THE ENCODING OF MODEL-BASED CONTROL SIGNALS IN RAT ANTERIOR CINGULATE CORTEX

ALI MASHHOORI
Master of Science, **Shiraz University, 2013**

A thesis submitted
in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

NEUROSCIENCE

Department of Neuroscience
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

THE ENCODING OF MODEL-BASED CONTROL SIGNALS IN RAT ANTERIOR CINGULATE CORTEX

Ali Mashhoori

Date of Defence: February 22, 2021

Dr. Aaron J. Gruber Thesis Supervisor	Associate Professor	Ph.D.
Dr. Matthew Tata Thesis Examination Committee Member	Professor	Ph.D.
Dr. Masami Tatsuno Thesis Examination Committee Member	Associate Professor	Ph.D.
Dr. David Euston Internal External Examiner	Associate Professor	Ph.D.
Dr. Stephen Cowen External Examiner Department of Psychology University of Arizona	Associate Professor	Ph.D.
Dr. Ian Whishaw Chair, Thesis Examination Committee	Professor	Ph. D.

DEDICATION

To my family for giving me all the warmth, care, and love I could ever wish for.

ABSTRACT

The anterior cingulate cortex (ACC) has been implicated in a wide variety of behaviours. Yet, a comprehensive theory of the ACC function is lacking. A promising theory is that ACC uses information from past experience to create predictive mental models that guide future response selection – that is the ACC is a model-based controller. This thesis tests the key hypotheses supporting this theory in two separate tasks. The results of the first experiment show that ACC tracks a rat's state with high spatiotemporal resolution during a binary-choice task. The results of the second task demonstrate that ACC neurons encode several abstract features of a task, including information about the block structure of the task. Together the results suggests that the schema in the ACC is to reconfigure in order to reflect abstract rules of the task, which is essential for optimizing future response selection.

ACKNOWLEDGEMENT

First and foremost, I would like to express my gratitude to my supervisor Dr. Aaron Gruber for his support and mentorship during my studies. I would also like to thank my supervisory committee members, Dr. Masami Tatsuno and Dr. Matthew Tata, as well as members of my thesis examination committee, Dr. David Euston, Dr. Stephen Cowen, and Dr. Ian Whishaw, for their valuable suggestions and comments on my thesis.

I am very grateful to Dr. Saeedeh Hashemnia and Victorita Ivan for providing the data that was used in my research. This thesis would not have been possible without their contribution.

Special appreciation goes to Megan Sholomiski for her love and encouragement. I am also thankful to all of my colleagues and friends at the Canadian Center for Behavioural Neuroscience who indirectly contributed to my research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACC	Anterior cingulate cortex
PFC	Prefrontal cortex
mPFC	Medial prefrontal cortex
vmPFC	Ventromedial prefrontal cortex
dlPFC	Dorsolateral prefrontal cortex
HPC	Hippocampus
OFC	Orbitofrontal cortex
ERN	Error related negativity
RPE	Reward prediction error
RMSE	Root mean square error
ANN	Artificial neural network
AUC	Area under curve
PCA	Principal component analysis
SVM	Support vector machine

1 General Introduction

In the human brain, the anterior cingulate cortex (ACC) is the frontal part of the cingulate cortex (Brodmann areas 24, 32, and 33) neighboring the genu of the corpus callosum. In rodents, the ACC is located on the medial surface of the prefrontal cortex, caudal to the genu of the corpus callosum (Palkovits 1983). Although there is no exact equivalent of primate ACC in rodents, there are strong similarities between the anatomy of primate area 24, of which ACC is a part, and area 24 in rodent brain (Murray, Wise, and Graham 2017).

The ACC has broad connections to other areas in the brain (Barbas and Pandya 1989). These connections include prominent projections to and from the brain systems associated with emotions, decisions, and cognition, such as amygdala, hypothalamus, ventromedial prefrontal cortex (vmPFC), insula, ventral striatum, nucleus accumbens, hippocampus, prefrontal cortex, ventrolateral prefrontal cortex, frontal pole, and parietal cortex (Van Hoesen, Morecraft, and Vogt 1993; Haber et al. 2006). The ACC is also connected with areas involved in motor control including the motor cortex, premotor cortex, and the spinal cord (Picard and Strick 1996).

Consistent with the diversity of connections, the ACC appears to be involved in a wide range of processes important for regulating responses in a situationally-appropriate manner, including assessing the salience of environmental features, signaling motivational information, assigning control to other areas in the brain, conflict detection, reward-based learning, and decision making (Heilbronner and Hayden 2016). A concise theory for the function of ACC that accounts for its role in these scenarios is presently missing.

An early theory proposed that ACC's main function was error monitoring (Clay B. Holroyd and Coles 2002). This idea was formed based on the prevalence of the error-related-negativity signal in the ACC (Gehring et al. 1993; Gehring and Willoughby 2002). Error-related negativity (ERN) is one type of event-related-potential (electrical activity in the brain time-locked to an external event) that is observed after the commission of an error during a choice task (Gehring et al. 1990), even when the subject is not explicitly aware of the error (Nieuwenhuis et al. 2001). Correlates of the ERN signal have been found in the activity of primate ACC in both single-cell studies (Gemba, Sasaki, and Brooks 1986; S. Ito et al. 2003; Narayanan et al. 2013; Shen et al. 2015) and imaging studies (Ullsperger and Von Cramon 2001). In addition, extensive empirical research points to the ACC as the brain area that generates the ERN (S. Ito et al. 2003; C B Holroyd et al. 2004; Stemmer et al. 2004).

Later studies reported observations that were inconsistent with the error monitoring theory. One example is that the ACC's activity is influenced by situations where errors are likely, but do not actually happen (J. W. Brown and Braver 2005). Another example is the effect of the cues indicating the need to change the response strategy, irrespective of error commission (Céline Amiez, Joseph, and Procyk 2005). Other studies found the presence of strong control and comparison signals in ACC activity (Johnston et al. 2007; Shenhav, Botvinick, and Cohen 2013; Seo and Lee 2007). Consequently, the idea of error monitoring as the exclusive role of the ACC has been largely rejected (Wallis and Rich 2011).

To address the shortcomings of the error-monitoring theory in explaining the data, researchers proposed the conflict monitoring theory of ACC function (M. M. Botvinick et al. 2001; Van Veen et al. 2001; Kerns et al. 2004). In this view, activity in the ACC encodes the level of conflict or competition between possible responses. The ACC then generates a

signal that triggers strategic adjustments in cognitive control and allocation of additional cognitive resources to prevent a conflict in subsequent trials. The conflict monitoring theory explains the ACC activity in three behavioural contexts: overriding the dominant response in a task (M. Botvinick et al. 1999; M. M. Botvinick et al. 2001), selecting a response among a set of equally permissible responses (Palmer et al. 2001; Thompson-Schill et al. 1997), and mediating situations pertaining to error commission (Yeung, Botvinick, and Cohen 2004). Although numerous neuroimaging studies support the conflict monitoring theory, several studies at the single-unit level have failed to find evidence for it (C. Amiez, Joseph, and Procyk 2006; Cai and Padoa-Schioppa 2012; Hayden et al. 2011; Quilodran, Rothé, and Procyk 2008). One suggestion to resolve this issue is that conflict influences the ACC by increasing the number of active cells, not by increasing the firing rates of conflict-sensitive neurons (Nakamura, Roesch, and Olson 2005; Alexander and Brown 2011). In this view, the conflict signal is a by-product of the ACC processing, but not its main function.

The ACC appears to be involved in the processing of positive and negative outcomes. Neurophysiological studies in humans have pointed to the ACC's role in the processing of monetary gains and losses (Gehring and Willoughby 2002). Studies on monkeys have shown that single neurons in the ACC are sensitive to the values of the obtained rewards (C. Amiez, Joseph, and Procyk 2006). Furthermore, lesion or inactivation of the ACC impairs a monkey's ability to select the appropriate response in a reward-based selection task (Keisetsu Shima and Tanji 1998; Hadland et al. 2003). Encoding of the value of the obtained reward by the ACC has also been reported in rodent studies (Hashemnia, Euston, and Gruber 2020; Mashhoori et al. 2018). Despite the ACC's apparent involvement in processing reward, the way that reward influences the activity of the neurons in the ACC

seems to be context-dependent. For instance, a larger reward can evoke either a reduction or an increase in the firing rate of individual ACC neurons (Hayden, Pearson, and Platt 2011; Kennerley, Behrens, and Wallis 2011; Hayden, Pearson, and Platt 2009; Hillman and Bilkey 2010). The direction of tuning can change even within a session (Hayden et al. 2011; C. H. Luk and Wallis 2009; M. Matsumoto et al. 2007).

Some researchers have interpreted the post-outcome activity of the ACC as a reward prediction error (RPE), representing the difference between an expected and an actual outcome (Kennerley, Behrens, and Wallis 2011; M. Matsumoto et al. 2007; Seo and Lee 2007). Other researchers have found a representation of the unsigned RPE in the ACC (Bryden et al. 2011; Hayden et al. 2011), which can be considered a measure of surprise. To reconcile these inconsistencies, it is proposed that the ACC might carry a control signal, correlated with reward, and modulated by surprise, which triggers a change in strategy, and depending on the context, may reflect signed or unsigned RPE (Duncan 2001).

In addition to the encoding of a received reward, the ACC also signals rewards available in alternate states of the task. These include the value of the unchosen options (Blanchard and Hayden 2014), the value of fictive rewards; i.e., rewards that could have been obtained but were not (Hayden, Pearson, and Platt 2009), and the average value of alternative options (Kolling et al. 2012). In one study, the values of possible choice options varied in three different dimensions: payoff, probability, and effort cost (effort). Neurons in the ACC represented the value of each option by a coding scheme that integrated the three value dimensions (Kennerley et al. 2009). Such integration of cost-benefit information is typically formulated in the concept of 'utility', which is the key quantity for nearly all theories of economic choice (Glimcher and Fehr 2013). The same study showed

that the utility representation in the ACC is stronger than other areas associated with value, such as OFC and dlPFC, suggesting that utility computation may be a key function of ACC.

Some evidence has suggested that the ACC encodes the value of a non-default option, or the value in switching away from the standard action (Blanchard and Hayden 2014; Boorman, Rushworth, and Behrens 2013; Kolling et al. 2012; Strait, Blanchard, and Hayden 2014). Supporting this idea, imaging studies show a decline in the activity of the ACC as the value of the difference between chosen and unchosen value increases (Boorman et al. 2009; FitzGerald, Seymour, and Dolan 2009; Lim, O'Doherty, and Rangel 2011). ACC activity seems to be greatest when the non-default option becomes a better choice, requiring an update of the choice strategy (Shenhav, Botvinick, and Cohen 2013).

Contrary to these findings, Cai and Padoa-Schioppa (2012) did not find any representation of option value other than the subjective value of the chosen option in monkeys. Another study showed that the ACC encoding of value is context-dependent (C.-H. Luk and Wallis 2013). One solution to reconcile these contradictory findings is that ACC represents a variable correlated with the value, not the value itself (Heilbronner and Hayden 2016). In sum, at least some ACC neurons appear to generate signals related to the utility of chosen or un-chosen options, but this computation is either highly dependent on abstract task features (e.g., context) or is mixed with populations of cells encoding other features such that random selection of ACC neurons (as with electrode recordings) does not always sufficiently sample the utility-encoding signals. For instance, it is possible (but speculative) that the type of encoding varies by cortical layer, and that different research groups do not sample equally across them.

A large body of evidence implicates the ACC with control, defined as a psychological process involved in direct regulation of an action, a cognitive process, or

another control process. Both single-unit and neuroimaging studies show greater ACC activation when control is valuable or necessary (Johnston et al. 2007; Shenhav, Botvinick, and Cohen 2013). A recent theory of ACC function proposes that the ACC estimates the net value associated with allocating control to a given task, a quantity referred to as the expected value of control (EVC), by integrating information about the expected rewards and costs in a control-demanding task (Shenhav, Botvinick, and Cohen 2013). This signal is then used to select among competing tasks and allocate the appropriate amount of control. This view is closely related to the conflict monitoring theory of the ACC, as recruitment of control is necessary in a conflicting situation to resolve the conflict and improve future performance (M. M. Botvinick et al. 2001).

The relation between the ACC and motor control is featured by connections between the ACC and brain areas involved in motor control such as the motor cortex and the spinal cord (Akkal et al. 2002; Morecraft and Van Hoesen 1998; K. Shima et al. 1991). Studies report the sensitivity of ACC neurons to saccade direction (Hayden and Platt 2010; Cai and Padoa-Schioppa 2012) and arm movement direction in monkeys (Kennerley et al. 2009), movement and head direction in rodents (D. R. Euston and McNaughton 2006), and specific elements of compound actions (Hoshi, Sawamura, and Tanji 2005). However, not all results are consistent with these reports, as several studies found little ACC selectivity to spatially oriented features (Hoshi, Sawamura, and Tanji 2005; Kennerley and Wallis 2009a; M. Matsumoto et al. 2007). To reconcile these different results, it is proposed that encoding of spatially oriented features appears in the ACC only when the information is relevant to decision-making.

Several studies have implicated the ACC in triggering adjustments in action plans or strategies. In this view, ACC activity is linked to the receipt of information that leads to

changes in a subject's internal model of a task (Hayden et al. 2011; Quilodran, Rothé, and Procyk 2008). For instance, monkeys performed a task with two possible actions, one of which was rewarded in each block. The neurons in the ACC responded most strongly when the reward associated with one action was reduced, and the optimal action changed (Keisetsu Shima and Tanji 1998). Supporting this view, lesions to the ACC have been shown to impair animals' ability to switch their action in response to a reduction in reward (M. F.S. Rushworth et al. 2003; Keisetsu Shima and Tanji 1998; Williams et al. 2004) and to maintain a new strategy after a switch (Chudasama et al. 2013; Kennerley et al. 2006).

There is a close relationship between outcomes and events that signal an adjustment in strategy. Therefore, it is difficult to distinguish the outcome and control-related activity in the ACC. One study showed that the activity in the ACC correlated with the outcome when the strategic adjustment was held constant, and it predicted the adjustment when the outcome was held constant. (Hayden et al. 2009). These results suggest that the ACC may serve as both a monitor for outcomes and controller of other cognitive processes.

The ACC has been associated with persistence and self-control. This theory is supported by rodent studies using a delay task (Narayanan and Laubach 2006) and forced swim task (Warden et al. 2012). ACC-lesioned rats shift their behaviour away from investing physical effort to reach the high-cost, high-reward options (Walton et al. 2003; Schweimer, Saft, and Hauber 2005; Floresco and Ghods-Sharifi 2007), despite being able to discriminate reward size (Peter H. Rudebeck et al. 2006) and reward location (Kesner and Ragozzino 2003). It is also supported by human studies using an intertemporal choice task (Peters and Büchel 2010) and a response inhibition task (Floden and Stuss 2006). In humans, activation of the ACC produces a strong desire to persevere against challenges (Parvizi et al. 2013). Therefore, ACC activation might bias the choice toward the higher

cost option if the outcome is sufficiently greater (Johnston et al. 2007; P. H. Rudebeck et al. 2006). The ACC might promote self-control by carrying a value signal that allows decision-makers to overcome temptations (Blanchard, Strait, and Hayden 2015; Hillman and Bilkey 2010) or encoding the cost of failing to persist (Blanchard and Hayden 2014; Kurzban et al. 2013).

The ACC and OFC are considered to play a role in stimulus-outcome and action-outcome associations, respectively. This dissociation is backed by findings in humans (Camille, Tsuchida, and Fellows 2011), monkeys (Peter H. Rudebeck et al. 2008), and rodents (Balleine and Dickinson 1998; Pickens et al. 2003; Ostlund and Balleine 2007) at the population level. The results, however, are inconsistent at the single-neuron level. Robust encoding of actions in the OFC has been reported in non-human primates (Tsujimoto, Genovesio, and Wise 2009) and rodents (Feierstein et al. 2006; Furuyashiki, Holland, and Gallagher 2008; Sul et al. 2010; Van Wingerden et al. 2010). Also, although numerous studies have emphasized the role of ACC in predicting the outcome associated with a given action (S. Ito et al. 2003; K. Matsumoto, Suzuki, and Tanaka 2003; Williams et al. 2004; C. H. Luk and Wallis 2009; Hayden and Platt 2010), there have been studies showing encoding of the rewards predicted by stimuli by ACC neurons (Hyojung Seo and Lee 2007; Kennerley et al. 2009; Cai and Padoa-Schioppa 2012). C.-H. Luk and Wallis (2013) investigated this issue in a task comprised of action-outcome and stimulus-outcome trials and found a small population of neurons encoded the specific action-outcome or stimulus-outcome associations. In contrast, they found robust encoding of the action expressing the choice. Furthermore, the encoding of this action was stronger in the OFC in the stimulus-outcome trials and stronger in the ACC in the action-outcome trials. They concluded that the main function of these two areas is not the encoding of specific action-

outcome or stimulus-outcome associations, but rather using those associations to guide decision-making.

To select the best course of action in an environment, a comparison between expected utilities of the available options is necessary. A primary component of the expected utility is the cost associated with each option. Previous studies have reported ACC's sensitivity to both physical (Croxson et al. 2009; Klein-Flugge et al. 2016; Prevost et al. 2010; Hashemnia, Euston, and Gruber 2020; Porter, Hillman, and Bilkey 2019) and mental (Engström, Landtblom, and Karlsson 2013; Hosking, Cocker, and Winstanley 2014; Vassena et al. 2014) effort. Besides, several studies have reported direct encoding of the utility by the ACC neurons in rodents and primates (Kennerley and Walton 2011; Wallis and Kennerley 2011; Hashemnia 2019). For instance, in primates, ACC neurons are able to encode the overall value of options by taking into accounts factors such as reward probability, reward value, and costs, including delays and physical effort (Hosokawa et al. 2013; Kennerley and Wallis 2009b; Kennerley et al. 2009). In rodents, ACC neurons show selectivity to the most valuable option based on the integration of effort costs and reward values (Cowen, Davis, and Nitz 2012; Hillman and Bilkey 2010). Taken together, these studies suggest that ACC neurons represent the net utilities of options by integrating the cost and reward associated with each one and bias choice toward the most valuable option. This view is supported by studies that report impairment of optimal decision-making in ACC-lesioned animals (Walton et al. 2003; Schweimer, Saft, and Hauber 2005; Floresco and Ghods-Sharifi 2007), especially in the situations involving balancing effort and reward.

Conventionally, economic choice is considered to be comprised of evaluation and comparison stages. The evaluation stage is often linked to the orbital surface, and the comparison stage is suspected to be restricted to the ACC (Hare et al. 2011; Seo and Lee

2007; Wunderlich, Range, and O'Doherty 2009). This view is backed by studies that show ACC-lesioned animals switch away from the high-cost high-reward option (P. H. Rudebeck et al. 2006). Nevertheless, a large number of studies associate comparison with other regions of the brain such as OFC (Padoa-Schioppa and Assad 2006), vmPFC (Strait, Blanchard, and Hayden 2014), and ventral striatum (Strait et al. 2016). Strong evidence against this view comes from studies that show no significant impairment of the economic choice in ACC-lesioned animals (Chudasama et al. 2013; Matthew F.S. Rushworth et al. 2011), suggesting that the ACC is a part of a distributed choice process rather than the main center for value comparison (Heilbronner and Hayden 2016).

Overall, ACC neurons encode many task variables at both the population level and the single-unit level. These variables can be grouped into ones that define the task state, or the context, and ones that derive action. Here, action is a general term that could mean abstract actions, such as adopting a new strategy or enhancing cognitive control. Heilbronner and Hayden (2016) propose that the ACC encodes all and only the variables that are relevant to decision-making, and the strength and prevalence of the encoding of each variable reflect its importance. Therefore, depending on the task and factors relevant to decision making, the ACC tracks different variables, such as errors, conflicts, offers' values, outcomes, spatial representation. This view can explain the inconsistent observations regarding the encoding of different variables such as conflict, space, and reward in the ACC. Similar ideas have been proposed by other researchers (Clay B. Holroyd and Yeung 2012; Clay B. Holroyd and McClure 2015).

In this thesis, I propose that the ACC plays a role in model-based control. In this view, the ACC forms schematic knowledge of the task (e.g., rules, spatial layout, outcome values and probabilities) based on experience, and utilizes it to predict likely states and

outcomes as a function of context and response options. Such model-based predictions can be used to guide response selection and strategy shift when such adjustments are likely to yield better returns. This theory generates a number of key predictions:

- **Hypothesis 1.** To form a predictive model of the task, the ACC must represent an abstract map of task states and track an animal's state as it performs the task. This, in turn, raises the possibility of mental navigation of the task state space by the brain.
- **Hypothesis 2.** The ACC needs to track task variables that define and represent context, stimulus, responses, and outcomes.
- **Hypothesis 3.** In addition, the ACC needs to create abstract states that are inferred from the history of events and their relationship that are critical for future predictions.
- **Hypothesis 4.** Finally, the ACC must detect deviations between the predictions and actual outcomes to promote model update when the previous model is not valid anymore.

In the following chapters, I test the key hypotheses presented above using state-of-the-art methods to decode information from ensembles of simultaneously recorded neurons from rat ACC. In chapter 2, I show that the ACC tracks the state of a rat as it performs a binary-choice task by demonstrating that the position of the rat can be decoded accurately from ACC activity on a figure-8 maze. I further provide evidence of counterfactual processing in rodent ACC, a process that is likely to play a role in shifting strategy in future choices. In chapter 3, I show that ACC encodes main task variables and represents abstract task states that are useful for creating a predictive model of the task. In chapter 4, I demonstrate a novel deep neural network that transforms data into a new space in which variations informative about one variable is removed, and the remaining variation is retained. This network can be used to eliminate the confounding effect of one variable on

decoding another variable, and is a key tool to dissociate encoding of abstract features (e.g. task rules) from confounding signals (e.g. reward). In the last chapter, I discuss how the results presented in this thesis support the role of the ACC in model-based control, and how it accounts for several previously proposed functions of this structure.

2 Rat Anterior Cingulate Cortex Recalls Features of Remote Reward Locations after Disfavoured Reinforcements

Abstract

The anterior cingulate cortex (ACC) encodes information supporting mnemonic and cognitive processes. We show here that a rat's position can be decoded with high spatiotemporal resolution from ACC activity. ACC neurons encoded the current state of the animal and task, except for brief excursions that sometimes occurred at target feeders. During excursions, the decoded position became more similar to a remote target feeder than the rat's physical position. Excursions recruited activation of neurons encoding choice and reward, and the likelihood of excursions at a feeder was inversely correlated with feeder preference. These data suggest that the excursion phenomenon was related to evaluating real or fictive choice outcomes, particularly after disfavoured reinforcements. We propose that the multiplexing of position with choice-related information forms a mental model isomorphic with the task space, which can be mentally navigated via excursions to recall multimodal information about the utility of remote locations.

2.1 Introduction

The ACC and other nearby structures in the medial prefrontal cortex (mPFC) play an important role in the control of both memories and decisions (David R. Euston, Gruber, and McNaughton 2012). These structures influence memory retrieval via connectivity with the hippocampus (H. T. Ito et al. 2015; Rajasethupathy et al. 2015), and are thought to utilize hippocampal output to form semantic or schematic knowledge of the world from past experience (McClelland, McNaughton, and O'Reilly 1995). Activation of patterned neural activity in the mPFC may thus play an important role in utilizing experiential or schematic knowledge to plan or control behaviour (Tse et al. 2007; Wang, Tse, and Morris 2012). Neurons in the ACC and nearby regions encode a variety of task features related to reinforcement and decisions (Kennerley et al. 2006; A. J. Gruber et al. 2010; Sul et al. 2011), and many are also selectively active over large regions of the task space (Jung et al. 1998; D. R. Euston and McNaughton 2006; Fujisawa et al. 2008; Jadhav et al. 2016). The function of this broad and distributed spatial mapping by individual ACC units has remained more contentious than the sparse encoding of location by neurons in the hippocampus (Burton et al. 2009; Hyman et al. 2012).

The hippocampus contains 'place cells' that provide information about the position of an animal in an environment (O'Keefe and Dostrovsky 1971; M. A. Wilson and McNaughton 1993), which is utilized by the ACC (Burton et al. 2009). Although ACC neurons had generally been thought not to generate place fields (Poucet 1997), recent work has revealed neurons in mPFC with spatially specific firing that typically span over 50cm, and that are distributed over the task environment (Fujisawa et al. 2008). These are properties expected of a place code. Nonetheless, the broad spatial encoding in the mPFC

has often been interpreted as signaling contextual features, such as the environment (Hyman et al. 2012) and task (L. Ma et al. 2016). The ACC also appears to utilize reinforcement information from past actions to engage action strategies that improve cost-benefit outcomes (M. M. Botvinick, Cohen, and Carter 2004; C. Amiez, Joseph, and Procyk 2006; Kennerley et al. 2006; Matthew F.S. Rushworth et al. 2011; Heilbronner and Hayden 2016), particularly when rapid shifts in strategy are needed to optimize reward acquisition (Posner et al. 2007; Passingham and Wise 2015). One of its specific functions is to encode unattained rewards (Hayden, Pearson, and Platt 2009), which may contribute to its role in signalling regret when outcomes do not meet expectations (Coricelli et al. 2005). It is possible that the ACC uses a fine-grained spatial map as a mnemonic scheme to recall possible alternate outcomes at other locations for such processing, but we are unaware of any direct evidence with sufficient spatiotemporal resolution to accurately decode such shifts.

2.2 Results

2.2.1 Precise spatio-temporal encoding of position by ACC

Because populations of broadly tuned cells can encode quantities with more precision than the encoding of individual cells (Kim, Ganguli, and Frank 2012), we first sought to determine whether population activity in the ACC accurately encodes the position of an animal. We recorded ensembles of ACC neurons while rats performed a binary choice task on a figure-8 track (Figure 2.1.A). Two target feeders at the north corners of the track could be reached after turning right or left at the choice point, and a third feeder in the central segment was used to motivate rats to return to the starting position. Rats received reinforcement at the central feeder on every trial. The effort-reward utility of each choice was independently controlled by elevating the target feeders to one of three heights and by providing either a small or large reward volume. The utilities were held fixed for 16 consecutive trials and the animals were forced to alternate between the right and left options on the first 10 trials before being allowed free selection of either option for the remaining 6 trials of the block. The task was run in a 6-block sequence of 96 trials (Figure 2.1.B), and the same sequence would restart upon completion within each session. The reward contingencies were counterbalanced among right and left feeders between sessions. Rats completed 162-256 trials per session. Animals received the same small reward at the central feeder. We simultaneously recorded position and neural activity in the ACC (Figure 2.1.C-D) and found that most cells activated over large areas of the track (Figure 2.2), consistent with previous reports.

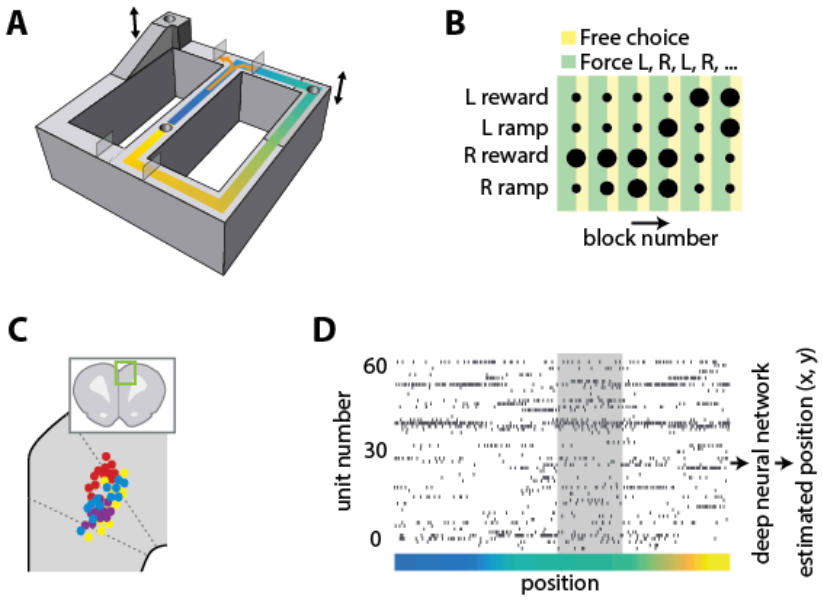


Figure 2.1: Task and neural recording.
(A) Schematic illustration of the figure-8 track, showing the locations of the feeders (cylindrical depressions), height-adjustable platforms (indicated by arrows), and movable gates (translucent rectangles). **(B)** Graphical representation of the choice reward-effort utilities (dot size) and choice option (color) structure of one task session. The effort-reward utility of each choice was constant during each block of 16 laps. **(C)** Illustration of estimated recording locations in the dorsal medial prefrontal cortex (inset), showing that most fell in the ACC. **(D)** Representative example of simultaneously recorded ACC ensemble activity during one lap of the task. The color indicates the position on the track as coded in panel **A,** and the grey shaded region is corresponds to the target feeder location.

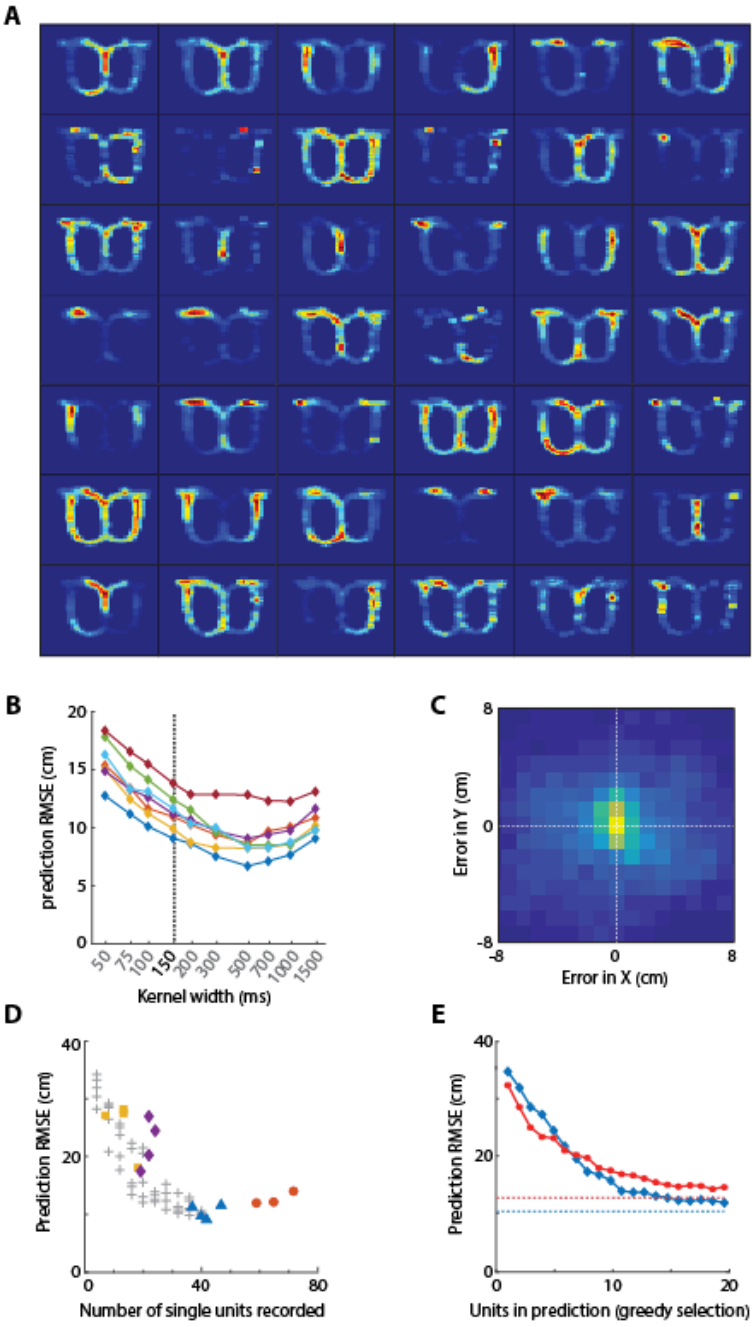


Figure 2.2: Spatial attributes of ACC unit firing and decoding error.
(A) Maps of the spatial attributes of neural firing for each of 42 simultaneously recorded neurons in the ACC. Red indicates highest likelihood of unit firing. **(B)** Prediction error versus the width of the smoothing kernel for all seven sessions analyzed. The dashed line indicates the width used for decoding analyses in Figure 2.3. **(C)** Histogram of the prediction errors in Cartesian coordinates for one representative session. **(D)** Prediction error (RMSE) as a function of the number of qualified units recorded for several sessions from each animal (symbols and colors). The prediction error increased with fewer units. To test if the poor decoding performance for some subjects was due to the low number of recorded units, we tested if downsampling the number units from another animal

with 40 units would result in similar poor performance. The prediction error increased super-linearly as the downsampling was increased (grey '+') and is in-line with the poor predictions from the two animals with few units. **(E)** The prediction error achieved by using the most informative units, as determined by a greedy selection process for the two animals with the most units. Colors correspond to those in panel **D**.

The currently predominant neural decoding model for position is Bayesian reconstruction (E. N. Brown et al. 1998; Zhang et al. 1998; Carr, Jadhav, and Frank 2011). We found, however, that we could achieve significantly lower decoding error (36% reduction; $t(6) = 9.0$, $p= 0.0001$, power = 1) than the Bayesian method by using a deep artificial neural network (dANN) to decode location from patterns of neural activity in bins of 20-50ms (Figure 2.3.B; see METHODS). The dANN was memoryless in that it only used information from the present time bin for the predictions. It could exploit higher-order statistical relationships among inputs than could the Bayesian method, and could learn to ignore spurious information. Its superior performance therefore likely demonstrates that either these higher-order statistics carry a significant amount of information about spatial position, as previously predicted (Fujisawa et al. 2008), or that the representation of non-spatial features hampers Bayesian reconstruction. Our analysis of multiple sessions from four animals revealed that approximately 30 randomly selected ACC units are required for good reconstruction accuracy, whereas near asymptotic accuracy can be achieved by using the 17 most informative cells (Figure 2.2). We therefore focused subsequent analysis on seven sessions from two animals with at least 40 simultaneously recorded cells so as to achieve decoding error close to the apparent asymptotic limit.

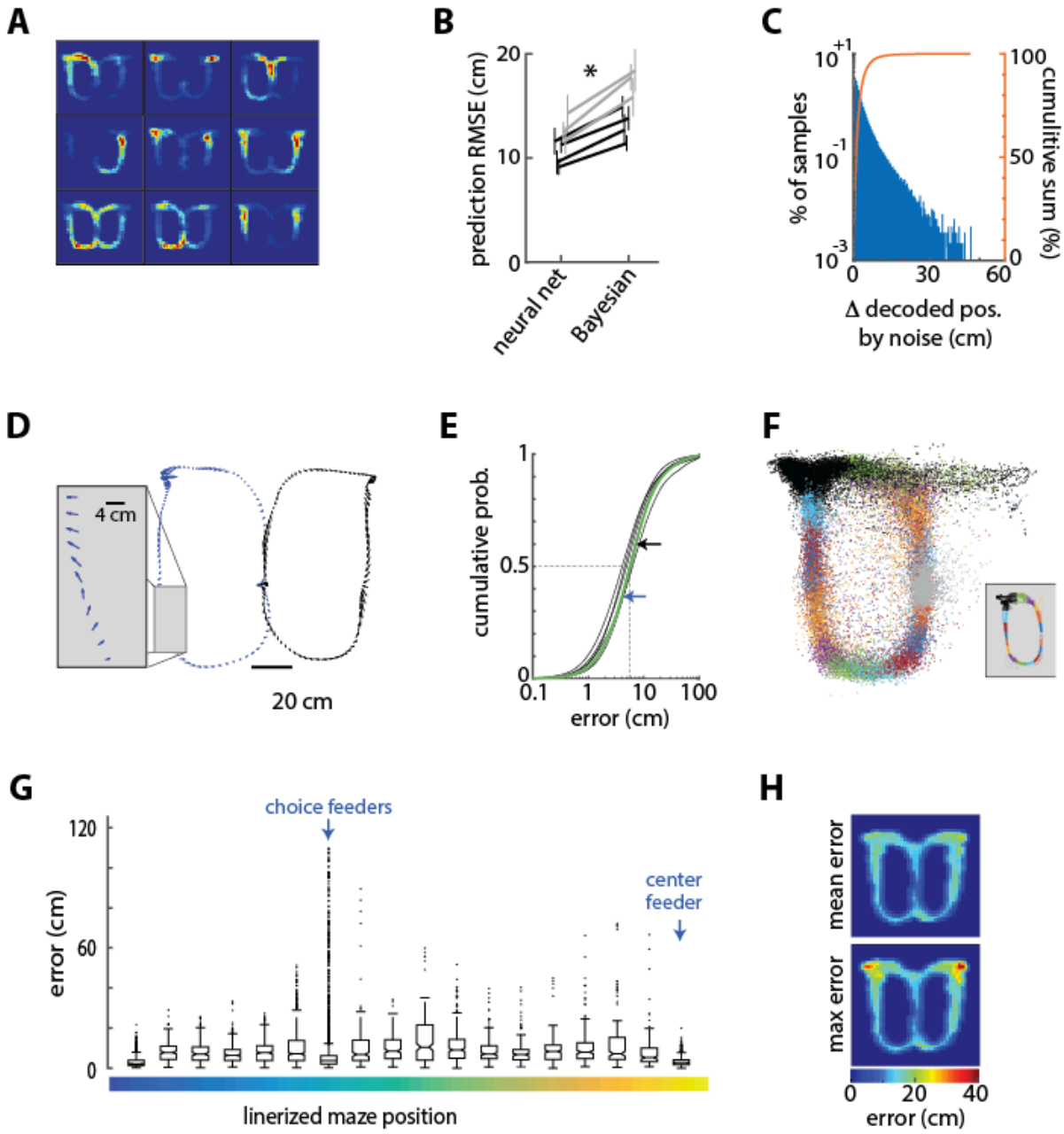


Figure 2.3: Decoding position from ensemble ACC activity.

(A) The spatial selectivity of the top 9 most informative ACC cells for decoding position in one session chosen by the decoding network. Spike density on the track is coded by color, from blue to red. The units are rank ordered by importance from top left to bottom right. **(B)** The root mean squared error (RMSE) of the position decoded from the ACC activity as compared to the actual position of the rat for each session (line) and rat (shade), showing that the deep artificial neural network generates lower prediction error than does a Bayesian decoder for each of the seven sessions tested. These session-averaged errors are inflated by occasional large errors around the reward zones, as described below. Error bars show standard deviation of 20 randomly selected training and test sets for each session and method. **(C)** The distribution of changes in decoded

position by noise**. (D)** Error vectors for two representative laps of the task. The arrows indicate the magnitude and direction of the decoding error every 50 ms. **(E)** Cumulative probabilities of the prediction error magnitude for the seven sessions. The dotted lines indicate the median, and the arrows indicate the median error for the left (blue) and right (black) laps session shown in panel **D** (green curve). **(F)** Decoded position for test data from one session, color coded by the actual position (inset) **(G)** The error computed every 50 ms in one representative session, represented as a box plot according to track position as shown in Figure 2.1. The box plot shows the median (horizontal lines in boxes), 95% confidence intervals (notches), 1st and 3rd quartiles (boxes ends), and outliers (dots). There are a disproportionate number of outliers in the bin corresponding to the target feeder locations, but the median prediction accuracy is as good at these feeders as anywhere else on the track. **(H)** The mean (top) and maximal (bottom) prediction error for discretized positions on the track, showing that the very large errors occurred exclusively at the location of the target feeders. These show mean of means computed from all sessions.

We next sought to determine the robustness of the decoder by quantifying its performance in the presence of noise. We shifted each spike time by a randomly drawn value from a distribution that was proportional to the variance of the cell's firing interval (25% of its STD). The dANN was trained on some trials of the uncorrupted data and tested on different trials with the noisy data (mean spike shift was 25 ms). This increased the error (RMSE) by 1.3% (0.14 cm), which was a small yet statistically significant difference ($t(6) = 4.8$; $p = 0.003$; power = 0.95). This indicates that spike timing on the order of 25 ms contained useful information for the decoding with the present neuronal sample size. Furthermore, the noise did not cause any large deviations in decoding. Only 0.07% of samples deviated more than 25 cm from the original decoding, and none were more than 50 cm (Figure 2.3.C). These data indicate that the decoder is robust against moderate levels of spike jitter. The median decoding accuracy of the rat's position on the task was less than 10 cm (Figure 2.3.D-E). This is much less than the length of the animal's body, and much less than the spatial selectivity of individual cells (Fujisawa et al. 2008).

2.2.2 Invariance of spatial encoding during task sessions

We analyzed the stability of the spatial information over the entire maze through time. For this, we separated all the trials that belonged to the trial configuration (reward, effort, direction) with the maximum number of trials in each session. Then, we performed two different tests. In one, we created the training set by selecting every other trial in the list and used the remaining trials as the test set. In the second test, we used the first half of the trials as the training set and used the second half of the trials as the test set. If the spatial encoding shifted as a result of the intervening configurations (block types), or drifted in time, then the decoding accuracy in these two cases should be different. We found that decoding errors were not significantly different in these two cases (0.6 cm change in error, $t(6) = 1.28$, $p=0.24$). This analysis provides some evidence that the position-related features of neural activity used by the decoding network are stable within the session, despite changes in effort and reward contingencies that occur during intervening blocks.

2.2.3 Excursions of spatial encoding from the physical position to a feeder

The mean spatial decoding error was low in all positions on the track (Figure 2.3.F-H). The decoded position, however, sometimes deviated from the actual position by up to nearly 100 cm, and these large excursions occurred almost exclusively at the two choice-option feeders (Figure 2.3.G-H). These excursions were not random in time or decoded position. Rather, the excursions consisted of several consecutive points encoding the alternate target feeder before returning to the present location of the rat (Figure 2.4.A-B; online video). The localization of these excursion endpoints was particularly striking because our decoder network output only an x and y coordinate, with no constraints that

the decoded position lie on the track. This pattern of endpoints is therefore exceedingly unlikely by chance (χ^2 = 484, p=1 E-40, power = 1). These excursions did not originate or terminate at the center feeder, even though the animal received the same reward type at all feeders, and the volume at the center feeder was equivalent to the small volume at the choice feeders. These features strongly suggest that the excursion phenomenon is involved in computations related to evaluating the choice options, and not related to general qualities of the reward or to planning the immediately next action, which was always a return to the center feeder as enforced by gates on the track.

The median duration of the excursion events was 400 ms (20 consecutive bins of 20 ms; Figure 2.4.C) and occurred almost exclusively while the animal was stationary at the feeders. Although excursions occasionally occurred during the reinforcement, most occurred after the reinforcement and prior to locomotion away from the feeder (Figure 2.4.D). The duration of the excursions was appreciably longer than the width of the smoothing kernel (120 ms), so the end-points of the excursions were not affected by the smoothing. The dynamics of the transition, however, occurred on a time scale less than the kernel width and was therefore strongly affected. The intermediate points of excursion between the feeder sites appear to be a blend of the encoding of the two feeder sites coming from the neural dynamics and/or the smoothing kernel in our analysis. We therefore suggest that the phenomenon is more accurately conceptualized as a shift rather than a replay of the true trajectory in the physical space.

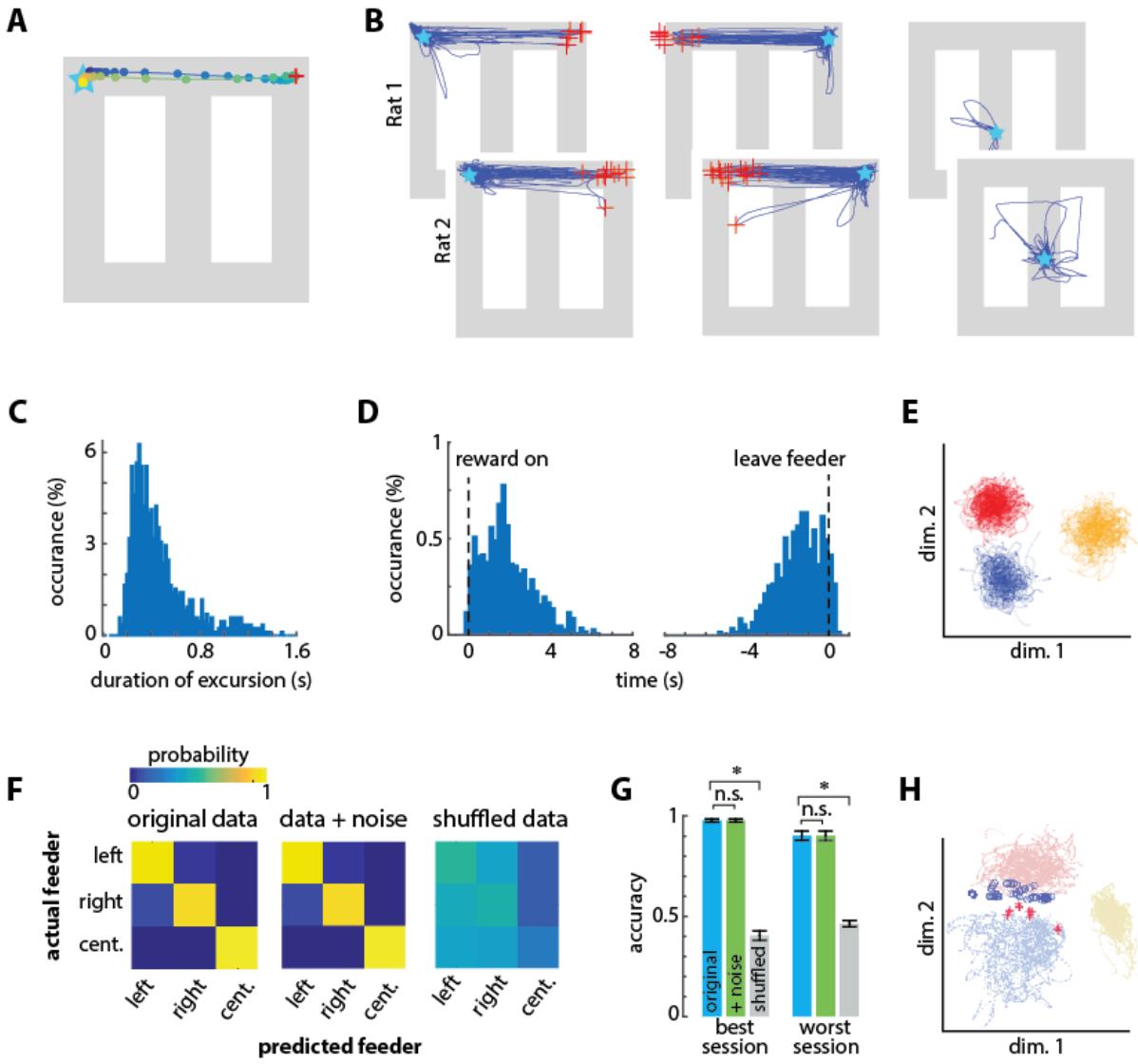


Figure 2.4: Excursions of encoded position sometimes shift to the alternate choice feeder

(A) Example of one excursion episode. The decoded position is indicated by circles plotted every 20ms. The excursion sweeps from the actual position of the rat (star) to the right-side feeder, and then returns. The red '+' indicates the maximal prediction error. **(B)** All excursions from the three feeders in the test set from one session in each rat. The red '+' indicate the maximum error distance from the occupied feeder of at least 70cm. The excursions from the right/left target feeders generate a trajectory to the alternate target feeder location. **(C)** Frequency distribution of excursion duration from all sessions. **(D)** Timing of excursion onset aligned to feeder activation or feeder zone exit, showing that excursions occur predominantly between these events**. (E)** Neural activity is distinct at the three feeder zones, as shown here by linear discriminant analysis of the smoothed neural data. Dots are the smoothed and binned neural patterns at feeders in the absence of excursions, and each color indicates data from one feeder (center feeder is yellow). **(F)** Confusion matrices for identifying the three feeders based on neural activity. Shifting each spike time by a random value (indicated as +noise) had little effect on the ability to correctly identify the feeders, whereas fully shuffling the inter spike intervals eliminated discriminability. **(G)** The prediction accuracy of classifying feeders for the best and worst sessions. Data in panels F and G include excursion events, which degrade performance. Asterisks indicate significantly different means with a $p<0.001$. **(H)**

PCA of activation in the middle layer of the spatial decoding network for non-excursion (dots) and excursion (+ and o) patterns at target feeders (blue, red) and center feeder (yellow).

We next sought further independent evidence as to whether the excursion events were an artifact of misclassification rather than a neurophysiological phenomenon. One possibility is that both target feeders are encoded by similar ensemble activity because of shared reward encoding, or some other feature, so that the network might confuse left and right feeders such that the decoded position might appear to jump from one to the other. To address this concern, we first used linear discriminant analysis to test if the neural patterns at the feeders are distinct from one another. Linear discriminant analysis of the input activity patterns in the absence of excursion did form distinct clusters for each feeder, indicative of unique pattern features at each feeder (Figure 2.4.E). This method is independent of our decoder, and therefore provides graphical validation. In order to quantify the pattern separation, we trained a new network with 3 hidden layers to classify patterns of neural activity from each of the three feeders, and then tested the classification accuracy (via cross-validation) on either samples of the original data, original data corrupted by noise, or fully shuffled data (Figure 2.4.F-G). These tests include all trials, both with and without excursion events. If the excursion events were due to a classification error arising from similarity of neural patterns at the feeders, then adding noise should decrease accuracy. The noise did not significantly reduce feeder decoding accuracy ($t(6) = 1.02$; $p= 0.35$, power = 0.87), but fully shuffling the spike time intervals did ($t(6) = 32.6$; $p= 3e\text{-}8$, power = 1). These data indicate that the activity patterns at the feeders are sufficiently distinct such that the introduction of noise does not cause misclassification, suggesting that the excursions are not a consequence of small random variation of the inputs.

The analyses above do not rule out the possibility that the excursions arise from brief shifts from the unique activity features at each feeder to an activity state common to both feeders. For instance, the reward encoding neurons could strongly activate to overshadow the position information in some instances, and this could produce a pattern that emerges at both feeders, but is distinct from the normal encoding at the feeders, and thereby confuses the decoder. This should be apparent in the variance of patterns at target feeders represented by the decoder network. PCA of activation in layer 3 of the position decoder shows distinct clusters for the non-excursion patterns at the target feeders (Fig 3.H). Moreover, the excursion patterns do not overlap completely or form their own cluster, and instead tend to overlap with the unoccupied target feeder cluster. We next conducted an independent and quantitative test for a common state by computing the classification accuracy of untransformed ACC patterns (input to the network) among the four classes: feeder A during an excursion (A'); feeder B during an excursion (B'); feeder A not during excursion (A); and feeder B not during excursion (B). If the excursions are because of a transition to a common state from both feeders, then the excursion patterns should be highly discriminable from the non-excursion patterns at the same feeder (A' from A, and B' from B), but not discriminable from each other (A' from B'). We found strong evidence for the former, but not the latter. We used the area under the curve (AUC) of the receiver operator characteristic (ROC) to quantify discriminability of samples from pairs of these conditions. An AUC value close to 0.5 indicates that the patterns from two classes are indiscriminable, whereas an AUC value close to 1 indicates that patterns are highly discriminable by the classifier. AUC values between these limits indicate that features of the patterns are sometimes similar and sometimes dissimilar in at least some dimensions. The discrimination of excursions from non-excursions at the same feeder (A' from A, and B'

from B) was very high (AUC = 0.94). On the other hand, the patterns during the excursions from the two feeders (A' from B') were discriminable at a moderate level (mean AUC = 0.75). This latter analysis was limited in power because of the low number of excursions at the preferred feeder relative to the dimensionality of the patterns. Nonetheless, these data indicate that the untransformed input patterns during excursions can often be distinguished based on the position of the rat. The non-linear transform of the input by the middle layers of the decoding network apparently separates the excursions from one another, but not from the typical feeder patterns (Fig 3.H). In sum, the non-excursion patterns (A, B) are highly discriminable (e.g., Fig 3.E-G), as are the excursion patterns from the non-excursion patterns at the target feeder (A, A', and B, B'). The excursion patterns sometimes overlap with each other (A', B'), and with the pattern from the unselected feeder (A', B and B', A; Fig 3.B,H). It thus appears that some features of the encoding shift to be more similar to the unselected feeder during excursion events. Because reward and location were confounded in the experimental design, we cannot rule out the possibility that reward encoding contributes to the phenomena. The inability to fully discriminate the excursion patterns from one another could involve some feature of the reward, such as volume, which flips between the choice feeders during the session. We therefore next investigated if units encoding reward were activated during the excursions.

2.2.4 Excursions were more likely to occur at less-preferred feeders, and encoded reward & choice information

Rats developed a strong feeder preference during free-choice trials within each block because of the unequal effort-reward utilities. Within each block of 6 free-choice

trials, the preferred feeder was chosen an average of 5.3 ± 0.5 times. We used this choice bias as a measure of revealed preference among feeders and analyzed the 10 forced-choice trials that preceded it. The selective occurrence of the excursions at the right-side feeder site in these forced-choice trials was strongly anti-correlated with the revealed preference for this feeder on free-choice trials (Figure 2.5.A; $r^2 = 0.86$; $F(6) = 31.5$; $p=0.002$; power = 0.97). In other words, the excursion was much more likely to occur when the rat was forced to select the less-preferred feeder. The excursions also emerged more frequently in free-choice trials when rats chose the less-preferred option. This dependence suggests that excursions are related to disfavored outcomes, consistent with proposals that primate ACC is involved in regret or signalling other outcomes that could have occurred (Coricelli et al. 2005; Hayden, Pearson, and Platt 2009). This supposition predicts that excursions should contain information related to the value of choice options. We thus next sought to determine if cells encoding choice or reward become activated during excursions. We first used logistic regression with norm-1 regularization to determine the degree to which cells encoded reward or choice information at the feeder locations in the absence of excursions (see METHODS). We found that 30% and 36% of cells were strongly predictive of reward and choice, respectively. We next independently determined which cells in the population significantly increased firing during excursions. This analysis revealed that some cells predictive or reward and/or choice activated during the excursions (Figure 2.5.B).

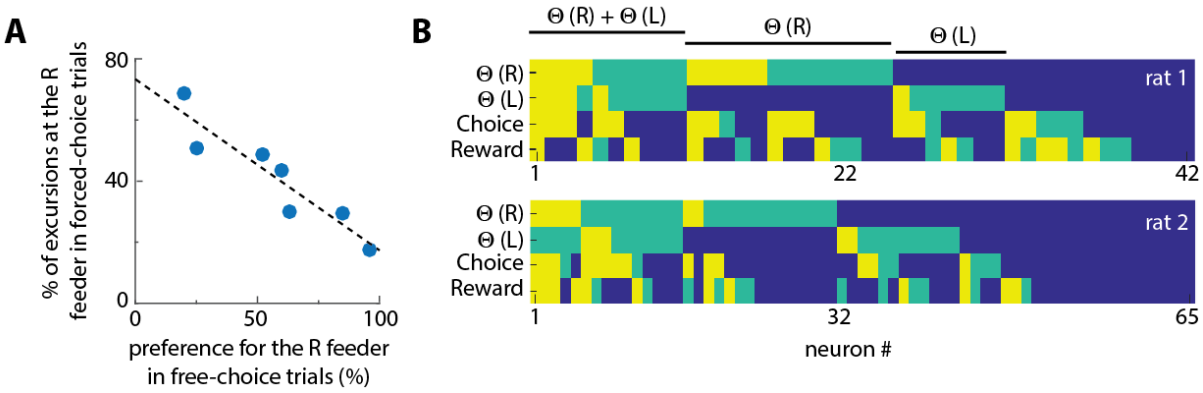


Figure 2.5: Excursions are more likely at non-preferred feeders, and encode choice-reward information

(A) The relationship between excursions frequency and feeder preference. Each dot is the session-averaged occurrence of excursion events at the right-hand (R) target feeder (as % of all events) during forced-selection trials, plotted against the revealed preference for that same feeder computed by the choice bias to the right-hand feeder in free selection trials (% of all choices). The negative correlation reveals that the excursion phenomenon is more likely to occur at disfavored feeders. **(B)** Encoding of information related to excursion, choice, and reward among neurons. The relative information of each neuron for discriminating excursions at the left target feeder (Θ(L)), excursions at the right target feeder (Θ(R)), the feeder choice, and the reward level. The level of information was determined by the frequency it was used by a Lasso classifier to discriminate this information, and is categorized as very informative (yellow), somewhat informative (green), or uninformative (blue). This analysis shows that some cells involved in excursions also encode information about choice and/or reward.

2.3 Discussion

We have shown here that the head position of a rat on a track can be decoded from the activity of several dozen ACC neurons with an accuracy of about 10 cm. This is a much finer spatial scale than the very broad spatial sensitivity of individual neurons in the ACC and nearby mPFC (Fujisawa et al. 2008). This raises the possibility that ACC ensembles may represent environmental features on a fine spatial resolution. Moreover, we found that the encoded position normally tracked the current state of the rat, but sometimes dissociated from its physical position at the target feeders. These excursions did not occur at the central feeder, were more likely to occur at the disfavoured target feeder and involved the activation of neurons encoding reward and choice. We suggest that this is functional evidence for the evaluation of choice outcomes that is more likely to occur following disappointing reinforcements. Moreover, the preponderance of evidence suggests that the ACC encoding during most excursions became more similar to the unselected target feeder than the selected one. If so, the data provide evidence that the ACC evaluates unrealized choice outcomes at locations remote from the animal's position.

These data are consistent with many previous findings and proposals. First, indirect evidence has previously suggested that ACC and adjacent areas in mPFC encode the position of the animal and objects on a fine spatial scale. For instance, small deviations in running path has been shown to explain a significant amount of variance in rat ACC activity (D. R. Euston and McNaughton 2006; Cowen, Davis, and Nitz 2012), and mPFC lesions impair object-in-place memory but not object memory (Barker et al. 2007). Second, ACC encodes a variety of task-related information such as reward, choice, and effort, which is thought to support choices among options with differing costs and benefits (M. M.

Botvinick, Cohen, and Carter 2004; C. Amiez, Joseph, and Procyk 2006; D. R. Euston and McNaughton 2006; Cowen, Davis, and Nitz 2012; Heilbronner and Hayden 2016). A specific function of individual ACC units in monkeys is the signalling of fictive outcomes, which are potential reinforcements that did not occur (Hayden, Pearson, and Platt 2009). This could be analogous to the activation of reward-encoding units during excursions, although the difference in species, spatial component, and explicit cues for the unattained location/reward limit the comparison to general features. The task used here also involves choice conflict between reward and effort. The ACC appears to play a role in resolving such conflicts (Hillman and Bilkey 2010), and could account for the high frequency of excursions in the present data. If the excursions support the comparison of realized and fictive outcomes more generally, this suggests that the broad post-reward activation of neurons in the ACC and nearby structures observed in several species (C. Amiez, Joseph, and Procyk 2006; Kennerley et al. 2006; A. J. Gruber et al. 2010) may involve similar excursions and the recall of information in multiple modes.

Our present data extend previous work demonstrating that the activity of mPFC is sufficient to decode spatial position on a track (Fujisawa et al. 2008). Whether the encoded information is a pure spatial signal or due to encoding of spatially-locked actions, stimuli, and/or events remains unclear. The task is repetitive and rats' movements tends to be stereotyped, so that specific actions (e.g., turns) and task events (e.g., approach to barrier) occur at the same location on every trial. Whatever its nature, our analysis reveals that the ensemble activity is sufficient to predict where an animal is with an unprecedented level of spatial resolution. It should be emphasized that this spatial information is multiplexed with many other task features encoded by this region, such as reward and effort, but can be extracted by an appropriately trained neural network decoder. This work

extends on the work of others who have shown that the medial prefrontal region encodes a trajectory through task space (Lapish et al. 2008; Durstewitz et al. 2010) and shows that, at least under some circumstances, such a state-space trajectory is isomorphic with real world spatial coordinates.

Why is spatial encoding prevalent in the ACC? We propose that the ACC may form a topographically-organized representational space, based on real space or action/events at particular locations on the maze, which can serve as a scaffold for the encoding of behaviourally-relevant events. For example, if the rat is attacked by a neighbor near its nest, the ACC may encode the event and trigger avoidance on subsequent visits to that vicinity. This is similar to a recent proposal that the orbitofrontal cortex uses a map of abstract task-states to facilitate reinforcement-based behavioural adaptation (R. C. Wilson et al. 2014), except that in our case, the representation has real-world spatial correlates. Our proposal is also closely related to past proposals that the medial PFC likely forms and stores schema which map context and events onto appropriate actions (Jung et al. 1998; Miller and Cohen 2001; Alexander and Brown 2011), which serves to engage appropriate emotional or motoric responses to a given set of events in light of past experience (Bechara and Damasio 2005; David R. Euston, Gruber, and McNaughton 2012). Again, the differentiating feature of our proposal is a higher spatial resolution. The spatial representation in mPFC often smoothly varies as an animal navigates the task space, but it can also drastically shift its response profile over the same task space in some circumstances, such as a switch of task rules (Rich and Shapiro 2009; Durstewitz et al. 2010; L. Ma et al. 2016). These investigators proposed that this provides a shift in context so as to facilitate learning or utilizing different sets of associations (e.g., action-outcome). It remains unclear whether these shifts are due to a global remapping of the entire ensemble

or only a subset of task-relevant cells. The decoding algorithms demonstrated here may be useful for determining if schemas (a.k.a. mental models or cognitive maps) retain associative information about space or other features across such shifts, or if ACC wipes the slate clean in some conditions.

The proposal of a mnemonic schema organized around position does not preclude its role in flexibly encoding other information to support decisions. Rather, it is a framework for the integration of information over several time scales, from consolidated memories to short-term 'working' memory, which is well supported by a large body of evidence in rodents and primates (David R. Euston, Gruber, and McNaughton 2012). The ACC thus uses information gleaned over both recent and remote experiences to form a model of the world organized around the spatial feature of the environment that also includes features useful for decisions that impact affective state, such as finding food and avoiding pain.

A novel aspect of the present data is the excursions from the present state at the choice feeders. This raises the possibility that the brain can mentally navigate the ACC map to recall information, or even generate hypothetical states consistent with the world model. Such prospection is consistent with the limited evidence available in other rodent PFC regions (Steiner and Redish 2014), and is consistent with evidence in primates, which we discuss later. In our study, these shifts may have been due to encoding of 1) the spatial location of the alternate feeder, 2) the expected reward at that location and/or 3) the sensory features (e.g. proximity to a ramp) of the alternate location. These factors were partially confounded in our study. The mPFC is well known to encode reward amount (Pratt and Mizumori 2001; Kargo, Szatmary, and Nitz 2007; Horst and Laubach 2013; Insel and Barnes 2015) and may plausibly encode sensory features, but our evidence weighs in favor

of a spatial shift. First, the excursion patterns were distinct from the patterns normally observed at the feeder, but overlapped with the patterns at the remote feeder. Second, excursions did not originate or terminate at the central feeder, even though the reward type and volume was comparable to that at the target feeders. Ultimately, whether the shifts are based on space, reward, or sensory features, our data still suggest that excursions involved a shift away from the present target feeder to encode features of the unselected target feeder, thus processing information related to choices and outcomes. The partial discrimination among excursion patterns may result from confounds of reward and locations, the similarity of sensory features of the two target feeders or could reflect processing of latent information (e.g. affective state). We also note that spontaneous reactivation of ensemble neural activity during replay events often differs in the number or timing of spikes as compared to the patterns during behaviour (Foster and Wilson 2006; David R. Euston, Tatsuno, and McNaughton 2007). The excursions in our data were relatively brief as compared to the time of feeder occupancy, raising the possibility of temporal compression as observed during replay, and likely introducing additional confounds for the classification analysis. We made no attempt to optimize the pre-processing of the input signal, such as the smoothing kernel width or normalization/convolution, which likely would have partially accounted for these effects.

We speculate that the brain dynamics involved in the excursions are not isolated to the ACC but are likely coordinated with that in other brain regions. The hippocampus sometimes also generates replay events after reward consumption (Foster and Wilson 2006; Carr, Jadhav, and Frank 2011). These occur during large amplitude fluctuations of the field potential called sharp wave ripples and occur in bouts lasting several hundred milliseconds. Task-related cells in the mPFC are modulated by these ripples following reward

consumption, suggesting that this is a period of communication between the hippocampus and mPFC (Jadhav et al. 2016). A similar post-reward replay in the sensory domain has been reported in the orbitofrontal cortex (Steiner and Redish 2014). Coordination of such events in ACC, orbitofrontal cortex, and hippocampus during pauses of directed behaviour following reinforcement would account nicely for the activity pattern and timing of the so-called default mode network (Buckner, Andrews-Hanna, and Schacter 2008). Activation of this network in humans occurs during pauses in directed action, typically after reinforcement, and is associated with 'mind wandering' often involving imagined shifts in time and place (Buckner and Carroll 2007). Analogous, and possibly homologous, default mode networks have been reported in non-human primates and rodents (Hayden, Smith, and Platt 2009; Mantini et al. 2011; Lu et al. 2012). The ACC, hippocampus, and other structures comprising the telencephalon emerged early in vertebrate evolution hundreds of millions of years ago, and likely supported a predatory foraging habit (Murray, Wise, and Graham 2017). The widespread conservation of the telencephalon among modern vertebrates suggests that it has functions useful in many situations and natural environments. It is therefore possible that a proposed human homologue (area 24) of rodent ACC (Uylings, Groenewegen, and Kolb 2003; Seamans, Lapish, and Durstewitz 2008) may also employ spatial associations to organize and navigate a schematic world model (Kaplan et al. 2017). It may, therefore, not be a coincidence that space-based imagery is one of the most prevalent top-down mnemonic strategies employed by humans, which has been used throughout recorded human history (O'Keefe and Nadel 1987). For instance, a person may imagine being in a particular restaurant in order to recall food options and quality, which is useful for making future dinner plans. The large expansion of granular prefrontal cortex, much of which connects extensively with ACC, likely endows primates with a greater

ability to abstract problems (Seamans, Lapish, and Durstewitz 2008; Murray, Wise, and Graham 2017), and possibly a greater ability to exert top-down control over ACC dynamics. Therefore, if primate ACC has a schematic world model organized similarly to that shown here, the dynamics of excursions and any shifted perception associated with them are likely different than those in rodents. In other words, the neurophysiology that leads to excursions might be similar in rats and humans, but we make no claims that the prospective representation of information via excursions in rats is perceived or controlled similarly to prospection in humans. Along the same lines, the strong correlation of excursions with disfavour of a feeder is consistent with the activation of human ACC in regretful situations (Coricelli et al. 2005), but we have no independent means to assess if rats perceive regret in the present data.

We propose that the excursion events represent navigation of a schematic world model organized around spatial position for some purpose related to task performance, such as comparing the utility of the obtained reward to an unattained one. The emergence of excursions exclusively at the target feeders, and not the center feeder, suggests a role in outcome comparison or future choice. Excursions did not terminate at the center feeder, suggesting they do not encode the subsequent action from the target feeder, which is always a return to the center feeder as enforced by gates on the track. It is possible that the excursions reflect an unexecuted plan to move from the occupied feeder to the other. If this were the case, however, we would expect to occasionally observe excursions when the rat is at the center feeder or other location on the track. A possible alternative is that the excursions reflect a mechanism for shifting strategies. The rodent ACC is involved in shifting responses (Joel, Weiner, and Feldon 1997; Birrell and Brown 2000), and appears to sustain information over time (Dalley, Cardinal, and Robbins 2004; Takehara-Nishiuchi

and McNaughton 2008). Although speculative, it is therefore possible that the excursions trigger a memory trace in ACC that promotes a response shift at the next visit to the choice point on the track. In other words, the ACC may have made a decision for the next choice while at the target feeder, which could preclude excursions at the center feeder or other intermediate point. The ACC is only one of several dissociated circuits that influence binary choice (Aaron J. Gruber and McDonald 2012), and is posited to bias competition among these other systems (Murray, Wise, and Graham 2017). Excursions may therefore have a probabilistic influence on future choice rather than fully determining it. The present data show only a correlation between revealed feeder preference and the likelihood of excursion. The design of the present task (e.g. forced alternation and relatively short blocks) prevents us from rigorously testing whether the excursion events influence future choice. We note that other evidence of spontaneous activation of task-related neural ensembles in the cortex and hippocampus has similarly shown correlation with general features of behaviour, such as learning, but most have not yet been shown to accurately predict future actions on a trial-by-trial (M. A. Wilson and McNaughton 1994; David R. Euston, Tatsuno, and McNaughton 2007; Fujisawa et al. 2008; Dragoi and Tonegawa 2013; Steiner and Redish 2014; but see Johnson and Redish 2007). We anticipate that advances in collecting and decoding ensemble neural activity will reveal such linkages between retrospective or prospective encoding and future behaviour.

2.4 Experimental Procedures

2.4.1 Behavioural apparatus

We constructed a running track 15 cm wide, with 36 cm high walls on both sides. It was configured into a 'figure-8' track measuring 102 cm long, 114 cm wide, and 60 cm height from the floor. Reward was delivered via three conical plastic feeders (24mm diameter). One was located on the central arm, and two others on 6 x 15 cm platforms at the north corners of the track. The reinforcement was a chocolate-flavored beverage (Ensure®, Abbott laboratories, Brockville, ON). The platforms were elevated 0-48 cm above the track. The ascent to the platform was by a vertical wire mesh (1.6 cm thick galvanized steel wire with a 1.25 cm square spacing). The descent was by a ramp made from the same material, but with a solid opaque plastic immediately under the mesh to provide support. The elevation of each platform was independently controlled by a stepper motor (Model 23Y9, Anaheim Automation, Anaheim, CA) driven by a stepper motor controller (Model G251X, Gecko drive, Tustin CA). A rack and pinion gear system was used to carry the platform up and down. A programmable digital input/output board (National Instruments PCIe-7841R, Toronto, ON) and custom software written in Microsoft Visual Basic and Labview (National Instruments, Toronto, ON) were used to automatically control and store the time of track events.

2.4.2 Data collection

We used Fisher-Brown Norway or Long-Evans rats born and raised on-site. Rats were habituated to handling for two weeks prior to the experiment. We surgically implanted a recording drive prior to any training. The drive and implantation were carried out as

described previously (D. R. Euston and McNaughton 2006). The position of the animal and neural signals were recorded simultaneously with a digital acquisition system (Cheetah SX, Neuralynx, Tucson, AZ). Neural signals were amplified with a unity gain headstage (HS-54, Neuralynx, Tucson, AZ), amplified with a gain of 1000, and band pass filtered between 600 and 6000Hz. Voltage waveforms exceeding a manually set threshold were recorded during behaviour, and then sorted into distinct clusters offline.

Animals began training on the figure-8 track following a one-week recovery from surgery. Behaviour was shaped by allowing rats to navigate the track for 7-10 days with no variation of reward volume or barrier height. All subsequent sessions followed a fixed reward/effort schedule in which the task was organized into 6 blocks of 16 trials. Gates on the track forced the rat to alternate between the left and right loops on the first 10 trials of the block. The rat was free to choose either side for the remaining 6 trials of the block. The barrier height (0-46 cm) and/or the reward volume (30 or 120 µL of chocolate beverage) at one or both of the feeders changed across each block. The block order was: [R_S,E_0, R_B,E_0], [R_S,E_0, R_B,E_1], [R_S,E_0, R_B,E_2], [R_S,E_2, R_B,E_2], [R_B,E_0, R_S,E_0], [R_B,E_2, R_S,E_0], where R_i (i codes Small, or Big) is the reward, and E_i (i codes height level) is effort. The block sequence is repeated until the animal stops performing trials. The side with initially large reward is counterbalanced over consecutive sessions. Animals were reinforced at the central feeder on every lap with the same chocolate beverage and small reward volume as for the choice feeders.

2.4.3 Decoding ensemble activity

We first removed individual trials with durations longer than 1.5 times the median trial length in each session so as to reduce neural correlates of non-task behaviour (e.g.

grooming, rearing) in the data. We further eliminated neurons with an average firing rate below 0.5Hz because our initial tests indicated that these cells did not improve decoding accuracy. We then used a Gaussian kernel with a standard deviation of 150 ms to smooth the spike data, and then binned the resultant signal by 50 ms. The position of the animal assigned to each bin was the average of the coordinates of the video tracker system for the corresponding time window. We next applied the square root transformation to the binned firing rates to normalize the activity distribution (e.g. make them more Gaussian) among neurons. We then used the z-transform so that the activity of each neuron would have zero mean and unit variance. For decoding using the Bayesian method, we discretized space into regions of 4 cm by 3 cm. The dANN method operated at the pixel resolution of the video tracker (0.27 x 0.24 cm). To ensure that the difference in spatial resolution among decoding methods did not present confounds, we applied the dANN method to the data discretized in the same way as for the Bayesian method. This did not affect the prediction error, so we only show the results from the finer resolution in the figures.

The position decoder was a multi-layer feedforward artificial neural network with three hidden layers. The number of units in the input layer equals the number of recorded neurons in each session. The number of units in the first, second, and third hidden layer is 100, 50, and 25, respectively. The output layer consists of two units, which represent the coordinates of the position on the track. The activation function of the first two hidden layers is rectified linear, and the activation function of the third layer is hyperbolic tangent. The number of units and the activation function for each layer were selected via cross validation. The quality of predictions was quantified with the mean squared error (MSE) between the actual position of the rat and the predicted location:

$$MSE = \frac{1}{2N}\sum_{i=1}^{N}\left\|\begin{bmatrix}\hat{x}_i\\ \hat{y}_i\end{bmatrix} - \begin{bmatrix}x_i\\ y_i\end{bmatrix}\right\|^2$$

where $\hat{x}_i$ and $\hat{y}_i$ are the predicted coordinates for the ith test sample, and x_i and y_i are the target coordinates. $\|\quad\|$ represents L^2-norm, and N is the number of samples in the test set. For the Bayesian decoder, we computed the MSE between the center of the target region and the predicted region.

We used mini-batch gradient descent with weight decay and momentum to train the network. The batch size was 100 and the learning was stopped after 100 epochs. For each session, we created 10 sets of trials, each consisting of an equal number of left and right choice trials. We used 75% of the trials in each set to train the network and the remaining 25% to evaluate the model. The reported prediction error for each session was the average of the errors computed for each of the 10 test sets.

2.4.4 Downsampling and greedy selection

To compute the reconstruction error versus the number of cells in the data, we created different datasets, each containing a subset of cells. The number of cells in these sets ranged from 4 to 40 by increments of 4. For each set size, we created 5 different datasets by sampling cells randomly. These steps resulted in 50 datasets, consisting of 10 different set sizes and 5 different cell sets for each size. The process for evaluating each set is the same as that described above. In order to determine the best set of cells for the decoding, we used the forward feature selection algorithm to choose the best set of 20 cells for decoding position. In this approach, the best set of cells is initially empty. We iterated over all the cells to find the cell that will result in the lowest error if is added to the current best set of cells. This cell is then added to the current best set. These steps are repeated until the

best set reaches the desired cardinality. In order to determine the error for each candidate set, we used the same procedure described in the previous section (average of 10 sets of randomly selected trials).

2.4.5 Error map construction

For each session, we created 10 different datasets, consisting of an equal number of right and left choice trials. We used 75% of the trials in each dataset as the training set and the remaining 25% as the test set. After training the network on each training set, we decoded the location of the rat for the corresponding test set. In the next step, we discretized the space into tiles of 4 by 3 cm, and for each test trial computed the maximum reconstruction error when the rat was in a particular tile. The error map for each session was constructed by taking the average of the maximum reconstruction error values over trials. The final map is constructed by averaging error maps of all sessions.

2.4.6 Determining the reward and choice encoding cells

To determine which cells encoded information about reward or choice, we computed the average firing rate of all cells per trial in a 1.5s window beginning immediately after reward delivery at the north feeders, so as to create an N by T matrix, where N is the number of cells and T is the number of trials. The target associated with each column of this matrix was a binary value that represented different conditions of the parameter of interest (right/left for choice, high/low for reward). For each parameter of interest, we created 50 random sets of trials, each one containing an equal number of trials with different conditions. We used 75% of the trials in each set as the training data and the remaining 25% as the test data. We fit a logistic regression model with norm-1

regularization (Lasso) and a maximum degree of freedom of 20 to each training set. Models with different degrees of freedom were trained using 5-fold cross-validation. For each set, we picked the model that resulted in the best prediction accuracy of the target based on the neural vectors, and then logged the cells that were assigned non-zero weights (e.g. they provide useful information for discrimination). To ensure that the selected cells were meaningful, we evaluated the accuracy of the model on the test data. In our experiments, the accuracy of the final model was always above 90%. Finally, we computed the percentage of times (out of the 50 sets) each cell was assigned a non-zero weight. For a value less than or equal to 30%, the cell was classified as not informative. For values greater than 30% and less than 70%, the cell was classified as relatively informative. For values greater than 70%, the cell was classified as very informative.

2.4.7 Excursion detection

We increased the temporal resolution by reducing the smoothing kernel width to 120 ms and binning the resultant signal by 20 ms. All other features of the decoding network were as described above. For each session, we created 10 sets of trials by selecting all the trials performed on the less preferred side of the maze and randomly selecting an equal number of trials from the opposite side of the maze. The number of selected free choice trials for both sides were equal. We divided the data in each set into two groups. First, we randomly selected half of the trials from the set. The first group contained data points that fell into a time window of 1.5s after reward delivery in the selected trials. All the other samples from the original set were assigned to the second group. We used the samples in the second group to train the decoder network, and then used the network to predict the location of the rat from the samples in the first group. We marked each test trial

as an excursion trial if the maximum error between the decoded position and the actual location of the rat was greater than 70 cm during that trial. The statistics we report in this paper were computed by taking the average of the results we obtained on each set of test trials.

2.4.8 Identifying excursion-related cells

We divided the samples in each test set created for excursion detection into two categories. The first category contained all excursion samples that fell into a window of 80 ms, centered at time at which the distance between the actual location of the rat and the predicted location was at its peak. The second category contained all the other samples. We computed the average firing rate of the cells per category and applied a square root transformation to convert the activity distribution into a normal distribution. We then used a t-test with a significance level of 0.01 to detect the cells that had significantly different firing rates under the two conditions. We computed the percentage of time each cell was selected among cells with significantly different firing rate and used the same procedure described above to classify cells into three levels of importance.

2.4.9 Adding noise, shuffling data, and constructing confusion matrices

Confusion matrices show the probabilities of assigning patterns to correct and incorrect categories. We preprocessed neural data as for the detection of excursions. The spike data was smoothed for each session using a Gaussian kernel of width 120 ms, and then binned by 20 ms. We selected two sets of data from each trial of the task: the set containing all the samples that fell into a window of 750 ms after the off-set of the center feeder; and the set of all data points that fell into a window of 1500 ms after the off-set of

the selected side feeder. Each sample was labeled based on the feeder associated with it (right/left/center).

To create a noisy dataset, we concatenated the sets associated with each feeder and computed the standard deviation of the inter-spike intervals of each neuron for each feeder. Then, we shifted spike times in each feeder set by values that were randomly drawn from Gaussian distribution with zero mean and a variance equal to 25% of the cell's firing interval variance in that set.

To create the shuffled dataset, we concatenated all the selected time windows for all the feeders and computed the inter-spike intervals for each neuron. Then we created a new spike time-series by randomly permuting the inter-spike intervals. The label of each time-point in the shuffled dataset was the same as the label of that time point in the original dataset.

To obtain a confusion matrix for each dataset, we performed the following procedure: we randomly selected an equal number of left and right trials and used the samples from the selected trials in the original data set to train a neural network classifier. The architecture of the classifier was identical to the network that was used for decoding position, except for the output layer, which had 3 units (to represent the three feeders) and a softmax activation function. For each dataset (original/ noisy/ shuffled) we used the samples from the remaining trials as the test set and computed the normalized confusion matrix. We repeated this process 10 times for each session and computed the average classification accuracy of these repetitions. The reported results represent the average over all sessions.

2.4.10 Discrimination analysis of excursion patterns

From each dataset created for excursion detection, we created a new dataset by performing the following steps. For the trials without excursion, we computed the mean firing rate of each neuron during a window of 0-1.5 s after the feeder was closed. For the trials with excursion, we computed the average firing rate of the neurons in a window of 100 ms, centered at the time at which the excursion was at maximum departure from the actual location of the rat. Then, we assigned a label to each one of these vectors: normal pattern at feeder A (A), excursion pattern at feeder A (A'), normal pattern at feeder B (B), excursion pattern at feeder B (B'). To differentiate between normal and excursion patterns at each feeder (A vs A' and B vs B'), we used a decision tree with a maximum number of splits of 10. To separate the excursion patterns at different feeders, we used an SVM classifier with a linear kernel. Because the number of samples was unequal among the classes, we used the area under the curve (AUC) of the receiver operator characteristic (ROC) as a measure of pattern separability.

3 Rat Anterior Cingulate Cortex Represents Perceptible and Abstract Task Variables Essential for Creating a Predictive Model

Abstract

The anterior cingulate cortex (ACC) has been proposed to use information from past experience to create schemas or mental models that can be used to estimate the cost and benefit of possible responses in a task. This information is useful for guiding choices in order to collect necessary resources according to need, risk, effort, and other factors. To date, little direct evidence of the encoding or use of schemas by neural ensembles has been reported. Here, I present novel evidence that ACC neurons encode several abstract task features of the task state, some of which are determined by un-cued changes in reward probability. Rats performed a binary choice task with probabilistic rewards. The reward probability was higher at one feeder for a block of trials, after which the other feeder became the high-probability feeder for the subsequent block. The ACC distinguished blocks when the animal was stationary at the beginning of each trial. It also recalled past reward. Once the animal began the trial, information about block and reward diminished, and information about the future choice emerged. I show that encoding of block is independent of previous reward and choice. Interestingly, block encoding was not different on trials when animals chose the lower probability option. Together, these data indicate that ACC encodes abstract features of the task pertinent to decisions, but that other brain regions appear to select actions on most trials. This is consistent with proposals that the ACC plays a supervisory role useful for rapidly shifting strategies when task contingencies change.

3.1 Introduction

The Anterior Cingulate Cortex (ACC) is thought to use information from past experience to create predictive models that map context and events onto appropriate responses, thereby improving the cost-benefit of outcomes (M. M. Botvinick, Cohen, and Carter 2004; C. Amiez, Joseph, and Procyk 2006; Kennerley et al. 2006; Matthew F.S. Rushworth et al. 2011; Heilbronner and Hayden 2016; David R. Euston, Gruber, and McNaughton 2012; Shenhav, Cohen, and Botvinick 2016). To accomplish this, the ACC tracks many task variables relevant to decision making. These variables include positive and negative outcomes (Gehring and Willoughby 2002; Hayden and Platt 2010; C. Amiez, Joseph, and Procyk 2006; Kennerley and Wallis 2009a; Mashhoori et al. 2018; Hashemnia, Euston, and Gruber 2020), the value of unchosen options (Blanchard and Hayden 2014), fictive reward (Hayden, Pearson, and Platt 2009), and the average value of options available in the environment (Kolling et al. 2012). The ACC is also shown to be sensitive to the cost of options, including physical effort (Croxson et al. 2009; Klein-Flugge et al. 2016; Prevost et al. 2010) and mental effort (Engström, Landtblom, and Karlsson 2013; Hosking, Cocker, and Winstanley 2014). The ACC integrates cost and benefit information to estimate option utility, and biases decisions toward the optimal choice (Hillman and Bilkey 2010; Shenhav, Botvinick, and Cohen 2013). Some studies have reported direct encoding of the utility by the ACC neurons in rodents and primates (Kennerley and Walton 2011; Wallis and Kennerley 2011; Hashemnia 2019).

In addition to ACC's role in estimating option value and creating associations between actions and outcomes (S. Ito et al. 2003; K. Matsumoto, Suzuki, and Tanaka 2003; Williams et al. 2004; C. H. Luk and Wallis 2009; Hayden and Platt 2010), several studies

have implicated the ACC in triggering adjustments in action plans or strategies when they are no longer beneficial. In this view, ACC activity is linked to the receipt of information that leads to changes in a subject's internal model of a task (Hayden et al. 2011; Quilodran, Rothé, and Procyk 2008). This claim is further supported by ACC's signaling of reward prediction errors (Kennerley, Behrens, and Wallis 2011; M. Matsumoto et al. 2007; Seo and Lee 2007) and surprise (Cavanagh et al. 2012; Wessel et al. 2012; Nee, Kastner, and Brown 2011). A related apparent role of the ACC is to supress a default response when shift in strategy is likely to yield better returns (Paus et al. 1993; Shenhav, Botvinick, and Cohen 2013).

In addition to neural recordings, the role of the ACC in guiding decisions is also informed by lesion studies. ACC lesions or inactivation impair animals' ability to select the appropriate response in a reward-based selection task (Keisetsu Shima and Tanji 1998; Hadland et al. 2003), evaluate the costs and benefits of actions (Amemori and Graybiel 2012; Friedman et al. 2015), change strategy as the value of options change (Kennerley et al. 2006; Kawai et al. 2015), or explore alternative choices (Tervo et al. 2014).

Although evidence supporting the ACC's involvement in cognitive and executive control is plentiful, the specific purpose it subserves and a mechanistic explanation of its function are still under debate. I advocate here that the ACC creates and maintains models related to multiple strategies that the brain can use to solve problems. These strategies compete with each other and compete with other brain circuits utilizing distinct strategies to influence choice.

3.2 Results

3.2.1 Free and forced-choice trials are distinguishable in the middle arm of the maze

The ACC is proposed to help optimize choice by shifting responses more quickly than is possible by reinforcement-driven systems when environmental contexts change. If so, it stands to reason that the ACC will play a larger role in guiding behaviour when the animal is free to choose among several options, as compared to conditions in which only one course of action is possible. We evaluated this by testing whether ACC activity was different in free-choice trials, in which the rats were able to choose either of two target feeders, and forced-choice trials, in which animals could only access one feeder. We did this by computing the accuracy by which we could determine the trial type based on ACC activity. Specifically, we computed classification accuracy of free/forced trials in different regions (spatial bins) of the maze. The highest accuracies were obtained in the bins in the middle arm of the maze (Figure 3.1), which are the regions before and approaching the point at which the animal must turn. The highest obtained accuracy was 75% (std = 5.75; $t_4 = 9.69$; $p < 0.001$ vs null hypothesis of no encoding; $power_{0.01} = 1.00$) in the second bin of the middle arm.

Because free and forced trials were presented randomly, the high accuracy of trial type detection prior to the turn point is likely due to unplanned auditory and visual information. Forced-choice trials were implemented by the closing of a motorized gate so as to block access to one of the gates. This typically occurred while rats were exiting the central feeder, and the motor actuation was clearly audible. Moreover, gates in the closed state were visible as rats approached the ‘T’ junction. Therefore, visual observation of a gate was indicative of a forced-choice trial.

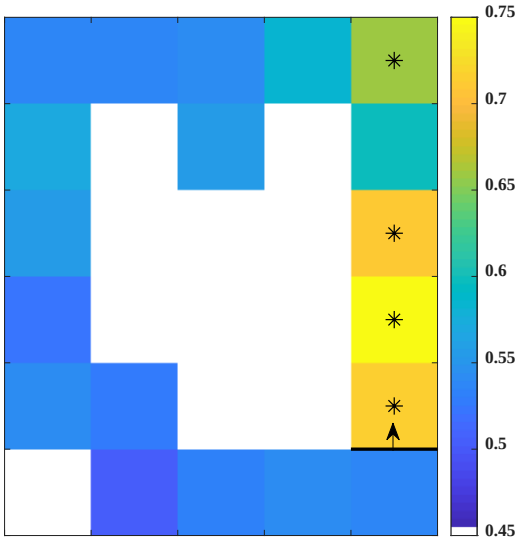


Figure 3.1. Accuracy of decoding trial type (free vs forced-choice). '*' represents $p < 0.01$. The accuracies are highest in the middle arm of the maze, which correspond to the time/position at which the animal has auditory and visual information about the barrier state.

Furthermore, the average classification accuracies in other regions of the maze are close to chance. The cues indicative of trial type were less pronounced in these other regions. Therefore, it is unclear whether the prominence of free/forced encoding in ACC is higher on approach to the turn point because it is useful for guiding behaviour, or because this is the region in which the relevant sensory input is greatest.

3.2.2 Block type influences animals' choices

The reward probabilities among the target feeders were always unequal, but remained constant within blocks of 40 or 50 consecutive trials, before reversing. The high utility feeder delivered reward on 80% of trials, whereas the other feeder paid out on 20% of trials. The high utility feeder alternated several times between the right and the left

feeders during each session. Knowledge about this block structure is useful for maximizing reward because animals can exploit the temporal correlations to facilitate choice shifts shortly after block switches and prevent shifts at inappropriate times.

The choices of the animals indicated that they typically detected block changes after the first 10 trials in a block and changed their choices accordingly (Figure 3.2). The ratio of optimal to sub-optimal choices are near or below chance level during the first 10 trials but increase sharply afterward. We therefore considered the first 10 trials in each block as adaptation trials.

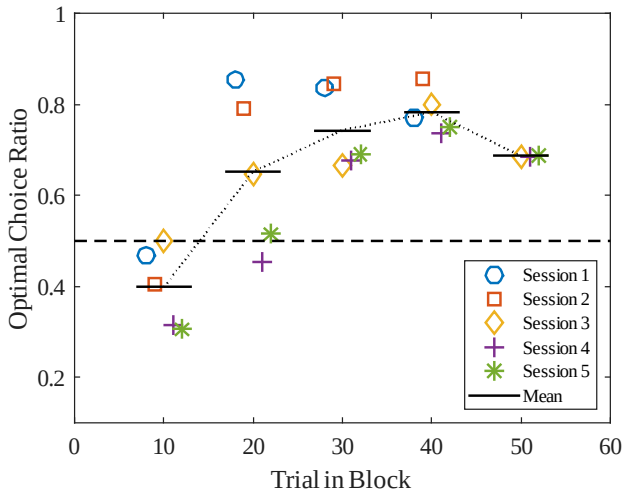


Figure 3.2. Choice adaptation within blocks. Ratios of optimal choices to all choices made in consecutive groups of 10 trials, averaged over blocks in each session. The performance during the first 10 trials were close or below the chance level. After 10 trials, however, the rats detected the change in high utility feeder location and changed their choice preference accordingly.

We note that both animals had a bias toward the right-side feeder. The bias was stronger in rat #2, which tended to go to the right-side feeder even when the optimal choice was the feeder on the left. This bias explains why the ratios of optimal choices in the second trial group in sessions 4 and 5 (rat #2 sessions) were lower than the other three sessions (rat #1 sessions).

We next sought to identify the decision-making strategy used by each subject to generate the pattern of choices we observed in free trials. We considered three strategies: choosing the high-utility feeder in each block (strategy name: block); choosing a preferred side independent of reward or block (strategy name: preference), and choosing the same side as the previous trial if the animal was rewarded, and switching to the other side if no reward was delivered (strategy name: win stay, lose shift). We excluded the first 10 trials (adaptation trials) in each block for this analysis.

The block strategy was the best predictor of choice in 4 of the 5 sessions (Figure 3.3). In session 5, due to the animal's strong bias for one feeder, the preference strategy was the best predictor of choice.

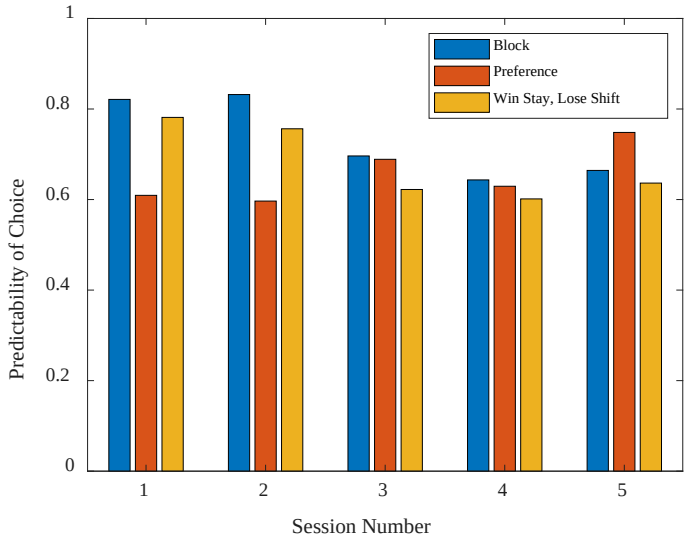


Figure 3.3. Predictability of choice by three different strategies. 1) choosing the optimal feeder in each block (block strategy), 2) choosing the preferred feeder each time (preference strategy), and 3) win stay, lose shift strategy. In 4 out of 5 sessions, the block strategy is the best predictor of the behaviour.

3.2.3 Block type is decodable at the center feeder

The analysis of behaviour indicated that the predominant choice strategy was to select the high-utility feeder in each block. This suggests that rats do utilize information about blocks. We thus next sought to determine if ACC neurons encoded information about block. Because block identity is strongly correlated with reward probabilities and choice (consequently location) in well-performing rats, we restricted the analysis to the middle arm of the maze, which is prior to the decision execution and reward.

For this analysis, we excluded the forced trials and the first 10 trials in each block (i.e., the adaptation trials). The highest mean accuracy was 64.2% (std = 5.76; $t_4 = 5.51$; $p = 0.002$ vs chance; $power_{0.01} = 0.886$), achieved in the bin associated with the center feeder. The accuracy was not significantly higher than chance in other spatial bins in the middle arm (Figure 3.4).

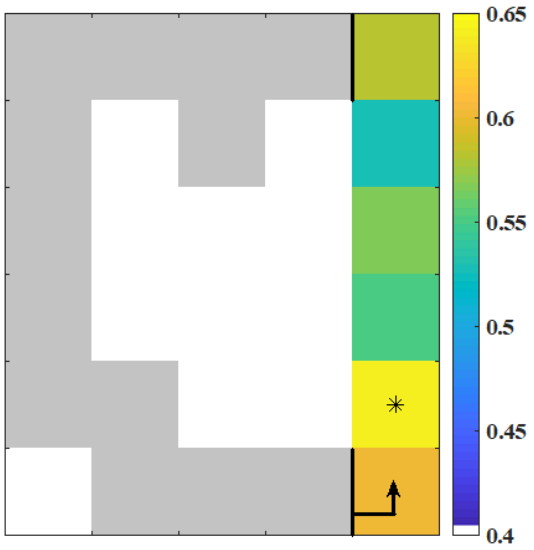


Figure 3.4. Accuracy of decoding block type in spatially-binned regions of the middle arm. '*' indicates $p < 0.01$. The highest accuracy was achieved in the bin associated with the center feeder (64.2%, $p = 0.002$)

The results suggest that neurons in the ACC were most sensitive to block type at the start of each trial when the rats were stationary at the center feeder. However, ACC activity did not represent the block type once the animal moved away from the center feeder.

Next, we computed the block decoding accuracy at the center feeder for the adaptation trials (10 first trials of each block). The average accuracy was 47.59 and not significantly different from chance (std = 7.04; $t_4 = 0.766$; $p = 0.48$). The result confirms that during adaptation trials, a clear block signal was not present.

3.2.4 The choice signal is weak in the middle arm but strong in the side arms

Next, we investigated if ACC encoded the animal's choice in specific regions of the maze. We computed the accuracy of choice decoding in two stages. The first stage, which we refer to as the prospective stage, covered the period from the beginning of a trial until arrival at the one of the side feeders. The second stage, referred to as the retrospective stage, covered the period from departure at the selected target feeder until arrival to the 'T' junction of the maze in the next trial. For this analysis, we excluded the first 10 trials in each block.

In the prospective stage, the accuracies in the bins located in the middle arm of the maze were low (Figure 3.5.A). The lowest accuracy was 54.8% (std = 9.41; $t_4 = 1.13$; $p = 0.16$; $power_{0.01} = 0.06$), obtained in the center feeder bin. The accuracy increased as the animal moved toward the 'T' junction, and it approached 100% as they transitioned into the side portion of the maze.

This increase in the accuracy of choice decoding as the animals move to the sides of the maze is predicted by previous studies showing that ACC activity is highly sensitive

to the chosen path (Cowen, Davis, and Nitz 2012; D. R. Euston and McNaughton 2006) and the position of the animal in similar tasks (Mashhoori et al. 2018). In the present task, the running path of the rat diverges as it approaches the T junction for trials to the right as compared to left; therefore, the accuracy for choice classification improves. This suggests that in the middle arm of the maze, the ACC is encoding position rather than choice.

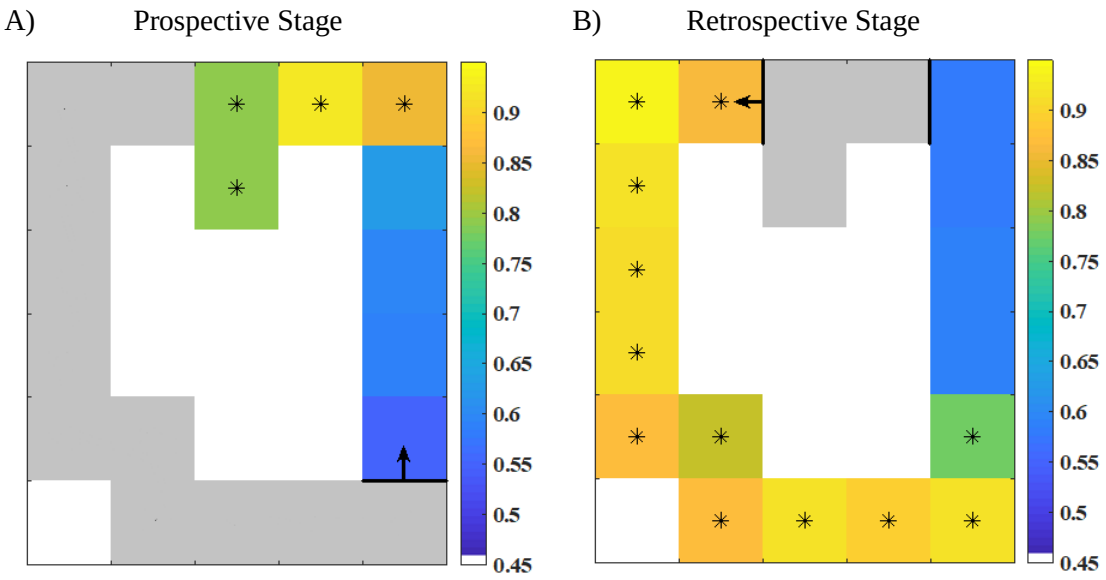


Figure 3.5. Accuracy of decoding choice. (A) Accuracy of decoding choice in the prospective stage. '*' represents $p < 0.01$. The accuracies are low in the middle arm of the maze. The lowest accuracy is at the center feeder bin (54.8%, $p = 0.16$). The accuracy increases as the animals move toward on side of the maze, and it approaches 100% as they enter one of the side arms. **(B)** Accuracy of decoding choice in the retrospective stage. The accuracy is high in all regions of the side and bottom parts of the maze, and it stays high when the rat is at the center feeder, in the beginning of the next trial.

The low accuracy for choice detection at the center feeder contrasts with the relatively high accuracy for block type detection at the same location. This observation indicates that the block signal is not merely a correlate of the choice signal in the ACC. The relationship between the choice and block type signals at the center feeder will be investigated in more detail in the next section.

In the retrospective stage, the accuracy remained high in all spatial bins in the side and bottom parts of the maze. The accuracy for decoding previous choice was 77.4% (std = 13.09; $t_4 = 4.69$; $p = 0.004$; $power_{0.01} = 0.768$) at the beginning of the next trial, when the rat was at the center feeder (Figure 3.5.B). This high accuracy can be attributed to the difference in the animal's body orientation based on its returning path. The retrospective choice decoding accuracy dropped to chance as the rat moved away from the center feeder.

3.2.5 Encoding of block and choice are temporally locked to the departure of the start feeder

We next sought to determine if the signaling of block and choice were temporally locked to particular phases of the task. We therefore used a sliding window method, time-locked to arrival at and departure from the center feeder, to partition data into bins of *time*. We computed the accuracy of decoding the two variables in each time bin. The width of the window and the step size were set to 250 ms and 50 ms, respectively. We first analyzed encoding with respect to arrival at the center feeder. The average velocity profile indicates that they have largely ceased running upon arrival to the center feeder (Figure 3.6.A). Time t = 0 represents the reach time to the center feeder. As shown, before reaching the center feeder, the velocity was high as the rats were moving towards the feeder. After the reach time, the velocity approached zero, and the animals were stationary for a period of time. Note that the animals' heads typically move during reward consumption, leading to velocity greater than zero even though they are not moving along the main axis of the track.

Figure 3.6.B and 3.6.C show the accuracy of the block and choice detection in each temporal bin, respectively. For block type detection, there is a cluster of high accuracy

points with an average accuracy of 61.1% (std = 3.58; t_4= 6.90; p = 0.001; $power_{0.01}$= 0.98) after the reach time to the center feeder, when the animals were stationary. During the same period, the average accuracy for choice detection is 54.4% (std = 3.88; t_4 = 2.50; p = 0.033; $power_{0.01}$ = 0.27), which is close to the chance level. After 0.5 s, the accuracy for block detection decreases.

Figure 3.6.D illustrates the average accuracies for block type superimposed with choice decoding, in order to facilitate comparison among them. The accuracy for block detection was higher than the choice detection for almost all temporal bins. After arrival at the center feeder, there are two point clusters where the differences are statically significant ($p < 0.05$). This suggests that as the animals reached the center feeder and became stationary, the strength of the block type signal in the ACC increased and became stronger than the choice signal.

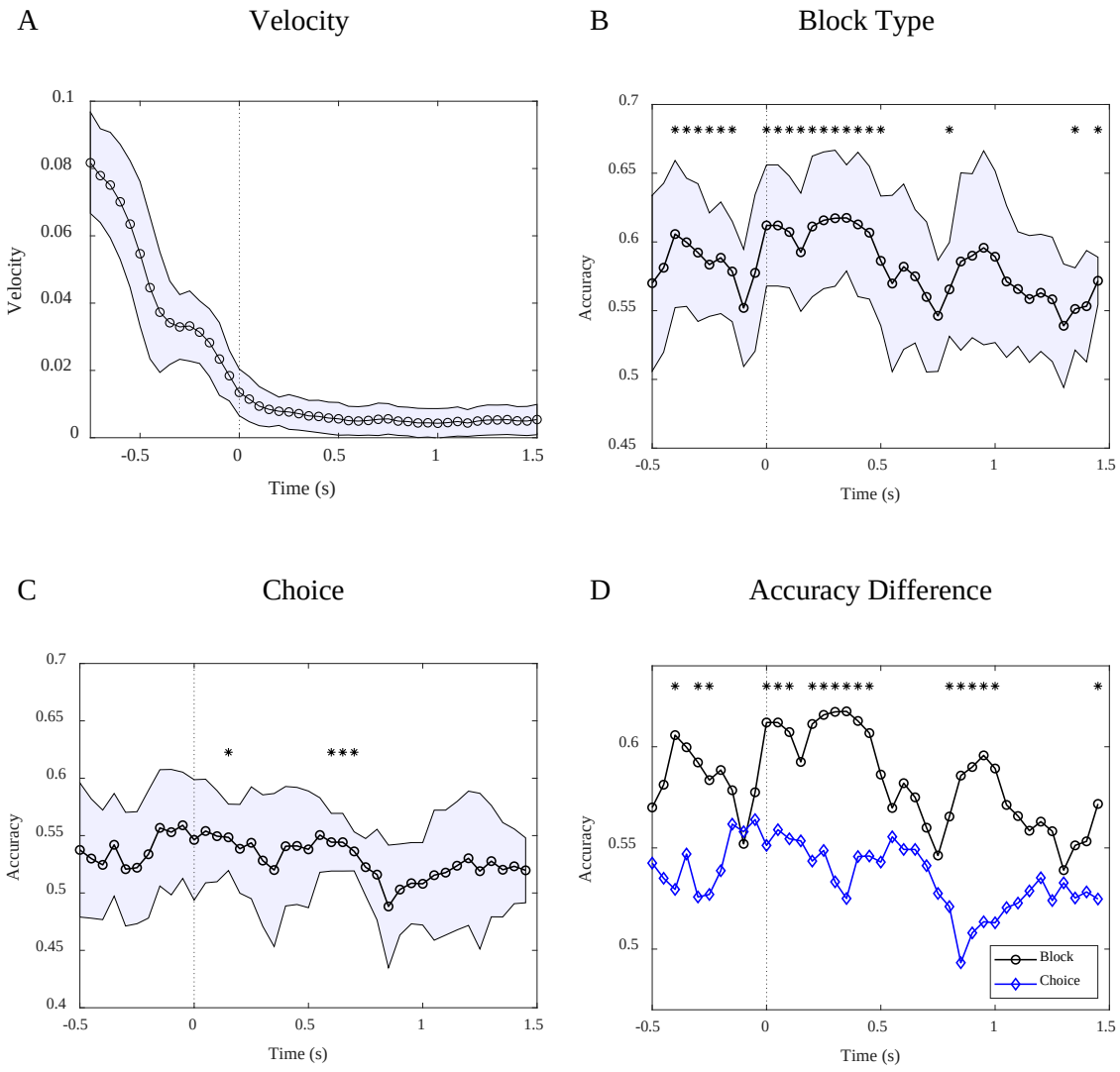


Figure 3.6. Dynamic changes in encoding of block and choice as the rat arrives at the center feeder. (A) Mean and standard deviation of head velocity as the rat reaches the center feeder (t=0). **(B)** Decoding of block type in temporal bins centered on arrival to the center feeder. Asterisks '*' indicate significant difference from chance levels at p=0.01. **(C)** Decoding of choice. Same format as in B. **(D)** Superposition of mean decoding accuracy of choice and block. Asterisks '*' indicates statistically significant ($p < 0.05$) difference of decoding of these variables. There are two point clusters during which the accuracy for block type detection is significantly higher than the choice detection.

We next repeated these analyses with respect to the departure from the center feeder. Time t = 0 represents the movement initiation time of departure from the center feeder (Figure 3.7.A). Figure 3.7.B and 3.7.C show the accuracy of block and choice detection in each temporal bin, respectively. For block type detection, there is a cluster of high accuracy points with an average accuracy of 59.5% (std = 2.42; $t_4 = 8.75$; $p < 0.001$; $power_{0.01} = 0.99$) shortly before the animal starts running. During this period, the animal was stationary. During the same period, the average accuracy for choice detection is 52.6%, which is not significantly above the chance level (std = 6.44; $t_4 = 0.91$; $p = 0.207$; $power_{0.01} = 0.05$). After the initiation of movement, the accuracy for block detection *decreases* while the accuracy for choice detection *increases*.

Figure 3.7.D illustrates the average accuracies for block type and choice detection superimposed with one another. The accuracy for block detection is significantly higher than the choice detection before movement initiation ($p < 0.05$). The increase in the strength of the block type signal right before the animals started running, and the weakness of the choice signal in the same period, raises the possibility that the ACC contains a signal that guides decision-making, but the ACC is not responsible for the choice itself. The increase in the accuracy of choice detection after the initiation of movement suggests that the representation of choice in the ACC follows action, and ACC activity is not predictive of choice.

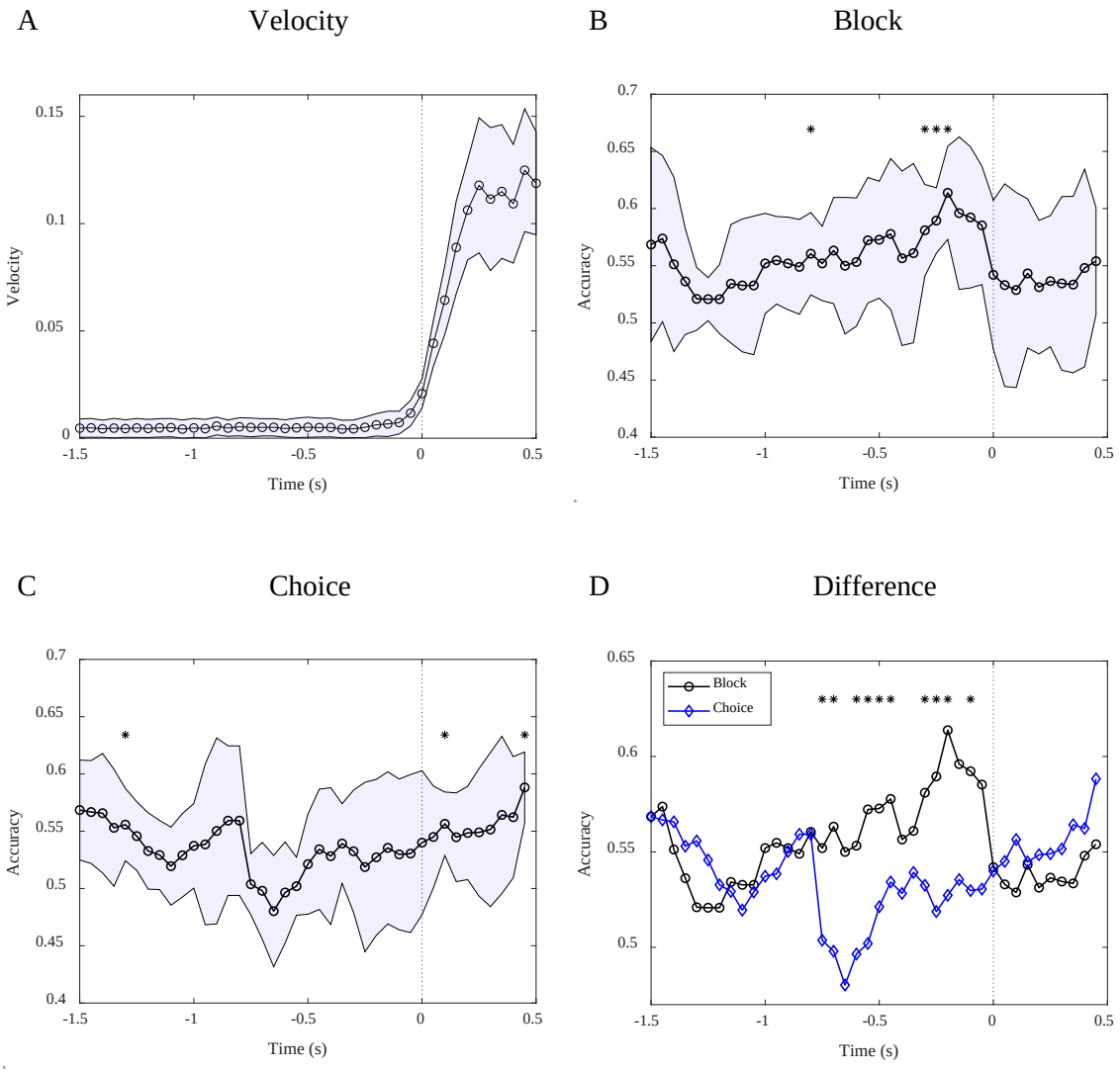


Figure 3.7. Dynamic changes in encoding of block and choice as the rat departs from the center feeder. (A) Mean and standard deviation of head velocity as the rat departs from the center feeder (t=0). **(B)** Decoding of block type in temporal bins centered on departure from to the center feeder. Asterisks '*' indicate significant difference from chance levels at p=0.01. **(C)** Decoding of choice. Same format as in B. **(D)** Superposition of mean decoding accuracy of choice and block. Asterisks '*' indicates statistically significant ($p < 0.05$) difference of decoding of these variables. The accuracy for block detections peaks before movement initiation, while the accuracy of choice detection increases after the animal starts moving.

3.2.6 No reward expectation signal in the middle arm of the maze

Next, we investigated if ACC neurons encoded a reward expectation signal in the middle arm of the maze. Because reward delivery by the high-utility and low-utility feeders in each block were probabilistic, even with perfect knowledge of the block type, it would not be possible for the rats to predict the outcome with 100% accuracy. To partially control for this unpredictability, we only used the data from trials in which the animals chose the optimal feeder and received the reward or chose the non-optimal feeder and received no reward. This should provide the largest contrast in reward prediction. We also excluded the first 10 trials in each block. Figure 3.8 illustrates the accuracies for predicting reward in the spatial bins before the animals reached the side feeders. The accuracies in all bins are around the chance level, which indicates that there was no reward expectation signal in the ACC when the rats were in the middle arm of the maze.

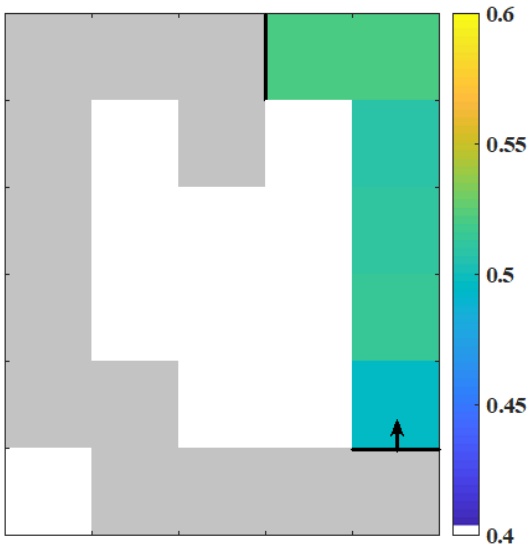


Figure 3.8. Accuracy of prospective reward prediction in the middle arm of the maze. No accuracy is significantly higher than the chance level.

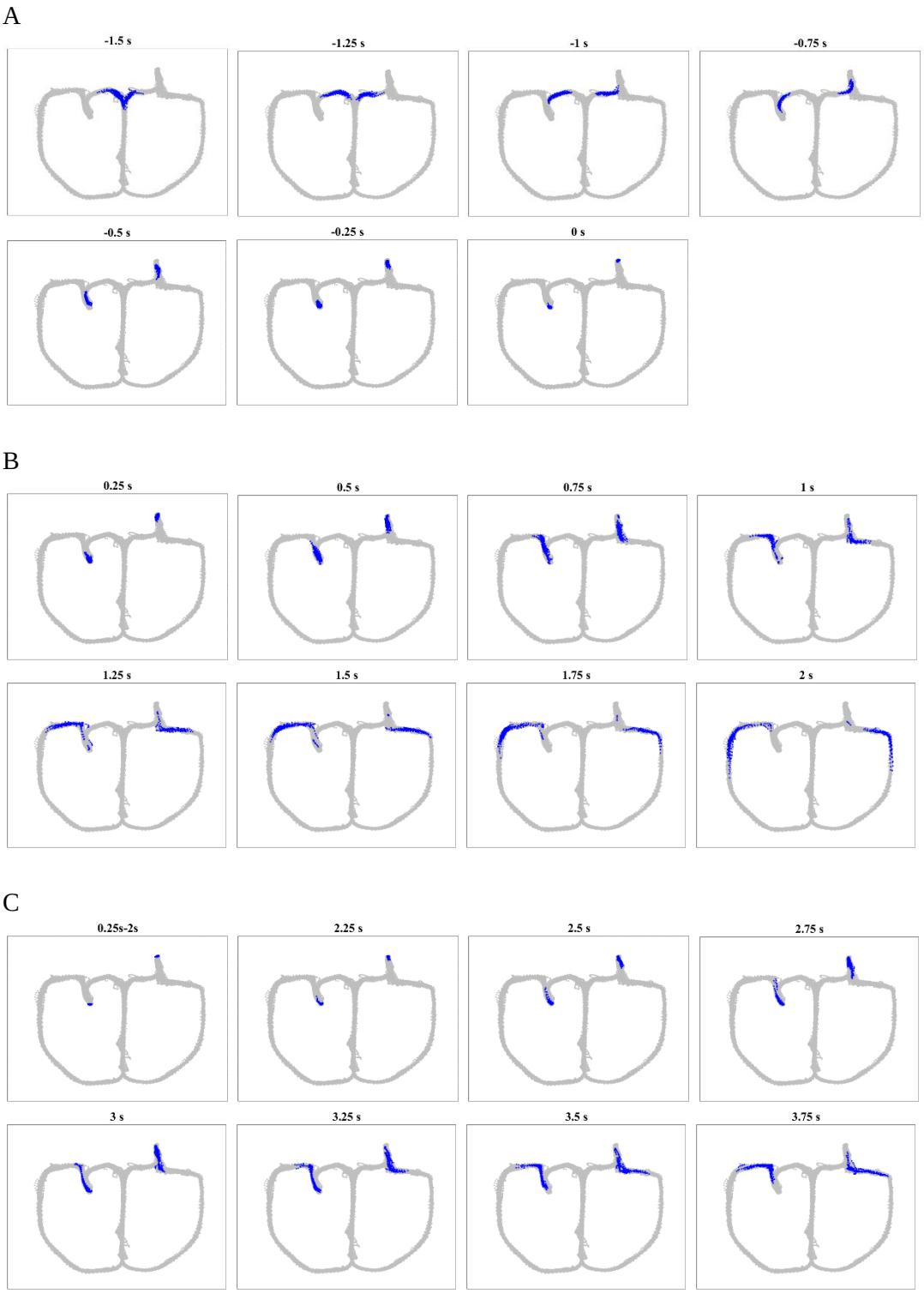


Figure 3.9. Position of the rat in one session for progressive time bins before and after arrival at side feeders. (A) The period before the reach time to the side feeders. **(B)** Position in 'no reward' trials. **(C)** Position in 'rewarded' trials. The animal left the feeder zone much earlier in 'no reward' trials compared with 'rewarded' trials.

Note that reward prediction is possible if information about choice and block type are both available. That is, the choice signal predicts the side feeder that the rat will choose in the current trial, and the block type signal specifies the expected outcome at that feeder. Thus, the low accuracy of the reward prediction in the center arm was expected because choice and block type were not simultaneously signaled in the middle arm.

3.2.7 ACC neurons encode the outcome at the side feeders

Next, we investigated the response of ACC neurons to reward outcomes (i.e., delivery or omission) at the side feeders. For this analysis, we used a sliding window method time-locked to the arrival at the side feeders and computed the accuracy of outcome decoding for each temporal bin.

Figure 3.10 illustrates the accuracy of the outcome detection for each temporal bin. Note that since the zero on the time axis represents the reach time to the side feeders, negative points in Figure 3.10 represent reward expectation, and positive points represent the outcome encoding. The positions of the animals in the no-reward and rewarded trials during different time windows relative to the reach time to the side feeders are shown in Figure 3.9.

The results showed no clear reward expectation signal as the animals approached the side feeders, evident by the low accuracy before reaching the feeders. However, the accuracy increased sharply and reached 85% after the animals arrived at the side feeders, indicating a clear encoding of the outcome after the outcome was revealed to the animals. Note that although the animals left the side feeders much earlier in the ‘no reward’ trials compared to ‘rewarded’ trials, during the period shown in Figure 3.10, the positions of the

animals in the 'rewarded' and 'no reward' trials were the same. Therefore, outcome decoding is not an artifact of the difference between the positions of the animals in the two groups.

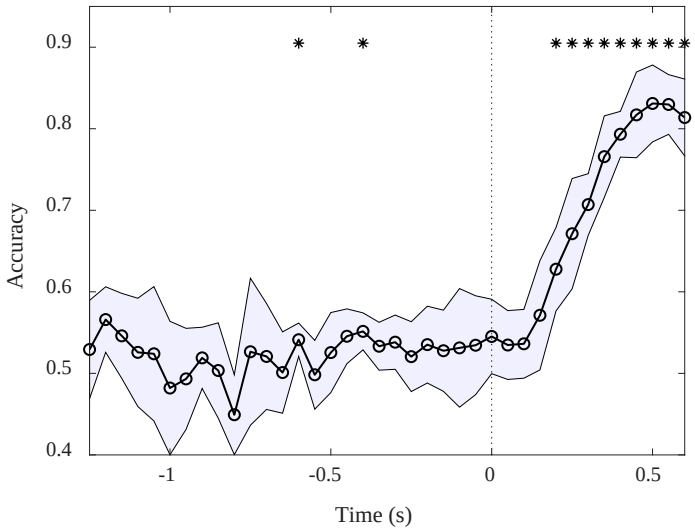


Figure 3.10. Accuracy of reward outcome detection in time bins locked to the arrival at the side feeders. The results show no clear reward expectation signal as the animals approached the feeders. However, there is clear encoding of outcome after they arrive at the feeder.

3.2.8 Information about the previous reward at target feeders is recalled at the starting feeder

To identify if/when the ACC recalled information about the outcome in a trial, we computed the accuracy of the reward decoding in the spatial bins after the side feeder zone, as the animals moved away from the side feeders and approached the starting feeder. In each session, we restricted our analysis to the trials in which the rat chose its favorite feeder to eliminate the confounding effect of choice. Figure 3.11 displays the accuracies for different spatial bins on the maze. The average accuracy in the vicinity of the side feeder is 60.1% (std = 5.78; $t_4 = 3.91$; $p = 0.009$; $power_{0.01} = 0.61$; $power_{0.05} = 0.94$). The accuracy

decreases as the animals move away from the feeder and reaches 53% (std = 4.94; t_4 = 1.43; p = 0.113; $power_{0.01}$ = 0.09) by the end of the trial. Despite the decreasing trend in the accuracy of reward decoding in a given trial, there is a spike in the accuracy when the animals arrive at the center feeder, at the beginning of the next trial (accuracy 58%; std = 4.20, t_4 = 4.02; p = 0.008; $power_{0.01}$ = 0.63; $power_{0.05}$ = 0.95). From there, the accuracy decreases again.

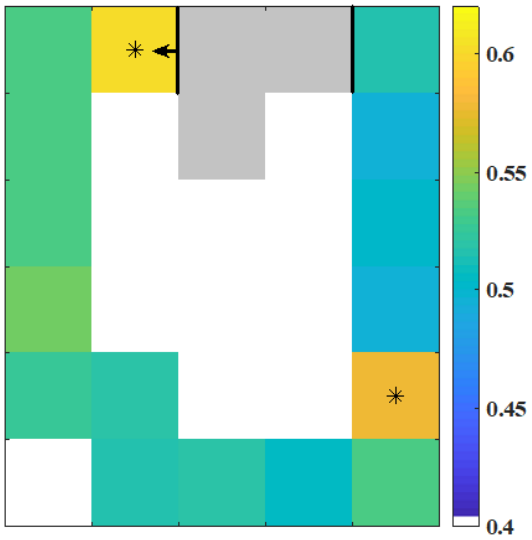
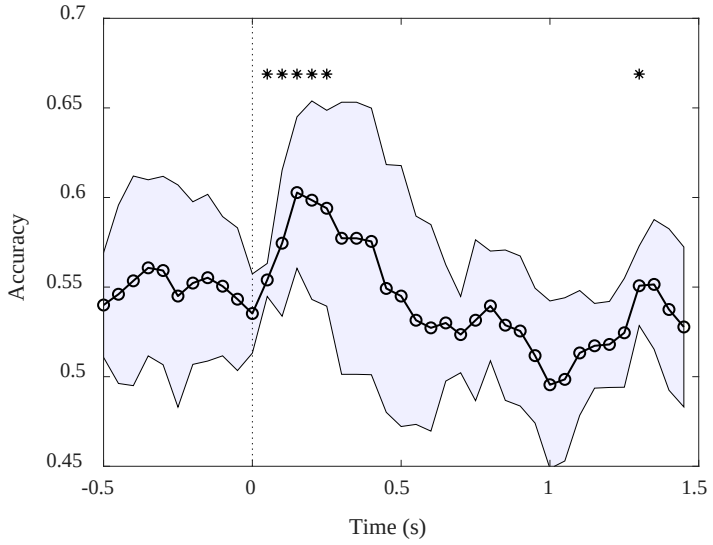


Figure 3.11. Accuracy of decoding retrospective outcome in different spatial bins after the animals move away from the side feeder area. The accuracy has a decreasing trend during the same trial. However, there is a peak in accuracy when the animals arrive at the center feeder at the beginning of the next trial. '*' indicates significant difference from chance at p = 0.01.

Analysis of the dynamic changes in the accuracy of outcome decoding revealed two peaks: one occurred after arrival at the center feeder (Figure 3.12.A), and the other, before departure from the feeder (Figure 3.12.B). This pattern is similar to what we observed in block detection accuracy at the center feeder.

A Arrival to the center feeder



B Departure from the center feeder

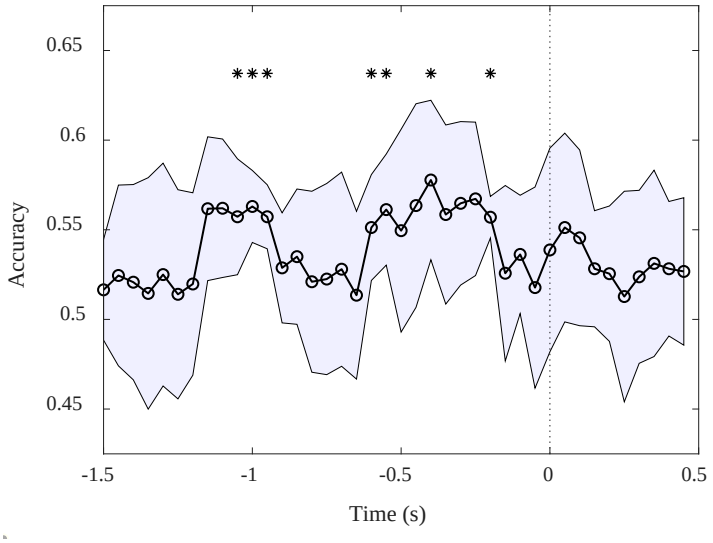


Figure 3.12. Accuracy of decoding the previous outcome in different temporal bins relative to the arrival/departure time at the center feeder.
(A) Decoding of the outcome in temporal bins centered on arrival at the center feeder. Asterisks '*' indicate significant difference from chance levels at p=0.01. **(B)** Decoding of the outcome in temporal bins centered on departure from the center feeder. There are two peaks in the accuracy after arrival at and (a weaker one) before departure from the feeder.

3.2.9 Reward history is distinguishable

The optimal choice in the task depended on the block type. Because the reward delivery was probabilistic, inferring the block type based on several past rewards was a better strategy than inference solely based on the immediately prior outcome. We examined the response of the ACC to identical outcomes with different histories to determine whether reward history affected the neural activity.

We divided the trials in each session into four groups based on the current and previous outcomes: A) both the current and previous trials were rewarded, B) the current trial was rewarded, but not the previous trial, C) neither the current trial nor the previous one was rewarded, D) the current trial was not rewarded, but the previous trial was. Accordingly, the trials in groups A (*reward* → *reward* sequence) and B (*no reward* → *reward* sequence) had the same outcome for the current trial (reward) but different outcome histories. The same applied to the trials in groups C (*reward* → *reward* sequence) and D (*reward* → *no reward* sequence).

To determine whether ACC response to identical outcomes differed based on the reward history, we assessed the discriminability of the trials in groups A versus B ('rewarded' condition) and C versus D ('no reward' condition) in different temporal bins between reaching the side feeder and moving away from it. We used a sliding window method, with a width of 250 ms and steps of 50 ms, time-locked to the reach time to the feeder to compute the classification accuracy for each temporal bin. For this analysis, we only used the trials in which the rats went to their preferred side feeder (both free and forced-choice trials) because the number of available trials for that feeder was higher, and limiting data to one feeder would eliminate to confounding effect of location and choice.

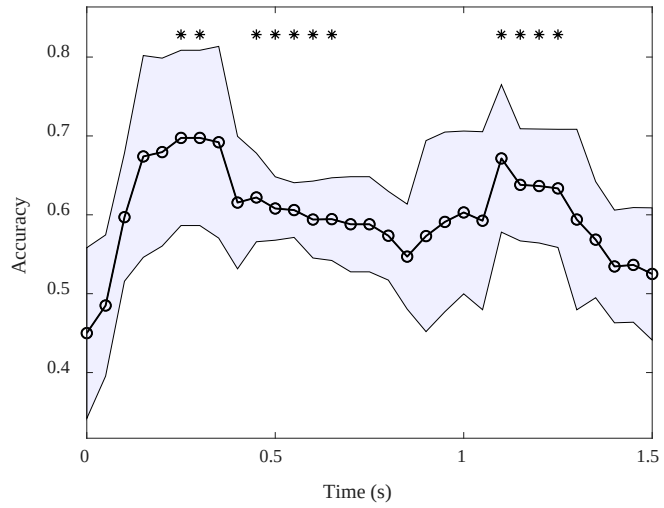
Figures 3.13.A and 3.13.B illustrate the results for the ‘no reward’ and ‘rewarded’ conditions, respectively. The duration of the time shown for the ‘no reward’ condition (Figure 3.13.A) is shorter than the ‘rewarded’ condition (Figure 3.13.B) because the animals moved away from the feeder earlier in the ‘no reward’ condition compared with to ‘rewarded’ condition.

The results show that, for each condition, there are two clusters of points during which the reward history is distinguishable from ACC activity with statistical significance (‘*’ in the figures represent $p < 0.01$). The peak of the first cluster in each condition occurs during the first 0.5 s of reaching the feeder, when the outcome for a trial is revealed. The second peak occurs after 1 s in the ‘no reward’ condition and 2 s in the ‘rewarded’ condition. In both conditions, the second peak is around the time that the animals moved away from the feeder (Figure 3.9).

Note that what is referred to as encoding of the reward history here is equivalent to a reward prediction error (RPE) signal discussed in many other papers. In this interpretation, the animal is presumed to expect the same outcome over subsequent choices to the same feeder. Therefore, an RPE signal appears in the ACC when the outcomes were incongruent. A comparison between the results for the two conditions shows that the peaks in the ‘no reward’ condition are higher than the peaks in the ‘rewarded’ condition, suggesting that the response to the unexpected negative outcome might be stronger than the unexpected positive outcome. This ‘negative surprise’ has been described in the ACC of primates (Alexander and Brown 2011; Egner 2011).

Regardless of whether the separation of groups in each condition results from the direct encoding of reward history or an RPE signal, the results show that ACC response to identical outcomes differed based on the recent outcome history.

A ‘no reward → no reward’ versus ‘reward → no reward’



B ‘no reward → reward’ versus ‘reward → reward’

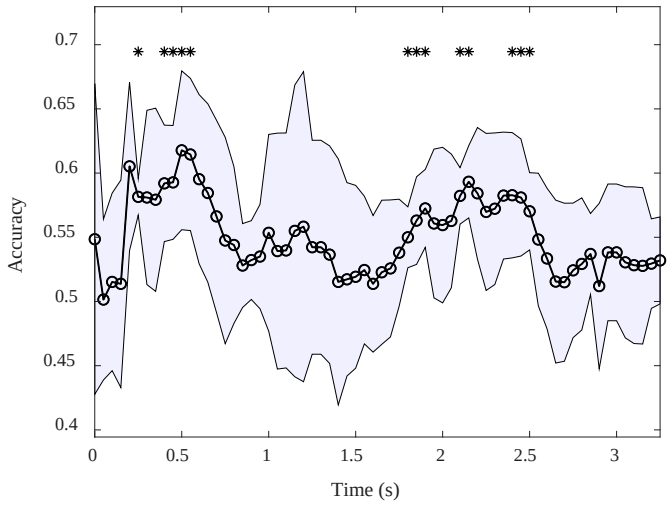


Figure 3.13. Accuracy of decoding outcome history for identical outcomes at a side feeder. (A) The accuracy of distinguishing between the ‘no reward → no reward’ versus ‘reward → no reward’ sequences (‘no reward’ condition). **(B)** The accuracy of distinguishing between the ‘no reward → reward’ versus ‘reward → reward’ sequences (‘rewarded’ condition). The highlighted areas show the standard deviation. ‘*’ represents $p < 0.01$. For each condition, two peaks in the accuracy can be observed. The first peaks occur during the first 0.5 s of reaching the feeder, when the outcome for a trial is revealed. The second peaks occurs as the animals moved away from the feeder.

3.2.10 Block type signal is independent of the prospective choice signal

Within each session, block type was correlated with choice and previous reward. This correlation was caused by the tendency of rats to choose the high-probability reward option for each block. To determine if the block type signal is independent of these other two variables, we developed a novel machine learning method to remove choice information from the neural activity and examined its effect on block decoding. A high degree of dependency between these two signals would entail a lower accuracy for block type detection after removal of choice information.

Because block decoding had the highest accuracy at the center feeder, the analysis was performed on the data from that region. Forced trials and the first 10 trials in each block were excluded. For every session, the average firing rates of neurons when the animals were at the spatial bin associated with the center feeder were used as the input samples. First, a filtering network was used to transform the samples into a new space where the information about the next choice (prospective choice) was removed. Then, 10 different classifiers were used to assess accuracies for the prospective choice and block type detection from the transformed data. The maximum accuracy for each variable was recorded as an estimation of the maximum achievable accuracy for that variable from the transformed data. The same classifiers were used to assess the maximum achievable accuracy for the prospective choice and block type from the original data. In the next experiment the same process was repeated for removal of variation informative about the retrospective choice (choice in the previous trial).

Figure 3.14 illustrates the results that were obtained by averaging the accuracies over all sessions. The bars labeled “Original” show the maximum accuracies for decoding

choice and block type from the original data. The bars labeled “Transformed” show the maximum accuracies for decoding the next choice and block type from the transformed, with information about choice removed.

In the first experiment, after transformation, the maximum accuracy obtained by the classifiers for decoding the next choice is not above the chance level (accuracy 49.40%; std = 0.4, t_4 = -3.2071; p = 0.984; power = 0.00), indicating that the network was able to eliminate the variation related to the prospective choice in the data. In contrast, the accuracy for decoding block type from the transformed data (68%) is not significantly different from the accuracy obtained on the original data (std = 2.70; t_4 = 0.54; p = 0.62; $power_{0.01}$ = 0.02), indicating that information about block type was retained after transformation (Figure 3.14.A).

In the second experiment, the maximum accuracy obtained by the classifiers for decoding the previous choice from the transformed data is not above the chance level (accuracy 50.76%; std = 2.68, t_4 = 0.634; p = 0.280; power = 0.031), indicating that the network was able to eliminate the variation informative about the retrospective choice. In contrast, the accuracy for decoding block type from the transformed data is 66.6% which is not significantly different from the accuracy obtained on the original data (std = 4.32; t_4 = 1.43; p = 0.226; $power_{0.01}$ = 0.0519), indicating that information about block type was retained after transformation (Figure 3.14.B).

Being able to decode the block type from the transformed data, after removal of prospective and retrospective choice information, with the same accuracy as the original data suggests that the block type signal in the ACC is independent of the choice signal.

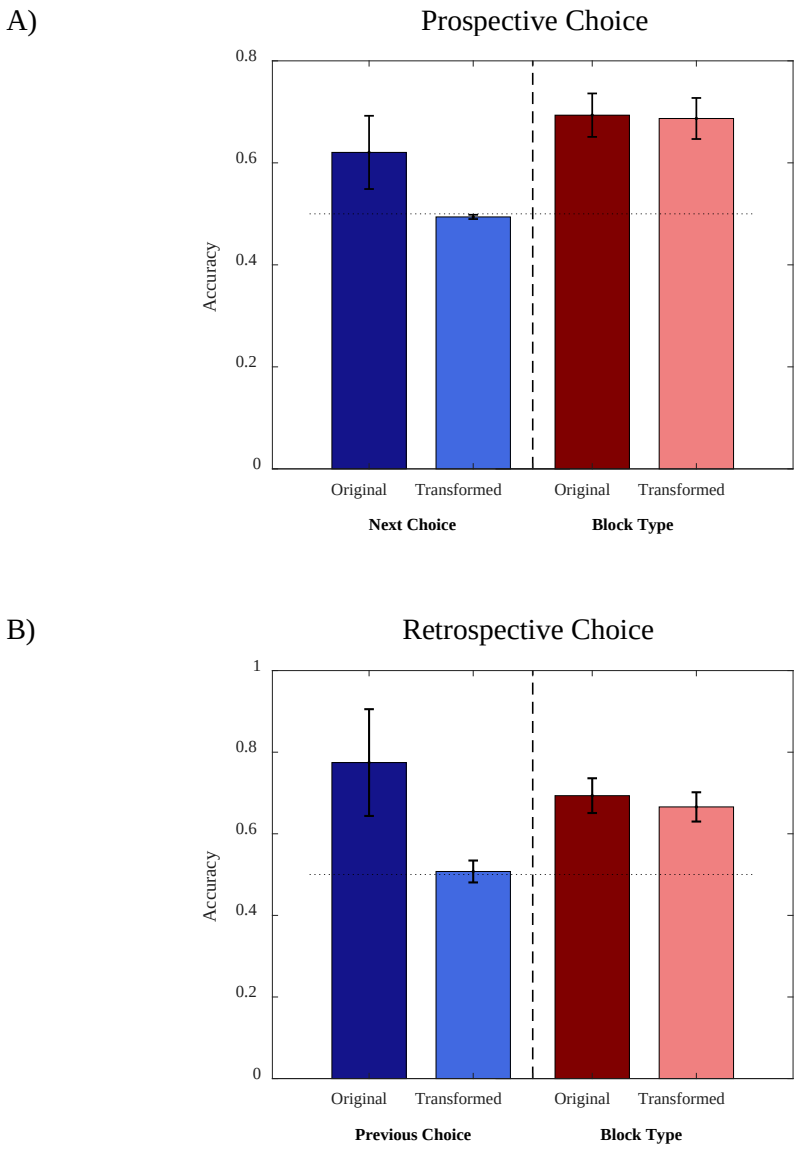


Figure 3.14. Comparison between accuracies for choice and block type detection before and after removal of choice information. (A) Removal of prospective choice information. After transformation, the accuracy for choice detection is not significantly higher than chance level ($p = 0.984$), while the accuracy for block type detection is not significantly different from the one achieved on the original data ($p = 0.62$). **(B)** Removal of retrospective choice information. After transformation, the accuracy for choice detection is not significantly higher than chance level ($p = 0.280$), while the accuracy for block type detection is not significantly different from the one achieved on the original data ($p = 0.23$).

3.2.11 Block type signal is independent of the previous reward signal

Next, we used the same procedure explained in the previous section to investigate the dependency between the block type and previous outcome signals in the ACC at the center feeder. We eliminated the information about the previous outcome from the ACC activity at the center feeder and examined its effect on the accuracy of decoding the block type. A high degree of dependency between these two signals would result in a lower accuracy for block type detection from the transformed data.

For this analysis, forced trials and the first 10 trials in each block were excluded. For every session, the average firing rates of neurons when the animals were at the spatial bin associated with the center feeder were used as the input samples. The proposed network was then used to transform the samples into a new space where the information about the previous outcome was removed. Then, 10 different classifiers were used to assess the accuracies for the previous outcome and block type detection from the transformed data. The maximum accuracy for each variable was recorded as an estimation of the maximum achievable accuracy for that variable from the transformed data. The same 10 classifiers were used to assess the maximum achievable accuracy for the task variables from the original data.

Figure 3.15 illustrates the average accuracies obtained in each condition. The bars labeled “Original” show the maximum accuracies for decoding the two variables from the original data. The bars labeled “Transformed” show the maximum decoding accuracies from the transformed data, with information about the outcome removed. The results show that after transformation, the maximum accuracy obtained by the classifiers for decoding the outcome is not above the chance level (accuracy 50.7%; std = 2.25; $t_4 = 0.65$; $p = 0.28$;

power = 0.03), indicating that the variation related to the outcome was successfully eliminated from the data by the network. On the other hand, the accuracy for decoding block type from the transformed data (67%) is not significantly different from the accuracy obtained on the original data (std = 4.02; t_4 = 1.08; p = 0.34; power = 0.03), indicating that block type information was retained in the transformed data. Achieving the same decoding accuracy for block type from the transformed data, after removing the variation informative about the outcome, as the accuracy on the original data suggests that, at the center feeder, the block type signal in the ACC is independent of the outcome signal.

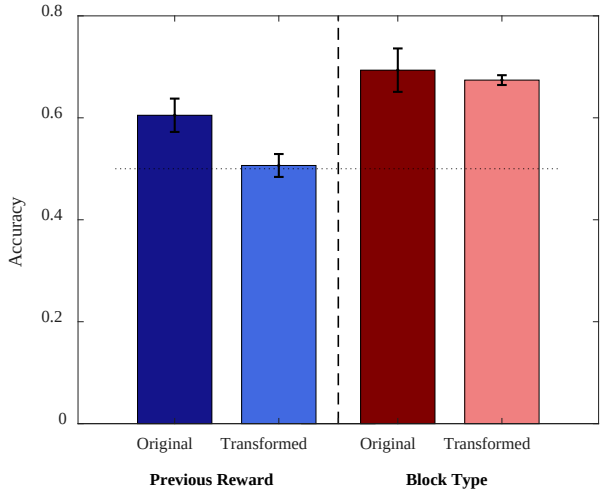


Figure 3.15. Comparison between accuracies for outcome and block type detection before and after removal of retrospective outcome information. After transformation, the accuracy for outcome detection is not significantly higher than the chance level (p = 0.28), while the accuracy for block type detection is not significantly different from the one achieved on the original data (p = 0.34).

3.2.12 The block type signal is stable even when the choice is not optimal

As shown before, the animals were usually able to notice the shift in the block type after about 10 trials (adaptation trials), and they changed their choices accordingly. However, they occasionally chose the non-optimal option even after the adaptation trials.

Because the block type signal had the highest strength at the center feeder, it was possible that making the non-optimal choice was the result of a defective block signal at that location.

To test this hypothesis, first, we eliminated the variation related to choice from ACC activity at the center feeder. Then, we divided the trials in each session into two groups, depending on whether the choice in a trial was optimal or not (defined by block type). We excluded the forced and the first 10 trials in each block. A support vector machine classifier with a polynomial kernel of degree 1 then was trained on the transformed data of the trials in the optimal group to decode the block type. The classifier was used to detect the block type for the trials in the non-optimal group. If making a non-optimal choice was due to a defective block type signal at the center feeder, the accuracy obtained by training on the optimal group and testing on the non-optimal group would be lower compared with the one obtained by training and testing the classifier on a mixed set of optimal and non-optimal trials. However, the results, illustrated in Figure 3.16, showed no significant difference between the accuracies in the two conditions (std = 7.98; $t_4 = 0.58$; $p = 0.59$; $power_{0.01} = 0.02$).

The results suggest that representation of block type in the ACC was stable even when the animals did not make the optimal choice.

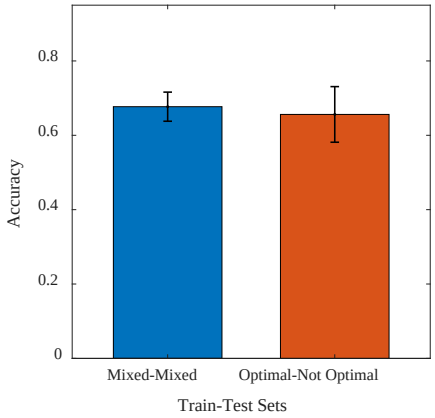


Figure 3.16. Comparison between accuracy of block detection when a classifier is trained and tested on a mixed dataset of optimal and non-optimal trials, versus when the classifier was trained on the optimal trials and testes on non-optimal trials. The result show no significant difference ($p = 0.59$).

3.3 Discussion

We recorded neural activity in the ACC of freely moving rats performing a two-choice task on a figure 8-maze. The task was organized into blocks of 40-50 consecutive trials in which one feeder had a higher probability of reward (pro=0.8) than the other. The reward probabilities reversed among feeders between subsequent blocks. Therefore, the optimal choice for each trial depended on the block type. Free and forced trials were interspersed randomly throughout each session.

We found that ACC activity was most discriminative of free versus forced trials in the middle arm of the maze, shortly after departing the starting feeder. This is well before the 'T' intersection that forced a turn to reach one of the feeders. These were also the regions where the animal was exposed to the unplanned auditory (motor noise) and visual (presence of gate) cues indicative of trial type. We found no evidence of ACC selectivity for the trial type in other regions of the maze where the cues indicative of trial type were less pronounced.

Based on these results alone, it is unclear whether the prominence of free/forced trial type encoding in the ACC while the rat approached the turning point is due to its use in guiding behaviour, or simply a consequence that the sensory cues are most perceptible in this region. In the context of other studies, however, it appears that this encoding is part of a cognitive process. Specifically, prior papers from our group and an unrelated laboratory (Cowen, Davis, and Nitz 2012; Hashemnia 2019) reported weak or no encoding of trial type (free vs forced) by the ACC in tasks with slightly different designs that the one used here. Specifically, in the aforementioned studies, rewards were deterministic. Rewards were given on every trial regardless of choice, and reward magnitude delivered by each

feeder was fixed during trial blocks. Consequently, the rat could repeat its choice once the optimal option was found without re-evaluating the options. In contrast, in the current task, reward delivery and trial type were probabilistic. Therefore, optimal behaviour required constant evaluation of options based on the outcome history. The uncertainty of the outcomes and frequent evaluations of the options in the current task may have caused a higher cognitive load, resulting in distinct patterns of activity in free versus forced trials.

The behavioural analysis showed that choices in each session were influenced by shifts in the block type. Rats usually detected changes in the reward delivery probabilities during the first 10 trials of each block and shifted its choice bias in the subsequent trials accordingly. Also, the variability in the animals' choices in most sessions could be better explained by a strategy based on choosing the high utility feeder in each block, as compared with a win-stay-lose-shift strategy or selecting a preferred feeder in every free trial. These observations indicate that the animals were able to track the shifts in the block type and utilized this information to optimize their choices.

We found the highest discriminability of the block type in the ACC when rats were at the center feeder. Examination of the dynamic changes in the accuracy of block decoding revealed two peaks: after arrival to the center feeder, and just prior to departing from it to start the next trial. This result indicates that ACC activity signals the optimal choice at the beginning of each trial when the animal is still stationary. The strength of the block signal before movement initiation and its decline afterwards suggests that the block signal plays a role in decision-making. This finding is consistent with the role of ACC in evaluating the value of options with probabilistic outcomes (Kennerley et al. 2006) and biasing choices towards the higher utility option (Hillman and Bilkey 2010). Interestingly, during both peaks in block-type encoding, the accuracy of decoding the animal's future *choice* was

significantly lower than at other times. This raises the possibility that the block type information is recalled before and during the initial phases of choice generation.

The block type signal disappears as the animal starts running away from the start feeder. A potential factor contributing to the weakness of the block signal after movement onset is that ACC activity becomes dominated by encoding of other task variables, such as position, motor responses, and environmental features. This is primarily supported by the observation that both peaks in the block type detection accuracy occur when the animal is stationary. The encoding of position, in particular, may be analogous to the activity of place cells in the hippocampus, which predominate encoding during movement, but is weak when the animal is stationary (O'Keefe and Dostrovsky 1971; M. A. Wilson and McNaughton 1993). Therefore, as the animal starts moving, the activity of place encoding cells in the ACC might obscure the block.

Our results revealed a low accuracy for detecting choice in the middle arm of the maze. In contrast to what we observed for block type, the choice decoding accuracy was the lowest at the center feeder, and it increased only after the rat departed from the center feeder. This observation indicates that choice signal in the ACC follows action and is not predictive of it. This claim is consistent with findings in previous studies (Cowen, Davis, and Nitz 2012). We note that ACC firing rate biases were reported to be predictive of choice behaviour when at least one option involved competitive effort, for example climbing a barrier (Hillman and Bilkey 2012). However, competitive effort was not a factor in the present task, and so likely accounts for the differences in findings.

The accuracy of choice detection increased as the animal approached the T-junction and reached to approximately 100% upon the entrance to one side of the maze. The accuracy remained high for the remainder of the trial. The increase in the accuracy of choice

detection can be explained by the divergence of the running trajectory in trials to the left compared with the right side of the maze. Previous studies show that ACC activity is highly sensitive to the route taken (Cowen, Davis, and Nitz 2012) and the rat's position on a maze (Mashhoori et al. 2018; D. R. Euston and McNaughton 2006). Therefore, the apparent choice encoding is likely to be a correlate of the animal's position. On the other hand, numerous studies have reported encoding of action by the ACC and have pointed to its role in creating associations between actions and outcomes (Hayden and Platt 2010; C. Amiez, Joseph, and Procyk 2006; C. H. Luk and Wallis 2009; Quilodran, Rothé, and Procyk 2008). Therefore, it is possible that, at least in some phases of the task (e.g., when the animal is at a side feeder), ACC activity represents choice. Due to the perfect dependency between choice and position, it is impossible to disentangle the choice and position signals in the current dataset.

Our results showed no evidence for the presence of a reward expectation signal when the animal was in the in the middle arm of the maze, or as it approached the side feeders. The absence of the reward expectation signal in the middle arm of the maze is in line with the results of previous studies (Cowen, Davis, and Nitz 2012). Although we did not find a strong reward expectation signal before arrival at the side feeders, we did find a strong encoding of reward outcomes (reward delivery or omission) by the ACC at the side feeders after the outcome was revealed, which is consistent with previous reports (Hayden and Platt 2010; C. Amiez, Joseph, and Procyk 2006; Kennerley and Wallis 2009a; Mashhoori et al. 2018; Hashemnia, Euston, and Gruber 2020).

Furthermore, we found that the ACC's response to identical outcomes differed depending on recent outcome history. Attention to reward history was particularly important in the current task because reward delivery during each block was probabilistic.

Therefore, inferring the block type based on the outcome history rather than a single outcome was the optimal strategy. This is in line with studies in both human and non-human primates demonstrating the critical role of the ACC in reinforcement-guided behaviour based on the history of actions and outcomes, as opposed to detecting or correcting errors based on a single outcome (Kennerley et al. 2006; Clay B. Holroyd and Coles 2008).

We note that encoding of outcome history, as defined here, is equivalent to a reward prediction error (RPE) signal, representing the difference between the expected and obtained reward. The probability of two consecutive trials having the same outcome was 0.68, which is much higher than the probability of encountering incongruent outcomes (0.32). An RPE signal appears in the ACC when the outcome is incongruent from expectation. Evidence of the RPE in the post-outcome activity of the ACC has been found in both electrophysiological recordings of single units (Kennerley, Behrens, and Wallis 2011; M. Matsumoto et al. 2007; Seo and Lee 2007; Quilodran, Rothé, and Procyk 2008) and imaging studies (Céline Amiez et al. 2012; Kahnt et al. 2011). More generally, studies have shown elevated ACC activity in response to surprising outcomes (Cavanagh et al. 2012; Wessel et al. 2012; Nee, Kastner, and Brown 2011) and unanticipated shifts in task contingencies (Alexander and Brown 2011; Behrens et al. 2007). This response can trigger a model update and shift choice bias in subsequent trials. This view is supported by previous studies linking ACC activity to the receipt of information that leads to changes in a subject's internal model of a task (Hayden et al. 2011; Quilodran, Rothé, and Procyk 2008) or choice strategy (Shenhav, Botvinick, and Cohen 2013). In the present study, the ACC's response to omission following delivery on consecutive trials was stronger than the response to delivery following omission. This observation suggests that ACC's response to an unexpected negative outcome is stronger than an unexpected positive outcome.

The signal representing reinforcement outcome weakened as the animal moved away from the side feeders. However, there was an increase in the signal when the rat reached the center feeder, suggesting a recall of the outcome at that point. The dynamic pattern of the accuracy was similar to the block type signal, with peaks upon arrival at and before departure from the feeder.

The presence of both the block type and previous outcome signals at the center feeder, along with the similarity between their temporal changes, raises two possibilities. First, the representation of the block type may be highly dependent on the representation of the immediately previous outcome. In this view, the representation of the block type signal is mostly a correlate of recalling the previous outcome rather than a weighted average over many past outcomes. Based on the reward delivery probabilities in the task, the accuracy for inferring the block type from a single outcome is 80%. Additionally, most forms of applied reinforcement learning (i.e. those including a temporal discounting factor) give the most weight to recent reinforcements. Thus, the outcome value estimation is expected to depend strongly on recent outcomes. The block type only differs in outcome value expectation, so what we label as block type information may actually be a form of tracking expected outcome value.

Alternatively, despite the dependency between block type and recent reinforcement, their representation in the ACC may be independent, such that each signal is used for a different purpose. Both the block type and previous reward outcome are critical information for two different decision-making strategies. The previous outcome is the main factor in a win-stay-lose-switch strategy, which is a short horizon policy based on the last outcome. Although win-stay-lose-switch is not optimum in the current task, it still provides a reasonable strategy. This strategy is a good predictor of choice in some sessions, suggesting

that rats may utilize it in some instances of the task. On the other hand, block type is the main factor for decision making based on the outcome history over several trials, which is the optimal strategy for this task. It is possible that these strategies, along with other alternatives, compete to influence choice. Our findings support this interpretation, as discussed below.

Within each session, block type was correlated with choice and previous reward – that is, rats tended to choose the high-probability reward option for each block, and therefore gain reward more often than chance. Also, the representations of these task variables in the ACC might be dependent on one another due to inherent features of the cognitive processes involved. To determine if the block type signal is independent of these other two variables, we developed and used a novel method to removed information about choice or previous outcome from the neural activity prior to decoding block type. We found that removal of this information did not affect the accuracy of block decoding. This indicates that the representation of the block type in the ACC is independent from choice (both prospective and retrospective) and previous reward, and its decoding is not an artifact of its correlations with other variables.

It has been shown that successful behaviour depends on the representation of coherent and distinct patterns in the ACC, and that a breakdown of these patterns leads to errors (Lapish et al. 2008). In the current task, the block type signal was most pronounced at the center feeder. Therefore, it is conceivable that non-optimal choices were due to defective block representation at that location. However, we found that block type representation in the ACC was not impaired on trials in which rats chose the low-probability feeder. This finding is consistent with previous observations of a bias in ACC activity towards high-utility choices, even when the information is not immediately

reflected in behaviour (Hillman and Bilkey 2010). In other words, the animal can choose the sub-optimal option even when ACC encodes the optimal one. Our finding is also in line with reports of frequent updates of the task sets by the ACC, despite available options not being used instantly (Hyafil, Summerfield, and Koechlin 2009).

In summary, we found that: 1) ACC activity distinguished between free and forced trials. 2) ACC neurons encoded the outcome at the target feeders. 3) ACC responses to identical outcomes differed based on the outcome history. 4) The block type and previous outcome signals were stronger at the center feeder compared with other regions of the maze. However, this strength coincided with the weakness of the choice signal at that location. 5) The block type and previous reward are both main factors in two decision-making strategies that the rats employ in a session. 6) The representation of the block type is independent of both choice and the previous reward. 7) The block type signal is stable even when the information is not reflected in the behaviour in a trial.

Our findings support the hypothesis that the brain relies on multiple strategies to solve problems, and these likely depend on dissociated brain circuits. The ACC uses information about previous choices and their outcomes to create predictive models relevant to these strategies. The models include information about the predicted cost and value associated with each strategy. The ACC is also responsible for updating these internal models upon the receipt of new data. These strategies compete to influence choice. However, the ACC itself is not responsible for choosing between these strategies. Indeed, the information relevant to some alternative strategies might be represented by other brain areas. An example is the representation of default option value, which is considered the value of the default strategy in our view, by the vmPFC (Strait, Blanchard, and Hayden 2014; Blanchard and Hayden 2014; Boorman, Rushworth, and Behrens 2013). This

information is utilized by another brain region that chooses the winning strategy, and consequently, the action. This is related to, but different from, the view that the ACC represents the cost and value of single actions and their outcomes.

3.4 Experimental Procedures

3.4.1 Subjects

In this study, two male Fisher-Brown Norway rats, aged 9 months old, were used. The rats were born on-site and housed individually in a 12h-12h reverse light cycle. One rat performed the task over three sessions and the other rat for two sessions. Animals were food restricted for the duration of the behavioural testing, and their weight was monitored to not drop below 85% of their baseline weight.

3.4.2 Ensemble recording

Established techniques for surgery and hyper-drive implantation were used to record ensemble activity with high-density electrophysiology (D. R. Euston and McNaughton 2006). Rats were habituated to handling for two weeks prior to the experiment. We used aseptic surgical techniques to implant onto each rat's head a hyper-drive device with 16 tetrodes that were independently lowered into the ACC. The tetrodes were electroplated with gold to an impedance of 200-400 kΩ. Neural signals and the positions of subjects were recorded with a digital acquisition system (Cheetah SX & Cube 2, Neuralynx). Neural signals were amplified with a gain of 1000, digitized at 30 kHz, and recorded to a hard drive. The waveforms of action potentials were detected offline using semi-automated techniques and the clusters were checked manually using MClust. All procedures were performed in accordance with the Canadian Council of Animal Care and the Animal Welfare Committee at the University of Lethbridge

3.4.3 Behavioural apparatus

The task was implemented with a figure-8 track. There were three feeders on the maze that delivered chocolate beverage as the reward: the center feeder (located at the bottom of the center arm of the track), and two side feeders (one on each side of the maze). Figure 3.17.A illustrates the trajectory of a rat in one session and the positions of the feeders on the maze. Each trial began with the animal's arrival at the center feeder. After receiving food at the center feeder, rats ran up the middle arm of the track and then made a right or left turn to reach one of the two target feeders.

A B

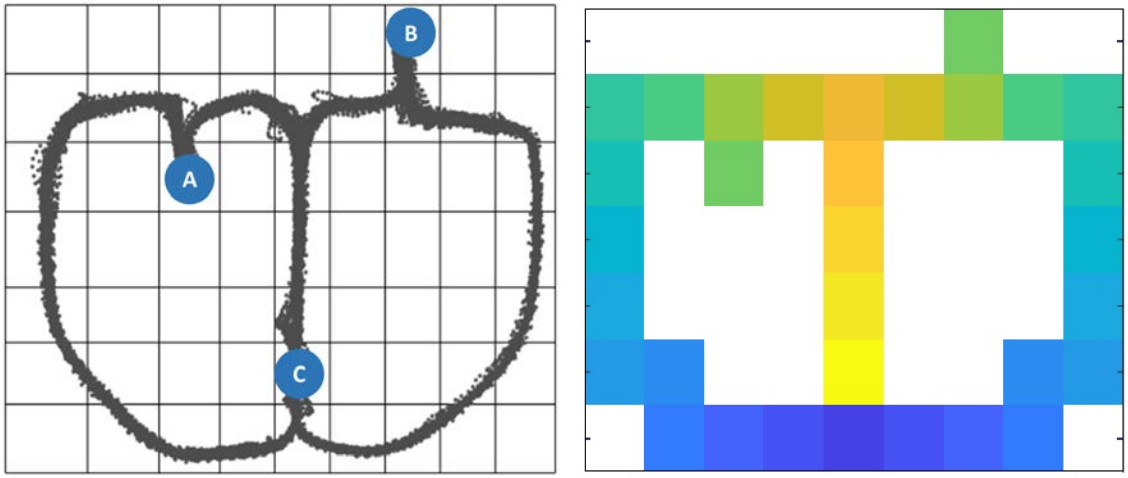


Figure 3.17. Schematic illustration of the maze and the rat's trajectory.
(A) The trajectory of the rat on the maze in one session. The circles show the locations of the feeders. The circles labeled A and B represent the side feeders. The one labeled C represents the center feeder. The grids depict how the maze was divided into a number of spatial bins. **(B)** The spatial bins on the maze. The bins having the same color on both side of the maze were considered identical in our analysis.

In 70% of trials of each session, the animal was free to go to either side of the maze (free-choice trials). In the rest of the trials, computer-controlled gates forced the animal to

go to one side, with equal probability for the left and the right side (forced-choice trials). Free and forced-choice trials were distributed randomly in each session.

In each trial, when the animal was at one of the side feeders, it either received a specific amount of food as reward or no reward at all. The amounts of reward for both side feeders were fixed and equal. Rats then ran along the lateral arms of the track to return to the start feeder. The center feeder delivered food in every trial to motivate rats to return to the start position.

Each session consisted of 250 to 280 trials, divided into blocks of size 40 or 50. There were two block types, B1 and B2. The rewarding schedule was fixed during each block. For B1, the probability of delivering food by feeder A and B was 0.8 and 0.2, respectively. For the other block type, the probabilities changed to 0.2 for feeder A, and 0.8 for feeder B. Therefore, during each block, one of the side feeders was the optimal choice. Block type alternated in each session. Forced-choice trials ensured that the animal sampled both options to detect the shift in block type.

3.4.4 Data preprocessing

A Gaussian kernel with a standard deviation of 50 ms was used to smooth the spike data. The resulting signal was binned by 50 ms. A square root transformation was then applied to the binned firing rates to normalize the activity distribution (e.g., make them more Gaussian) among neurons. Finally, a z-transform was applied to the activity of each neuron to make it have zero mean and unit variance.

3.4.5 The procedure for computing the accuracy of decoding a task variable

Given a data set of input samples and their target values (i.e., labels), two samples were selected from each class randomly as the test set (4 samples in total). The remaining samples were used as the training set. The training set was balanced by up-sampling the instanced in the minority class. A classifier was trained on the training set and evaluated on the test set. This process was repeated 100 times. The average accuracy of the classifier over these 100 randomly constructed training and test sets was recorded.

3.4.6 Computing the classification accuracy of a task variable in different spatial bins

To compute the classification accuracy of a given task variable (e.g., choice) in a spatial bin on the maze, a dataset of samples was made by computing the average firing rate of neurons when the animal's position was within that spatial bin in each trial. This process resulted in a dataset of N samples, each having D features, where N is the number of trials, and D is the number of neurons. The value of the variable of interest associated with each trial was assigned to that sample as the target value. A support vector machine classifier with a polynomial kernel of degree 1 or 3 was used to compute the accuracy of decoding the variable of choice in each spatial bin. The procedure for evaluating the accuracy was explained in section 4.5.

3.4.7 Computing the classification accuracy of a task variable, time-locked to an event

To compute the classification accuracy of a given task variable time-locked to an event (e.g., reach time to a feeder), a sliding window method with a width of 250 ms and steps of 50 ms was used. The window was placed at the beginning of the time range. For

that temporal bin, a dataset of samples was constructed by computing the average firing rates of the neurons during that time window (250 ms) in each trial. The value of the variable of interest associated with each trial was assigned to that sample as the target value. The procedure described in section 4.5 was used to assess the classification accuracy. A support vector machine with a polynomial kernel of order 1 or 3 was used as the classifier. The window was moved forward by the step size (50 ms). The same process was repeated for every temporal bin.

3.4.8 Estimating the maximum achievable accuracy for decoding a parameter

From the dataset, two samples were selected from each class randomly as the test set (4 samples in total). The remaining samples were used as the training set. The training set was balanced by up-sampling the instanced in the minority class. 10 different classifiers were used to estimate the maximum achievable accuracy for decoding the variable of interest. The 10 classifiers used were: logistic regression, k-nearest neighbors (k: 1, 3), support vector machines (kernel: linear and cubic), Gaussian process (kernel: RBF), decision tree (max depth: 5), random forest (max_depth:5, number of estimators: 10, maximum number of features: 10), neural network with one hidden layer (number of units: 16), AdaBoost, Gaussian Naïve Bayes, quadratic discriminant analysis. Each classifier was trained on the training set and evaluated on the test set. This process was repeated 100 times. The average accuracy for each classifier over these 100 runs was recorded. After the average accuracies for the 10 classifiers were obtained, the maximum accuracy was used in the figures.

3.4.9 Removing variation related to one task variable from the data

The goal was to transform the data into a new space in which the variation related to variable V is removed and test the accuracy that was achievable for decoding variable W from the transformed data. The network proposed in chapter 3 was used to transform the data. The dataset for the task was divided into a training set and a test set. The test set comprised of 4 samples from each class of variable V. The training set was made balanced by up-sampling the instances in the minority class of variable V. The labels of variable V were used as the target for both classifiers in the two branches of the network. The input samples were used as the target for the decoder part of the network. The network was trained for 300 epochs on mini-batches of size 20. The performance of the network was monitored based on the results on the test set. After training was completed, the original training data, before the up-sampling step, was transformed using the encoder in the filter branch of the network. The maximum achievable accuracies for variable V and W in the new space were assessed using the procedure explained in the previous section.

4 A Deep Neural Network for Generating a Representation Invariant to a Task Variable

Abstract

When analyzing neural data, we often aim to determine if neural activity is sensitive to a particular variable. One way to tackle this problem is to compute the decoding accuracy of the variable of interest from the neural activity. However, this decoding may be confounded with the effect of correlations between different task variables. We propose a deep neural network, based on the recent advances in adversarial training of artificial neural networks, to transform a dataset into a new space in which the variation informative about a given variable is eliminated while the remaining variation is retained. The network can be used to compute the decoding accuracy of one variable without the confounding effect of another variable.

4.1 Introduction

When analyzing neural recordings, the aim is often to measure the degree of sensitivity of the activity in a brain region to a particular task variable. A solution is to compute the decoding accuracy of that variable from the neural activity of the cells in that region. However, a brain region can simultaneously encode several task variables. Consequently, the decoding of one variable can become confounded with the effect of its correlations with the other task variables. These dependencies might emerge in the data due to task design, subject behaviour, or brain internal mechanisms for encoding task variables. Therefore, it is desirable to measure the accuracy of decoding a task variable without the confounding effect of another one.

A strategy to resolve this issue is to create a balanced dataset with respect to both variables. For instance, in the case of two binary variables A and B, one can create a dataset in which $P(B = 1|A = 1) = P(B = 1|A = 0) = 0.5$ to eliminate the confounding effect variable B on decoding variable A. However, while this approach removes the dependencies that arise due to the correlation between class labels in a given dataset, it fails to address the issue of dependencies resulting from the shared features that are informative about different classes.

A potential solution to the problem of shared informative features is to simply detect and remove the cells that are selective for a given variable and compute the decoding accuracy of the variable of interest from the activity of the remaining cells. However, not only neurons in a region might multiplex information about several task variables, the encoding for any given variable might happen at the ensemble level, through possibly nonlinear interactions. For instance, numerous studies have shown that ACC neurons

multiplex information about reward, choice, and effort (Hayden and Platt 2010; Mashhoori et al. 2018; Hashemnia, Euston, and Gruber 2020). Also, in a previous study, we found that the accuracy of detecting a rat's position on a maze from ACC activity using a deep neural network was significantly higher than the accuracy achieved by a Bayesian decoder, supporting the role of nonlinear interactions between the activity of neurons in the encoding of position (Mashhoori et al. 2018). This distributed representation and the many-to-many relationship between neurons and task variables render the mere removal of selective cells an ineffective solution.

Finally, conventional unsupervised dimensionality reduction, feature extraction, and decomposition methods such as principal component analysis (Jolliffe 1986), independent component analysis (Comon 1994), or typical autoencoders (Makhzani and Frey 2014; Vincent et al. 2010) are not proper tools for this purpose because they either have a different objective (e.g. retaining most variation in the data in a lower dimension) or cannot directly exploit the labels associated with the samples.

Here, we propose an autoencoder (I. Goodfellow, Bengio, and Courville 2016) with a novel architecture that employs the idea of adversarial training of artificial neural networks (I. J. Goodfellow et al. 2014) to transform a dataset into a new space in which the variation informative about a given task variable is eliminated while the remaining variation is retained.

The following sections are organized as follows. Section 4.2 presents the mathematical formulation of the method and the architecture of the proposed network. Section 4.3 demonstrates test cases for evaluating the effectiveness of the method in eliminating variation related to a task variable while retaining the remaining variation in the data on both synthetic and real neural recording datasets.

4.2 Method

4.2.1 Mathematical formulation

Let D be a data set of samples and V a task variable that induces variation in the data. Each sample in the dataset belongs to a category of V. The objective is to decompose the variation in D into three components: variation informative about the task variable V, variation non-informative about V, and noise. This is equivalent to decomposing each sample x into three terms:

$$x = x_v + x_{-v} + \varepsilon \tag{1}$$

where x_v is the component related to variable V, x_{-v} is the component unrelated to V, and ε is a random error term.

Note that Formula 1 assumes that an additive decomposition is feasible in the original input space. However, in general this might not be the case. To address this issue, it is possible to use a function ψ to transform the data and perform the decomposition in a new space where the components can be simply added together. Specifically:

$$\psi(x) = \psi(x)_v + \psi(x)_{-v} + \varepsilon \tag{2}$$

By replacing $\psi(x)_v$ and $\psi(x)_{-v}$ with $f(x)$ and $g(x)$, respectively, we get:

$$\psi(x) = f(x) + g(x) + \varepsilon \tag{3}$$

In Formula 3, $f(x)$ is a function that maps x to the component informative about variable V in the transformed space. Likewise, $g(x)$ is a function that maps x to the component uninformative about variable V in the transformed space.

To reconstruct x from its component in the transformed space, we have:

$$x = \psi^{-1}(f(x) + g(x)) + \varepsilon \tag{4}$$

Now, the objective becomes finding the three functions f, g, and ψ^{-1}. The functions f and g are learned separately but their outputs need to have identical dimensionalities. Also, because the final code in the transformed space is the sum of the outputs of these two functions, restricting the output elements to a predefined range can increase learning efficiency. We choose to restrict the range of each element in $f(x)$ and $g(x)$ to [-1, 1]. One way to accomplish this is by subtracting the mean of f over the dataset (denoted by $\bar{f}$) from $f(x)$ and then dividing each element of result by the maximum absolute value of the corresponding element over the whole dataset:

$$w_f = \max_{\mathrm{x}}(\left|\left(f(x) - \bar{f}\right)\right|) \tag{5}$$

$$f'(x) = \left(f(x) - \bar{f}\right) \oslash w_f \tag{6}$$

$$f'(x) = f(x) \oslash w_f - \bar{f} \oslash w_f \tag{7}$$

$$f'(x) = f(x) \oslash w_f - b_f \tag{8}$$

Where $\oslash$ represents element-wise division. The same process is repeated for $g(x)$, resulting in:

$$g'(x) = g(x) \oslash w_g - b_g \tag{9}$$

By rearranging the terms in (8) and (9) and using the resulting formulas in (4), we get:

$$x = \psi^{-1}(f'(x) \otimes w_f + b_f + g'(x) \otimes w_g + b_g) + \varepsilon \tag{10}$$

$$x = \psi^{-1}\{w_f \otimes (f'(x) + g'(x) \otimes w_g \oslash w_f + (b_f + b_g) \oslash w_f)\} + \varepsilon \tag{11}$$

Where $\otimes$ represents the element-wise multiplication. By incorporating w_f to the reverse transformation, replacing $w_g \oslash w_f$ by w, and replacing $(b_f + b_g) \oslash w_f)$ by b, we get:

$$x = \psi^{-1}(f'(x) + g'(x) \otimes w + b) + \varepsilon \tag{12}$$

The final objective is to find the encoding function $f'(x)$ that transforms each input sample to the corresponding component in the transformed space that depends on the task variable V, the encoding function $g'(x)$ that transforms each input sample to the corresponding component in the transformed space that does not depend on the variable V, the element-wise mixing weight w, the bias vector b, and finally the decoding function ψ^{-1} that reconstruct x from final code.

4.2.2 Network architecture

The proposed network is essentially an autoencoder consisting of two encoding branches, an aggregator that combines the codes produced by these two branches to form the final code, and a decoder that reconstruct the original input from the final code. The first encoding branch extracts features that are informative about a particular task variable (denoted by V). This is accomplished by transforming the data into a new space where variations in the samples are maximally informative about variable V. We refer to this branch as the discriminator branch. The second branch transforms the data into a new space in such a way that all information about variable V is removed while the remaining variation in the original data is retained. We refer to the second branch as the filter branch. The encoder in each branch is attached to a classifier. Each classifier assigns the input samples into different categories of variable V based on the codes produced by the encoder attached to it. Ideally, we want to observe a high accuracy for the classifier in the

discriminator branch and an accuracy not significantly higher than chance for the classifier in the filter branch. The first outcome shows that the transformed data in the discriminator branch is maximally informative about variable V. The second outcome indicates that variations informative about variable V have been removed from the transformed data in the filter branch.

The final code in the network is computed by a weighted sum of the codes produced by each branch (see Formula 10). The decoder uses this code to reconstruct the original input. The network is trained to minimize the reconstruction error. Because the discriminator branch is heavily regularized to only keep variations informative about variable V, minimizing the reconstruction error is equivalent to maximizing the retained variation in the filter branch (excluding variations informative about variable V).

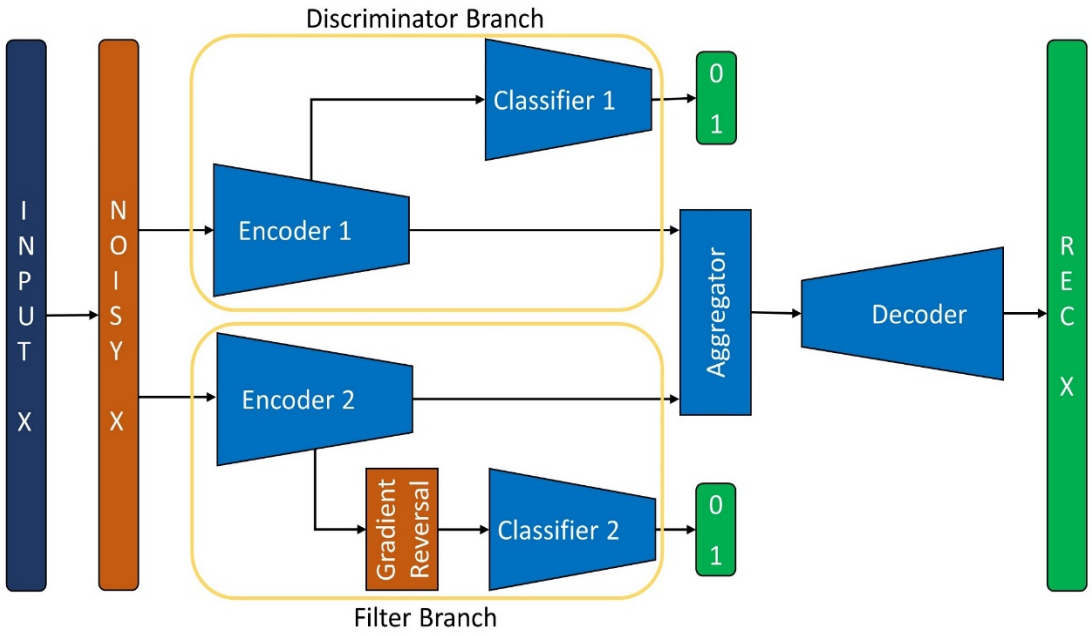


Figure 4.1. The high level architecture of the network

The main components of the network and its architecture are illustrated in Figure 4.1. Each component of the network is described below:

Noise Layer: selects a subset of cells according to a binomial distribution with parameter ρ and adds zero-mean Gaussian noise with standard deviation σ to the activity of those cells. In our experiments, adding large variance noise ($\sigma = 0.5$) to a few cells ($\rho = 0.05$) or adding low variance noise ($\sigma = 0.05$) to a higher ratio of cells ($\rho = 0.5$) produced the best results. The two encoders get the noisy samples as the input. To keep the notation uncluttered, the noisy version of x is referred to x itself in the rest of the document.

Encoder 1: transforms the input into a space in which variations are maximally informative for classification of variable V. Because we want only the variation informative about variable V to pass through this encoder, the parameters and the kernel weights of this component are regularized heavily. This ensures that the filter branch is a less costly way for the variations not informative about variable V to pass through. The output of encoder 1 is denoted as $enc_1(x)$. Learning the parameters of this encoder is equivalent to learning the function $f'(x)$ in Formula 12. The activation of the last layer of the encoder is a *tanh* function to restrict its output to [-1, 1].

Encoder 2: transforms the data into a space in which all information about variable V is removed while the rest of the structured variation in the data is retained. The architecture of the encoder 2 can be different from the architecture of encoder 1. However, the output layers of both encoders must have the same size and activation function. The output of encoder 2 is denoted as $enc_2(x)$. Learning the parameters of this encoder is equivalent to learning the function $g'(x)$ in Formula 12. Similar to the other encoder, the activation of the last layer of the encoder is a *tanh* function to restrict its output to [-1, 1].

Classifier 1: is attached to the encoder 1 and receives the output of that component as its input directly. Classifier 1 tries to detect the corresponding value of variable V for each input sample x from the encoded vector $enc_1(x)$. Backpropagation of the error signal from classifier 1 to encoder 1 improves the encoder's ability to extract features that are more discriminative between different classes of variable V. The output of classifier 1 is denoted as $cl_1(x)$.

Classifier 2: is attached to encoder 2 via a gradient reversal layer and receives the output of that component as its input after the input passes through the gradient reversal layer. Classifier 2 tries to detect the value of variable V for each input sample x from the encoded vector $enc_2(x)$. The output of the classifier 2 is denoted as $cl_2(x)$.

Gradient Reversal Layer: In the forward pass, the gradient reversal layer (Ganin and Lempitsky 2015) acts as the identity function. In the backward pass (backpropagation), it reverses the sign of the gradient it receives from the next layer and propagates the result to the previous layer. During training, the parameters of classifier 2 get updated by the normal gradient signal computed based on the cross entropy loss between predicted values and the true labels of the samples. These updates result in a decrease in the classification loss value and an increase in the accuracy of the classifier. However, the sign of the gradient changes when the gradient is back-propagated to encoder 2. When the parameters of encoder 2 get updated using this new gradient, the value of the loss function increases. As the result, after each update, classifier 2 improves in classifying samples based on the output of the encoder 2, however, encoder 2 gets better at removing information about variable V. Ideally, after the training is complete, there is no information about variable V in the output of encoder 2 and the accuracy of the classifier 2 is closed to chance.

Aggregator receives the outputs of the two encoders and computes a weighted sum of the codes according to:

$$agg(x) = enc_1(x) + w.enc_2(x) + b \quad (13)$$

Where both w and b are parameter vectors with the same size as the outputs of the two encoders. The output of the aggregator can be considered the final representation of the data in the transformed space. In this space, the signal can be decomposed into two parts: the part produced by encoder1 in the discriminating branch that explains the variation associated with variable V, and the second part, produced by the encoder in the filter branch that contains the remaining variation in the data. Learning the parameters of the aggregator is equivalent to learning the parameters w and b in Formula 12.

Note that there is no explicit constraint on the output of the two encoders to be in the same space. The architectures of the two networks can be different. The only constraints are that the output layers have the same number of units and identical activation functions. However, because the final code in the network is the weighted sum of the codes produced by the two encoders, the sum provides useful information for the decoder to reconstruct the original signal.

Decoder: is attached to the aggregator and reconstructs the original input x from the encoded vector $agg(x)$. The output of the decoder network is denoted as $dec(x)$. Learning the parameters of the decoder is equivalent of learning the function ψ^{-1} in Formula 12.

The reconstructed vector $\tilde{x}$ is obtained according to:

$$\tilde{x} = dec(enc_1(x) + w \otimes enc_2(x) + b) \quad (14)$$

Where $\otimes$ represents the element-wise multiplication. Note that in this formula x is actually the noisy version of the original input. This formula is equivalent to Formula 12 presented earlier.

The full objective function of the network is:

$$L = MSE(X, \tilde{X}) + BXE\left(Y, cl_1\left(enc_1(X)\right)\right) + BXE\left(Y, cl_2\left(enc_2(X)\right)\right) + Reg \quad (15)$$

Where MSE is the mean square error, BXE is the binary cross-entropy, and Reg is the term that results from the combination of all the regularization in the network. The network is trained end to end using an Adam optimizer.

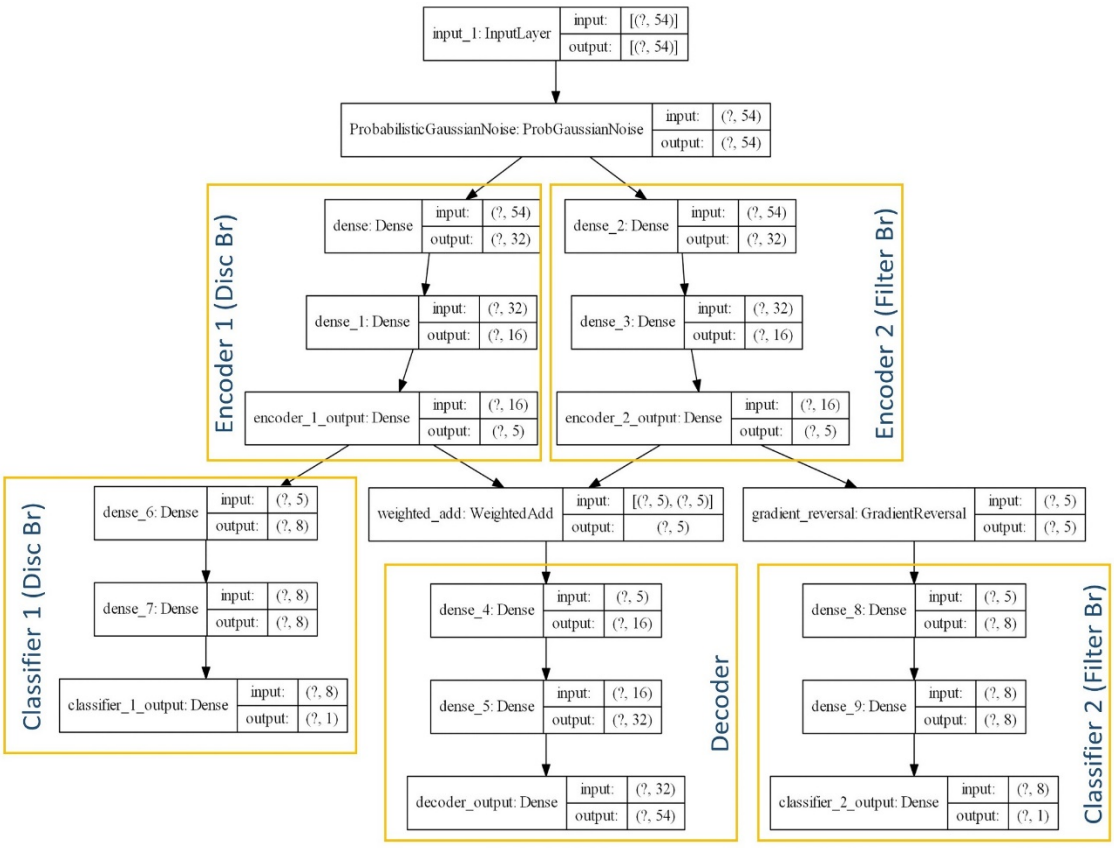


Figure 4.11. Detailed network architecture

The activation function of the last layers of both encoders is the *tanh* function. The activation function of the last layer of both classifiers is the *sigmoid* function. The activation function for all the other layers is Exponential Linear Unit (ELU), defined as:

$$ELU(x) = \begin{cases} x, & x \geq 0 \\ \alpha(e^x - 1), & x < 0 \end{cases} \quad (16)$$

With $\alpha = 1$. Using ELU instead of RELU produced superior results, perhaps because ELU can produce negative results and its derivative changes gradually from 1 to 0 as the input moves from positive values to negative values.

After the training is complete the combination of encoder and the classifier attached to it in the discriminator branch can be used as a standalone classifier for variable V. This feature allows us to train the network in a semi-supervised manner, by assigning weights to samples in different terms in the objective function. In this case, the weight of all samples in the reconstruction part would be 1, however, for the unlabeled samples, the weights in the classification term would be 0.

4.3 Results

4.3.1 Test on synthetic data

We created two synthetic datasets to evaluate the effectiveness of the proposed method. Each dataset comprised of 5000 samples, 3 features, and 2 labels associated with each sample, representing the categories in variables A and B. The first two features were randomly selected from a uniform distribution in the range of [-1, 1]. The third feature was randomly selected from a normal distribution with mean zero and unit variance.

In the first dataset, both labels of each sample were assigned based on the value of the first feature. The ranges of values for labeling were selected in such a way that classifying the instances based on feature 1 would result in an accuracy of 90% for variable A and an accuracy of 75% for variable B. We refer to this dataset as the "shared feature" case (Figure 4.3.A).

In the second dataset, variable A labels were assigned based on the value of the first feature. But, variable B labels were assigned based on the second feature. Again, the ranges of values for labeling were selected in such a way that classifying the instances based on feature 1 would result in an accuracy of 90% for variable A, and classification based on feature 2 would result in an accuracy of 75% for variable B. We refer to this dataset as the "independent feature" case (Figure 4.3.B).

We used the network to eliminate variation informative about variable B and computed the classification accuracy of variable A on both datasets from the filter branch of the netowrk. The architecture of the network for this task was identical to the network shown in Figure 4.2 except for the number of units in the last layer of each decoder which was set to 2. The network for 100 epochs on mini-batches of size 50. A 10-fold cross

validation was used to assess the average accuracies. For assessing classification accuracy on the filter branch, we used a logistic regression classifier.

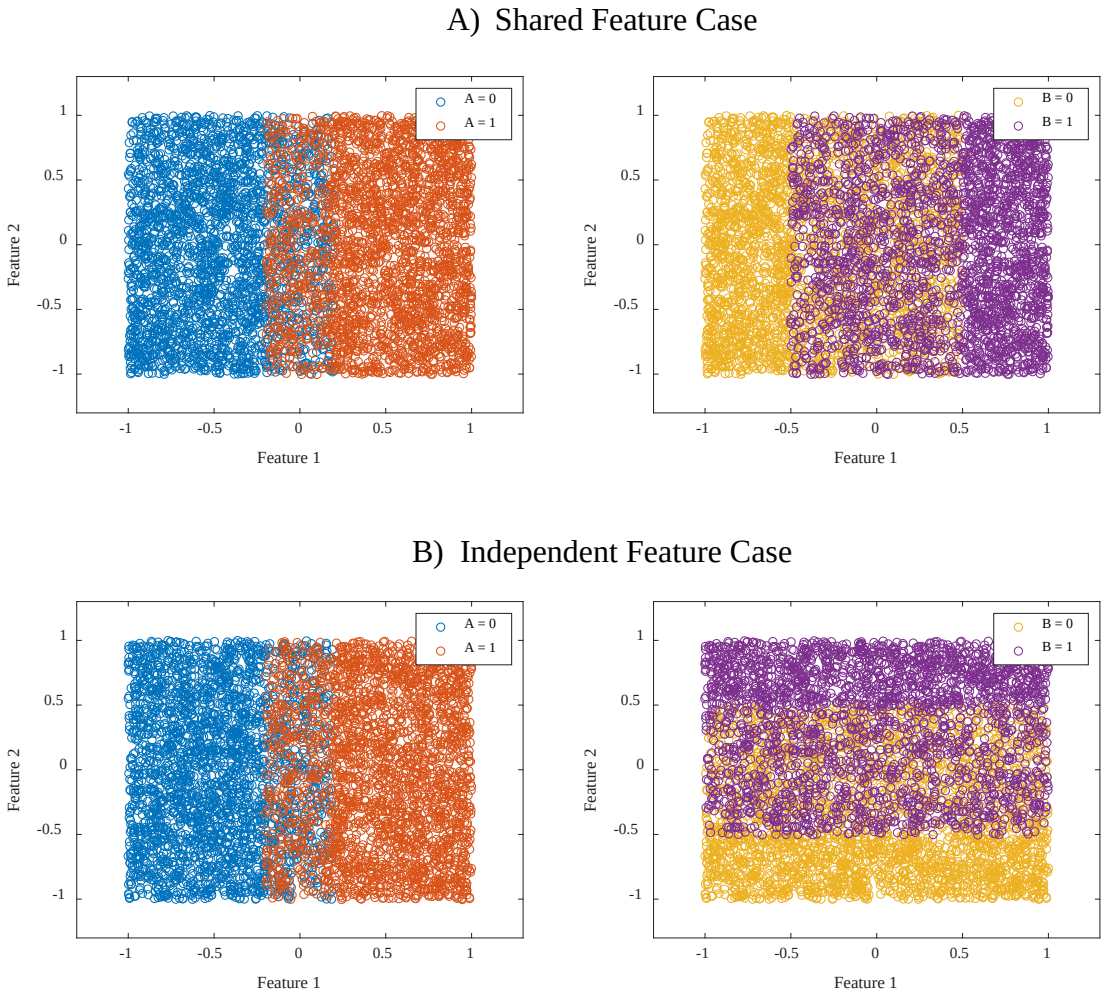


Figure 4.3. Illustration of the relationship between the first two features and class labels in each dataset.

In each case, the accuracy for decoding variable *B* from the discriminator branch of the network was not significantly lower than the accuracy achieved on the original data (p = 0.405; t(9) = 0.2475), indicating that information about variable *B* was retained in the code produced by this branch of the network. In contrast, for each condition, the decoding accuracy of variable *B* from the filter branch of the network was not significantly above

chance (p = 0.9845; t(9) = 2.5539), suggesting that variation informative about variable B was successfully removed from the data in the filter branch (Figure 4.4).

In the shared feature case, the average accuracy of decoding variable A from the filter branch was 54.3% (std = 2.97) which is significantly lower than the accuracy achieved on the original data (p = 0; t(9) = -37.925) (Figure 4.4.A). This behaviour was expected because in this case the same feature was informative about both variables A and B. Therefore, removing variation informative about variable B would inevitably result in a significant reduction in accuracy of decoding variable A.

In contrast, in the independent feature case, the average accuracy of decoding variable A from the filter branch was 88.3% (std = 1.34) which is very close to the maximum achievable accuracy (90%) on the original data (Figure 4.4.B). This results shows while the network removed the information about variable B in the filter branch, it was able to retain variation informative about variable A. This behaviour was expected because in this case the features informative about the two variables were separate.

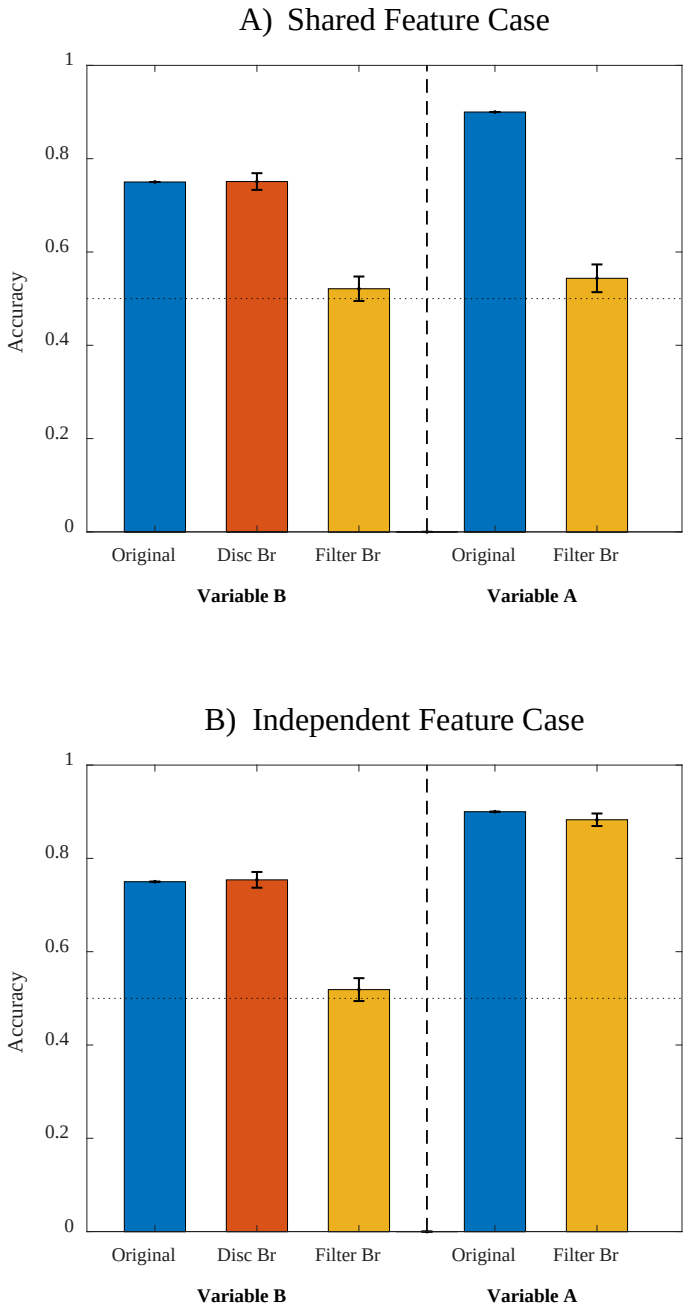


Figure 4.4. Accuracy of decoding the variables A and B from the original versus the transformed data. The bars labeled Original show the accuracies obtained on the original data. The bars labeled Disc Br and Filter Br show the accuracies obtained on the transformed data in the discriminator branch and the filter branch of the network, respectively.

4.3.2 Test on real neural data

We used the neural data recorded for the experiment in Chapter 3 to test the effectiveness of the method. We used the network to eliminate the variation informative about choice when the animals were in the bins associated with the side feeders and tested if it was possible to decode reward outcome from the transformed activity.

The average firing rates of the cells when the animals were at the side feeder were fed to the network as input. The choice label associated with each trial was provided as the target value in both branches of the network. The network was trained for 300 epochs on mini-batches of size 20. We expected the encoder in the discriminator branch to extract features informative about choice and the encoder in the filter branch to transform the data in such a way that choice information was eliminated while all the remaining variations, including the reward outcome information, were retained.

After training was complete, we computed the decoding accuracy of choice from the codes produced in both branches of the network and the decoding accuracy of reward outcome from the code in the second branch. We estimated the maximum achievable accuracy in each case using the same procedure explained in section 3.4.8.

We found that the accuracy for decoding choice from the discriminator branch of the network was not significantly lower than the accuracy achieved on the original data ($p = 0.283$; $t(4) = 0.629$), indicating that choice information was retained in the code produced by this branch of the network (Figure 4.5). In contrast, the decoding accuracy of choice from the filter branch of the network was not significantly above chance ($p = 0.895$; $t(4) = -1.499$), suggesting that choice information was removed from the data in that branch, and the transformation had made the data invariant to choice. Finally, the decoding accuracy of

reward outcome from the filter branch of the network was not significantly lower than the accuracy achieved on the original data ($p = 0.988$; $t(4) = -3.526$), showing that the outcome information was retained in the data after transformation. Overall, these results indicate that the network was able to eliminate variation related to choice while retaining outcome information in the data.

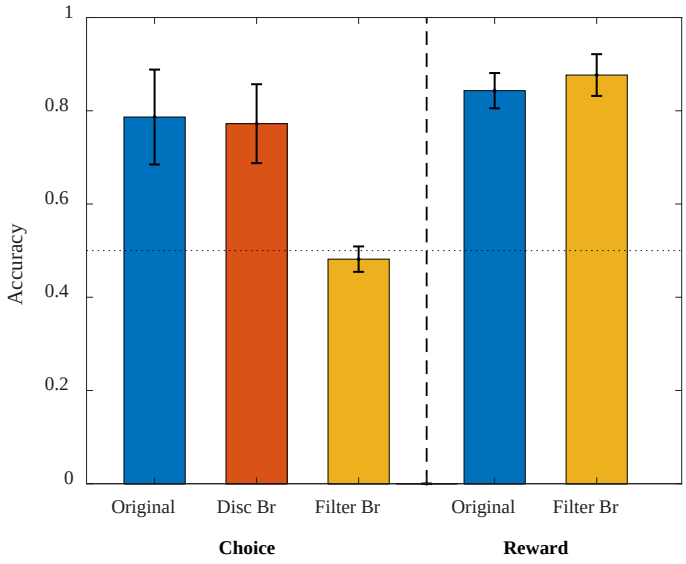


Figure 4.5. Accuracy of decoding choice and reward outcome from original data versus transformed data. The bars labeled Original show the accuracies obtained on the original data. The bars labeled Disc Br and Filter Br show the accuracies obtained on the transformed data in the discriminator branch and the filter branch of the network, respectively.

Figure 4.6 illustrates examples of the codes that were produced by the two branches of the network. The dimensionality of the code was reduced to two using PCA for better visualization. Samples with different colors have different choice labels. As shown, samples from the two classes are completely separated in the code produced in the discriminator branch. In contrast, the samples are mixed and indistinguishable in codes produced in the filter branch.

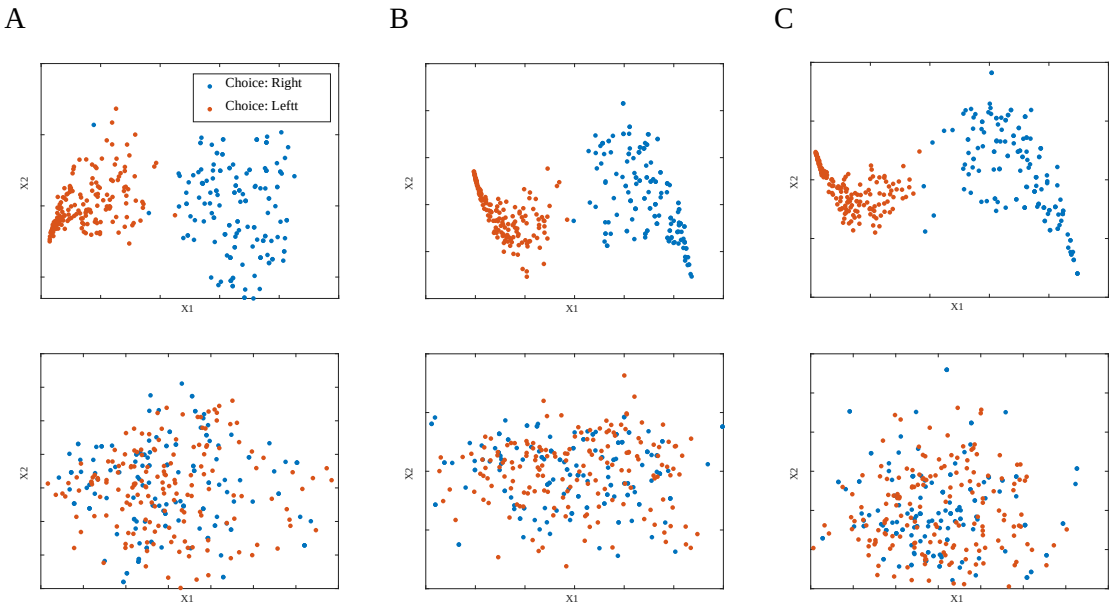


Figure 4.6. Examples of the codes produced in the two branches of the network. Top row: the code in the discriminator branch. Samples having different choice labels are completely separable. Bottom row: The code in the filter branch. Samples having different choice labels are not distinguishable.

5 General Discussion

In this thesis, I analyzed electrophysiological recording from rat ACC in two separate binary-choice tasks to test multiple hypotheses supporting the role of the ACC in model-based control. Specifically, I tested: 1) whether the ACC represents a map of the task space, which in turn raises the possibility of mental navigation of the map, 2) whether the ACC encodes variables that define actions, outcomes, and task states, 3) the ACC's ability to represent abstract task states that are critical for future predictions, and finally, 4) the presence of signals that indicate a deviation between the predictions and actual outcome that may promote model update.

5.1 The ACC represents the task space in a spatial framework

A model-based approach for tackling a task requires knowledge about task states and their relationships. In addition, an agent needs to locate itself within this abstract state space as it performs the task. In an experiment in which task states correlate with spatial coordinates, this may translate into a representation of the current position of the animal. Our results showed that the position of a rat on a track could be decoded accurately from the population activity of ACC neurons. This finding accords with other reports of spatial encoding in the mPFC (Fujisawa et al. 2008; Lindsay et al. 2018; Zielinski, Shin, and Jadhav 2019; D. R. Euston and McNaughton 2006; Cowen, Davis, and Nitz 2012). However, by using a deep neural network as the decoder and analyzing population activity instead of single-cell analysis, we were able to obtain significantly more accurate results.

Spatial encoding in the ACC is facilitated through the connections to the hippocampus and retrosplenial cortex (Burton et al. 2009; Zielinski, Shin, and Jadhav 2019). However, our results along with others show that the representation of spatial information in the ACC differs from that of the hippocampus. In the dorsal hippocampus, representation of space is largely encoded by the activity of so-called 'place cells', with each cell generating action potentials only in a small region of the environment (O'Keefe and Dostrovsky 1971; M. A. Wilson and McNaughton 1993). In contrast, individual neurons in the ACC have much broader spatial sensitivity (Fujisawa et al. 2008). For example, we found cells that were active when an animal was on one side arm of the maze and cells that were active over multiple disconnected regions of the maze. Despite the low spatial selectivity of individual ACC neurons, we found that the position of an animal can be decoded accurately from population activity. This highly distributed representation in the ACC has been reported by a recent study that investigated the encoding of limb movements, poses, and spatial locations of rats (Lindsay et al. 2018).

Representation of position in the ACC is most likely an abstraction (or a correlate) of the task space rather than an absolute specification of the position in Euclidean coordinates. This claim is supported by our observation that the accuracy of position decoding was significantly lower in the trials in which rats became distracted and disengaged from the task for a brief period. This view is also consistent with other reports of generalized representation of space in the rodent ACC while performing a task (Kaefer et al. 2020) and during awake memory reactivation (Yu et al. 2018). Kaefer et al. (2020) found a generalized representation of space in the ACC, where instead of coding for specific locations, neurons encode the relative position between the start and the goal. It is, therefore, plausible that one use of forming a state-space map in a high-resolution spatial

framework is to track progress towards a goal state, thereby keeping the animal motivated and preventing disengaging from the task.

A model of the task organized around space might facilitate the formation of knowledge structures or schemas. Once such knowledge structures are formed, new information that fits prior knowledge can be encoded and consolidated more efficiently than when information is novel (Van Kesteren et al. 2012; Gilboa and Marlatte 2017). For example, rats were found to more rapidly encode and consolidate new information when that information fit within a spatial schema (Tse et al. 2007; Hasan et al. 2019).

Overall, being able to decode the position of the animal as it performs the task indicates that the ACC creates a map of the task space and tracks the current state of the animal as it performs the task. This idea is consistent with previous studies that show that the medial prefrontal region encodes a trajectory through the task space (Lapish et al. 2008; Durstewitz et al. 2010) and demonstrates that, in some situations, such a state-space trajectory is strongly correlated with the real spatial coordinates. The representation of all states in a two-step task in mice ACC has been reported a recent study (Akam et al. 2020). This is similar to the proposal that the orbitofrontal cortex encapsulates the available information by creating abstract task-states to facilitate reinforcement learning in other brain regions. This process is particularly important when task states include unobservable information (R. C. Wilson et al. 2014).

5.2 The ACC mentally navigates the map of the task

Having a model of the task state-space opens up the possibility that the brain mentally navigates the ACC map to recall information or even simulate states sequences

consistent with a world model. We found that while the decoded position normally tracked the current state of the rat, it sometimes dissociated from its physical position when the rat was at one of the target feeders. These excursions represented the state associated with the other target feeder and were more likely at the disfavored target feeder. This pattern of activity suggests a role for the ACC in the evaluation of outcomes following disappointing reinforcements and can reflect a mechanism for shifting strategies. Supporting this view, a recent study found reactivation of neural patterns associated with the sequences of positions distant from the actual position in rat mPFC during awake immobility in a rule-switching task (Kaefer et al. 2020). The occurrence rate of mPFC replay at the goal predicted how fast the animal switched to the new rule. Although, a single replay event was not immediately followed by improved performance on the next trial.

The excursions reported in this research could represent a more general form of the counterfactual outcome processing in primates. ACC neurons in monkeys have been shown to signal fictive outcomes, which are potential rewards that were not achieved (Hayden, Pearson, and Platt 2009), and the value of counter factual choice, which is the best value among available or unavailable alternatives (Fouragnan et al. 2019). The excursions could support the comparison of realized and fictive outcomes in a multimodal way where instead of signaling only the counter factual value, the ACC represents and recalls the counter factual states that includes value and other features (choice, cost, etc.).

5.3 The ACC tracks task variables relevant to creation of a predictive model

In order to create a predictive model for a task, the ACC needs to track task variables that define and represent context, stimulus, responses, and outcomes. Our results

demonstrated that the ACC cells distinguished between reward outcomes in both low versus high reward (in Chapter 2) and omission versus delivery categories (Chapter 3). ACC neurons discriminated between free and forced trials when sensory cues about trial type were present. Moreover, although ACC activity was not predictive of choice during immobility at the beginning of trials, it showed a clear representation of choice after animals departed from the start feeder. These results are consistent with previous reports of ACC's encoding of task variables such as outcome (Hayden and Platt 2010; C. Amiez, Joseph, and Procyk 2006; Kennerley and Wallis 2009a; Hashemnia, Euston, and Gruber 2020), effort level (Cowen, Davis, and Nitz 2012; Hashemnia, Euston, and Gruber 2020; Croxson et al. 2009; Klein-Flugge et al. 2016; Prevost et al. 2010; Porter, Hillman, and Bilkey 2019), and actions (Hayden and Platt 2010; C. Amiez, Joseph, and Procyk 2006; C. H. Luk and Wallis 2009; Quilodran, Rothé, and Procyk 2008). Overall, a large body of evidence indicate that the ACC encodes the value of relevant task variables directly perceivable in a task.

The ACC encodes task variables in a dynamic and context-dependent manner. For instance, ACC's representation of identical actions have been shown to differ depending on the context (Hyman et al. 2012) and order of that action in a sequence within the same context (L. Ma et al. 2016). Also, ACC's response to the receipt of a reward might vary depending on the value of the alternatives (Hillman and Bilkey 2010). Furthermore, whether ACC shows sensitivity to a particular variable at all or not is context-dependent and is contingent on the relevance of that variable to the current task (Heilbronner and Hayden 2016). Therefore, ACC's response to stimulus, outcomes, action, and other task variables is an abstract context-dependent representation linked to higher level cognitive

processes rather than a representation of the sensory inputs and motoric responses associates with those task variables.

5.4 The ACC infers abstract task states relevant to creation of a predictive model

In the real world, a single event rarely provides a clear picture of what is happening in the environment. Instead, it is a combination of the relationship between the events, their context-dependent implications, and their significance in light of past experience that is critical for successful behaviour. To create a predictive model, the ACC needs to create abstract states that are inferred from the history of events and their relationship. Our result showed that in a binary-choice task with probabilistic reward delivery, rats were able to track the high-utility option by integrating the outcomes over multiple trials. The signal representing the high utility feeder (block type signal) was most pronounced in the ACC during immobility at the start feeder and weakened as the animals started running. The coincidence of block type signal's strength with the weakness of choice signal suggests that the block type signal is used in the process of decision making. The presence of signals representing abstract states in a task has been reported by other studies. The ACC has been shown to represent the value of the counter factual choice (Fouragnan et al. 2019), the order of an action in a sequence (L. Ma et al. 2016), progress towards a goal (Liya Ma et al. 2014; Shidara and Richmond 2002), utility of options (Kennerley and Wallis 2009b; Hashemnia 2019), and task rules (Kaefer et al. 2020).

5.5 ACC represent surprise signals that could trigger model update

As the contingencies in an environment change, the predictive model formed by the ACC may not remain valid. Thus, the model needs to be updated. The necessity of such adjustments can be detected when actual outcomes deviate from predictions of the model. Such events lead to prediction errors and surprises. Our results showed evidence of the presence of such surprise signals in the ACC. Specifically, we found that ACC's response to identical outcomes differed depending on the previous outcome. In the task, the probability of congruent outcomes in consecutive trials was much higher than incongruent outcomes, therefore the latter was surprising. Also, such events increased the probability that the high utility option had changed, promoting a shift in an animal's choice bias. This finding accords with other reports of the prominence of reward prediction error (RPE) and surprise signals (unsigned RPE) in the ACC in studies on both humans (O'Reilly et al. 2013; Vassena, Deraeve, and Alexander 2020; Cavanagh et al. 2012; Wessel et al. 2012) and non-human primates (Hayden et al. 2011; Kennerley, Behrens, and Wallis 2011; M. Matsumoto et al. 2007; Seo and Lee 2007). These signals are proposed to trigger model update and shifts in strategies (Hayden et al. 2011; Quilodran, Rothé, and Procyk 2008).

5.6 ACC as a model-based controller

Taken together, our results support the role of the ACC in model-based control. In this view, the ACC creates a model of the task based on experience and uses the model to predict the future states and outcomes given the context and response options. These predictions will be used to guide decision making. This view is similar to the previous proposals that assign a role to the mPFC in creating models which map context and events

onto appropriate emotional or motoric responses (Miller and Cohen 2001; Alexander and Brown 2011; David R. Euston, Gruber, and McNaughton 2012). This model provides a different, but compatible, perspective on the ACC function compared with other theories that assign the ACC a main role in conflict monitoring (M. M. Botvinick, Cohen, and Carter 2004), representing the volatility of environment (Behrens et al. 2007), coding for choice difficulty (Shenhav et al. 2014), and effort allocation (Verguts, Vassena, and Silvetti 2015).

The role of the ACC in model-based control is supported in a very recent study that showed that, in a sequential decision task, mouse ACC predicts the state that actions will lead to, rather than estimating whether the actions are rewarding or not, and monitors for deviations between these predictions and actual outcomes (Akam et al. 2020). Also, using optogenetic inhibition of the ACC, they showed that the ACC is necessary for updating model-based strategies.

The view of the ACC as a model-based controller is closely related to the model-based reinforcement learning in the field of machine learning (Sutton and Barto 1998). The predicted response outcome theory (Alexander and Brown 2011) proposes ACC to be a model-based leaner with an actor-critic architecture. In this proposal, the critic component predicts the likelihood of future actions and outcomes in a context regardless of their affective valence based on stimulus input from the environment. When these predictions deviate from actual outcomes a prediction error is produced, resulting in the update of the model for future reference. The actor is responsible for response selection. This model has been used to simulate a wide variety of empirical findings, including the effects of conflict, errors, and reward prediction errors (Alexander and Brown 2014; Jahn et al. 2014). A variant of the same model has been proposed to explain effortful control (Vassena, Deraeve, and Alexander 2017).

Although, some proposals assign ACC a role in actions selection (K. Matsumoto, Suzuki, and Tanaka 2003; M. F.S. Rushworth et al. 2004), our results did not show ACC activity to be predictive of choice. One potential explanation is that different regions of the ACC are responsible for prediction and action selection. Supporting this view, a human fMRI study linked evaluative and predictive functions to anterior rostral cingulate zone and response selection to more causal and rostral regions of the ACC (Nee, Kastner, and Brown 2011). The alternative explanation is that the ACC is a part of a distributed choice process rather than a main center for response selection, and its involvement in action selection is task dependent (Heilbronner and Hayden 2016b).

Although most research focuses on the role of the ACC in a single task, the view of the ACC as a model-based controller opens up the possibility that this region plays a role in learning long-horizon behaviours by simulating scenarios in a high-level abstract state-space using the world model formed based on previous direct or indirect experiences (e.g., observation). Such learning models have been proposed in the machine learning field and shown to be very promising in planning long-term action plans (Hafner et al. 2019). This proposal is consistent with the role of the ACC in maintaining long-term goals (Clay B. Holroyd and Yeung 2012), but further research is needed to elucidate ACC's role in long-term planning.

REFERENCES

Akam, Thomas, Ines Rodrigues-Vaz, Ivo Marcelo, Xiangyu Zhang, Michael Pereira, Rodrigo Freire Oliveira, Peter Dayan, and Rui M. Costa. 2020. "The Anterior Cingulate Cortex Predicts Future States to Mediate Model-Based Action Selection." *Neuron*. https://doi.org/10.1016/j.neuron.2020.10.013.

Akkal, D., B. Bioulac, J. Audin, and Pierre Burbaud. 2002. "Comparison of Neuronal Activity in the Rostral Supplementary and Cingulate Motor Areas during a Task with Cognitive and Motor Demands." *European Journal of Neuroscience*. https://doi.org/10.1046/j.1460-9568.2002.01920.x.

Alexander, William H., and Joshua W. Brown. 2011. "Medial Prefrontal Cortex as an Action-Outcome Predictor." *Nature Neuroscience*. https://doi.org/10.1038/nn.2921.

———. 2014. "A General Role for Medial Prefrontal Cortex in Event Prediction." *Frontiers in Computational Neuroscience*. https://doi.org/10.3389/fncom.2014.00069.

Amemori, Ken Ichi, and Ann M. Graybiel. 2012. "Localized Microstimulation of Primate Pregenual Cingulate Cortex Induces Negative Decision-Making." *Nature Neuroscience*. https://doi.org/10.1038/nn.3088.

Amiez, C., J. P. Joseph, and E. Procyk. 2006. "Reward Encoding in the Monkey Anterior Cingulate Cortex." *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhj046.

Amiez, Céline, Jean Paul Joseph, and Emmanuel Procyk. 2005. "Anterior Cingulate Error-Related Activity Is Modulated by Predicted Reward." *European Journal of Neuroscience*. https://doi.org/10.1111/j.1460-9568.2005.04170.x.

Amiez, Céline, Jérôme Sallet, Emmanuel Procyk, and Michael Petrides. 2012. "Modulation of Feedback Related Activity in the Rostral Anterior Cingulate Cortex during Trial and Error Exploration." *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2012.06.023.

Balleine, Bernard W., and Anthony Dickinson. 1998. "Goal-Directed Instrumental Action: Contingency and Incentive Learning and Their Cortical Substrates." In *Neuropharmacology*. https://doi.org/10.1016/S0028-3908(98)00033-1.

Barbas, H., and D. N. Pandya. 1989. “Architecture and Intrinsic Connections of the Prefrontal Cortex in the Rhesus Monkey.” *Journal of Comparative Neurology*. https://doi.org/10.1002/cne.902860306.

Barker, Gareth R.I., Flora Bird, Victoria Alexander, and E. Clea Warburton. 2007. “Recognition Memory for Objects, Place, and Temporal Order: A Disconnection Analysis of the Role of the Medial Prefrontal Cortex and Perirhinal Cortex.” *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.5289-06.2007.

Bechara, Antoine, and Antonio R. Damasio. 2005. “The Somatic Marker Hypothesis: A Neural Theory of Economic Decision.” *Games and Economic Behavior*. https://doi.org/10.1016/j.geb.2004.06.010.

Behrens, Timothy E.J., Mark W. Woolrich, Mark E. Walton, and Matthew F.S. Rushworth. 2007. “Learning the Value of Information in an Uncertain World.” *Nature Neuroscience*. https://doi.org/10.1038/nn1954.

Birrell, Jennifer M., and Verity J. Brown. 2000. “Medial Frontal Cortex Mediates Perceptual Attentional Set Shifting in the Rat.” *Journal of Neuroscience*. https://doi.org/10.1523/jneurosci.20-11-04320.2000.

Blanchard, Tommy C., and Benjamin Y. Hayden. 2014. “Neurons in Dorsal Anterior Cingulate Cortex Signal Postdecisional Variables in a Foraging Task.” *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.3151-13.2014.

Blanchard, Tommy C., Caleb E. Strait, and Benjamin Y. Hayden. 2015. “Ramping Ensemble Activity in Dorsal Anterior Cingulate Neurons during Persistent Commitment to a Decision.” *Journal of Neurophysiology*. https://doi.org/10.1152/jn.00711.2015.

Boorman, Erie D., Timothy E.J. Behrens, Mark W. Woolrich, and Matthew F.S. Rushworth. 2009. “How Green Is the Grass on the Other Side? Frontopolar Cortex and the Evidence in Favor of Alternative Courses of Action.” *Neuron*. https://doi.org/10.1016/j.neuron.2009.05.014.

Boorman, Erie D., Matthew F. Rushworth, and Tim E. Behrens. 2013. “Ventromedial Prefrontal and Anterior Cingulate Cortex Adopt Choice and Default Reference Frames during Sequential Multi-Alternative Choice.” *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.3022-12.2013.

Botvinick, Matthew M., Cameron S. Carter, Todd S. Braver, Deanna M. Barch, and Jonathan D. Cohen. 2001. “Conflict Monitoring and Cognitive Control.” *Psychological Review*. https://doi.org/10.1037/0033-295X.108.3.624.

Botvinick, Matthew M., Jonathan D. Cohen, and Cameron S. Carter. 2004. “Conflict Monitoring and Anterior Cingulate Cortex: An Update.” *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2004.10.003.

Botvinick, Matthew, Leigh E. Nystrom, Kate Fissell, Cameron S. Carter, and Jonathan D. Cohen. 1999. “Conflict Monitoring versus Selection For-Action in Anterior Cingulate Cortex.” *Nature*. https://doi.org/10.1038/46035.

Brown, Emery N., Loren M. Frank, Dengda Tang, Michael C. Quirk, and Matthew A. Wilson. 1998. “A Statistical Paradigm for Neural Spike Train Decoding Applied to Position Prediction from Ensemble Firing Patterns of Rat Hippocampal Place Cells.” *Journal of Neuroscience*. https://doi.org/10.1523/jneurosci.18-18-07411.1998.

Brown, Joshua W., and Todd S. Braver. 2005. “Learned Predictions of Error Likelihood in the Anterior Cingulate Cortex.” *Science*. https://doi.org/10.1126/science.1105783.

Bryden, Daniel W., Emily E. Johnson, Steven C. Tobia, Vadim Kashtelyan, and Matthew R. Roesch. 2011. “Attention for Learning Signals in Anterior Cingulate Cortex.” *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.4715-11.2011.

Buckner, Randy L., Jessica R. Andrews-Hanna, and Daniel L. Schacter. 2008. “The Brain’s Default Network: Anatomy, Function, and Relevance to Disease.” *Annals of the New York Academy of Sciences*. https://doi.org/10.1196/annals.1440.011.

Buckner, Randy L., and Daniel C. Carroll. 2007. “Self-Projection and the Brain.” *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2006.11.004.

Burton, Brian G., Vincent Hok, Etienne Save, and Bruno Poucet. 2009. “Lesion of the Ventral and Intermediate Hippocampus Abolishes Anticipatory Activity in the Medial Prefrontal Cortex of the Rat.” *Behavioural Brain Research*. https://doi.org/10.1016/j.bbr.2008.11.045.

Cai, X., and C. Padoa-Schioppa. 2012. “Neuronal Encoding of Subjective Value in Dorsal and Ventral Anterior Cingulate Cortex.” *Journal of Neuroscience*. https://doi.org/10.1523/jneurosci.3864-11.2012.

Camille, Nathalie, Ami Tsuchida, and Lesley K. Fellows. 2011. “Double Dissociation of Stimulus-Value and Action-Value Learning in Humans with Orbitofrontal or Anterior Cingulate Cortex Damage.” *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.3164-11.2011.

Carr, Margaret F., Shantanu P. Jadhav, and Loren M. Frank. 2011. “Hippocampal Replay in the Awake State: A Potential Substrate for Memory Consolidation and Retrieval.” In *Nature Neuroscience*. https://doi.org/10.1038/nn.2732.

Cavanagh, James F., Christina M. Figueroa, Michael X. Cohen, and Michael J. Frank. 2012. “Frontal Theta Reflects Uncertainty and Unexpectedness during Exploration and Exploitation.” *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhr332.

Chudasama, Yogita, Teresa E. Daniels, Daniel P. Gorrin, Sarah E.V. Rhodes, Peter H. Rudebeck, and Elisabeth A. Murray. 2013. “The Role of the Anterior Cingulate Cortex in Choices Based on Reward Value and Reward Contingency.” *Cerebral Cortex (New York, N.Y. : 1991)*. https://doi.org/10.1093/cercor/bhs266.

Comon, Pierre. 1994. “Independent Component Analysis, A New Concept?” *Signal Processing*. https://doi.org/10.1016/0165-1684(94)90029-9.

Coricelli, Giorgio, Hugo D. Critchley, Mateus Joffily, John P. O’Doherty, Angela Sirigu, and Raymond J. Dolan. 2005. “Regret and Its Avoidance: A Neuroimaging Study of Choice Behavior.” *Nature Neuroscience*. https://doi.org/10.1038/nn1514.

Cowen, Stephen L., Glen A. Davis, and Douglas A. Nitz. 2012. “Anterior Cingulate Neurons in the Rat Map Anticipated Effort and Reward to Their Associated Action Sequences.” *Journal of Neurophysiology*. https://doi.org/10.1152/jn.01012.2011.

Croxson, P. L., M. E. Walton, J. X. O’Reilly, T. E. J. Behrens, and M. F. S. Rushworth. 2009. “Effort-Based Cost-Benefit Valuation and the Human Brain.” *Journal of Neuroscience*. https://doi.org/10.1523/jneurosci.4515-08.2009.

Dalley, Jeffrey W., Rudolf N. Cardinal, and Trevor W. Robbins. 2004. “Prefrontal Executive and Cognitive Functions in Rodents: Neural and Neurochemical Substrates.” In *Neuroscience and Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2004.09.006.

Dragoi, George, and Susumu Tonegawa. 2013. “Distinct Preplay of Multiple Novel Spatial Experiences in the Rat.” *Proceedings of the National Academy of Sciences of*

the United States of America. https://doi.org/10.1073/pnas.1306031110.

Duncan, John. 2001. "An Adaptive Coding Model of Neural Function in Prefrontal Cortex." *Nature Reviews Neuroscience*. https://doi.org/10.1038/35097575.

Durstewitz, Daniel, Nicole M. Vittoz, Stan B. Floresco, and Jeremy K. Seamans. 2010. "Abrupt Transitions between Prefrontal Neural Ensemble States Accompany Behavioral Transitions during Rule Learning." *Neuron*. https://doi.org/10.1016/j.neuron.2010.03.029.

Egner, Tobias. 2011. "Surprise! A Unifying Model of Dorsal Anterior Cingulate Function?" *Nature Neuroscience*. https://doi.org/10.1038/nn.2932.

Engström, Maria, Anne-Marie Landtblom, and Thomas Karlsson. 2013. "Brain and Effort: Brain Activation and Effort-Related Working Memory in Healthy Participants and Patients with Working Memory Deficits." *Frontiers in Human Neuroscience*. https://doi.org/10.3389/fnhum.2013.00140.

Euston, D. R., and B. L. McNaughton. 2006. "Apparent Encoding of Sequential Context in Rat Medial Prefrontal Cortex Is Accounted for by Behavioral Variability." *Journal of Neuroscience*. https://doi.org/10.1523/jneurosci.3803-06.2006.

Euston, David R., Aaron J. Gruber, and Bruce L. McNaughton. 2012. "The Role of Medial Prefrontal Cortex in Memory and Decision Making." *Neuron*. https://doi.org/10.1016/j.neuron.2012.12.002.

Euston, David R., Masami Tatsuno, and Bruce L. McNaughton. 2007. "Fast-Forward Playback of Recent Memory Sequences in Prefrontal Cortex during Sleep." *Science*. https://doi.org/10.1126/science.1148979.

Feierstein, Claudia E., Michael C. Quirk, Naoshige Uchida, Dara L. Sosulski, and Zachary F. Mainen. 2006. "Representation of Spatial Goals in Rat Orbitofrontal Cortex." *Neuron*. https://doi.org/10.1016/j.neuron.2006.06.032.

FitzGerald, Thomas H.B., Ben Seymour, and Raymond J. Dolan. 2009. "The Role of Human Orbitofrontal Cortex in Value Comparison for Incommensurable Objects." *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.0717-09.2009.

Floden, Darlene, and Donald T. Stuss. 2006. "Inhibitory Control Is Slowed in Patients

with Right Superior Medial Frontal Damage." *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn.2006.18.11.1843.

Floresco, Stan B., and Sarvin Ghods-Sharifi. 2007. "Amygdala-Prefrontal Cortical Circuitry Regulates Effort-Based Decision Making." *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhj143.

Foster, David J., and Matthew A. Wilson. 2006. "Reverse Replay of Behavioural Sequences in Hippocampal Place Cells during the Awake State." *Nature*. https://doi.org/10.1038/nature04587.

Fouragnan, Elsa F., Bolton K.H. Chau, Davide Folloni, Nils Kolling, Lennart Verhagen, Miriam Klein-Flügge, Lev Tankelevitch, et al. 2019. "The Macaque Anterior Cingulate Cortex Translates Counterfactual Choice Value into Actual Behavioral Change." *Nature Neuroscience*. https://doi.org/10.1038/s41593-019-0375-6.

Friedman, Alexander, Daigo Homma, Leif G. Gibb, Ken Ichi Amemori, Samuel J. Rubin, Adam S. Hood, Michael H. Riad, and Ann M. Graybiel. 2015. "A Corticostriatal Path Targeting Striosomes Controls Decision-Making under Conflict." *Cell*. https://doi.org/10.1016/j.cell.2015.04.049.

Fujisawa, Shigeyoshi, Asohan Amarasingham, Matthew T. Harrison, and György Buzsáki. 2008. "Behavior-Dependent Short-Term Assembly Dynamics in the Medial Prefrontal Cortex." *Nature Neuroscience*. https://doi.org/10.1038/nn.2134.

Furuyashiki, Tomoyuki, Peter C. Holland, and Michela Gallagher. 2008. "Rat Orbitofrontal Cortex Separately Encodes Response and Outcome Information during Performance of Goal-Directed Behavior." *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.0319-08.2008.

Ganin, Yaroslav, and Victor Lempitsky. 2015. "Unsupervised Domain Adaptation by Backpropagation." In *32nd International Conference on Machine Learning, ICML 2015*.

Gehring, William J., Brian Goss, Michael G.H. Coles, David E. Meyer, and Emanuel Donchin. 1993. "A Neural System for Error Detection and Compensation." *Psychological Science*. https://doi.org/10.1111/j.1467-9280.1993.tb00586.x.

Gehring, William J., and Adrian R. Willoughby. 2002. "The Medial Frontal Cortex and the Rapid Processing of Monetary Gains and Losses." *Science*.

https://doi.org/10.1126/science.1066893.

Gehring, William J, Michael G H Coles, David E Meyer, and Emanuel Donchin. 1990. "The Error-Related Negativity: An Event-Related Brain Potential Accompanying Errors." *Psychophysiology*.

Gemba, H., K. Sasaki, and V. B. Brooks. 1986. "'Error' Potentials in Limbic Cortex (Anterior Cingulate Area 24) of Monkeys during Motor Learning." *Neuroscience Letters*. https://doi.org/10.1016/0304-3940(86)90467-2.

Gilboa, Asaf, and Hannah Marlatte. 2017. "Neurobiology of Schemas and Schema-Mediated Memory." *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2017.04.013.

Glimcher, Paul W., and Ernst Fehr. 2013. *Neuroeconomics: Decision Making and the Brain: Second Edition*. *Neuroeconomics: Decision Making and the Brain: Second Edition*. https://doi.org/10.1016/C2011-0-05512-6.

Goodfellow, Ian J., Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. "Generative Adversarial Nets." In *Advances in Neural Information Processing Systems*. https://doi.org/10.3156/jsoft.29.5_177_2.

Gruber, A. J., G. G. Calhoon, I. Shusterman, G. Schoenbaum, M. R. Roesch, and P. O'Donnell. 2010. "More Is Less: A Disinhibited Prefrontal Cortex Impairs Cognitive Flexibility." *Journal of Neuroscience*. https://doi.org/10.1523/jneurosci.4623-10.2010.

Gruber, Aaron J., and Robert J. McDonald. 2012. "Context, Emotion, and the Strategic Pursuit of Goals: Interactions among Multiple Brain Systems Controlling Motivated Behavior." *Frontiers in Behavioral Neuroscience*. https://doi.org/10.3389/fnbeh.2012.00050.

Haber, Suzanne N., Ki Sok Kim, Philippe Mailly, and Roberta Calzavara. 2006. "Reward-Related Cortical Inputs Define a Large Striatal Region in Primates That Interface with Associative Cortical Connections, Providing a Substrate for Incentive-Based Learning." *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.0271-06.2006.

Hadland, K. A., M. F.S. Rushworth, D. Gaffan, and R. E. Passingham. 2003. "The

Anterior Cingulate and Reward-Guided Selection of Actions." *Journal of Neurophysiology*. https://doi.org/10.1152/jn.00634.2002.

Hafner, Danijar, Timothy Lillicrap, Jimmy Ba, and Mohammad Norouzi. 2019. "Dream to Control: Learning Behaviors by Latent Imagination." *ArXiv*.

Hare, Todd A., Wolfram Schultz, Colin F. Camerer, John P. O'Doherty, and Antonio Rangel. 2011. "Transformation of Stimulus Value Signals into Motor Commands during Simple Choice." *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1109322108.

Hasan, Mahadi, Murugappan Suresh Kanna, Wang Jun, Aruna Surendran Ramkrishnan, Zafar Iqbal, Youngjin Lee, and Ying Li. 2019. "Schema-like Learning and Memory Consolidation Acting through Myelination." *FASEB Journal*. https://doi.org/10.1096/fj.201900910R.

Hashemnia, Saeedeh. 2019. "The Encoding of Utility and Its Modulation by Psychostimulants in Rat Anterior Cingulate Cortex." University of Lethbridge. https://hdl.handle.net/10133/5550.

Hashemnia, Saeedeh, David R. Euston, and Aaron J. Gruber. 2020. "Amphetamine Reduces Reward Encoding and Stabilizes Neural Dynamics in Rat Anterior Cingulate Cortex." *ELife*. https://doi.org/10.7554/ELIFE.56755.

Hayden, Benjamin Y., Sarah R. Heilbronner, John M. Pearson, and Michael L. Platt. 2011. "Surprise Signals in Anterior Cingulate Cortex: Neuronal Encoding of Unsigned Reward Prediction Errors Driving Adjustment in Behavior." *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.4652-10.2011.

Hayden, Benjamin Y., John M. Pearson, and Michael L. Platt. 2009. "Fictive Reward Signals in the Anterior Cingulate Cortex." *Science*. https://doi.org/10.1126/science.1168488.

———. 2011. "Neuronal Basis of Sequential Foraging Decisions in a Patchy Environment." *Nature Neuroscience*. https://doi.org/10.1038/nn.2856.

Hayden, Benjamin Y., and Michael L. Platt. 2010. "Neurons in Anterior Cingulate Cortex Multiplex Information about Reward and Action." *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.4874-09.2010.

Hayden, Benjamin Y., David V. Smith, and Michael L. Platt. 2009. “Electrophysiological Correlates of Default-Mode Processing in Macaque Posterior Cingulate Cortex.” *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.0812035106.

Heilbronner, Sarah R., and Benjamin Y. Hayden. 2016. “Dorsal Anterior Cingulate Cortex: A Bottom-Up View.” *Annual Review of Neuroscience*. https://doi.org/10.1146/annurev-neuro-070815-013952.

Hillman, Kristin L., and David K. Bilkey. 2010. “Neurons in the Rat Anterior Cingulate Cortex Dynamically Encode Cost-Benefit in a Spatial Decision-Making Task.” *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.1273-10.2010.

———. 2012. “Neural Encoding of Competitive Effort in the Anterior Cingulate Cortex.” *Nature Neuroscience*. https://doi.org/10.1038/nn.3187.

Hoesen, Gary W. Van, Robert J. Morecraft, and Brent A. Vogt. 1993. “Connections of the Monkey Cingulate Cortex.” In *Neurobiology of Cingulate Cortex and Limbic Thalamus*. https://doi.org/10.1007/978-1-4899-6704-6_9.

Holroyd, C B, S Nieuwenhuis, R Mars, and M G H Coles. 2004. “Anterior Cingulate Cortex, Selection for Action, and Error Processing.” In *Cognitive Neuroscience of Attention*.

Holroyd, Clay B., and Michael G.H. Coles. 2002. “The Neural Basis of Human Error Processing: Reinforcement Learning, Dopamine, and the Error-Related Negativity.” *Psychological Review*. https://doi.org/10.1037/0033-295X.109.4.679.

———. 2008. “Dorsal Anterior Cingulate Cortex Integrates Reinforcement History to Guide Voluntary Behavior.” *Cortex*. https://doi.org/10.1016/j.cortex.2007.08.013.

Holroyd, Clay B., and Samuel M. McClure. 2015. “Hierarchical Control over Effortful Behavior by Rodent Medial Frontal Cortex: A Computational Model.” *Psychological Review*. https://doi.org/10.1037/a0038339.

Holroyd, Clay B., and Nick Yeung. 2012. “Motivation of Extended Behaviors by Anterior Cingulate Cortex.” *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2011.12.008.

Horst, Nicole K., and Mark Laubach. 2013. “Reward-Related Activity in the Medial Prefrontal Cortex Is Driven by Consumption.” *Frontiers in Neuroscience*. https://doi.org/10.3389/fnins.2013.00056.

Hoshi, Eiji, Hiromasa Sawamura, and Jun Tanji. 2005. “Neurons in the Rostral Cingulate Motor Area Monitor Multiple Phases of Visuomotor Behavior with Modest Parametric Selectivity.” *Journal of Neurophysiology*. https://doi.org/10.1152/jn.01201.2004.

Hosking, Jay G., Paul J. Cocker, and Catharine A. Winstanley. 2014. “Dissociable Contributions of Anterior Cingulate Cortex and Basolateral Amygdala on a Rodent Cost/Benefit Decision-Making Task of Cognitive Effort.” *Neuropsychopharmacology*. https://doi.org/10.1038/npp.2014.27.

Hosokawa, T., S. W. Kennerley, J. Sloan, and J. D. Wallis. 2013. “Single-Neuron Mechanisms Underlying Cost-Benefit Analysis in Frontal Cortex.” *Journal of Neuroscience*. https://doi.org/10.1523/jneurosci.2221-13.2013.

Hyafil, Alexandre, Christopher Summerfield, and Etienne Koechlin. 2009. “Two Mechanisms for Task Switching in the Prefrontal Cortex.” *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.2828-08.2009.

Hyman, J. M., L. Ma, E. Balaguer-Ballester, D. Durstewitz, and J. K. Seamans. 2012. “Contextual Encoding by Ensembles of Medial Prefrontal Cortex Neurons.” *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1114415109.

Insel, Nathan, and Carol A. Barnes. 2015. “Differential Activation of Fast-Spiking and Regular-Firing Neuron Populations during Movement and Reward in the Dorsal Medial Frontal Cortex.” *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhu062.

Ito, Hiroshi T., Sheng Jia Zhang, Menno P. Witter, Edvard I. Moser, and May Britt Moser. 2015. “A Prefrontal-Thalamo-Hippocampal Circuit for Goal-Directed Spatial Navigation.” *Nature*. https://doi.org/10.1038/nature14396.

Ito, Shigehiko, Veit Stuphorn, Joshua W. Brown, and Jeffrey D. Schall. 2003. “Performance Monitoring by the Anterior Cingulate Cortex during Saccade Countermanding.” *Science*. https://doi.org/10.1126/science.1087847.

Jadhav, Shantanu P P., Gideon Rothschild, Demetris K K. Roumis, and Loren M M.

Frank. 2016. “Coordinated Excitation and Inhibition of Prefrontal Ensembles during Awake Hippocampal Sharp-Wave Ripple Events.” *Neuron*. https://doi.org/10.1016/j.neuron.2016.02.010.

Jahn, Andrew, Derek Evan Nee, William H. Alexander, and Joshua W. Brown. 2014. “Distinct Regions of Anterior Cingulate Cortex Signal Prediction and Outcome Evaluation.” *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2014.03.050.

Joel, Daphna, Ina Weiner, and Joram Feldon. 1997. “Electrolytic Lesions of the Medial Prefrontal Cortex in Rats Disrupt Performance on an Analog of the Wisconsin Card Sorting Test, but Do Not Disrupt Latent Inhibition: Implications for Animal Models of Schizophrenia.” *Behavioural Brain Research*. https://doi.org/10.1016/S0166-4328(97)87583-3.

Johnson, Adam, and A. David Redish. 2007. “Neural Ensembles in CA3 Transiently Encode Paths Forward of the Animal at a Decision Point.” *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.3761-07.2007.

Johnston, Kevin, Helen M. Levin, Michael J. Koval, and Stefan Everling. 2007. “Top-Down Control-Signal Dynamics in Anterior Cingulate and Prefrontal Cortex Neurons Following Task Switching.” *Neuron*. https://doi.org/10.1016/j.neuron.2006.12.023.

Jolliffe, I. T. 1986. “Principal Components in Regression Analysis.” In . https://doi.org/10.1007/978-1-4757-1904-8_8.

Jung, Min W., Yulin Qin, Bruce L. McNaughton, and Carol A. Barnes. 1998. “Firing Characteristics of Deep Layer Neurons in Prefrontal Cortex in Rats Performing Spatial Working Memory Tasks.” *Cerebral Cortex*. https://doi.org/10.1093/cercor/8.5.437.

Kaefer, Karola, Michele Nardin, Karel Blahna, and Jozsef Csicsvari. 2020. “Replay of Behavioral Sequences in the Medial Prefrontal Cortex during Rule Switching.” *Neuron*. https://doi.org/10.1016/j.neuron.2020.01.015.

Kahnt, Thorsten, Marcus Grueschow, Oliver Speck, and John Dylan Haynes. 2011. “Perceptual Learning and Decision-Making in Human Medial Frontal Cortex.” *Neuron*. https://doi.org/10.1016/j.neuron.2011.02.054.

Kaplan, Raphael, John King, Raphael Koster, William D. Penny, Neil Burgess, and Karl

J. Friston. 2017. “The Neural Representation of Prospective Choice during Spatial Planning and Decisions.” *PLoS Biology*. https://doi.org/10.1371/journal.pbio.1002588.

Kargo, William J., Botond Szatmary, and Douglas A. Nitz. 2007. “Adaptation of Prefrontal Cortical Firing Patterns and Their Fidelity to Changes in Action-Reward Contingencies.” *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.3604-06.2007.

Kawai, Takashi, Hiroshi Yamada, Nobuya Sato, Masahiko Takada, and Masayuki Matsumoto. 2015. “Roles of the Lateral Habenula and Anterior Cingulate Cortex in Negative Outcome Monitoring and Behavioral Adjustment in Nonhuman Primates.” *Neuron*. https://doi.org/10.1016/j.neuron.2015.09.030.

Kennerley, Steven W., Timothy E.J. Behrens, and Jonathan D. Wallis. 2011. “Double Dissociation of Value Computations in Orbitofrontal and Anterior Cingulate Neurons.” *Nature Neuroscience*. https://doi.org/10.1038/nn.2961.

Kennerley, Steven W., Aspandiar F. Dahmubed, Antonio H. Lara, and Jonathan D. Wallis. 2009. “Neurons in the Frontal Lobe Encode the Value of Multiple Decision Variables.” *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn.2009.21100.

Kennerley, Steven W., and Jonathan D. Wallis. 2009a. “Encoding of Reward and Space during a Working Memory Task in the Orbitofrontal Cortex and Anterior Cingulate Sulcus.” *Journal of Neurophysiology*. https://doi.org/10.1152/jn.00273.2009.

———. 2009b. “Evaluating Choices by Single Neurons in the Frontal Lobe: Outcome Value Encoded across Multiple Decision Variables.” *European Journal of Neuroscience*. https://doi.org/10.1111/j.1460-9568.2009.06743.x.

Kennerley, Steven W., and Mark E. Walton. 2011. “Decision Making and Reward in Frontal Cortex: Complementary Evidence from Neurophysiological and Neuropsychological Studies.” *Behavioral Neuroscience*. https://doi.org/10.1037/a0023575.

Kennerley, Steven W., Mark E. Walton, Timothy E.J. Behrens, Mark J. Buckley, and Matthew F.S. Rushworth. 2006. “Optimal Decision Making and the Anterior Cingulate Cortex.” *Nature Neuroscience*. https://doi.org/10.1038/nn1724.

Kerns, John G., Jonathan D. Cohen, Angus W. MacDonald, Raymond Y. Cho, V. Andrew Stenger, and Cameron S. Carter. 2004. "Anterior Cingulate Conflict Monitoring and Adjustments in Control." *Science*. https://doi.org/10.1126/science.1089910.

Kesner, Raymond P., and Michael E. Ragozzino. 2003. "The Role of the Prefrontal Cortex in Object-Place Learning: A Test of the Attribute Specificity Model." *Behavioural Brain Research*. https://doi.org/10.1016/j.bbr.2003.09.024.

Kesteren, Marlieke T.R. Van, Dirk J. Ruiter, Guillén Fernández, and Richard N. Henson. 2012. "How Schema and Novelty Augment Memory Formation." *Trends in Neurosciences*. https://doi.org/10.1016/j.tins.2012.02.001.

Kim, Steve M., Surya Ganguli, and Loren M. Frank. 2012. "Spatial Information Outflow from the Hippocampal Circuit: Distributed Spatial Coding and Phase Precession in the Subiculum." *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.5942-11.2012.

Klein-Flugge, M. C., S. W. Kennerley, K. Friston, and S. Bestmann. 2016. "Neural Signatures of Value Comparison in Human Cingulate Cortex during Decisions Requiring an Effort-Reward Trade-Off." *Journal of Neuroscience*. https://doi.org/10.1523/jneurosci.0292-16.2016.

Kolling, Nils, Timothy E. J. Behrens, Rogier B. Mars, and Matthew F. S. Rushworth. 2012. "Neural Mechanisms of Foraging." *Science*. https://doi.org/10.1126/science.1216930.

Kurzban, Robert, Angela Duckworth, Joseph W. Kable, and Justus Myers. 2013. "An Opportunity Cost Model of Subjective Effort and Task Performance." *Behavioral and Brain Sciences*. https://doi.org/10.1017/S0140525X12003196.

Lapish, Christopher C., Daniel Durstewitz, L. Judson Chandler, and Jeremy K. Seamans. 2008. "Successful Choice Behavior Is Associated with Distinct and Coherent Network States in Anterior Cingulate Cortex." *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.0804045105.

Lim, Seung Lark, John P. O'Doherty, and Antonio Rangel. 2011. "The Decision Value Computations in the VmPFC and Striatum Use a Relative Value Code That Is Guided by Visual Attention." *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.1246-11.2011.

Lindsay, Adrian J., Barak F. Caracheo, Jamie J. Grewal, Daniel Leibovitz, and Jeremy K. Seamans. 2018. “How Much Does Movement and Location Encoding Impact Prefrontal Cortex Activity? An Algorithmic Decoding Approach in Freely Moving Rats.” *ENeuro*. https://doi.org/10.1523/ENEURO.0023-18.2018.

Lu, Hanbing, Qihong Zou, Hong Gu, Marcus E. Raichle, Elliot A. Stein, and Yihong Yang. 2012. “Rat Brains Also Have a Default Mode Network.” *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1200506109.

Luk, C.-H., and J. D. Wallis. 2013. “Choice Coding in Frontal Cortex during Stimulus-Guided or Action-Guided Decision-Making.” *Journal of Neuroscience*. https://doi.org/10.1523/jneurosci.4920-12.2013.

Luk, Chung Hay, and Jonathan D. Wallis. 2009. “Dynamic Encoding of Responses and Outcomes by Neurons in Medial Prefrontal Cortex.” *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.0386-09.2009.

Ma, L., J. M. Hyman, D. Durstewitz, A. G. Phillips, and J. K. Seamans. 2016. “A Quantitative Analysis of Context-Dependent Remapping of Medial Frontal Cortex Neurons and Ensembles.” *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.3176-15.2016.

Ma, Liya, James M. Hyman, Anthony G. Phillips, and Jeremy K. Seamans. 2014. “Tracking Progress toward a Goal in Corticostriatal Ensembles.” *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.3834-13.2014.

Makhzani, Alireza, and Brendan Frey. 2014. “K-Sparse Autoencoders.” In *2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings*.

Mantini, Dante, Annelis Gerits, Koen Nelissen, Jean Baptiste Durand, Olivier Joly, Luciano Simone, Hiromasa Sawamura, et al. 2011. “Default Mode of Brain Function in Monkeys.” *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.2318-11.2011.

Mashhoori, Ali, Saeedeh Hashemnia, Bruce L McNaughton, David R Euston, and Aaron J Gruber. 2018. “Rat Anterior Cingulate Cortex Recalls Features of Remote Reward Locations after Disfavoured Reinforcements.” *ELife*. https://doi.org/10.7554/elife.29793.

Matsumoto, Kenji, Wataru Suzuki, and Keiji Tanaka. 2003. “Neuronal Correlates of Goal-Based Motor Selection in the Prefrontal Cortex.” *Science*. https://doi.org/10.1126/science.1084204.

Matsumoto, Madoka, Kenji Matsumoto, Hiroshi Abe, and Keiji Tanaka. 2007. “Medial Prefrontal Cell Activity Signaling Prediction Errors of Action Values.” *Nature Neuroscience*. https://doi.org/10.1038/nn1890.

McClelland, James L, Bruce L McNaughton, and Randall C O’Reilly. 1995. “Why There Are Complementary Learning Systems in the Hippocampus and Neocortex: Insights from the Successes and Failures of Connectionist Models of Learning and Memory.” *Psychological Review*. https://doi.org/10.1037/0033-295X.102.3.419.

Miller, E. K., and J. D. Cohen. 2001. “An Integrative Theory of Prefrontal Cortex Function.” *Annual Review of Neuroscience*. https://doi.org/10.1146/annurev.neuro.24.1.167.

Morecraft, Robert J., and Gary W. Van Hoesen. 1998. “Convergence of Limbic Input to the Cingulate Motor Cortex in the Rhesus Monkey.” *Brain Research Bulletin*. https://doi.org/10.1016/S0361-9230(97)00344-4.

Murray, Elisabeth A, Steven P Wise, and Kim S Graham. 2017. *The Evolution of Memory Systems: Ancestors, Anatomy, and Adaptations*. *The Evolution of Memory Systems: Ancestors, Anatomy, and Adaptations.*

Nakamura, Kae, Matthew R. Roesch, and Carl R. Olson. 2005. “Neuronal Activity in Macaque SEF and ACC during Performance of Tasks Involving Conflict.” *Journal of Neurophysiology*. https://doi.org/10.1152/jn.00305.2004.

Narayanan, Nandakumar S., James F. Cavanagh, Michael J. Frank, and Mark Laubach. 2013. “Common Medial Frontal Mechanisms of Adaptive Control in Humans and Rodents.” *Nature Neuroscience*. https://doi.org/10.1038/nn.3549.

Narayanan, Nandakumar S., and Mark Laubach. 2006. “Top-Down Control of Motor Cortex Ensembles by Dorsomedial Prefrontal Cortex.” *Neuron*. https://doi.org/10.1016/j.neuron.2006.10.021.

Nee, Derek Evan, Sabine Kastner, and Joshua W. Brown. 2011. “Functional Heterogeneity of Conflict, Error, Task-Switching, and Unexpectedness Effects within Medial Prefrontal Cortex.” *NeuroImage*.

https://doi.org/10.1016/j.neuroimage.2010.08.027.

Nieuwenhuis, Sander, K. Richard Ridderinkhof, Jos Blom, Guido P.H. Band, and Albert Kok. 2001. “Error-Related Brain Potentials Are Differentially Related to Awareness of Response Errors: Evidence from an Antisaccade Task.” *Psychophysiology*. https://doi.org/10.1017/S0048577201001111.

O’Keefe, J., and J. Dostrovsky. 1971. “The Hippocampus as a Spatial Map. Preliminary Evidence from Unit Activity in the Freely-Moving Rat.” *Brain Research*. https://doi.org/10.1016/0006-8993(71)90358-1.

O’keefe, John, and Lynn Nadel. 1987. *The Hippocampus as a Cognitive Map.* Oxford: Clarendon Press.

O’Reilly, Jill X, Urs Schüffelgen, Steven F Cuell, Timothy E J Behrens, Rogier B Mars, and Matthew F S Rushworth. 2013. “Dissociable Effects of Surprise and Model Update in Parietal and Anterior Cingulate Cortex.” *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1305373110.

Ostlund, Sean B., and Bernard W. Balleine. 2007. “Orbitofrontal Cortex Mediates Outcome Encoding in Pavlovian but Not Instrumental Conditioning.” *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.5443-06.2007.

Padoa-Schioppa, Camillo, and John A. Assad. 2006. “Neurons in the Orbitofrontal Cortex Encode Economic Value.” *Nature*. https://doi.org/10.1038/nature04676.

Palkovits, Miklos. 1983. “The Rat Brain in Stereotaxic Coordinates.” *Neuropeptides*. https://doi.org/10.1016/0143-4179(83)90049-5.

Palmer, Erica D., Howard J. Rosen, Jeffrey G. Ojemann, Randy L. Buckner, William M. Kelley, and Steven E. Petersen. 2001. “An Event-Related FMRI Study of Overt and Covert Word Stem Completion.” *NeuroImage*. https://doi.org/10.1006/nimg.2001.0779.

Parvizi, Josef, Vinitha Rangarajan, William R. Shirer, Nikita Desai, and Michael D. Greicius. 2013. “The Will to Persevere Induced by Electrical Stimulation of the Human Cingulate Gyrus.” *Neuron*. https://doi.org/10.1016/j.neuron.2013.10.057.

Passingham, Richard E., and Steven P. Wise. 2015. *The Neurobiology of the Prefrontal Cortex*. *The Neurobiology of the Prefrontal Cortex*. https://doi.org/10.1093/acprof:osobl/9780199552917.001.0001.

Paus, T., M. Petrides, A. C. Evans, and E. Meyer. 1993. "Role of the Human Anterior Cingulate Cortex in the Control of Oculomotor, Manual, and Speech Responses: A Positron Emission Tomography Study." *Journal of Neurophysiology*. https://doi.org/10.1152/jn.1993.70.2.453.

Peters, Jan, and Christian Büchel. 2010. "Episodic Future Thinking Reduces Reward Delay Discounting through an Enhancement of Prefrontal-Mediotemporal Interactions." *Neuron*. https://doi.org/10.1016/j.neuron.2010.03.026.

Picard, Nathalie, and Peter L. Strick. 1996. "Motor Areas of the Medial Wall: A Review of Their Location and Functional Activation." *Cerebral Cortex*. https://doi.org/10.1093/cercor/6.3.342.

Pickens, Charles L., Michael P. Saddoris, Barry Setlow, Michela Gallagher, Peter C. Holland, and Geoffrey Schoenbaum. 2003. "Different Roles for Orbitofrontal Cortex and Basolateral Amygdala in a Reinforcer Devaluation Task." *Journal of Neuroscience*. https://doi.org/10.1523/jneurosci.23-35-11078.2003.

Porter, Blake S., Kristin L. Hillman, and David K. Bilkey. 2019. "Anterior Cingulate Cortex Encoding of Effortful Behavior." *Journal of Neurophysiology*. https://doi.org/10.1152/jn.00654.2018.

Posner, Michael I., Mary K. Rothbart, Brad E. Sheese, and Yiyuan Tang. 2007. "The Anterior Cingulate Gyrus and the Mechanism of Self-Regulation." *Cognitive, Affective and Behavioral Neuroscience*. https://doi.org/10.3758/CABN.7.4.391.

Poucet, Bruno. 1997. "Searching for Spatial Unit Firing in the Prelimbic Area of the Rat Medial Prefrontal Cortex." *Behavioural Brain Research*. https://doi.org/10.1016/S0166-4328(96)00144-1.

Pratt, Wayne E., and Sheri J.Y. Mizumori. 2001. "Neurons in Rat Medial Prefrontal Cortex Show Anticipatory Rate Changes to Predictable Differential Rewards in a Spatial Memory Task." *Behavioural Brain Research*. https://doi.org/10.1016/S0166-4328(01)00204-2.

Prevost, C., M. Pessiglione, E. Metereau, M.-L. Clery-Melin, and J.-C. Dreher. 2010.

"Separate Valuation Subsystems for Delay and Effort Decision Costs." *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.2752-10.2010.

Quilodran, René, Marie Rothé, and Emmanuel Procyk. 2008. "Behavioral Shifts and Action Valuation in the Anterior Cingulate Cortex." *Neuron*. https://doi.org/10.1016/j.neuron.2007.11.031.

Rajasethupathy, Priyamvada, Sethuraman Sankaran, James H. Marshel, Christina K. Kim, Emily Ferenczi, Soo Yeun Lee, Andre Berndt, et al. 2015. "Projections from Neocortex Mediate Top-down Control of Memory Retrieval." *Nature*. https://doi.org/10.1038/nature15389.

Rich, Erin L., and Matthew Shapiro. 2009. "Rat Prefrontal Cortical Neurons Selectively Code Strategy Switches." *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.6068-08.2009.

Rudebeck, P. H., M. J. Buckley, M. E. Walton, and M. F.S. Rushworth. 2006. "A Role for the Macaque Anterior Cingulate Gyrus in Social Valuation." *Science*. https://doi.org/10.1126/science.1128197.

Rudebeck, Peter H., Timothy E. Behrens, Steven W. Kennerley, Mark G. Baxter, Mark J. Buckley, Mark E. Walton, and Matthew F.S. Rushworth. 2008. "Frontal Cortex Subregions Play Distinct Roles in Choices between Actions and Stimuli." *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.3541-08.2008.

Rudebeck, Peter H., Mark E. Walton, Angharad N. Smyth, David M. Bannerman, and Matthew F.S. Rushworth. 2006. "Separate Neural Pathways Process Different Decision Costs." *Nature Neuroscience*. https://doi.org/10.1038/nn1756.

Rushworth, M. F.S., K. A. Hadland, D. Gaffan, and R. E. Passingham. 2003. "The Effect of Cingulate Cortex Lesions on Task Switching and Working Memory." *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/089892903321593072.

Rushworth, M. F.S., M. E. Walton, S. W. Kennerley, and D. M. Bannerman. 2004. "Action Sets and Decisions in the Medial Frontal Cortex." *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2004.07.009.

Rushworth, Matthew F.S., Mary Ann P. Noonan, Erie D. Boorman, Mark E. Walton, and Timothy E. Behrens. 2011. "Frontal Cortex and Reward-Guided Learning and Decision-Making." *Neuron*. https://doi.org/10.1016/j.neuron.2011.05.014.

Schweimer, Judith, Simone Saft, and Wolfgang Hauber. 2005. “Involvement of Catecholamine Neurotransmission in the Rat Anterior Cingulate in Effort-Related Decision Making.” *Behavioral Neuroscience*. https://doi.org/10.1037/0735-7044.119.6.1687.

Seamans, Jeremy K., Christopher C. Lapish, and Daniel Durstewitz. 2008. “Comparing the Prefrontal Cortex of Rats and Primates: Insights from Electrophysiology.” *Neurotoxicity Research*. https://doi.org/10.1007/BF03033814.

Seo, Hyojung, and Daeyeol Lee. 2007. “Temporal Filtering of Reward Signals in the Dorsal Anterior Cingulate Cortex during a Mixed-Strategy Game.” *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.2369-07.2007.

Shen, Chen, Salva Ardid, Daniel Kaping, Stephanie Westendorff, Stefan Everling, and Thilo Womelsdorf. 2015. “Anterior Cingulate Cortex Cells Identify Process-Specific Errors of Attentional Control Prior to Transient Prefrontal-Cingulate Inhibition.” *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhu028.

Shenhav, Amitai, Matthew M. Botvinick, and Jonathan D. Cohen. 2013. “The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function.” *Neuron*. https://doi.org/10.1016/j.neuron.2013.07.007.

Shenhav, Amitai, Jonathan D. Cohen, and Matthew M. Botvinick. 2016. “Dorsal Anterior Cingulate Cortex and the Value of Control.” *Nature Neuroscience*. https://doi.org/10.1038/nn.4384.

Shenhav, Amitai, Mark A. Straccia, Jonathan D. Cohen, and Matthew M. Botvinick. 2014. “Anterior Cingulate Engagement in a Foraging Context Reflects Choice Difficulty, Not Foraging Value.” *Nature Neuroscience*. https://doi.org/10.1038/nn.3771.

Shidara, Munetaka, and Barry J. Richmond. 2002. “Anterior Cingulate: Single Neuronal Signals Related to Degree of Reward Expectancy.” *Science*. https://doi.org/10.1126/science.1069504.

Shima, K., K. Aya, H. Mushiake, M. Inase, H. Aizawa, and J. Tanji. 1991. “Two Movement-Related Foci in the Primate Cingulate Cortex Observed in Signal-Triggered and Self-Paced Forelimb Movements.” *Journal of Neurophysiology*. https://doi.org/10.1152/jn.1991.65.2.188.

Shima, Keisetsu, and Jun Tanji. 1998. “Role for Cingulate Motor Area Cells in Voluntary Movement Selection Based on Reward.” *Science*. https://doi.org/10.1126/science.282.5392.1335.

Steiner, Adam P., and A. David Redish. 2014. “Behavioral and Neurophysiological Correlates of Regret in Rat Decision-Making on a Neuroeconomic Task.” *Nature Neuroscience*. https://doi.org/10.1038/nn.3740.

Stemmer, Brigitte, Sidney J. Segalowitz, Wolfgang Witzke, and Paul Walter Schönle. 2004. “Error Detection in Patients with Lesions to the Medial Prefrontal Cortex: An ERP Study.” *Neuropsychologia*. https://doi.org/10.1016/S0028-3932(03)00121-0.

Strait, Caleb E., Tommy C. Blanchard, and Benjamin Y. Hayden. 2014. “Reward Value Comparison via Mutual Inhibition in Ventromedial Prefrontal Cortex.” *Neuron*. https://doi.org/10.1016/j.neuron.2014.04.032.

Strait, Caleb E., Brianna J. Sleezer, Tommy C. Blanchard, Habiba Azab, Meghan D. Castagno, and Benjamin Y. Hayden. 2016. “Neuronal Selectivity for Spatial Positions of Offers and Choices in Five Reward Regions.” *Journal of Neurophysiology*. https://doi.org/10.1152/jn.00325.2015.

Sul, Jung Hoon, Suhyun Jo, Daeyeol Lee, and Min Whan Jung. 2011. “Role of Rodent Secondary Motor Cortex in Value-Based Action Selection.” *Nature Neuroscience*. https://doi.org/10.1038/nn.2881.

Sul, Jung Hoon, Hoseok Kim, Namjung Huh, Daeyeol Lee, and Min Whan Jung. 2010. “Distinct Roles of Rodent Orbitofrontal and Medial Prefrontal Cortex in Decision Making.” *Neuron*. https://doi.org/10.1016/j.neuron.2010.03.033.

Sutton, R.S., and A.G. Barto. 1998. “Reinforcement Learning: An Introduction.” *IEEE Transactions on Neural Networks*. https://doi.org/10.1109/tnn.1998.712192.

Takehara-Nishiuchi, Kaori, and Bruce L. McNaughton. 2008. “Spontaneous Changes of Neocortical Code for Associative Memory during Consolidation.” *Science*. https://doi.org/10.1126/science.1161299.

Tervo, Dougal G.R., Mikhail Proskurin, Maxim Manakov, Mayank Kabra, Alison Vollmer, Kristin Branson, and Alla Y. Karpova. 2014. “Behavioral Variability through Stochastic Choice and Its Gating by Anterior Cingulate Cortex.” *Cell*. https://doi.org/10.1016/j.cell.2014.08.037.

Thompson-Schill, Sharon L., Mark D'Esposito, Geoffrey K. Aguirre, and Martha J. Farah. 1997. "Role of Left Inferior Prefrontal Cortex in Retrieval of Semantic Knowledge: A Reevaluation." *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.94.26.14792.

Tse, Dorothy, Rosamund F. Langston, Masaki Kakeyama, Ingrid Bethus, Patrick A. Spooner, Emma R. Wood, Menno P. Witter, and Richard G.M. Morris. 2007. "Schemas and Memory Consolidation." *Science*. https://doi.org/10.1126/science.1135935.

Tsujimoto, Satoshi, Aldo Genovesio, and Steven P. Wise. 2009. "Monkey Orbitofrontal Cortex Encodes Response Choices near Feedback Time." *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.5777-08.2009.

Ullsperger, Markus, and D. Yves Von Cramon. 2001. "Subprocesses of Performance Monitoring: A Dissociation of Error Processing and Response Competition Revealed by Event-Related FMRI and ERPs." *NeuroImage*. https://doi.org/10.1006/nimg.2001.0935.

Uylings, Harry B.M., Henk J. Groenewegen, and Bryan Kolb. 2003. "Do Rats Have a Prefrontal Cortex?" *Behavioural Brain Research*. https://doi.org/10.1016/j.bbr.2003.09.028.

Vassena, Eliana, James Deraeve, and William H. Alexander. 2017. "Predicting Motivation: Computational Models of PFC Can Explain Neural Coding of Motivation and Effort-Based Decision-Making in Health and Disease." *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn_a_01160.

———. 2020. "Surprise, Value and Control in Anterior Cingulate Cortex during Speeded Decision-Making." *Nature Human Behaviour*. https://doi.org/10.1038/s41562-019-0801-5.

Vassena, Eliana, Massimo Silvetti, Carsten N. Boehler, Eric Achten, Wim Fias, and Tom Verguts. 2014. "Overlapping Neural Systems Represent Cognitive Effort and Reward Anticipation." *PLoS ONE*. https://doi.org/10.1371/journal.pone.0091008.

Veen, Vincent Van, Jonathan D. Cohen, Matthew M. Botvinick, V. Andrew Stenger, and Cameron S. Carter. 2001. "Anterior Cingulate Cortex, Conflict Monitoring, and Levels of Processing." *NeuroImage*. https://doi.org/10.1006/nimg.2001.0923.

Verguts, Tom, Eliana Vassena, and Massimo Silvetti. 2015. “Adaptive Effort Investment in Cognitive and Physical Tasks: A Neurocomputational Model.” *Frontiers in Behavioral Neuroscience*. https://doi.org/10.3389/fnbeh.2015.00057.

Vincent, Pascal, Hugo Larochelle, Isabelle Lajoie, Yoshua Bengio, and Pierre Antoine Manzagol. 2010. “Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion.” *Journal of Machine Learning Research*.

Wallis, Jonathan D., and Steven W. Kennerley. 2011. “Contrasting Reward Signals in the Orbitofrontal Cortex and Anterior Cingulate Cortex.” *Annals of the New York Academy of Sciences*. https://doi.org/10.1111/j.1749-6632.2011.06277.x.

Wallis, Jonathan D., and Erin L. Rich. 2011. “Challenges of Interpreting Frontal Neurons during Value-Based Decision-Making.” *Frontiers in Neuroscience*. https://doi.org/10.3389/fnins.2011.00124.

Walton, Mark E., David M. Bannerman, Karin Alterescu, and Matthew F.S. Rushworth. 2003. “Functional Specialization within Medial Frontal Cortex of the Anterior Cingulate for Evaluating Effort-Related Decisions.” *Journal of Neuroscience*. https://doi.org/10.1523/jneurosci.23-16-06475.2003.

Wang, Szu Han, Dorothy Tse, and Richard G M Morris. 2012. “Anterior Cingulate Cortex in Schema Assimilation and Expression.” *Learning and Memory*. https://doi.org/10.1101/lm.026336.112.

Warden, Melissa R., Aslihan Selimbeyoglu, Julie J. Mirzabekov, Maisie Lo, Kimberly R. Thompson, Sung Yon Kim, Avishek Adhikari, Kay M. Tye, Loren M. Frank, and Karl Deisseroth. 2012. “A Prefrontal Cortex-Brainstem Neuronal Projection That Controls Response to Behavioural Challenge.” *Nature*. https://doi.org/10.1038/nature11617.

Wessel, Jan R., Claudia Danielmeier, J. Bruce Morton, and Markus Ullsperger. 2012. “Surprise and Error: Common Neuronal Architecture for the Processing of Errors and Novelty.” *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.6352-11.2012.

Williams, Ziv M., George Bush, Scott L. Rauch, G. Rees Cosgrove, and Emad N. Eskandar. 2004. “Human Anterior Cingulate Neurons and the Integration of Monetary Reward with Motor Responses.” *Nature Neuroscience*. https://doi.org/10.1038/nn1354.

Wilson, Matthew A., and Bruce L. McNaughton. 1993. “Dynamics of the Hippocampal Ensemble Code for Space.” *Science*. https://doi.org/10.1126/science.8351520.

———. 1994. “Reactivation of Hippocampal Ensemble Memories during Sleep.” *Science*. https://doi.org/10.1126/science.8036517.

Wilson, Robert C., Yuji K. Takahashi, Geoffrey Schoenbaum, and Yael Niv. 2014. “Orbitofrontal Cortex as a Cognitive Map of Task Space.” *Neuron*. https://doi.org/10.1016/j.neuron.2013.11.005.

Wingerden, Marijn Van, Martin Vinck, Jan V. Lankelma, and Cyriel M.A. Pennartz. 2010. “Learning-Associated Gamma-Band Phase-Locking of Action-Outcome Selective Neurons in Orbitofrontal Cortex.” *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.0222-10.2010.

Wunderlich, Klaus, Antonio Range, and John P. O’Doherty. 2009. “Neural Computations Underlying Action-Based Decision Making in the Human Brain.” *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.0901077106.

Yeung, Nick, Matthew M. Botvinick, and Jonathan D. Cohen. 2004. “The Neural Basis of Error Detection: Conflict Monitoring and the Error-Related Negativity.” *Psychological Review*. https://doi.org/10.1037/0033-295X.111.4.931.

Yu, Jai Y., Daniel F. Liu, Adrianna Loback, Irene Grossrubatscher, and Loren M. Frank. 2018. “Specific Hippocampal Representations Are Linked to Generalized Cortical Representations in Memory.” *Nature Communications*. https://doi.org/10.1038/s41467-018-04498-w.

Zhang, Kechen, Iris Ginzburg, Bruce L. McNaughton, and Terrence J. Sejnowski. 1998. “Interpreting Neuronal Population Activity by Reconstruction: Unified Framework with Application to Hippocampal Place Cells.” *Journal of Neurophysiology*. https://doi.org/10.1152/jn.1998.79.2.1017.

Zielinski, Mark C., Justin D. Shin, and Shantanu P. Jadhav. 2019. “Coherent Coding of Spatial Position Mediated by Theta Oscillations in the Hippocampus and Prefrontal Cortex.” *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.0106-19.2019.