MPFC and LPFC control areas.

Importantly, lesion experiments in non-human primates and humans show mixed evidence on whether MPFC, LPFC, and other regions such as OFC play causal roles in utilizing conflict to adapt behavior (Mansouri et al., 2017), leaving the specific interpretation of conflict effects in each region uncertain. The widespread occurrence of conflict mirrors observations of other cognitive variables such as reward and surprise signals across much of the brain and highlights the distributed, interactive nature of computations in the brain (Fouragnan et al., 2018; Hunt & Hayden, 2017). As in the case of value computations, these results call into question the computations underlying these conflict effects and thus the nature of conflict itself (Cisek & Kalaska, 2010; Eisenreich et al., 2017). As mentioned in the introduction, the operational definition of conflict as differences across conditions with varying levels of response competition include a variety of relevant control processes such as detection, resolution, or monitoring, but this general criterion is also sensitive to non-control functions. For example, the early conflict effects in TMP cortex could be explained by sensory or linguistic processing of word stimuli in congruent and neutral conditions compared to the "XXXXX" stimulus in the neutral condition. Although not a confound per se, these results underscore the importance of careful theoretical models that capture cognitive computations of interest, which should be paired with task designs and modeling approaches that further distinguish between competing hypotheses (Freund, Etzel, et al., 2021; Fu et al., 2021).

In order to delineate some of the factors underlying our conflict effects, we developed a dynamic sliding window approach to separate stimulus, decision, and response phases of the trial. In addition to avoiding analysis confounds of mis-aligned trial averaging across conditions with different RTs, this temporal structure enabled us to assess networks underlying conflict detection, resolution, and monitoring. Each of these phases elicited conflict-sensitive HFB activity in networks of distributed regions, but decision and response conflict networks recruited more than twice as many sites as the stimulus phase. These networks were spatially intermingled in each region but displayed minimal overlap, with around half of stimulus processing sites overlapping with the decision network and approximately one third of sites overlapping between decision and response phases. Similar functional diversity has been observed amongst LPFC single units in seminal non-human primate studies of working memory, which also reported around half of cue-selective neurons participated in a larger delay period ensemble (Funahashi et al., 1990). Also, previous work has shown that detailed anatomical parcellations of sub-regions in MPFC, insula, and other areas can explain differences in activity patterns (Amiez et al., 2013, 2016; Jahn et al., 2014; Lopez-Persem et al., 2019), so future studies may elucidate the role of anatomical and functional connectivity or other principles in determining membership of any given site in the different processing stages.

One clue to the relationship between these networks is the temporal dynamics of conflict processing. We observed heterogeneous conflict onset latencies within each region, suggesting these brain areas operate in parallel. These findings contradict the serial hypothesis of conflict monitoring theory, which was largely based on between-trial adaptation effects in non-invasive studies (Egner, 2007; Kerns, 2006; Kerns et al., 2004; MacDonald et al., 2000). This serial hypothesis has also received empirical support from another report of iEEG conflict onset latencies in the Stroop task (Tang et al., 2016), but our study included a larger sample size,

which was particularly valuable for estimating within-subject pair-wise latency differences between regions. The parallel processing argument supported by our results matches general theories of brain computations in action and value (Cisek & Kalaska, 2010; Eisenreich et al., 2017; Hunt & Hayden, 2017; Pezzulo & Cisek, 2016), but the lack of clear differences in onset latencies does not rule out directional influences between regions. Indeed, recent microstimulation experiments have confirmed asymmetric influences from MPFC to LPFC (Nácher et al., 2019), suggesting alternative metrics may still support some central tenets of conflict monitoring theory (Botvinick et al., 2004; Shenhav et al., 2013). Future studies should further examine interactions between the networks identified here and use directed connectivity analyses and causal stimulation techniques to dissect their relationships. These interactions may also change depending on different mechanisms for control allocation (Braver, 2012; Egner et al., 2007), so future investigations should also assess how between-trial congruence sequence and between-block proportion congruence adaptation effects modulate within-trial conflict processing.

Overall, our results provide novel insights into the spatiotemporal evolution of different phases of conflict processing that emphasize parallel computations across multiple distributed networks. These results call into question serial processing claims implied by classical conflict monitoring theory in favor of more complex models of control circuits (Eisenreich et al., 2017; Heilbronner & Hayden, 2016; Shenhav et al., 2013). In sum, these findings emphasize heterogeneous functions within classical prefrontal control regions, contributions of non-classical regions such as OFC, INS, and TMP, and coordination of distributed networks in multiple stages of conflict processing.

Acknowledgements:

We thank the patients for their time and selfless dedication and Vitoria Piai for contributions to study design.

Chapter 5: Conclusions

"Just do what must be done. This may not be happiness, but it is greatness." -George Bernard Shaw

Cognitive control is the fundamental component of mental life that lets us bridge knowledge and action to get things done and achieve our goals. It frees us from instincts and reflexes and allows us to become better versions of ourselves. Cognitive control depends on the precise coordination of information across many neural circuits distributed throughout the brain at the speed of thought. The projects in this thesis combined careful analyses of behavior with non-invasive scalp EEG and invasive iEEG recordings to characterize these complex dynamics and hopefully shed light on how our brains enable successful navigation through an ever-changing environment.

In **Chapter 2**, we tackled the intertwined methodological and theoretical debates around reward and surprise components of feedback ERPs in scalp EEG. Our experimental design and behavioral modeling separated valenced RPE value from two non-valenced sources of salience: RPE magnitude and probability. Single-trial regression across time, space, and frequency domains disentangled different EEG signatures for each of these PEs, but our original analyses reproduced the persistent mistake in the field of combining wins and losses, which fuels interpretation of the FRN as a valenced RPE value signal. Thanks to helpful input from reviewer James Cavanagh, we conducted follow up analyses on wins and losses separately, which revealed that the FRN actually tracks non-valenced RPE magnitude for negative outcomes while the RewP represents non-valenced RPE magnitude for positive outcomes. These results also confirmed our novel finding of a late frontal P3 component tracking probability holds regardless of outcome valence. These results demonstrate the power of modeling individual participant behavior and analyzing multiple dimensions of the EEG signal together, which is often necessary to unravel the overlapping components at the scalp level.

In **Chapter 3**, we used iEEG recordings to investigate the neural sources of these scalp signals. We found large swaths of MPFC and LPFC encoded RPE value and magnitude both separately and with a substantial number of sites showing mixed selectivity for both signals. The close proximity of these reward and surprise signals mere millimeters away from each other clearly demonstrates the necessity of invasive recordings to elucidate the neural mechanisms of cognitive control, which could not be disentangled in FRN and RewP scalp ERPs. Interestingly, these iEEG recordings revealed insular cortex had the largest proportion of RPE coding sites and also had the largest proportion of mixed selectivity sites. In addition to this different coding scheme, RPE value signals onset in insula before MPFC, suggesting a potential leading role for the insula in reward processing. Notably, the location of insular cortex buried underneath the frontal, parietal, and temporal lobes makes it a difficult region to identify using source localization algorithms with non-invasive EEG data, again emphasizing the utility of iEEG to reveal the nature of neural computations.

In **Chapter 4**, we examine the temporal evolution of conflict effects using iEEG recordings in a color-word Stroop task. Based on between-trial adaptation effects in fMRI studies and recent iEEG studies (Tang et al., 2016), the highly influential (though dated) conflict

monitoring theory proposed a serial flow of information from MPFC to LPFC in conflict processing (Botvinick et al., 2001; Kerns et al., 2004). We leveraged the high spatiotemporal resolution of the iEEG HFB signal to separate conflict effects into stimulus, decision, and response phases of the trial associated with conflict detection, resolution, and monitoring. Our results showed each of these stages of conflict processing recruits a distributed network of regions extending well beyond classic MPFC and LPFC control regions and into sensorimotor, insular, orbitofrontal, and temporal cortices. Furthermore, the majority of the conflict-sensitive sites participating in these stimulus, decision, and response networks were distinct, suggesting sub-circuits in these regions perform different functions throughout the trial. Lastly, conflict onsets were variable within each region and showed no clear latency differences between regions, implying conflict processing occurs in parallel across these regions.

Grounding interpretation of neural data in cognitive and computational frameworks is particularly important when analyzing high dimensional iEEG data. One major strength of the work in **Chapters 2 and 3** is using behavioral modeling to estimate latent cognitive variables in a reinforcement learning framework, which dissociated the latent RPE value and magnitude variables underlying the cognitive computations of the behavioral problem and integrated reward learning and salience processing perspectives on MPFC and insula activity. Similarly, separating conflict effects into different stages of the trial in **Chapter 4** constrained their interpretation relative to conflict detection, resolution, and monitoring functions. However, the ANOVA model used in the Stroop analyses does not account for important condition differences such as the non-word stimuli in the neutral condition. Future studies will be needed to clarify these distinctions by modeling these behavioral task features or investigating how the conflict dynamics are modulated by proactive and reactive adaptation to congruence sequence or proportion congruence effects (Braem et al., 2019; Braver, 2012; Bugg & Crump, 2012; Egner, 2007).

Another consistent theme across the iEEG findings in **Chapters 3 and 4** is the distributed representations and parallel processing of cognitive variables across multiple regions. RPE value, RPE magnitude, and different conflict representations were spatially intermingled in neighboring electrodes, and individual electrodes displayed mixed selectivity to RPE value and magnitude or to different stages of conflict processing. Onset latencies highlighted the early emergence of RPE value information in the insula, but otherwise the dominant feature of control signal timing was heterogeneity within and across regions, which is indicative of continuous, simultaneous engagement of these distributed networks. These findings support modern theories of cognitive function that deemphasize functional specialization by region in favor of dynamical systems approaches that highlight distributed, recurrent neural computations orchestrated by large populations of neurons embedded in cortical and subcortical hierarchies (Cisek & Kalaska, 2010; Hunt & Hayden, 2017; Pessoa, 2008; Pezzulo & Cisek, 2016).

Given the complexity of these circuit dynamics, mechanistic accounts of the neural computations underlying cognitive control will require more advanced analyses than the onset latency metrics used in this thesis. Future studies should incorporate functional and effective connectivity analyses to determine the direction of information flow between regions. In particular, measures of low frequency power, phase, and phase amplitude coupling may provide insight into information gating and integration (Womelsdorf et al., 2014). For example,

such analyses could bridge concurrent scalp and intracranial EEG recordings to link the scalp ERP signatures from **Chapter 2** to HFB activity in **Chapter 3** and provide a mechanistic account of RPE value and magnitude coding in the brain. However, time series analyses are still limited to correlational inferences. Ultimately, causal interventions such as brain stimulation will be required to elucidate the neural mechanisms of cognition, but overall, I am optimistic that clever applications of theoretical models and careful measurements and manipulations of behavior and neural activity will illuminate the elusive relationship between brain and mind.

References

Aarts, E., Verhage, M., Veenvliet, J. V., Dolan, C. V. & Sluis, S. van der. (2014). A solution to dependency: using multilevel analysis to accommodate nested data. *Nature Neuroscience*, *17*(4), 491–496. https://doi.org/10.1038/nn.3648

Adams, R. A., Shipp, S. & Friston, K. J. (2013). Predictions not commands: active inference in the motor system. *Brain Structure and Function*, *218*(3), 611–643. https://doi.org/10.1007/s00429-012-0475-5

Akaishi, R., Kolling, N., Brown, J. W. & Rushworth, M. (2016). Neural Mechanisms of Credit Assignment in a Multicue Environment. *Journal of Neuroscience*, *36*(4), 1096–1112. https://doi.org/10.1523/jneurosci.3159-15.2016

Alexander, W. H. & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*, *14*(10), 1338–1344. https://doi.org/10.1038/nn.2921

Alexander, W. H. & Brown, J. W. (2014). A general role for medial prefrontal cortex in event prediction. *Frontiers in Computational Neuroscience*, *8*, 69. https://doi.org/10.3389/fncom.2014.00069

Alexander, W. H. & Brown, J. W. (2015). Hierarchical Error Representation: A Computational Model of Anterior Cingulate and Dorsolateral Prefrontal Cortex. *Neural Computation*, *27*(11), 2354–2410. https://doi.org/10.1162/neco_a_00779

Alexander, W. H. & Brown, J. W. (2017). The Role of the Anterior Cingulate Cortex in Prediction Error and Signaling Surprise. *Topics in Cognitive Science*, *2*(10), 658–17. https://doi.org/10.1111/tops.12307

Alexander, W. H. & Brown, J. W. (2018). Frontal cortex function as derived from hierarchical predictive coding. *Scientific Reports*, *8*(1), 3843. https://doi.org/10.1038/s41598-018-21407-9

Alexander, W. H., Brown, J. W., Collins, A. G. E., Hayden, B. Y. & Vassena, E. (2017). Prefrontal Cortex in Control: Broadening the Scope to Identify Mechanisms. *Journal of Cognitive Neuroscience*, *30*(8), 1061–1065. https://doi.org/10.1162/jocn_a_01154

Amiez, C., Neveu, R., Warrot, D., Petrides, M., Knoblauch, K. & Procyk, E. (2013). The location of feedback-related activity in the midcingulate cortex is predicted by local morphology. *Journal of Neuroscience*, *33*(5), 2217–2228. https://doi.org/10.1523/jneurosci.2779-12.2013

Amiez, C. & Petrides, M. (2014). Neuroimaging Evidence of the Anatomo-Functional Organization of the Human Cingulate Motor Areas. *Cerebral Cortex*, *24*(3), 563–578. https://doi.org/10.1093/cercor/bhs329

Amiez, C., Wutte, M. G., Faillenot, I., Petrides, M., Burle, B. & Procyk, E. (2016). Single subject analyses reveal consistent recruitment of frontal operculum in performance monitoring. *NeuroImage*, *133*, 266–278. https://doi.org/10.1016/j.neuroimage.2016.03.003

Asaad, W. F. & Eskandar, E. N. (2011). Encoding of Both Positive and Negative Reward Prediction Errors by Neurons of the Primate Lateral Prefrontal Cortex and Caudate Nucleus. *Journal of Neuroscience*, *31*(49), 17772–17787. https://doi.org/10.1523/jneurosci.3793-11.2011

Asaad, W. F., Lauro, P. M., Perge, J. A. & Eskandar, E. N. (2017). Prefrontal Neurons Encode a Solution to the Credit-Assignment Problem. *Journal of Neuroscience*, *37*(29), 6995–7007. https://doi.org/10.1523/jneurosci.3311-16.2017

Badre, D. & D'Esposito, M. (2007). Functional Magnetic Resonance Imaging Evidence for a Hierarchical Organization of the Prefrontal Cortex. *Journal of Cognitive Neuroscience*, *19*(12), 2082–2099. https://doi.org/10.1162/jocn.2007.19.12.2082

Badre, D. & Nee, D. E. (2017). Frontal Cortex and the Hierarchical Control of Behavior. *Trends in Cognitive Sciences*, *22*(2), 1–19. https://doi.org/10.1016/j.tics.2017.11.005

Bai, Y., Katahira, K. & Ohira, H. (2015). Valence-separated representation of reward prediction error in feedback-related negativity and positivity. *NeuroReport*, *26*(3), 157–162. https://doi.org/10.1097/wnr.0000000000000318

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

Baker, T. E., Stockwell, T., Barnes, G., Haesevoets, R. & Holroyd, C. B. (2016). Reward Sensitivity of ACC as an Intermediate Phenotype between DRD4-521T and Substance Misuse. *Journal of Cognitive Neuroscience*, *28*(3), 460–471. https://doi.org/10.1162/jocn_a_00905

Barrett, L. F. & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience*, *16*(7), 419–429. https://doi.org/10.1038/nrn3950

Bartoli, E., Conner, C. R., Kadipasaoglu, C. M., Yellapantula, S., Rollo, M. J., Carter, C. S. & Tandon, N. (2017). Temporal Dynamics of Human Frontal and Cingulate Neural Activity During Conflict and Cognitive Control. *Cerebral Cortex (New York, N.Y.: 1991)*, *28*(11), 1–15. https://doi.org/10.1093/cercor/bhx245

Bartolo, R. & Averbeck, B. B. (2020). Prefrontal Cortex Predicts State Switches during Reversal Learning. *Neuron*, *106*(6), 1044-1054.e4. https://doi.org/10.1016/j.neuron.2020.03.024

Bastin, J., Deman, P., David, O., Gueguen, M., Benis, D., Minotti, L., Hoffman, D., Combrisson, E., Kujala, J., Perrone-Bertolotti, M., Kahane, P., Lachaux, J.-P. & Jerbi, K. (2017). Direct Recordings from Human Anterior Insula Reveal its Leading Role within the Error-Monitoring Network. *Cerebral Cortex*, *27*(2), 1545–1557. https://doi.org/10.1093/cercor/bhv352

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P. & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, *76*(4), 695–711. https://doi.org/10.1016/j.neuron.2012.10.038

Bayer, H. M. & Glimcher, P. W. (2005). Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal. *Neuron*, *47*(1), 129–141. https://doi.org/10.1016/j.neuron.2005.05.020

Becker, M. P. I., Nitsch, A. M., Miltner, W. H. R. & Straube, T. (2014). A Single-Trial Estimation of the Feedback-Related Negativity and Its Relation to BOLD Responses in a Time-Estimation Task. *Journal of Neuroscience*, *34*(8), 3005–3012. https://doi.org/10.1523/jneurosci.3684-13.2014

Benjamini, Y. & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bernat, E. M., Malone, S. M., Williams, W. J., Patrick, C. J. & Iacono, W. G. (2007). Decomposing delta, theta, and alpha time-frequency ERP activity from a visual oddball task using PCA.

International Journal of Psychophysiology, *64*(1), 62–74. https://doi.org/10.1016/j.ijpsycho.2006.07.015

Bernat, E. M., Nelson, L. D. & Baskin-Sommers, A. R. (2015). Time-frequency theta and delta measures index separable components of feedback processing in a gambling task. *Psychophysiology*, *52*(5), 626–637. https://doi.org/10.1111/psyp.12390

Bernat, E. M., Nelson, L. D., Steele, V. R., Gehring, W. J. & Patrick, C. J. (2011). Externalizing Psychopathology and Gain–Loss Feedback in a Simulated Gambling Task: Dissociable Components of Brain Response Revealed by Time-Frequency Analysis. *Journal of Abnormal Psychology*, *120*(2), 352–364. https://doi.org/10.1037/a0022124

Billeke, P., Ossandon, T., Perrone-Bertolotti, M., Kahane, P., Bastin, J., Jerbi, K., Lachaux, J.-P. & Fuentealba, P. (2020). Human Anterior Insula Encodes Performance Feedback and Relays Prediction Error to the Medial Prefrontal Cortex. *Cerebral Cortex*, *30*(7), 4011–4025. https://doi.org/10.1093/cercor/bhaa017

Bjellvi, J., Olsson, I., Malmgren, K. & Ramsay, K. W. (2019). Epilepsy duration and seizure outcome in epilepsy surgery: A systematic review and meta-analysis. *Neurology*, *93*(2), e159–e166. https://doi.org/10.1212/wnl.0000000000007753

Boroujeni, K. B., Tiesinga, P. & Womelsdorf, T. (2021). Interneuron-specific gamma synchronization indexes cue uncertainty and prediction errors in lateral prefrontal and anterior cingulate cortex. *eLife*, *10*, e69111. https://doi.org/10.7554/elife.69111

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S. & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652. https://doi.org/10.1037//0033-295x.108.3.624

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S. & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652.

Botvinick, M. M., Cohen, J. D. & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences*, *8*(12), 539–546. https://doi.org/10.1016/j.tics.2004.10.003

Braem, S., Bugg, J. M., Schmidt, J. R., Crump, M. J. C., Weissman, D. H., Notebaert, W. & Egner, T. (2019). Measuring Adaptive Control in Conflict Tasks. *Trends in Cognitive Sciences*, *23*(9), 769–783. https://doi.org/10.1016/j.tics.2019.07.002

Braem, S., Abrahamse, E. L., Duthoo, W. & Notebaert, W. (2014). What determines the specificity of conflict adaptation? A review, critical analysis, and proposed synthesis. *Auditory Cognitive Neuroscience*, *5*(1001), 1134. https://doi.org/10.3389/fpsyg.2014.01134

Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends in Cognitive Sciences*, *16*(2), 106–113. https://doi.org/10.1016/j.tics.2011.12.010

Bress, J. N., Foti, D., Kotov, R., Klein, D. N. & Hajcak, G. (2013). Blunted neural response to rewards prospectively predicts depression in adolescent girls. *Psychophysiology*, *50*(1), 74–81. https://doi.org/10.1111/j.1469-8986.2012.01485.x

Bridwell, D. A., Cavanagh, J. F., Collins, A. G. E., Nunez, M. D., Srinivasan, R., Stober, S. & Calhoun, V. D. (2018). Moving Beyond ERP Components: A Selective Review of Approaches to Integrate EEG and Behavior. *Frontiers in Human Neuroscience*, *12*, 1494–17. https://doi.org/10.3389/fnhum.2018.00106

Bromberg-Martin, E. S., Matsumoto, M. & Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron*, *68*(5), 815–834. https://doi.org/10.1016/j.neuron.2010.11.022

Brown, J. W. (2014). The tale of the neuroscientists and the computer: why mechanistic theory matters. *Frontiers in Neuroscience*, 1–3. https://doi.org/10.3389/fnins.2014.00349/full

Brown, J. W. & Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, *307*(5712), 1118–1121. https://doi.org/10.1126/science.1105783

Brush, C. J., Ehmann, P. J., Hajcak, G., Selby, E. A. & Alderman, B. L. (2018). Using Multilevel Modeling to Examine Blunted Neural Responses to Reward in Major Depression. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(12), 1032–1039. https://doi.org/10.1016/j.bpsc.2018.04.003

Bugg, J. M. & Crump, M. J. C. (2012). In support of a distinction between voluntary and stimulus-driven control: A review of the literature on proportion congruent effects. *Auditory Cognitive Neuroscience*, *3*(SEP), 367. https://doi.org/10.3389/fpsyg.2012.00367

Buzsáki, G. & Mizuseki, K. (2014). The log-dynamic brain: how skewed distributions affect network operations. *Nature Reviews Neuroscience*, *15*(4), 264–278. https://doi.org/10.1038/nrn3687

Cai, W., Ryali, S., Pasumarthy, R., Talasila, V. & Menon, V. (2021). Dynamic causal brain circuits during working memory and their functional controllability. *Nature Communications*, *12*(1), 3314. https://doi.org/10.1038/s41467-021-23509-x

Carlson, J. M., Foti, D., Mujica-Parodi, L. R., Harmon-Jones, E. & Hajcak, G. (2011). Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: A combined ERP and fMRI study. *NeuroImage*, *57*(4), 1608–1616. https://doi.org/10.1016/j.neuroimage.2011.05.037

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., Bismark, A. W., Frank, M. J. & Allen, J. J. B. (2019). Multiple Dissociations Between Comorbid Depression and Anxiety on Reward and Punishment Processing: Evidence From Computationally Informed EEG. *Computational Psychiatry*, *3*, 1–17. https://doi.org/10.1162/cpsy_a_00024

Cavanagh, J. F., Figueroa, C. M., Cohen, M. X. & Frank, M. J. (2012). Frontal Theta Reflects Uncertainty and Unexpectedness during Exploration and Exploitation. *Cerebral Cortex*, *22*(11), 2575–2586. https://doi.org/10.1093/cercor/bhr332

Cavanagh, J. F. & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414–421. https://doi.org/10.1016/j.tics.2014.04.012

Cavanagh, J. F., Frank, M. J., Klein, T. J. & Allen, J. J. B. (2010). Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *NeuroImage*, *49*(4), 3198–3209. https://doi.org/10.1016/j.neuroimage.2009.11.080

Cavanagh, J. F., Zambrano-Vazquez, L. & Allen, J. J. B. (2011). Theta lingua franca: A common mid-frontal substrate for action monitoring processes. *Psychophysiology*, *49*(2), 220–238. https://doi.org/10.1111/j.1469-8986.2011.01293.x

Cisek, P. & Kalaska, J. F. (2005). Neural Correlates of Reaching Decisions in Dorsal Premotor Cortex: Specification of Multiple Direction Choices and Final Selection of Action. *Neuron*, *45*(5), 801–814. https://doi.org/10.1016/j.neuron.2005.01.027

Cisek, P. & Kalaska, J. F. (2010). Neural Mechanisms for Interacting with a World Full of Action Choices. *Annual Review of Neuroscience*, *33*(1), 269–298. https://doi.org/10.1146/annurev.neuro.051508.135409

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/s0140525x12000477

Cohen, M. X., Ridderinkhof, K. R., Haupt, S., Elger, C. E. & Fell, J. (2008). Medial frontal cortex and response conflict: evidence from human intracranial EEG and medial frontal cortex lesion. *Brain Research*, *1238*, 127–142. https://doi.org/10.1016/j.brainres.2008.07.114

Cole, M. W., Yeung, N., Freiwald, W. A. & Botvinick, M. (2009). Cingulate cortex: diverging data from humans and monkeys. *Trends in Neurosciences*, *32*(11), 566–574. https://doi.org/10.1016/j.tins.2009.07.001

Collins, A. G. E. & Frank, M. J. (2015). Surprise! Dopamine signals mix action, value and error. *Nature Neuroscience*, *19*(1), 3–5. https://doi.org/10.1038/nn.4207

Collins, A. G. E. & Frank, M. J. (2018). Within- and across-trial dynamics of human EEG reveal cooperative interplay between reinforcement learning and working memory. *Proceedings of the National Academy of Sciences*, *115*(10), 201720963. https://doi.org/10.1073/pnas.1720963115

Collins, A. G. E. (2019). Reinforcement learning: bringing together computation and cognition. *Current Opinion in Behavioral Sciences*, *29*, 63–68. https://doi.org/10.1016/j.cobeha.2019.04.011

Collins, A. G. E. & Frank, M. J. (2016). Neural signature of hierarchically structured expectations predicts clustering and transfer of rule sets in reinforcement learning. *Cognition*, *152*, 160–169. https://doi.org/10.1016/j.cognition.2016.04.002

Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nature Reviews Neuroscience*, *3*(8), 655–666. https://doi.org/10.1038/nrn894

Craig, A. D. B. (2009). How do you feel--now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, *10*(1), 59–70. https://doi.org/10.1038/nrn2555

Dale, A. M., Fischl, B. & Sereno, M. I. (1999). Cortical Surface-Based Analysis I. Segmentation and Surface Reconstruction. *NeuroImage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395

Dewar, S. R. & Pieters, H. C. (2015). Perceptions of epilepsy surgery: A systematic review and an explanatory model of decision-making. *Epilepsy & Behavior*, *44*, 171–178. https://doi.org/10.1016/j.yebeh.2014.12.027

Donchin, E., Ritter, W. & McCallum, W. C. (1978). Cognitive psychophysiology: The endogenous components of the ERP. In C. P, T. P & K. S (Eds.), *Event-Related Brain Potentials in Man* (pp. 349–411). Academic Press.

Donkers, F. C. L., Nieuwenhuis, S. & Boxtel, G. J. M. van. (2005). Mediofrontal negativities in the absence of responding. *Cognitive Brain Research*, *25*(3), 777–787. https://doi.org/10.1016/j.cogbrainres.2005.09.007

Dronkers, N. F. (1996). A new brain region for coordinating speech articulation. *Nature*, *384*(6605), 159–161. https://doi.org/10.1038/384159a0

Dubey, A. & Ray, S. (2019). Cortical Electrocorticogram (ECoG) is a local signal. *Journal of Neuroscience*, *39*(22), 4299–4311. https://doi.org/10.1523/jneurosci.2917-18.2019

Dürschmid, S., Reichert, C., Hinrichs, H., Heinze, H.-J., Kirsch, H. E., Knight, R. T. & Deouell, L. Y. (2018). Direct Evidence for Prediction Signals in Frontal Cortex Independent of Prediction Error. *Cerebral Cortex*, *29*(11), 4530–4538. https://doi.org/10.1093/cercor/bhy331

Duthoo, W., Abrahamse, E. L., Braem, S., Boehler, C. N. & Notebaert, W. (2014). The heterogeneous world of congruency sequence effects: an update. *Frontiers in Psychology*, *5*, 1001. https://doi.org/10.3389/fpsyg.2014.01001

Egner, T. (2007). Congruency sequence effects and cognitive control. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(4), 380–390. https://doi.org/10.3758/cabn.7.4.380

Egner, T., Delano, M. & Hirsch, J. (2007). Separate conflict-specific cognitive control mechanisms in the human brain. *NeuroImage*, *35*(2), 940–948. https://doi.org/10.1016/j.neuroimage.2006.11.061

Eisenreich, B. R., Akaishi, R. & Hayden, B. Y. (2017). Control without Controllers: Toward a Distributed Neuroscience of Executive Control. *Journal of Cognitive Neuroscience*, *29*(10), 1684–1698. https://doi.org/10.1162/jocn_a_01139

Engelhard, B., Finkelstein, J., Cox, J., Fleming, W., Jang, H. J., Ornelas, S., Koay, S. A., Thiberge, S. Y., Daw, N. D., Tank, D. W. & Witten, I. B. (2019). Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature*, 1–27. https://doi.org/10.1038/s41586-019-1261-9

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

Euser, A. S., Arends, L. R., Evans, B. E., Greaves-Lord, K., Huizink, A. C. & Franken, I. H. A. (2012). The P300 event-related brain potential as a neurobiological endophenotype for substance use disorders: A meta-analytic investigation. *Neuroscience and Biobehavioral Reviews*, *36*(1), 572–603. https://doi.org/10.1016/j.neubiorev.2011.09.002

Evrard, H. C. (2019). The Organization of the Primate Insular Cortex. *Frontiers in Neuroanatomy*, *13*, 43. https://doi.org/10.3389/fnana.2019.00043

Falkenstein, M., Hohnsbein, J., Hoormann, J. & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology*, *78*(6), 447–455. https://doi.org/10.1016/0013-4694(91)90062-9

Ferdinand, N. K., Mecklinger, A., Kray, J. & Gehring, W. J. (2012). The Processing of Unexpected Positive Response Outcomes in the Mediofrontal Cortex. *Journal of Neuroscience*, *32*(35), 12087–12092. https://doi.org/10.1523/jneurosci.1410-12.2012

Fiorillo, C. D. (2013). Two dimensions of value: dopamine neurons represent reward but not aversiveness. *Science (New York, N.Y.)*, *341*(6145), 546–549. https://doi.org/10.1126/science.1238699

Fischer, A. G. & Ullsperger, M. (2013). Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron*, *79*(6), 1243–1255. https://doi.org/10.1016/j.neuron.2013.07.006

Fischer, R., Lee, A. & Verzijden, M. N. (2018). Dopamine genes are linked to Extraversion and Neuroticism personality traits, but only in demanding climates. *Scientific Reports*, *8*(1), 1733–10. https://doi.org/10.1038/s41598-017-18784-y

Folstein, J. R. & Petten, C. V. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, *45*(1), 152–170. https://doi.org/10.1111/j.1469-8986.2007.00602.x

Foti, D. & Hajcak, G. (2009). Depression and reduced sensitivity to non-rewards versus rewards: Evidence from event-related potentials. *Biological Psychology*, *81*(1), 1–8. https://doi.org/10.1016/j.biopsycho.2008.12.004

Foti, D., Weinberg, A., Bernat, E. M. & Proudfit, G. H. (2015). Anterior cingulate activity to monetary loss and basal ganglia activity to monetary gain uniquely contribute to the feedback negativity. *Clinical Neurophysiology*, *126*(7), 1338–1347. https://doi.org/10.1016/j.clinph.2014.08.025

Foti, D., Weinberg, A., Dien, J. & Hajcak, G. (2011). Event-related potential activity in the basal ganglia differentiates rewards from nonrewards: Temporospatial Principal Components Analysis and Source Localization of the Feedback Negativity. *Human Brain Mapping*, *32*(12), 2267–2269. https://doi.org/10.1002/hbm.21357

Fouragnan, E., Retzler, C. & Philiastides, M. G. (2018). Separate neural representations of prediction error valence and surprise: Evidence from an fMRI meta-analysis. *Human Brain Mapping*, *39*(7), 2887–2906. https://doi.org/10.1002/hbm.24047

Frank, M. J. & Badre, D. (2015). How cognitive theory guides neuroscience. *Cognition*, *135*, 14–20. https://doi.org/10.1016/j.cognition.2014.11.009

Freund, M. C., Bugg, J. M. & Braver, T. S. (2021). A representational similarity analysis of cognitive control during color-word Stroop. *Journal of Neuroscience*, JN-RM-2956-20. https://doi.org/10.1523/jneurosci.2956-20.2021

Freund, M. C., Etzel, J. A. & Braver, T. S. (2021). Neural Coding of Cognitive Control: The Representational Similarity Analysis Approach. *Trends in Cognitive Sciences*, *25*(7), 622–638. https://doi.org/10.1016/j.tics.2021.03.011

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Frömer, R., Maier, M. & Rahman, R. A. (2018). Group-Level EEG-Processing Pipeline for Flexible Single Trial-Based Analyses Including Linear Mixed Models. *Frontiers in Neuroscience*, *12*, 970–15. https://doi.org/10.3389/fnins.2018.00048

Fu, Z., Beam, D., Chung, J. M., Reed, C. M., Mamelak, A. N., Adolphs, R. & Rutishauser, U. (2021). The geometry of domain-general performance monitoring representations in the human medial frontal cortex. *BioRxiv*, 2021.07.08.451594. https://doi.org/10.1101/2021.07.08.451594

Fu, Z., Wu, D.-A. J., Ross, I., Chung, J. M., Mamelak, A. N., Adolphs, R. & Rutishauser, U. (2018). Single-Neuron Correlates of Error Monitoring and Post-Error Adjustments in Human Medial Frontal Cortex. *Neuron*. https://doi.org/10.1016/j.neuron.2018.11.016

Funahashi, S., Bruce, C. J. & Goldman-Rakic, P. S. (1990). Visuospatial coding in primate prefrontal neurons revealed by oculomotor paradigms. *Journal of Neurophysiology*, *63*(4), 814–831. https://doi.org/10.1152/jn.1990.63.4.814

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E. & Donchin, E. (1993). A Neural System for Error Detection and Compensation. *Psychological Science*, *4*(6), 385–390. https://doi.org/10.1111/j.1467-9280.1993.tb00586.x

Gehring, W. J., Liu, Y., Orr, J. M. & Carp, J. (2012). The Error-Related Negativity (ERN/Ne). In E. S. Kappenman & S. J. Luck (Eds.), *The Oxford Handbook of Event-Related Potential Components* (pp. 1–111). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0120

Gehring, W. J. & Willoughby, A. R. (2002). The Medial Frontal Cortex and the Rapid Processing of Monetary Gains and Losses. *Science*, *295*(5563), 2279–2282. https://doi.org/10.1126/science.1066893

Gehrlach, D. A., Dolensek, N., Klein, A. S., Chowdhury, R. R., Matthys, A., Junghänel, M., Gaitanos, T. N., Podgornik, A., Black, T. D., Vaka, N. R., Conzelmann, K.-K. & Gogolla, N. (2019). Aversive state processing in the posterior insular cortex. *Nature Neuroscience*, *22*(9), 1424–1437. https://doi.org/10.1038/s41593-019-0469-1

Gershman, S. J. & Uchida, N. (2019). Believing in dopamine. *Nature Reviews Neuroscience*, *20*(11), 703–714. https://doi.org/10.1038/s41583-019-0220-7

Gjedde, A., Kumakura, Y., Cumming, P., Linnet, J. & Møller, A. (2010). Inverted-U-shaped correlation between dopamine receptor availability in striatum and sensation seeking. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(8), 3870–3875. https://doi.org/10.1073/pnas.0912319107

Glazer, J. E., Kelley, N. J., Pornpattananangkul, N., Mittal, V. A. & Nusslock, R. (2018). Beyond the FRN: Broadening the time-course of EEG and ERP components implicated in reward processing. *International Journal of Psychophysiology*, *132*(Pt B), 184–202. https://doi.org/10.1016/j.ijpsycho.2018.02.002

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences*, *108*(Supplement 3), 15647–15654. https://doi.org/10.1073/pnas.1014269108

Gogolla, N. (2017). The insular cortex. *Current Biology*, *27*(12), R580–R586. https://doi.org/10.1016/j.cub.2017.05.010

Gratton, C., Sun, H. & Petersen, S. E. (2018). Control networks and hubs. *Psychophysiology*, *55*(3), e13032. https://doi.org/10.1111/psyp.13032

Grinband, J., Savitskaya, J., Wager, T. D., Teichert, T., Ferrera, V. P. & Hirsch, J. (2011). The dorsal medial frontal cortex is sensitive to time on task, not response conflict or error likelihood. *NeuroImage*, *57*(2), 303–311. https://doi.org/10.1016/j.neuroimage.2010.12.027

Grohn, J., Schüffelgen, U., Neubert, F.-X., Bongioanni, A., Verhagen, L., Sallet, J., Kolling, N. & Rushworth, M. F. S. (2020). Multiple systems in macaques for tracking prediction errors and other types of surprise. *PLOS Biology*, *18*(10), e3000899. https://doi.org/10.1371/journal.pbio.3000899

Gruber, W. R., Klimesch, W., Sauseng, P. & Doppelmayr, M. (2005). Alpha Phase Synchronization Predicts P1 and N1 Latency and Amplitude Size. *Cerebral Cortex*, *15*(4), 371–377. https://doi.org/10.1093/cercor/bhh139

Haber, S. N., Fudge, J. L. & McFarland, N. R. (2000). Striatonigrostriatal Pathways in Primates Form an Ascending Spiral from the Shell to the Dorsolateral Striatum. *Journal of Neuroscience*, *20*(6), 2369–2382. https://doi.org/10.1523/jneurosci.20-06-02369.2000

Hajcak, G. & Foti, D. (2020). Significance?... Significance! Empirical, methodological, and theoretical connections between the late positive potential and P300 as neural responses to stimulus significance: An integrative review. *Psychophysiology*, *57*(7), e13570. https://doi.org/10.1111/psyp.13570

Hajcak, G., Moser, J. S., Holroyd, C. B. & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology*, *71*(2), 148–154. https://doi.org/10.1016/j.biopsycho.2005.04.001

Hajcak, G., Moser, J. S., Holroyd, C. B. & Simons, R. F. (2007). It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*, *44*(6), 905–912. https://doi.org/10.1111/j.1469-8986.2007.00567.x

Hajihosseini, A. & Holroyd, C. B. (2013). Frontal midline theta and N200 amplitude reflect complementary information about expectancy and outcome evaluation. *Psychophysiology*, *50*(6), 550–562. https://doi.org/10.1111/psyp.12040

Haupt, S., Axmacher, N., Cohen, M. X., Elger, C. E. & Fell, J. (2009). Activation of the caudal anterior cingulate cortex due to task-related interference in an auditory Stroop paradigm. *Human Brain Mapping*, *30*(9), 3043–3056. https://doi.org/10.1002/hbm.20731

Hauser, T. U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S. & Brem, S. (2014). The feedback-related negativity (FRN) revisited: new insights into the localization, meaning and network organization. *NeuroImage*, *84*, 159–168. https://doi.org/10.1016/j.neuroimage.2013.08.028

Hayden, B. Y. & Platt, M. L. (2010). Neurons in anterior cingulate cortex multiplex information about reward and action. *Journal of Neuroscience*, *30*(9), 3339–3346. https://doi.org/10.1523/jneurosci.4874-09.2010

Haynes, W. I. A. & Haber, S. N. (2013). The Organization of Prefrontal-Subthalamic Inputs in Primates Provides an Anatomical Substrate for Both Functional Specificity and Integration: Implications for Basal Ganglia Models and Deep Brain Stimulation. *Journal of Neuroscience*, *33*(11), 4804–4814. https://doi.org/10.1523/jneurosci.4674-12.2013

Heilbronner, S. R. & Hayden, B. Y. (2016). Dorsal Anterior Cingulate Cortex: A Bottom-Up View. *Annual Review of Neuroscience*, 1–22. https://doi.org/10.1146/annurev-neuro-070815-013952

Heukelum, S. van, Mars, R. B., Guthrie, M., Buitelaar, J. K., Beckmann, C. F., Tiesinga, P. H. E., Vogt, B. A., Glennon, J. C. & 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

Hoemann, K., Gendron, M. & Barrett, L. F. (2017). Mixed emotions in the predictive brain. *Current Opinion in Behavioral Sciences*, *15*, 51–57. https://doi.org/10.1016/j.cobeha.2017.05.013

Holroyd, C. B. & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. https://doi.org/10.1037/0033-295x.109.4.679

Holroyd, C. B. & Krigolson, O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology*, *44*(6), 913–917. https://doi.org/10.1111/j.1469-8986.2007.00561.x

Holroyd, C. B., Krigolson, O. E. & Lee, S. (2011). Reward positivity elicited by predictive cues. *Neuroreport*, *22*(5), 249–252. https://doi.org/10.1097/wnr.0b013e328345441d

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, C. 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

Holroyd, C. B. & Verguts, T. (2021). The Best Laid Plans: Computational Principles of Anterior Cingulate Cortex. *Trends in Cognitive Sciences*, *25*(4), 316–329. https://doi.org/10.1016/j.tics.2021.01.008

Huang, Y., Kakusa, B. W., Feng, A., Gattas, S., Shivacharan, R. S., Lee, E. B., Parker, J. J., Kuijper, F. M., Barbosa, D. A. N., Keller, C. J., Bohon, C., Mikhail, A. & Halpern, C. H. (2021). The insulo-opercular cortex encodes food-specific content under controlled and naturalistic conditions. *Nature Communications*, *12*(1), 3609. https://doi.org/10.1038/s41467-021-23885-4

Hunt, L. T. & Hayden, B. Y. (2017). A distributed, hierarchical and recurrent framework for reward-based choice. *Nature Reviews Neuroscience*, *18*(3), 172–182. https://doi.org/10.1038/nrn.2017.7

Ichikawa, N., Siegle, G. J., Dombrovski, A. & Ohira, H. (2010). Subjective and model-estimated reward prediction: Association with the feedback-related negativity (FRN) and reward prediction error in a reinforcement learning task. *International Journal of Psychophysiology*, *78*(3), 273–283. https://doi.org/10.1016/j.ijpsycho.2010.09.001

Jahn, A., Nee, D. E., Alexander, W. H. & Brown, J. W. (2014). Distinct regions of anterior cingulate cortex signal prediction and outcome evaluation. *NeuroImage*, *95*, 80–89. https://doi.org/10.1016/j.neuroimage.2014.03.050

Jeon, Y.-W. & Polich, J. (2003). Meta-analysis of P300 and schizophrenia: Patients, paradigms, and practical implications. *Psychophysiology*, *40*(5), 684–701. https://doi.org/10.1111/1469-8986.00070

Johnston, R., Jones, K. & Manley, D. (2018). Confounding and collinearity in regression analysis: a cautionary tale and an alternative procedure, illustrated by studies of British voting behaviour. *Quality & Quantity*, *52*(4), 1957–1976. https://doi.org/10.1007/s11135-017-0584-6

Jong, J. W. de, Afjei, S. A., Dorocic, I. P., Peck, J. R., Liu, C., Kim, C. K., Tian, L., Deisseroth, K. & Lammel, S. (2019). A Neural Circuit Mechanism for Encoding Aversive Stimuli in the Mesolimbic Dopamine System. *Neuron*, *101*(1), 133-151.e7. https://doi.org/10.1016/j.neuron.2018.11.005

Joyner, K. J., Bowyer, C. B., Yancey, J. R., Venables, N. C., Foell, J., Worthy, D. A., Hajcak, G., Bartholow, B. D. & Patrick, C. J. (2019). Blunted Reward Sensitivity and Trait Disinhibition Interact to Predict Substance Use Problems. *Clinical Psychological Science*, *7*(5), 1109–1124. https://doi.org/10.1177/2167702619838480

Kam, J., Szczepanski, S. M., Cerebral, R. C. & 2016. (2016). Differential sources for 2 neural signatures of target detection: An electrocorticography study. *Cerebral Cortex*, *28*(1), 9–20. https://doi.org/10.1093/cercor/bhw343

Kennerley, S. W. & Wallis, J. D. (2009). Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables. *European Journal of Neuroscience*, *29*(10), 2061–2073. https://doi.org/10.1111/j.1460-9568.2009.06743.x

Kerns, J. G. (2006). Anterior cingulate and prefrontal cortex activity in an FMRI study of trial-to-trial adjustments on the Simon task. *NeuroImage*, *33*(1), 399–405. https://doi.org/10.1016/j.neuroimage.2006.06.012

Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A. & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, *303*(5660), 1023–1026. https://doi.org/10.1126/science.1089910

Kleckner, I. R., Zhang, J., Touroutoglou, A., Chanes, L., Xia, C., Simmons, W. K., Quigley, K. S., Dickerson, B. C. & Barrett, L. F. (2017). Evidence for a large-scale brain system supporting allostasis and interoception in humans. *Nature Human Behaviour*, *1*(5), 0069. https://doi.org/10.1038/s41562-017-0069

Knight, R. T. (1984). Decreased response to novel stimuli after prefrontal lesions in man. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *59*(1), 9–20. https://doi.org/10.1016/0168-5597(84)90016-9

Koga, S., Rothermel, R., Juhász, C., Nagasawa, T., Sood, S. & Asano, E. (2010). Electrocorticographic correlates of cognitive control in a stroop task-intracranial recording in epileptic patients. *Human Brain Mapping*, *32*(10), 1580–1591. https://doi.org/10.1002/hbm.21129

Kolling, N., Behrens, T., Wittmann, M. K. & Rushworth, M. (2016). Multiple signals in anterior cingulate cortex. *Current Opinion in Neurobiology*, *37*, 36–43. https://doi.org/10.1016/j.conb.2015.12.007

Koster-Hale, J. & Saxe, R. (2013). Theory of Mind: A Neural Prediction Problem. *Neuron*, *79*(5), 836–848. https://doi.org/10.1016/j.neuron.2013.08.020

Krigolson, O. E. (2018). Event-related brain potentials and the study of reward processing: Methodological considerations. *International Journal of Psychophysiology*, *132*(Pt B), 175–183. https://doi.org/10.1016/j.ijpsycho.2017.11.007

Kucyi, A. & Parvizi, J. (2020). Pupillary Dynamics Link Spontaneous and Task-Evoked Activations Recorded Directly from Human Insula. *Journal of Neuroscience*, *40*(32), 6207–6218. https://doi.org/10.1523/jneurosci.0435-20.2020

Kurth, F., Zilles, K., Fox, P. T., Laird, A. R. & Eickhoff, S. B. (2010). A link between the systems: functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Structure and Function*, *214*(5–6), 519–534. https://doi.org/10.1007/s00429-010-0255-z

Kusumoto-Yoshida, I., Liu, H., Chen, B. T., Fontanini, A. & Bonci, A. (2015). Central role for the insular cortex in mediating conditioned responses to anticipatory cues. *Proceedings of the National Academy of Sciences*, *112*(4), 1190–1195. https://doi.org/10.1073/pnas.1416573112

Lachaux, J.-P., Axmacher, N., Mormann, F., Halgren, E. & Crone, N. E. (2012). High-frequency neural activity and human cognition: past, present and possible future of intracranial EEG

research. *Progress in Neurobiology*, *98*(3), 279–301. https://doi.org/10.1016/j.pneurobio.2012.06.008
Lammel, S., Tye, K. M. & Warden, M. R. (2014). Progress in understanding mood disorders: optogenetic dissection of neural circuits. *Genes, Brain and Behavior*, *13*(1), 38–51. https://doi.org/10.1111/gbb.12049
Laubach, M., Amarante, L. M., Swanson, K. & White, S. R. (2018). What, If Anything, Is Rodent Prefrontal Cortex? *ENeuro*, *5*(5). https://doi.org/10.1523/eneuro.0315-18.2018
Leszczyński, M., Barczak, A., Kajikawa, Y., Ulbert, I., Falchier, A. Y., Tal, I., Haegens, S., Melloni, L., Knight, R. T. & Schroeder, C. E. (2020). Dissociation of broadband high-frequency activity and neuronal firing in the neocortex. *Science Advances*, *6*(33), eabb0977. https://doi.org/10.1126/sciadv.abb0977
Li, X., Liang, Z., Kleiner, M. & Lu, Z.-L. (2010). RTbox: a device for highly accurate response time measurements. *Behavior Research Methods*, *42*(1), 212–225. https://doi.org/10.3758/brm.42.1.212
Litt, A., Plassmann, H., Shiv, B. & Rangel, A. (2011). Dissociating Valuation and Saliency Signals during Decision-Making. *Cerebral Cortex*, *21*(1), 95–102. https://doi.org/10.1093/cercor/bhq065
Livneh, Y., Sugden, A. U., Madara, J. C., Essner, R. A., Flores, V. I., Sugden, L. A., Resch, J. M., Lowell, B. B. & Andermann, M. L. (2020). Estimation of Current and Future Physiological States in Insular Cortex. *Neuron*, *105*(6), 1094-1111.e10. https://doi.org/10.1016/j.neuron.2019.12.027
Logothetis, N K, Pauls, J., Augath, M., Trinath, T. & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150–157. https://doi.org/10.1038/35084005
Logothetis, Nikos K. (2008). What we can do and what we cannot do with fMRI. *Nature*, *453*(7197), 869–878. https://doi.org/10.1038/nature06976
Lopez-Persem, A., Verhagen, L., Amiez, C., Petrides, M. & Sallet, J. (2019). The Human Ventromedial Prefrontal Cortex: Sulcal Morphology and Its Influence on Functional Organization. *Journal of Neuroscience*, *39*(19), 3627–3639. https://doi.org/10.1523/jneurosci.2060-18.2019
Loued-Khenissi, L., Pfeuffer, A., Einhäuser, W. & Preuschoff, K. (2020). Anterior insula reflects surprise in value-based decision-making and perception. *NeuroImage*, *210*, 116549. https://doi.org/10.1016/j.neuroimage.2020.116549
Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT Press.
MacDonald, A. W., Cohen, J. D., Stenger, V. A. & Carter, C. S. (2000). Dissociating the Role of the Dorsolateral Prefrontal and Anterior Cingulate Cortex in Cognitive Control. *Science*, *288*(5472), 1835–1838. https://doi.org/10.1126/science.288.5472.1835
Maia, T. V. & Frank, M. J. (2017). An Integrative Perspective on the Role of Dopamine in Schizophrenia. *Biological Psychiatry*, *81*(1), 52–66. https://doi.org/10.1016/j.biopsych.2016.05.021
Manning, J. R., Jacobs, J., Fried, I. & Kahana, M. J. (2009). Broadband shifts in local field potential power spectra are correlated with single-neuron spiking in humans. *Journal of Neuroscience*, *29*(43), 13613–13620. https://doi.org/10.1523/jneurosci.2041-09.2009

Mansouri, F. A., Egner, T. & Buckley, M. J. (2017). Monitoring Demands for Executive Control: Shared Functions between Human and Nonhuman Primates. *Trends in Neurosciences*, *40*(1), 15–27. https://doi.org/10.1016/j.tins.2016.11.001

Mars, R. B., Debener, S., Gladwin, T. E., Harrison, L. M., Haggard, P., Rothwell, J. C. & Bestmann, S. (2008). Trial-by-Trial Fluctuations in the Event-Related Electroencephalogram Reflect Dynamic Changes in the Degree of Surprise. *Journal of Neuroscience*, *28*(47), 12539–12545. https://doi.org/10.1523/jneurosci.2925-08.2008

Matsumoto, H., Tian, J., Uchida, N. & Watabe-Uchida, M. (2016). Midbrain dopamine neurons signal aversion in a reward-context-dependent manner. *eLife*, *5*, e17328. https://doi.org/10.7554/elife.17328

McGuire, J. T., Nassar, M. R., Gold, J. I. & Kable, J. W. (2014). Functionally Dissociable Influences on Learning Rate in a Dynamic Environment. *Neuron*, *84*(4), 870–881. https://doi.org/10.1016/j.neuron.2014.10.013

Meadows, C. C., Gable, P. A., Lohse, K. R. & Miller, M. W. (2016). The effects of reward magnitude on reward processing: An averaged and single trial event-related potential study. *Biological Psychology*, *118*, 154–160. https://doi.org/10.1016/j.biopsycho.2016.06.002

Merre, P. L., Ährlund-Richter, S. & Carlén, M. (2021). The mouse prefrontal cortex: Unity in diversity. *Neuron*, *109*(12), 1925–1944. https://doi.org/10.1016/j.neuron.2021.03.035

Miller, E. K. & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Miltner, W. H., Braun, C. H. & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, *9*(6), 788–798. https://doi.org/10.1162/jocn.1997.9.6.788

Morehead, J. R. & Xivry, J.-J. O. de. (2021). A Synthesis of the Many Errors and Learning Processes of Visuomotor Adaptation. *BioRxiv*, 2021.03.14.435278. https://doi.org/10.1101/2021.03.14.435278

Moser, J. S., Moran, T. P., Schroder, H. S., Donnellan, M. B. & Yeung, N. (2013). On the relationship between anxiety and error monitoring: a meta-analysis and conceptual framework. *Frontiers in Human Neuroscience*, *7*, 466. https://doi.org/10.3389/fnhum.2013.00466

Nácher, V., Hassani, S. A. & Womelsdorf, T. (2019). Asymmetric effective connectivity between primate anterior cingulate and lateral prefrontal cortex revealed by electrical microstimulation. *Brain Structure and Function*, *224*(2), 779–793. https://doi.org/10.1007/s00429-018-1806-y

Neftci, E. O. & Averbeck, B. B. (2019). Reinforcement learning in artificial and biological systems. *Nature Machine Intelligence*, *1*(3), 133–143. https://doi.org/10.1038/s42256-019-0025-4

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C. & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(2), 241–268. https://doi.org/10.3758/s13415-011-0083-5

Nieuwenhuis, S., Holroyd, C. B., Mol, N. & Coles, M. G. H. (2004). Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. *Neuroscience and Biobehavioral Reviews*, *28*(4), 441–448. https://doi.org/10.1016/j.neubiorev.2004.05.003

Nieuwenhuys, R. (2012). The insular cortex: a review. *Progress in Brain Research*, *195*, 123–163. https://doi.org/10.1016/b978-0-444-53860-4.00007-6

Norman, K. J., Riceberg, J. S., Koike, H., Bateh, J., McCraney, S. E., Caro, K., Kato, D., Liang, A., Yamamuro, K., Flanigan, M. E., Kam, K., Falk, E. N., Brady, D. M., Cho, C., Sadahiro, M., Yoshitake, K., Maccario, P., Demars, M. P., Waltrip, L., … Morishita, H. (2021). Post-error recruitment of frontal sensory cortical projections promotes attention in mice. *Neuron*, *109*(7), 1202-1213.e5. https://doi.org/10.1016/j.neuron.2021.02.001

Nusslock, R. & Alloy, L. B. (2017). Reward processing and mood-related symptoms: An RDoC and translational neuroscience perspective. *Journal of Affective Disorders*, *216*, 3–16. https://doi.org/10.1016/j.jad.2017.02.001

Nutt, D. J., Lingford-Hughes, A., Erritzoe, D. & Stokes, P. R. A. (2015). The dopamine theory of addiction: 40 years of highs and lows. *Nature Reviews Neuroscience*, *16*(5), 305–312. https://doi.org/10.1038/nrn3939

O'Connell, R. G., Dockree, P. M. & Kelly, S. P. (2012). A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nature Neuroscience*, *15*(12), 1729–1735. https://doi.org/10.1038/nn.3248

O'Doherty, J. P., Hampton, A. & Kim, H. (2007). Model-Based fMRI and Its Application to Reward Learning and Decision Making. *Annals of the New York Academy of Sciences*, *1104*(1), 35–53. https://doi.org/10.1196/annals.1390.022

Oehrn, C. R., Hanslmayr, S., Fell, J., Deuker, L., Kremers, N. A., Lam, A. T. D., Elger, C. E. & Axmacher, N. (2014). Neural communication patterns underlying conflict detection, resolution, and adaptation. *Journal of Neuroscience*, *34*(31), 10438–10452. https://doi.org/10.1523/jneurosci.3099-13.2014

Oemisch, M., Westendorff, S., Azimi, M., Hassani, S. A., Ardid, S., Tiesinga, P. & Womelsdorf, T. (2019). Feature-specific prediction errors and surprise across macaque fronto-striatal circuits. *Nature Communications*, *10*(1), 176–15. https://doi.org/10.1038/s41467-018-08184-9

Olejnik, S. & Algina, J. (2003). Generalized Eta and Omega Squared Statistics: Measures of Effect Size for Some Common Research Designs. *Psychological Methods*, *8*(4), 434–447. https://doi.org/10.1037/1082-989x.8.4.434

Oliveira, F. T. P., McDonald, J. J. & Goodman, D. (2007). Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. *Journal of Cognitive Neuroscience*, *19*(12), 1994–2004. https://doi.org/10.1162/jocn.2007.19.12.1994

Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, *2011*, 156869. https://doi.org/10.1155/2011/156869

Pastor-Bernier, A. & Cisek, P. (2011). Neural correlates of biased competition in premotor cortex. *Journal of Neuroscience*, *31*(19), 7083–7088. https://doi.org/10.1523/jneurosci.5681-10.2011

Peirce, J. W. (2008). Generating Stimuli for Neuroscience Using PsychoPy. *Frontiers in Neuroinformatics*, *2*, 10. https://doi.org/10.3389/neuro.11.010.2008
Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, *9*(2), 148–158. https://doi.org/10.1038/nrn2317
Pezzulo, G. & Cisek, P. (2016). Navigating the Affordance Landscape: Feedback Control as a Process Model of Behavior and Cognition. *Trends in Cognitive Sciences*, *20*(6), 414–424. https://doi.org/10.1016/j.tics.2016.03.013
Pezzulo, G., Rigoli, F. & Friston, K. (2015). Active Inference, homeostatic regulation and adaptive behavioural control. *Progress in Neurobiology*, *134*, 17–35. https://doi.org/10.1016/j.pneurobio.2015.09.001
Pezzulo, G., Rigoli, F. & Friston, K. J. (2018). Hierarchical Active Inference: A Theory of Motivated Control. *Trends in Cognitive Sciences*, *22*(4), 294–306. https://doi.org/10.1016/j.tics.2018.01.009
Philiastides, M. G., Biele, G., Vavatzanidis, N., Kazzer, P. & Heekeren, H. R. (2010). Temporal dynamics of prediction error processing during reward-based decision making. *NeuroImage*, *53*(1), 221–232. https://doi.org/10.1016/j.neuroimage.2010.05.052
Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D. & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*(2), 127–152. https://doi.org/10.1111/1469-8986.3720127
Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019
Potts, G. F., Martin, L. E., Burton, P. & Montague, P. R. (2006). When Things Are Better or Worse than Expected: The Medial Frontal Cortex and the Allocation of Processing Resources. *Journal of Cognitive Neuroscience*, *18*(7), 1112–1119. https://doi.org/10.1162/jocn.2006.18.7.1112
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
Rangel, A., Camerer, C. & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*(7), 545–556. https://doi.org/10.1038/nrn2357
Rao, R. P. N. & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580
Ray, S., Crone, N. E., Niebur, E., Franaszczuk, P. J. & Hsiao, S. S. (2008). Neural correlates of high-gamma oscillations (60-200 Hz) in macaque local field potentials and their potential implications in electrocorticography. *Journal of Neuroscience*, *28*(45), 11526–11536. https://doi.org/10.1523/jneurosci.2848-08.2008
Rich, E. L. & Wallis, J. D. (2017). Spatiotemporal dynamics of information encoding revealed in orbitofrontal high-gamma. *Nature Communications*, *8*(1), 1139. https://doi.org/10.1038/s41467-017-01253-5
Ridderinkhof, K. R., Ullsperger, M., Crone, E. A. & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, *306*(5695), 443–447. https://doi.org/10.1126/science.1100301

Riesel, A., Goldhahn, S. & Kathmann, N. (2017). Hyperactive performance monitoring as a transdiagnostic marker: Results from health anxiety in comparison to obsessive–compulsive disorder. *Neuropsychologia*, *96*, 1–8. https://doi.org/10.1016/j.neuropsychologia.2016.12.029

Rigotti, M., Barak, O., Warden, M. R., Wang, X.-J., Daw, N. D., Miller, E. K. & Fusi, S. (2014). The importance of mixed selectivity in complex cognitive tasks. *Nature*, *497*(7451), 585–590. https://doi.org/10.1038/nature12160

Rothenhoefer, K. M., Hong, T., Alikaya, A. & Stauffer, W. R. (2021). Rare rewards amplify dopamine responses. *Nature Neuroscience*, 1–5. https://doi.org/10.1038/s41593-021-00807-7

Rouhani, N. & Niv, Y. (2021). Signed and unsigned reward prediction errors dynamically enhance learning and memory. *eLife*, *10*, e61077. https://doi.org/10.7554/elife.61077

Rushworth, M. F., Kolling, N., Sallet, J. & Mars, R. B. (2012). Valuation and decision-making in frontal cortex: one or many serial or parallel systems? *Current Opinion in Neurobiology*, *22*(6), 946–955. https://doi.org/10.1016/j.conb.2012.04.011

Sambrook, T. D. & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin*, *141*(1), 213–235. https://doi.org/10.1037/bul0000006

Sauseng, P., Klimesch, W., Gruber, W. R., Hanslmayr, S., Freunberger, R. & Doppelmayr, M. (2007). Are event-related potential components generated by phase resetting of brain oscillations? A critical discussion. *Neuroscience*, *146*(4), 1435–1444. https://doi.org/10.1016/j.neuroscience.2007.03.014

Schneider, D. M. (2020). Reflections of action in sensory cortex. *Current Opinion in Neurobiology*, *64*, 53–59. https://doi.org/10.1016/j.conb.2020.02.004

Schneider, D. M. & Mooney, R. (2018). How Movement Modulates Hearing. *Annual Review of Neuroscience*, *41*(1), 553–572. https://doi.org/10.1146/annurev-neuro-072116-031215

Schneider, D. M., Nelson, A. & Mooney, R. (2014). A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature*, *513*(7517), 189–194. https://doi.org/10.1038/nature13724

Schneider, D. M., Sundararajan, J. & Mooney, R. (2018). A cortical filter that learns to suppress the acoustic consequences of movement. *Nature*, *561*(7723), 391–395. https://doi.org/10.1038/s41586-018-0520-5

Schultz, W., Dayan, P. & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Seer, C., Lange, F., Boos, M., Dengler, R. & Kopp, B. (2016). Prior probabilities modulate cortical surprise responses: A study of event-related potentials. *Brain and Cognition*, *106*, 78–89. https://doi.org/10.1016/j.bandc.2016.04.011

Shenhav, A., Botvinick, M. M. & Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217–240. https://doi.org/10.1016/j.neuron.2013.07.007

Sherman, S. M. (2007). The thalamus is more than just a relay. *Current Opinion in Neurobiology*, *17*(4), 417–422. https://doi.org/10.1016/j.conb.2007.07.003

Sheth, S. A., Mian, M. K., Patel, S. R., Asaad, W. F., Williams, Z. M., Dougherty, D. D., Bush, G. & Eskandar, E. N. (2012). Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. *Nature*, *488*(7410), 218–221. https://doi.org/10.1038/nature11239

Shipp, S., Adams, R. A. & Friston, K. J. (2013). Reflections on agranular architecture: predictive coding in the motor cortex. *Trends in Neurosciences*, *36*(12), 706–716. https://doi.org/10.1016/j.tins.2013.09.004

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Driessche, G. van den, Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T. & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*(7587), 484–489. https://doi.org/10.1038/nature16961

Smillie, L. D., Jach, H. K., Hughes, D. M., Wacker, J., Cooper, A. J. & Pickering, A. D. (2019). Extraversion and reward-processing: Consolidating evidence from an electroencephalographic index of reward-prediction-error. *Biological Psychology*, *146*, 107735. https://doi.org/10.1016/j.biopsycho.2019.107735

Smith, E. H., Banks, G. P., Mikell, C. B., Cash, S. S., Patel, S. R., Eskandar, E. N. & Sheth, S. A. (2015). Frequency-Dependent Representation of Reinforcement-Related Information in the Human Medial and Lateral Prefrontal Cortex. *Journal of Neuroscience*, *35*(48), 15827–15836. https://doi.org/10.1523/jneurosci.1864-15.2015

Smith, E. H., Horga, G., Yates, M. J., Mikell, C. B., Banks, G. P., Pathak, Y. J., Schevon, C. A., McKhann, G. M., Hayden, B. Y., Botvinick, M. M. & Sheth, S. A. (2019). Widespread temporal coding of cognitive control in the human prefrontal cortex. *Nature Neuroscience*, *66*, 83–89. https://doi.org/10.1038/s41593-019-0494-0

Smith, N. J. & Kutas, M. (2015). Regression-based estimation of ERP waveforms: I. The rERP framework. *Psychophysiology*, *52*(2), 157–168. https://doi.org/10.1111/psyp.12317

Soder, H. E. & Potts, G. F. (2018). Medial frontal cortex response to unexpected motivationally salient outcomes. *International Journal of Psychophysiology*, *132*(Pt B), 268–276. https://doi.org/10.1016/j.ijpsycho.2017.11.003

Sohn, H., Meirhaeghe, N., Rajalingham, R. & Jazayeri, M. (2020). A Network Perspective on Sensorimotor Learning. *Trends in Neurosciences*, *44*(3), 170–181. https://doi.org/10.1016/j.tins.2020.11.007

Sridharan, D., Levitin, D. J. & Menon, V. (2008). A Critical Role for the Right Fronto-Insular Cortex in Switching between Central-Executive and Default-Mode Networks. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(34), 12569–12574. https://doi.org/10.2307/25463900?ref=no-x-route:5fa05007f97621543b5447ed91c16e09

Stolk, A., Griffin, S., Meij, R. van der, Dewar, C., Saez, I., Lin, J. J., Piantoni, G., Schoffelen, J.-M., Knight, R. T. & Oostenveld, R. (2018). Integrated analysis of anatomical and electrophysiological human intracranial data. *Nature Protocols*, *13*(7), 1699–1723. https://doi.org/10.1038/s41596-018-0009-6

Sutton, R. S. & Barto, A. G. (1998). *Introduction to reinforcement learning* (Vol. 135). MIT Press.

Sutton, S., Braren, M., Zubin, J. & John, E. R. (1965). Evoked-Potential Correlates of Stimulus Uncertainty. *Science*, *150*, 1187–1188.

Talmi, D., Atkinson, R. & El-Deredy, W. (2013). The feedback-related negativity signals salience prediction errors, not reward prediction errors. *Journal of Neuroscience*, *33*(19), 8264–8269. https://doi.org/10.1523/jneurosci.5695-12.2013

Tang, H., Yu, H.-Y., Chou, C.-C., Crone, N. E., Madsen, J. R., Anderson, W. S. & Kreiman, G. (2016). Cascade of neural processing orchestrates cognitive control in human frontal cortex. *eLife*, *5*, 4671. https://doi.org/10.7554/elife.12352

Tomaiuolo, F., Campana, S., Voci, L., Lasaponara, S., Doricchi, F. & Petrides, M. (2021). The Precentral Insular Cortical Network for Speech Articulation. *Cerebral Cortex*, *31*(8), 3723-3731. https://doi.org/10.1093/cercor/bhab043

Uddin, L. Q. (2015). Salience processing and insular cortical function and dysfunction. *Nature Reviews Neuroscience*, *16*(1), 55–61. https://doi.org/10.1038/nrn3857

Ullsperger, M., Danielmeier, C. & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiological Reviews*, *94*(1), 35–79. https://doi.org/10.1152/physrev.00041.2012

Ullsperger, M., Fischer, A. G., Nigbur, R. & Endrass, T. (2014). Neural mechanisms and temporal dynamics of performance monitoring. *Trends in Cognitive Sciences*, *18*(5), 259–267. https://doi.org/10.1016/j.tics.2014.02.009

Vestergaard, M. D. & Schultz, W. (2020). Retrospective Valuation of Experienced Outcome Encoded in Distinct Reward Representations in the Anterior Insula and Amygdala. *Journal of Neuroscience*, *40*(46), 8938–8950. https://doi.org/10.1523/jneurosci.2130-19.2020

Vickery, T. J., Chun, M. M. & Lee, D. (2011). Ubiquity and Specificity of Reinforcement Signals throughout the Human Brain. *Neuron*, *72*(1), 166–177. https://doi.org/10.1016/j.neuron.2011.08.011

Vidal, F., Burle, B., Bonnet, M. & Grapperon, J. (2003). Error negativity on correct trials: a reexamination of available data. *Biological Psychology, 64(3), 265-282*. https://doi.org/10.1016/s0301-0511(03)00097-8

Wada, M., Kurose, S., Miyazaki, T., Nakajima, S., Masuda, F., Mimura, Y., Nishida, H., Ogyu, K., Tsugawa, S., Mashima, Y., Plitman, E., Chakravarty, M. M., Mimura, M. & Noda, Y. (2019). The P300 event-related potential in bipolar disorder: a systematic review and meta-analysis. *Journal of Affective Disorders*, *256*, 234–249. https://doi.org/10.1016/j.jad.2019.06.010

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

Wang, J. X., Kurth-Nelson, Z., Kumaran, D., Tirumala, D., Soyer, H., Leibo, J. Z., Hassabis, D. & Botvinick, M. (2018). Prefrontal cortex as a meta-reinforcement learning system. *Nature Neuroscience*, *21*(6), 1–14. https://doi.org/10.1038/s41593-018-0147-8

Watabe-Uchida, M., Eshel, N. & Uchida, N. (2017). Neural Circuitry of Reward Prediction Error. *Annual Review of Neuroscience*, *40*, 373–394. https://doi.org/10.1146/annurev-neuro-072116-031109

Weinberg, A., Dieterich, R. & Riesel, A. (2015). Error-related brain activity in the age of RDoC: A review of the literature. *International Journal of Psychophysiology*, *98*(2 Pt 2), 276–299. https://doi.org/10.1016/j.ijpsycho.2015.02.029

Wessel, J. R., Danielmeier, C., Morton, J. B. & Ullsperger, M. (2012). Surprise and Error: Common Neuronal Architecture for the Processing of Errors and Novelty. *Journal of Neuroscience*, *32*(22), 7528–7537. https://doi.org/10.1523/jneurosci.6352-11.2012

Wessel, J. R. (2018). An adaptive orienting theory of error processing. *Psychophysiology*, *55*(3). https://doi.org/10.1111/psyp.13041

Wessel, J. R. & Aron, A. R. (2017). On the Globality of Motor Suppression: Unexpected Events and Their Influence on Behavior and Cognition. *Neuron*, *93*(2), 259–280. https://doi.org/10.1016/j.neuron.2016.12.013

Wessel, J. R. & Huber, D. E. (2019). Frontal cortex tracks surprise separately for different sensory modalities but engages a common inhibitory control mechanism. *PLOS Computational Biology*, *15*(7), e1006927. https://doi.org/10.1371/journal.pcbi.1006927

Williams, C. C., Ferguson, T. D., Hassall, C. D., Abimbola, W. & Krigolson, O. E. (2021). The ERP, frequency, and time–frequency correlates of feedback processing: Insights from a large sample study. *Psychophysiology*, *58*(2), e13722. https://doi.org/10.1111/psyp.13722

Williams, C. C., Hassall, C. D., Trska, R., Holroyd, C. B. & Krigolson, O. E. (2017). When theory and biology differ: The relationship between reward prediction errors and expectancy. *Biological Psychology*, *129*, 265–272. https://doi.org/10.1016/j.biopsycho.2017.09.007

Wise, S. P. (2008). Forward frontal fields: phylogeny and fundamental function. *Trends in Neurosciences*, *31*(12), 599–608. https://doi.org/10.1016/j.tins.2008.08.008

Wolpert, D. M., Diedrichsen, J. & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews Neuroscience*, *12*(12), 739–751. https://doi.org/10.1038/nrn3112

Womelsdorf, T. & Everling, S. (2015). Long-Range Attention Networks: Circuit Motifs Underlying Endogenously Controlled Stimulus Selection. *Trends in Neurosciences*, *38*(11), 682–700. https://doi.org/10.1016/j.tins.2015.08.009

Womelsdorf, T., Valiante, T. A., Sahin, N. T., Miller, K. J. & Tiesinga, P. (2014). Dynamic circuit motifs underlying rhythmic gain control, gating and integration. *Nature Neuroscience*, *17*(8), 1031–1039. https://doi.org/10.1038/nn.3764

Woolnough, O., Forseth, K. J., Rollo, P. S. & Tandon, N. (2019). Uncovering the functional anatomy of the human insula during speech. *eLife*, *8*, e53086. https://doi.org/10.7554/elife.53086

Wu, Y. & Zhou, X. (2009). The P300 and reward valence, magnitude, and expectancy in outcome evaluation. *Brain Research*, *1286*, 114–122. https://doi.org/10.1016/j.brainres.2009.06.032

Yang, Y.-P., Li, X. & Stuphorn, V. (2020). Primate anterior insular cortex represents economic decision variables postulated by Prospect theory. *BioRxiv*, 2020.11.12.380758. https://doi.org/10.1101/2020.11.12.380758

Yeung, N., Botvinick, M. M. & Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychological Review*, *111*(4), 931–959. https://doi.org/10.1037/0033-295x.111.4.939

Yeung, N. & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience*, *24*(28), 6258–6264. https://doi.org/10.1523/jneurosci.4537-03.2004

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*, *323*(5920), 1496–1499. https://doi.org/10.1126/science.1167342

Zelano, C., Mohanty, A. & Gottfried, J. A. (2011). Olfactory Predictive Codes and Stimulus Templates in Piriform Cortex. *Neuron*, *72*(1), 178–187. https://doi.org/10.1016/j.neuron.2011.08.010