

Expectation Modulates Episodic Memory Formation via Dopaminergic Circuitry

by

Jessica Kate Stanek

Department of Psychology and Neuroscience
Duke University

Date: _____

Approved:

R. Alison Adcock, Supervisor

Scott A. Huettel

Roberto Cabeza

Nan-kuei Chen

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Psychology and Neuroscience in the Graduate School
of Duke University

2016

ABSTRACT

Expectation Modulates Episodic Memory Formation via Dopaminergic Circuitry

by

Jessica K Stanek

Department of Psychology and Neuroscience
Duke University

Date: _____

Approved:

R. Alison Adcock, Supervisor

Scott A. Huettel

Roberto Cabeza

Nan-kuei Chen

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the Department of
Psychology and Neuroscience in the Graduate School of
Duke University

2016

Copyright by
Jessica Kate Stanek
2016

Abstract

Events that generate, meet, or violate expectations have the capacity to influence episodic memory formation. However, clarification is still required to illuminate the circumstances and direction of memory modulation. In the brain, the mechanisms by which expectation modulates memory formation also require consideration. The dopamine system has been implicated in signaling events associated with different states of expectancy; it has also been shown to modulate episodic memory formation in the hippocampus. Thus, the studies included in this dissertation utilized both functional magnetic resonance imaging (fMRI) and behavioral testing to examine when and how the dopaminergic system supports the modulation of memory by expectation. The work aimed to characterize the activation of dopaminergic circuitry in response to cues that generate expectancy, during periods of anticipation, and in response to outcomes that resolve expectancy. The studies also examined how each of these event types influenced episodic memory formation.

The present findings demonstrated that novelty and expectancy violation both drive dopaminergic circuitry capable of contributing to memory formation. Consistent with elevated dopaminergic midbrain and hippocampus activation for each, expected versus expectancy violating novelty did not differentially impact memory success. We also showed that high curiosity expectancy states drive memory formation. This was

supported by activation in dopaminergic circuitry that was greater for subsequently remembered information only in the high curiosity state. Finally, we showed that cues that generate high expected reward value versus high reward uncertainty differentially modulate memory formation during reward anticipation. This behavioral result was consistent with distinct temporal profiles of dopaminergic action having differential downstream effects on episodic memory formation.

Integrating the present studies with previous research suggests that dopaminergic circuitry signals events that are unpredicted, whether cuing or resolving expectations. It also suggests that contextual differences change the contribution of the dopaminergic system during anticipation, depending on the nature of the expectation. And finally, this work is consistent with a model in which dopamine elevation in response to expectancy events positively modulates episodic memory formation.

Dedication

For all the people who inspire me to never stop learning.

Contents

Abstract.....	iv
List of Tables	xiv
List of Figures	xv
Acknowledgements	xviii
1. Introduction	1
1.1 Expectation	1
1.2 Expectation modulates memory formation.....	2
1.2.1 Summary of expectation modulates memory formation	6
1.3 Expectation in the mesolimbic dopamine system	6
1.3.1 Cued expectation.....	8
1.3.2 Resolving expectation.....	13
1.3.3 Summary of expectation in the mesolimbic dopamine system.....	17
1.4 Mesolimbic dopamine shapes memory formation.....	17
1.4.1 Dopamine and hippocampus-dependent memory	19
1.4.2 Behavioral and functional neuronal support for dopaminergic contributions to memory formation	21
1.4.3 Broad dopaminergic circuitry contributes to memory	23
1.4.4 Summary of mesolimbic dopamine shapes memory formation	25
1.5 A model integrating expectation, dopamine, and memory	27
1.6 Open questions and experimental aims.....	29
2. Mesolimbic and cortical networks differentiate novelty versus surprise.....	32

2.1 Introduction.....	32
2.2 Methods	34
2.2.1 Subjects.....	34
2.2.2 Stimuli	35
2.2.3 Tasks.....	35
2.2.3.1 Familiarization	35
2.2.3.2 Incidental encoding	36
2.2.3.3 Recognition memory test	38
2.2.4 Procedure.....	39
2.2.5 Behavioral analysis.....	39
2.2.6 MRI data preprocessing and acquisition	40
2.2.7 Definition of a priori ROIs	41
2.2.8 Functional MRI data analysis	42
2.2.8.1 General Linear Model: Task-Related Activations	42
2.2.8.2 Group-Level Analysis	43
2.2.8.3 A Priori ROIs Analysis	44
2.2.9 Memory Correlation	45
2.3 Results	45
2.3.1 Behavior: Reaction Times and Accuracy	45
2.3.2 fMRI: A Priori ROI Analyses	47
2.3.2.1 Hippocampus: Main Effects	47
2.3.2.2 Hippocampus: Paired Contrasts.....	47

2.3.2.3 VTA and NAcc: Main Effects	48
2.3.2.4 SubN: Main Effects	50
2.3.3 fMRI: Whole Brain Analyses	50
2.3.3.1 Main Effects	50
2.3.3.2 Conjunctions	52
2.3.4 Behavior and fMRI: Recognition Memory.....	52
2.4 Discussion.....	53
2.4.1 Summary	53
2.4.2 Notable Findings: A Priori Regions.....	54
2.4.3 Dopaminergic Circuits in Encoding Novelty and Surprise: Integrating with Prior Human Imaging.....	55
2.4.4 Theoretical implications	57
2.4.5 Novelty and Surprise in Distributed Brain Networks	58
2.4.6 Limitations.....	60
2.4.7 Conclusions	60
3. The Neural Characterization of Curiosity at Cue, Anticipation, and Outcome	61
3.1 Introduction.....	61
3.2 Methods	64
3.2.1 Subjects.....	64
3.2.2 Stimuli	64
3.2.3 Tasks.....	64
3.2.3.1 Screening	64

3.2.3.2 Incidental Encoding.....	66
3.2.3.3 Surprise Recall.....	67
3.2.3.4 Final Ratings	67
3.2.4 Procedure.....	68
3.2.5 Analysis	68
3.2.5.1 Behavioral Analysis.....	68
3.2.5.2 MRI Data Preprocessing and Acquisition.....	68
3.2.5.3 Definition of A Priori ROIs.....	70
3.2.5.4 fMRI Data Analysis	70
3.2.5.5 Correlations between ROI activation and behavioral personality scores..	73
3.3 Results	73
3.3.1 Behavior: Memory.....	73
3.3.2 Behavior: Reaction Times and Accuracy	74
3.3.3 fMRI: Cue-Related Activation	75
3.3.3.1 Main Effects of Curiosity	75
3.3.3.2 Individual Differences in Curiosity.....	76
3.3.3.3 Main Effects of Motor Contingency	76
3.3.3.4 Curiosity by Motor Contingency Interaction.....	76
3.3.3.5 Main Effects of Subsequent Memory	77
3.3.3.6 Memory by Curiosity Interaction	77
3.3.4 fMRI: Anticipatory Activation.....	77
3.3.4.1 Main Effects of Curiosity	77

3.3.4.2 Individual Differences in Curiosity.....	78
3.3.4.3 Main Effects of Motor Contingency	79
3.3.4.4 Curiosity by Motor Contingency Interaction.....	79
3.3.4.5 Main Effects of Subsequent Memory	80
3.3.4.6 Memory by Curiosity Interaction	80
3.3.5 fMRI: Outcome-Related Activation	80
3.3.5.1 Main Effects of Curiosity	80
3.3.5.2 Individual Differences in Curiosity.....	81
3.3.5.3 Main Effects of Action Contingency	81
3.3.5.4 Curiosity by Action Contingency Interaction.....	82
3.3.5.5 Main Effects of Subsequent Memory	82
3.3.5.6 Memory by Curiosity Interaction	82
3.4 Discussion.....	83
3.4.1 Summary	83
3.4.2 Novel findings	84
3.4.3 Caveats.....	86
3.4.4 Conclusion.....	86
4. Expected reward value and uncertainty have dissociable effects on memory formation	88
4.1 Introduction.....	88
4.2 Methods	90
4.2.1 Subjects.....	90

4.2.2 Design and Procedure	91
4.2.2.1 Reward Learning	91
4.2.2.2 Incidental Encoding.....	91
4.2.2.3 Recognition Memory Test.....	94
4.2.2.4 Experiment 1 – 24 hour	94
4.2.2.5 Experiment 2 – Immediate.....	94
4.2.3 Analysis	94
4.2.3.1 Within Experiments.....	94
4.2.3.2 Across Experiments.....	95
4.3 Results	96
4.3.1 Reward Learning.....	96
4.3.2 Experiment 1 – 24 hour retrieval.....	96
4.3.3 Experiment 2 – Immediate retrieval	99
4.3.4 Experiments 1 & 2: Contrasting Delayed and Immediate memory results	101
4.3.5 Experiments 1 & 2: Attentional Performance during encoding	102
4.4 Discussion.....	103
4.4.1 Summary	103
4.4.2 Integration with existing literature.....	104
4.4.3 Alternative Accounts	105
4.4.4 Caveats.....	106
4.4.5 Conclusions	107
5. Discussion	108

5.1 Summary of findings	108
5.2 Synthesizing findings into larger model.....	113
5.3 Other considerations	117
5.3.1 Limitations.....	117
5.4 Future Directions	123
5.5 Conclusions	125
References	127
Biography	150

List of Tables

Table 1: Behavioral performance. Reaction times and corrected recognition memory... 46

List of Figures

Figure 1: A model integrating expectation, the dopamine system, and memory. Unexpected and expectancy violating events drive dopaminergic circuitry that could result in phasic and ramping dopamine release in the hippocampus. Green triangles represent the putative pattern of dopamine release. Hippocampal dopamine has the capacity to modulate memory formation. 27

Figure 2: Experimental design of the incidental encoding task. A categorical decision cue indicated the response to be made on each trial. Color pictures of objects were subsequently presented in sequences of three objects, with a train of three pre-familiarized sequences per trial. Responses were made to the final object in each sequence. The first two sequences comprised the expectancy generation period whereas the last sequence was the outcome. Expectancy was generated for the novelty or familiarity of the final object in each sequence as follows: In half of all trials the final object of the first two sequences was part of the pre-familiarized sequence; this pattern predicted the same object in the final sequence eighty percent of the time. In the other half of trials, the final object of the first two sequences was a unique novel object; this pattern predicted a novel object in the final sequence eighty percent of the time. On twenty percent of trials, the final object violated expectations and was thus surprising. There were four possible conditions at outcome: expected familiar (EF), surprising novel (SN), expected novel (EN), and surprising familiar (SF). 37

Figure 3: Novelty and surprise are distributed across the hippocampus. Clusters of activation were identified in the hippocampus for all unpredicted events relative to those that were both expected and familiar (novelty, EN > EF; surprise, SF > EF; and surprising novelty, SN > EF). All activations were small volume cluster corrected for the anatomical bilateral hippocampus at a primary threshold of $p < 0.01$, such that 19 contiguous voxels resulted in an overall threshold of $p < 0.05$ 48

Figure 4: VTA signals both surprise and novelty while NAcc selectively signals surprise. Mean activations were extracted across our anatomical *a priori* ROIs. A two-way repeated measures ANOVA with novelty and surprise as within-subjects factors revealed main effects of novelty and surprise in the VTA, but only a main effect of surprise in NAcc. A three-way ANOVA with novelty, surprise, and region as within-subject factors revealed an interaction of region by surprise, confirming a larger surprise effect in NAcc than VTA (* $p < 0.05$). 49

Figure 5: Brain regions engaged by novelty (yellow), surprise (blue), and their overlap (green). A, main effects maps from a whole brain two-way repeated measures ANOVA with surprise and novelty as the within-subject factors. B, conjunction maps for novelty (combining contrasts EN > EF and SN > SF) and surprise (combining contrasts SF > EF and SN > EN). All activations were cluster corrected at a primary threshold of $p < 0.001$, such that 21 contiguous voxels resulted in an overall threshold of $p < 0.05$ 51

Figure 6: Experimental design screening and incidental encoding. Participants completed a screening phase (A) from which trivia questions were sorted into high- and low-curiosity designations. In the MRI scanner, participants completed a study phase (B) in which they saw a trivia question from screening, experienced an anticipatory delay, and saw the trivia answer. Receipt of the trivia answer was contingent on a button press for half of all trials. 65

Figure 7: Curiosity and action contingency impact memory. Not only were there main effects of curiosity and action contingency on memory performance, there was also a significant curiosity by action contingency interaction such that curiosity differences were greater when participants were not required to make a motor action. (* $p < 0.05$). 74

Figure 8: Cue-related activation. Main effects of curiosity (A), individual differences in curiosity (B), action contingency by curiosity interaction (C), and memory by curiosity interaction (D). 75

Figure 9: Anticipation-related activation. Main effects of curiosity (A), individual differences in curiosity (B), action contingency by curiosity interaction (C), and memory by curiosity interaction (D). 78

Figure 10: Outcome-related activation. Main effects of curiosity (A), action contingency by curiosity interaction (B), and memory by curiosity interaction (C). 81

Figure 11: Figure 1. Experimental design: incidental encoding task. The task was designed to dissociate two physiological profiles of a putative dopamine response during reward anticipation - a phasic response that occurs rapidly and scales with expected value and a sustained response that increases with uncertainty (Fiorillo et al. 2003). Shaded triangles model these dopamine profiles relative to the cue, encoding, outcome, and probe events in each trial. Cues associated with 100%, 50%, or 0% reward probability were presented for 400ms. Incidental encoding objects were presented for 1s either immediately following the cue or shortly before anticipated reward outcome (400ms). Finally, a probe asking participants whether or not they received a reward, with the yes/no location counterbalanced, followed for 1s. 93

Figure 12: Figure 2. Memory performance for early and late encoding items at 24-hour and immediate retrieval. A. In the 24-hour retrieval group, early memory linearly increased with expected value; late memory was greatest for items encoded during reward uncertainty. This pattern was supported by significant memory for guesses only at the highest expected value early and during reward uncertainty late. B. In the immediate retrieval group, early memory did not differ by expected value; late memory was greatest for items encoded during reward uncertainty. There was no significant memory for guesses at this retrieval period. (* $p < 0.05$)..... 98

Figure 13: Retrieval group by expected value interaction. Within the early encoding period, there was a significant interaction between reward probability and retrieval group. There was a smaller difference between the 24-hour and immediate retrieval groups at 100% reward probability than at 0% reward probability, consistent with increased consolidation for higher probability items encoded early during anticipation. (* $p < 0.05$)..... 101

Figure 14: Figure 4. Attention for early and late encoding items at 24-hour and immediate retrieval. A. In the 24-hour retrieval group, early accuracy and early reaction time did not differ by expected value. Late accuracy linearly increased with increasing expected value. Late reaction times were slower for the 50% trials than the 0% trials B. In the 15-minute retrieval group, the patterns of accuracy and reaction time were the same as for the 24-hour retrieval group. (* $p < 0.05$)..... 103

Figure 15: An updated model integrating expectation, the dopamine system and memory. Our data support claims from the model in the introduction. It additionally suggests NAcc scales with the salience or value of an event to an individual. During a sustained expectancy period, PFC may hold the expectation in working memory. Memory may scale with uncertainty or salience, which both influence dopamine release in the hippocampus. 114

Acknowledgements

I would like to thank all those who supported the work that went into completing this dissertation. First of all, I would like to acknowledge the guidance and support of my dissertation advisor, Alison Adcock, who spent countless hours working with me to hone my scientific mind. I would also like to thank Vishnu Murty, Betsy Sumner, and Kathryn Dickerson for their caring and technical mentorship. Fellow lab members Nathan Clement, Mai-Anh Vu, Kimberly Chiew, Jeff MacInnes, Courtnea Rainey, Taylor Jackson, Kristin Duffy, Rita Loiotile, Lee Gans, and Laura Lerebours were an amazing group of people to work with every day. Faculty members Tobias Egner and Joe McClernon were very supportive mentors and collaborators. I would also like to thank my committee members, Scott Huettel, Roberto Cabeza, and Nan-kuei Chen for their helpful feedback and encouragement.

Finally, I would like to thank my family for their unconditional love and support. My parents, Diane and Michael Wilson, always believed in me and pushed me to be my best. My sister Rachel Wilson has been my rock my entire life and I wouldn't be here without her as a role model. And a huge thank you to my husband Ted, who has made the world a more fun place to be.

1. Introduction

1.1 *Expectation*

Expectation is the belief that something will happen or is likely to happen. It arises via learning from previous experiences and allows organisms to anticipate and make predictions for the future. Balkwell put it succinctly when describing social behavior: he claimed that “processes exhibit statistical regularities” and “a person’s perceptions and actions are mediated by cognitive expectation states” that result from these regularities (Balkwell, 1991). A single experience or many repeated experiences can dictate future expectations, and what we expect to encounter influences our perception, behavior, and, as I will argue, our memory for the world around us.

On one level, relations between objects in space and time can be encoded from a single event (Staresina & Davachi, 2009). For instance, a specific plant may produce edible berries and that association may be rapidly encoded. However, over many repetitions, relationships across space and time can be extracted more robustly by implicit statistical learning (Turk-Browne, Jungé, & Scholl, 2005), e.g., the berries occur near leaves that have a three-pronged shape, but only in the summertime.

When an event is unexpected or violates expectations, the error between the expected event and the actual event may cause an organism to explore its environment to determine the cause of the error (Stahl & Feigenson, 2015) and create a learning signal that results in the updating of expectations (Cohen, Elger, & Ranganath, 2007).

Continuing the above example, an animal may now learn that the three-pronged leaves only cue the presence of berries in the absence of thorns. Formal learning theory – for instance, models describing classical conditioning or reinforcement learning – such as the Rescorla-Wagner model or Pearce-Hall model of associative learning, relies on such updating from errors caused by expectation (Pearce & Hall, 1980; Rescorla & Wagner, 1972). In these models, unexpected and surprising events serve as important learning signals by changing the associative strengths between events. Organisms are constantly building expectations and subsequently encountering either expected or surprising (expectancy violating) events, then updating their model of the world to further develop future expectations.

Building and maintaining expectations has important consequences for behavior and memory, and has even been proposed as a unifying theory of brain function (Friston & Stephan, 2007). In the upcoming section, I will specifically describe the relationship between expectation and episodic memory. I will relate explicit expectations and memory, implicit expectations and memory, and expectations of valenced outcomes and memory.

1.2 Expectation modulates memory formation

For better or for worse, expectations shape memory. Explicit expectation of a test encourages individuals to better remember information, utilizing strategies selected to maximize recognition or recall test performance (Naus, Ornstein, & Kreshtool, 1977;

Tversky, 1973). Expectation of a final, cumulative test enhances long-term retention of encoded material; conversely, expectation of no memory test results in worse memory performance (Szpunar, McDermott, & Roediger, 2007). In these instances, explicit expectations about the relevance of information during encoding improves memory outcomes.

Likewise, explicit stimulus category expectation (such as cued expectation for upcoming faces or scenes) benefits long-term memory formation (Bollinger, Rubens, Zanto, & Gazzaley, 2010). Fulfilled expectations are associated with facilitated object recognition and cognitive processing (Esterman & Yantis, 2009; for a review see Summerfield & Egnér, 2009), and additional research has shown support for the idea that engaging top-down cognitive control (which comes online when expectations are formed) benefits long-term memory (Richter & Yeung, 2014). Specifically, it has been argued that selective attention guides memory (Chun & Turk-Browne, 2007). Thus, when expectations explicitly orient attention to relevant events, memory is facilitated.

Even implicit expectations change memory formation. In 1933, von Restorff described a phenomenon in which distinctive items are better remembered on a later memory test (Restorff, 1933). Although this has been discussed as a novelty advantage (Kishiyama & Yonelinas, 2003), the absolute novelty of any given item is not different from the others. Rather, as noted by Green (Green, 1956), surprise is an important factor, in that unexpected change benefits later recall. Within a study session, participants

develop expectations based on the similarity of items, and when that expectation is violated, memory is enhanced.

The expectancy violation interpretation of the “von Restorff effect” is coherent with the “consistency effect,” which argues that items inconsistent with schema expectation are better remembered (Friedman, 1979; Light, Kayra-Stuart, & Hollander, 1979). For instance, participants encoding items in a setting such as an office or a classroom are more likely to later correctly recall items that would have been statistically less likely to be present (Pedzek, Whetstone, Reynolds, & Askari, 1989). Another distinct iteration of the consistency effect is that participants show better memory for weakly-related word pairs than strongly-related ones (Hirshman, 1988). Additionally, people are more likely to remember the presence of feedback when the feedback does not match expectations (Fazio & Marsh, 2009). In the social realm, study participants have better memory for partners that violate behavioral expectations (Chang & Sanfey, 2009). A recent study demonstrated that memory for expectancy violations is stronger when the violations pertain to humans or animals than plants or objects, suggesting that more relevant violations are more strongly encoded (Porubanova, Shaw, McKay, & Xygalatas, 2014). Thus, as events deviate from schema expectation and expectations are violated, memory is enhanced.

Early mixed results in this field led to a meta-analysis revealing that when guesses are not included, memory is better for schema-inconsistent information, but

when guesses are included, memory is better for schema-consistent information (Rojahn & Pettigrew, 2011). This is consistent with the finding of more false alarms for words semantically associated with a list (Roediger & McDermott, 1995); these words are items that participants likely expected to see based on the other words encountered in the list. Similarly, cultural expectations shape memory for sentences (Federmeier & Kutas, 1999) and music (Curtis & Bharucha, 2009), with more false alarms observed to culturally consistent events. Integrating across the above results suggests that individuals use schemata and expectations to make guesses about the content of memory, which can be beneficial if events were in fact consistent with expectations. However, violations of expectations are more strongly encoded (not merely guessed), with an enhancement of memory for these events beyond those that are expected.

In addition to expectation about the content of an upcoming event, expectation of valenced outcomes, such as reward or punishment, also influences memory formation. Memory is enhanced for items associated with high versus low reward outcomes (Adcock, Thangavel, Whitfield-Gabrieli, Knutson, & Gabrieli, 2006; Murty & Adcock, 2013; Wittmann et al., 2005). Memory is also better for information that is expected to satisfy high curiosity questions, rather than low-curiosity questions (Gruber, Gelman, & Ranganath, 2014). Likewise, expectation of possible punishment facilitates memory that will help avoid aversive events (Murty, LaBar, & Adcock, 2012). In future sections, I will discuss possible biological mechanisms of this enhancement.

1.2.1 Summary of expectation modulates memory formation

Expectation captures our beliefs about the likelihood of future events. It is established based on previous experience and dictates future perception, behavior and memory. As briefly reviewed here, the relationship between expectation and memory has been examined in many ways in the literature. Explicit expectations about the relevance or content of information enhance memory for that information. When implicit, schema-generated expectations are violated, memory for the violations is enhanced. However, events that occur consistent with expectations are more readily guessed; this suggests that at different times, expectation-consistent or expectation-violating experiences can be beneficial for memory. Finally, expectations of reward or punishment (or other motivationally significant events) modulate memory for associated events, and the mechanism of this enhancement will be discussed further in later sections.

1.3 Expectation in the mesolimbic dopamine system

Expectation not only impacts behavioral output such as memory, it also modulates brain activity (for a review, see Friston, 2010). The mechanism of modulation that will be focused on in this dissertation is the dopamine neuromodulatory system. Dopamine is a neuromodulator with the ability to excite or inhibit neuronal activity in many regions throughout the brain (for a review, see Seeman, 1980). It is produced in several nuclei, including the ventral tegmental area (VTA) and substantia nigra (SubN)

(Nair-Roberts et al., 2008), which are collectively referred to as the dopaminergic midbrain. The SubN projects primarily to the dorsal striatum via the nigrostriatal pathway (Beckstead, Domesick, & Nauta, 1979). Dopamine cell death in SubN is the cause of Parkinson's disease and is most closely associated with movement disorders (Hirsch, Graybiel, & Agid, 1988). The VTA projects to the prefrontal cortex via the mesocortical pathway and to the nucleus accumbens and hippocampus via the mesolimbic pathway (Swanson, 1982). The mesocortical pathway is associated with motivation, cognitive control, and working memory, whereas the mesolimbic system is associated with motivation (specifically reward) and long-term memory formation (Krebs, Heipertz, Schuetze, & Duzel, 2011; Masayuki Matsumoto & Takada, 2013). Here, I will further discuss how expectation is coded in the mesolimbic dopamine system.

Mesolimbic dopamine has long been associated with reward (for a review, see Schultz, 2002); however, dopamine also has a relationship with the cuing and resolving of expectations. Seminal work by Shultz and colleagues (Schultz, Dayan, & Montague, 1997) showed that dopamine neurons in the midbrain fire in response to unexpected rewards or reward cues, leading many subsequent researchers to examine the relationship between the VTA and "reward prediction errors" (for a review, see Schultz, 2016). However, even prior to this extremely impactful study, dopamine was implicated in the signaling of neutrally valenced, unexpected events such as loud, auditory clicks or flashes of light (Horvitz, Stewart, & Jacobs, 1997; Ljungberg, Apicella, & Schultz, 1992;

Steinfels, Heym, Strecker, & Jacobs, 1983). In particular, salient stimuli elicited a burst-firing response in the dopaminergic midbrain, with the activation habituating and decreasing over successive repetitions of the event (Ljungberg et al., 1992).

In the following sections, I will discuss what is known about the dopaminergic response to cues that induce expectancy of many different kinds, as well as the dopaminergic response to the resolution of expectations. I will argue that expectations are an important driver of the mesolimbic dopamine system and thus have meaningful impact on the downstream memories and behaviors that are adaptively modulated by dopamine.

1.3.1 Cued expectation

Cuing events that induce expectancy have consistently activated dopamine neurons in the dopaminergic midbrain, including the ventral tegmental area (VTA) in particular. Cues that predict an upcoming reward have received the most attention in the literature; however, there is also a VTA dopamine neuron response to cues that predict punishment, information, novelty, or movement.

Cues that generate expectations about future rewards elicit a burst-firing response in midbrain dopamine neurons (Fiorillo, Tobler, & Schultz, 2003; Schultz et al., 1997; Tobler et al., 2005). First demonstrated in non-human primates, animals that have learned the relationship between a non-rewarding cue and a primary reward outcome show this pattern of burst firing to the cue. Burst firing results in rapid, transient

dopamine release in the nucleus accumbens (Roitman, Stuber, Phillips, Wightman, & Carelli, 2004; Stuber, Roitman, Phillips, Carelli, & Wightman, 2005) and has demonstrated synaptic effects in the hippocampus (Rosen, Cheung, & Siegelbaum, 2015), among other downstream regions in the mesolimbic pathway. Dopaminergic firing has been shown to scale with the expected value of the reward, including both reward magnitude and probability (Tobler et al., 2005). Importantly, the organism does not receive the reward at the time of the cue. Further, when the cue is paired with the reward, but animals do not know the predictive relationship between the two, the dopamine response is absent. Together, these findings imply that phasic dopamine response at the cue signals the initiation of meaningful expectancy, in this case, of meaningful reward. The strength of the expectation contributes to the dopamine firing in the midbrain.

Not only do reward cues elicit phasic dopamine release, there is also evidence that reward cues elicit sustained dopamine release as well (Fiorillo et al., 2003; Howe, Tierney, Sandberg, Phillips, & Graybiel, 2013; Totah, Kim, & Moghaddam, 2013). This sustained signal lasts on the order of seconds and ramps up to the reward receipt. Ramping, sustained dopamine firing has been demonstrated to scale with reward uncertainty (Fiorillo et al., 2003). Although there has been some argument that the ramp in dopamine firing rate is the result of additive phasic responses (Niv, Duff, & Dayan, 2005), work by other research teams has also provided support for a sustained

dopamine response. Recently, Howe and colleagues showed increasing dopamine in the nucleus accumbens as rodents navigated a maze toward a reward (Howe et al., 2013). In this instance, the slope of the ramp scaled with reward magnitude. Finally, the sustained signal has also been shown to be greater on trials in which rodents are anticipating making the correct behavioral choice to receive reward, rather than anticipating the incorrect choice (Totah et al., 2013). It still remains to be resolved exactly what this sustained dopamine ramp during expectancy is signaling, and the studies in this dissertation will explore this question further.

Using functional magnetic resonance imaging (fMRI) techniques in humans, it has been demonstrated that the VTA is more strongly activated for cues that predict the possibility of high reward rather than low reward (Adcock et al., 2006). Simultaneous fMRI-PET has shown that the fMRI blood-oxygenation-level dependent (BOLD) response in the VTA does in fact scale with dopamine release (Schott et al., 2008). We can therefore say with some confidence that BOLD VTA activation is a correlate of VTA dopamine. However, it is still unclear whether VTA activation in fMRI is a correlate of phasic dopamine, sustained dopamine, or both.

Although dopamine is often associated with reward, there is research examining the relationship between dopamine and aversive event cues as well (for a review, see Bromberg-Martin, Matsumoto, & Hikosaka, 2010). The results in the literature are mixed, with some studies associating aversive cues with increased dopamine (Guarraci

& Kapp, 1999) and others associating these cues with inhibited dopamine (Ungless, 2004). Subsets of dopamine neurons may signal signed (i.e., positive or negative) motivational valence while others signal motivational salience, regardless of sign (Masayuki Matsumoto & Hikosaka, 2009). Within a single study, dopamine neurons either increased or decreased their firing rates in response to cues predicting aversive events (Cohen, Haesler, Vong, Lowell, & Uchida, 2012), although these changes in firing were more moderate at cue than at the aversive outcome. Thus, it would seem that dopamine has some role in setting expectations for aversive events, although the mechanism and detailed context remains unclear.

Cues that generate expectations about upcoming information, even without guaranteeing receipt of an extrinsic primary or monetary reward, are also capable of eliciting a response in the VTA. Midbrain dopamine neurons in non-human primates signal the expectation of information (Bromberg-Martin & Hikosaka, 2009). Although this is interpreted as information being “rewarding,” another perspective is that the cue is generating behaviorally relevant expectations about a future event. The human dopaminergic midbrain is also activated in response to cues that generate expectancy about upcoming information. Gruber and colleagues showed that the SubN/VTA response to a trivia question parametrically scales with individuals’ curiosity about the answer (Gruber et al., 2014). Thus, expectations about the content of future information drive the dopaminergic midbrain in humans as well.

Although novelty can be difficult to expect with a precise representation (given that we define novelty here as a previously unexperienced sensory event), the general expectation of experiencing something new can in fact be generated by previous events. For instance, experiencing a novel object in an unfamiliar environment could cue one to expect another novel object. And in fact, a human fMRI study has demonstrated that BOLD activation in SubN/VTA is stronger for cues that predict novel versus familiar neutrally valenced images (Wittmann, Bunzeck, Dolan, & Duzel, 2007). Novelty is not inherently rewarding – it can be aversive, appetitive, or neutral depending on the context. However, it implies a new sensory experience to learn about, and thus may be more behaviorally relevant than a known familiar event. Thus, cues that generate expectancy about a more behaviorally relevant outcome also generate activation in the dopaminergic midbrain.

Finally, many of these studies of motivation, information, and novelty are confounded in that the cues also generate expectations of movement. For instance, reward receipt may involve moving the mouth or making a button press to receive money. Generating expectations for action may be enough to trigger dopamine neurons in the VTA. However, dopamine neurons in the neighboring (and historically difficult to dissociate) SubN are more strongly tied to action (DeLong, Crutcher, & Georgopoulos, 1983). That said, cues that generate expectations of actions may activate the VTA as well, but more work remains to be done to say for certain.

What computation is dopamine performing in response to expectancy-generating cues? Dopamine has been proposed to generate an alerting signal and orient attention (Redgrave & Gurney, 2006). It may also provide the signal to initiate action plans downstream in NAcc (Murty, Stanek, & Heusser, 2013) or memory processing in the hippocampus (Duzel, Bunzeck, Guitart-Masip, & Duzel, 2010; Shohamy & Adcock, 2010). The VTA may also have discrete populations of dopamine neurons that encode motivational valence versus motivational salience (Bromberg-Martin et al., 2010). The dorsolateral-ventromedial axis of the dopaminergic midbrain may separate cognitive and motivational signals (Masayuki Matsumoto & Takada, 2013), however, both signals occur when some future event of behavior relevance is to occur. Thus, we suggest that a fundamental purpose of the mesolimbic dopamine system is to put the brain in a state of expectancy that dictates future behavior.

1.3.2 Resolving expectation

An event that occurs in the absence of any expectation can be deemed 'unexpected.' However, following the generation of expectancy, an event either meets or violates expectations. Thus, three events types may occur at any given time: events that are expected, are expectancy violations, or are unexpected. The relationships between dopamine and these events that resolve expectation are different, and help elucidate an important role of mesolimbic dopamine signaling.

Unexpected events, even those that are unrewarded, have long been associated with dopamine. Steinfels and colleagues (Steinfels et al., 1983) showed that auditory clicks elicit dopamine firing in cats as long as they cause an orienting response. As the animal habituates to the sound and no longer responds to it, dopamine firing ceases. This finding is among the first evidence that unexpected events trigger a dopamine signal because the animal interprets them to be behaviorally relevant: there may be something to learn about in the environment. Further evidence supporting this viewpoint comes from research showing that dopamine neurons in non-human primates fire initially in response to an unexpected door opening to an empty food box, but this response habituates over time as the animal learns what to expect from this event (Ljungberg et al., 1992). Additionally, research has demonstrated that brief, non-conditioned auditory and visual stimuli lead to burst firing when delivered with randomly varied inter-stimulus intervals from 10-30s (Horvitz et al., 1997). When considering motivationally valenced unexpected events, either rewards or learned reward cues that are unexpected elicit burst-firing in midbrain dopamine neurons (Schultz et al., 1997; Schultz & Hollerman, 1998). Integrating across all these works has led to the idea that dopamine neurons fire in response to salient and arousing changes in environmental conditions, regardless of motivational valence (Horvitz, 2000), and that this could be an important alerting signal (Redgrave & Gurney, 2006; Redgrave, Prescott, & Gurney, 1999). Work in humans has shown VTA activation in response to

unvalenced unexpected novel events but not familiar ones (Bunzeck & Duzel, 2006; Wittmann et al., 2007). Neurons in the human NAcc also fire more for unexpected novel events (Axmacher et al., 2010) suggesting that the mesolimbic pathway may also be playing a role in this alerting response. More work remains to be done to determine whether a VTA response to unexpected, familiar events can be detected in humans.

Conversely, as events become expected, both temporally and in content (e.g. a flash of light or a reward of a specific type and magnitude), the dopamine response disappears (Ljungberg et al., 1992; Schultz et al., 1997). Interestingly, even a cue predictive of reward that is overlearned no longer elicits a dopamine response (Schultz & Hollerman, 1998).

Perhaps most widely discussed in the literature is the relationship between dopamine and violations of expectation. Violations of expectation occur when an event is different than what was predicted. More than unexpected events, these can be argued to elicit feelings of surprise, which is associated with dopamine (Barto, Mirollo & Baldassarre, 2013). For instance, reward prediction errors are expectancy violations in which reward receipt is different than the expectation of reward generated by a cue (for a review, see Colombo, 2014). This can occur because reward magnitude is greater than expected (Tobler et al., 2005), with a 'pause' in firing when less than expected (Schultz et al., 1997), or because a reward occurs when the event is probabilistic (Fiorillo et al., 2003). The same dopamine neurons have been shown to be sensitive to errors of both

reward magnitude and reward probability; thus, these neurons are comparing the outcome to the overall integrated expected reward value (Tobler et al., 2005). This has been shown to occur across many dimensions (Lak, Stauffer, & Schultz, 2014), and has been interpreted as a dopamine response signaling an expectancy violation relative to subjective rather than physical reward (Fiorillo, Newsome, & Schultz, 2008).

A recent study examined whether expectancy signals and reward signals are computed by the same or distinct populations of VTA dopamine neurons (Eshel, Tian, Bukwich, & Uchida, 2016). The authors found that dopamine firing increased as a function of reward magnitude. However, the function linearly scaled with expectancy, such that the firing rate was consistently higher for unexpected rewards, across all magnitudes. Thus, while dopamine neurons do encode motivational valence, another important and consistent function is to signal potentially relevant unexpected events in the environment.

What then is the purported relationship between expectation and dopamine? What is the functional signal and what are the downstream implications? Mesolimbic dopamine signals either the generation of expectation or a breach in expectations. This may be either an alerting/orienting signal (Redgrave & Gurney, 2006) or a learning signal (Roesch, Esber, Li, Daw, & Schoenbaum, 2012; Waelti, Dickinson, & Schultz, 2001) or both (Schultz, 2016). In any case, events must be behaviorally relevant or dopamine is absent, as is motivationally salient expectation. This means that downstream,

expectations may have the capacity to modulate actions and learning via dopaminergic mechanisms. This will be discussed more in the next sections.

1.3.3 Summary of expectation in the mesolimbic dopamine system

Midbrain dopamine neurons fire in response to cues that generate expectations or behaviorally relevant outcomes that are either unexpected or violate expectations. Dopamine neurons are activated by cues that predict reward, novelty, information, or aversive events. Their burst-firing results in transient phasic dopamine release. However, several studies suggest that there may be sustained dopamine firing and release that ramps over the course of expectancy. Future studies are still needed to characterize this lasting dopamine signal. Events that are fully expected do not excite dopamine neurons nearly as much as those that are unexpected or violate expectations. Interestingly, dopamine signals the violation of the subjective predicted experience rather than the actual predicted experience. Thus, the dopamine response may provide either or both an alerting or learning signal that can modulate future actions and learning.

1.4 Mesolimbic dopamine shapes memory formation

Since patient H.M. first underwent bilateral resection of the medial temporal lobe (MTL) and demonstrated an impaired ability to create new episodic memories, the MTL and the memory system have been tightly linked (Scoville & Milner, 1957). In the years since then, a host of studies have deemed that, within the MTL, the hippocampus is

critically involved in episodic memory formation (for reviews, see Burgess, Maguire, & O'Keefe, 2002; Eichenbaum, 2000; Rolls, 2010; Tulving & Markowitsch, 1998; Wang & Morris, 2009), including both encoding and consolidation.

Recognition of the importance of dopamine in modulating hippocampus-dependent episodic memory formation was slower to emerge, but has garnered considerable empirical support. Initially, the lack of episodic memory impairment in dopamine-depleted Parkinson's disease patients seemed to indicate that dopamine was not necessary for this type of memory (for a review, see Shohamy, Myers, Kalanithi, & Gluck, 2008). However, dopamine depletion in Parkinson's disease primarily originates from the SubN in the midbrain, rather than the VTA, and dopamine release in the dorsal striatum is drastically reduced (affecting the nigrostriatal pathway, rather than the mesolimbic pathway) (Kish, Shannak, & Hornykiewicz, 1988). In fact, as will be discussed in the following sections, there is a strong relationship between mesolimbic dopamine originating in the VTA and hippocampus-dependent memory.

In the upcoming sections, I will discuss the anatomical and functional neuroarchitecture binding the VTA and the hippocampus and possible mechanisms of dopaminergic memory enhancement. I will then briefly review the literature demonstrating VTA activation in support of episodic memory. Finally, I will expand beyond the VTA and hippocampus and examine a broader dopaminergic network that modulates memory, including the nucleus accumbens and the prefrontal cortex.

1.4.1 Dopamine and hippocampus-dependent memory

Dopamine modulates hippocampus-dependent memory formation (Bethus, Tse, & Morris, 2010; O'Carroll, Martin, Sandin, Frenguelli, & Morris, 2006). Anatomically, VTA afferents project to the hippocampus (Gasbarri, Packard, Campana, & Pacitti, 1994; Gasbarri, Sulli, & Packard, 1997). This provides a means by which dopamine neurons in the VTA can cause dopamine release downstream in the hippocampus, modulating the extent to which episodic sensory information is stored in long-term memory. Within the hippocampus, dopamine terminals as well as dopamine D1, D2, and D5 receptors are all present, with dopamine D1/D5 receptors more prevalent than D2 receptors (for a review, see Shohamy & Adcock, 2010). Activity at dopamine receptors modulates synaptic strength, altering the threshold needed for induction of long-term potentiation (LTP) and long-term depression (LTD), which are critical for long-term memory formation (Lemon, 2006). In particular, D1/D5 receptors in CA1 and subiculum are necessary for dopamine-induced enhancements in memory (Lemon, 2006; Lisman & Otmakhova, 2001; Martig & Mizumori, 2011; Otmakhova & Lisman, 1996; Roggenhofer et al., 2010; Vago, Bevan, & Kesner, 2007) and inactivation or deficiency in D1/D5 receptors leads to impairments in learning and memory (Moraga-Amaro et al., 2016).

One putative mechanism of dopamine-induced memory enhancement is described by the synaptic tagging and capture hypothesis (for reviews, see Lisman, Grace, & Duzel, 2011; Redondo & Morris, 2011; Takeuchi, Duzkiewicz, & Morris, 2013).

In this model, an event (either tetanus in slice physiology, or an event of motivational significance such as novelty in behavior) triggers dopamine release in the hippocampus, which “tags” nearby synapses and creates plasticity-related proteins (Moncada & Viola, 2007; Sajikumar, 2004). These plasticity-related proteins are present when a temporally adjacent event occurs, strengthening the synapse for LTP. This LTP is critical for the consolidation of an event into long-term memory. Synaptic tagging and capture provides a mechanism by which phasic dopamine triggered at encoding stabilizes memory representations over time.

While the above model only includes contributions by rapid, phasic dopamine release in the hippocampus, another potential mechanism for memory enhancement takes into account the contributions of sustained or tonic dopamine in the hippocampus (Li, Cullen, Anwyl, & Rowan, 2003; Rosen et al., 2015; Shohamy & Adcock, 2010). It has been noted that there is some distance between dopamine receptors in the hippocampus and dopamine terminals from the VTA (for a review, see Shohamy & Adcock, 2010). Thus, in addition to synaptic dopamine, extrasynaptic dopamine may also modulate memory formation (Floresco, West, Ash, Moore, & Grace, 2003). There is evidence that tonic (putatively more similar to sustained) dopamine has downstream effects extrasynaptically (Floresco et al., 2003); thus, sustained dopamine firing and release may act via an extrasynaptic mechanism to modulate hippocampal memory. This sustained dopamine is likely to be influential during and following events of sustained

motivational significance, such as novelty exposure (Fenker et al., 2008; Li et al., 2003) or reward anticipation (Adcock et al., 2006; Fiorillo, 2003). Thus, both rapid and sustained dopamine may have the capacity to enhance memory formation around events that cause dopamine release.

1.4.2 Behavioral and functional neuronal support for dopaminergic contributions to memory formation

While anatomy and slice physiology provide support for the relationship between VTA dopamine release and hippocampal memory formation, many studies in rodents and humans have contributed to our knowledge about what behaviors activate the VTA and hippocampus in support of memory. Seminal work in this field demonstrated that reward anticipation co-activates the dopaminergic midbrain and the hippocampus, and that functional connectivity between these regions relates to subsequent memory encoding success (Adcock et al., 2006; Wittmann et al., 2005). High-resolution fMRI has implicated hippocampal regions dentate gyrus and CA2/CA3 (Wolosin, Zeithamova, & Preston, 2012) and caudal, medial portion in the dopaminergic midbrain (Krebs et al., 2011) as particularly important for reward-related memory and novelty encoding. Very recently, pattern separation (by which two perceptually similar events are identified as different) has been shown to be enhanced in neutral contexts perceptually similar to prior contexts in which reward was received (Loh et al., 2015). Additional neuroimaging studies in humans have also demonstrated VTA and hippocampus activation contributing to memory for novelty (Bunzeck & Duzel, 2006;

Wittmann et al., 2007), associative novelty (Schott, 2004), generalization (Shohamy & Wagner, 2008), expectancy violations (Murty & Adcock, 2013) and curiosity (Gruber et al., 2014). This wide range of phenomena for which memory is enhanced, concurrent with mesolimbic activation, is consistent with the proposal that dopamine neurons respond to salient and arousing changes in environmental conditions, independent of motivational valence (Horvitz, 2000; Redgrave et al., 1999). Thus, we henceforth consider events that have motivational significance to trigger dopamine release and modulate hippocampal memory formation.

With the advent of optogenetics in rodents, a causal relationship between midbrain dopamine and hippocampus-dependent memory is starting to emerge. Optogenetic stimulation of VTA dopamine neurons, simulating either tonic or phasic dopamine release, impacts CA1 responses in the hippocampus (Rosen et al., 2015). Optogenetic burst stimulation of hippocampal dopaminergic fibers from midbrain neurons in mice exploring novel environments has also been shown to enhance neural reactivation during subsequent sleep and also improve later memory performance (McNamara, Tejero-Cantero, Trouche, Campo-Urriza, & Dupret, 2014). Further work illuminating the causal relationship between VTA, hippocampus, and memory is sure to emerge in the near future.

1.4.3 Broad dopaminergic circuitry contributes to memory

While this introduction to dopamine and memory has focused on the relationship between the VTA and the hippocampus in isolation, these two regions are part of a dopaminergic circuit that both modifies and is updated by VTA and hippocampus. Several anatomical and functional models describe these neural networks (Lisman & Grace, 2005; Shohamy & Adcock, 2010) as follows. Although the VTA has direct afferent projections to the hippocampus, the hippocampus sends signal to the VTA through several key relays, including one path through the NAcc and globus pallidus, which combine to disinhibit the VTA (Floresco et al., 2003; Floresco, Todd, & Grace, 2001). Another connection is through PFC, which receives efferents from the hippocampus (Carr & Sesack, 1996; Jay & Witter, 1991) and has direct excitatory projections to the VTA (Frankle, Laruelle, & Haber, 2006). Both the routes through NAcc and PFC may also signal the pedunculopontine nucleus, which could also be a source of excitation for the VTA (Lisman & Grace, 2005). Additionally, the laterodorsal tegmentum and lateral habenula project directly to the VTA (Lammel, Ion, Roeper, & Malenka, 2011) and could serve as relay stations to excite or inhibit dopamine neurons. The VTA in turn has direct dopaminergic release in the nucleus accumbens and hippocampus, along with the amygdala (Shohamy & Adcock, 2010). Together, these anatomical regions combine to form the mesolimbic dopamine pathway (including PFC, arguably the mesocorticolimbic pathway). For simplicity and due to the constraints of

our fMRI methods, we will further discuss only two more regions in detail: the NAcc and the PFC.

The NAcc serves as an important relay between the hippocampus and VTA (Lisman & Grace, 2005) and is also a robust site of dopamine release by the VTA (Rebec, Christensen, Guerra, & Bardo, 1997a; Sombers, Beyene, Carelli, & Mark Wightman, 2009). Recent work has demonstrated intrinsic connectivity between the hippocampus, NAcc, and VTA at rest (Kahn & Shohamy, 2012). Behaviorally, as in the VTA and hippocampus, NAcc has also been implicated in processing salient events, including reward, novelty, and expectancy violations (Knutson, Adams, Fong, & Hommer, 2001; Rebec, Grabner, Johnson, Pierce, & Bardo, 1997b; Zaehle et al., 2013). My colleagues and I proposed in a recent commentary that the NAcc signals the initiation of adaptive behaviors (Murty et al., 2013). Tonic dopamine release from the VTA modulates hippocampal inputs through D1 receptors, whereas phasic dopamine release modulates prefrontal cortical inputs via D2 receptors (Goto & Grace, 2005). Importantly, disrupting tonic dopamine release in NAcc results in impaired learning (Goto & Grace, 2005). Thus, NAcc is an important part of the pathway by which VTA dopamine modulates memory.

Along with NAcc, PFC also links the VTA and hippocampus (Floresco et al., 2003). Dorsolateral PFC in particular has been shown to support reward representations and serve as an entry point for subsequent reward-anticipation-related activity in VTA and hippocampus (Ballard et al., 2011). VTA connectivity with PFC at a reward cue also

predicts hippocampus activation in response to a subsequent expectancy violation (Murty & Adcock, 2013). In addition, whereas hippocampus activation positively predicts slow shifts in VTA activation, PFC activation positively predicts transient, event-evoked activation in the VTA (Murty, Ballard, & Adcock, 2016a). Together, these studies provide evidence that the PFC can influence phasic dopamine release by the VTA and that the more connected PFC and VTA are, the better the hippocampus encodes salient events.

Thus, the neuroimaging studies included in this dissertation address the NAcc and PFC among other regions that are co-activated with the VTA and hippocampus in support of episodic memory encoding.

1.4.4 Summary of mesolimbic dopamine shapes memory formation

In summary, the VTA and hippocampus are part of a dopaminergic system important for modulating memory formation. Dopamine neurons in the VTA project directly to the hippocampus. Both synaptic and extrasynaptic dopamine have the capacity to modulate memory, as a consequence of phasic and sustained dopamine, respectively. The synaptic tagging and capture hypothesis has been proposed as a mechanism for translating increased synaptic strength into long-term, consolidated memories. Many neuroimaging studies have reported the contribution of the VTA to memory enhancements, and VTA-hippocampus connectivity is important as well. Optogenetic studies are starting to demonstrate a causal link between VTA dopamine

and hippocampus-dependent memory and are expected to add further evidence in support of this relationship. Disruption of the NAcc or PFC also disrupts the ability of dopamine to modulate memory; thus, these regions are a critical part of the dopamine system involved in forming episodic memories. Work in this dissertation will focus on the contributions of the VTA dopamine system and the hippocampus to episodic memory, with some description of NAcc and PFC contributions as well.

1.5 A model integrating expectation, dopamine, and memory

Thus far, I have made three arguments. First, expectation modulates memory. Second, expectation is coded by the mesolimbic dopamine system. Third, the mesolimbic dopamine system provides a mechanism for modulating memory. Here, I integrate across these three arguments and propose a model by which expectation modulates memory formation via dopaminergic mechanisms. The empirical work discussed in this dissertation will be used to update and elaborate on the proposed model.

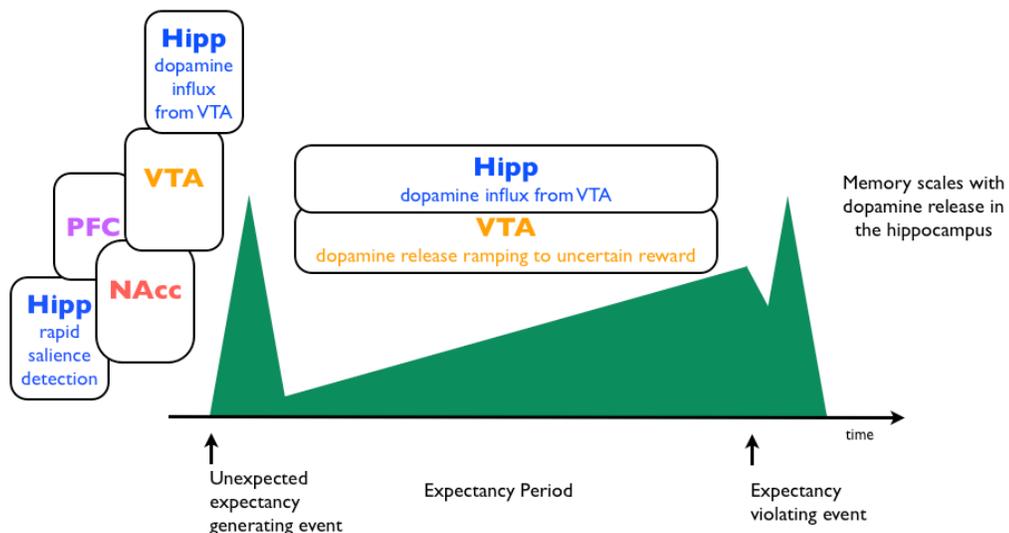


Figure 1: A model integrating expectation, the dopamine system, and memory. Unexpected and expectancy violating events drive dopaminergic circuitry that could result in phasic and ramping dopamine release in the hippocampus. Green triangles represent the putative pattern of dopamine release. Hippocampal dopamine has the capacity to modulate memory formation.

Events that unexpectedly generate motivationally significant expectancy use the mesolimbic dopamine neuromodulatory system to signal attention and orienting to another upcoming event. This can include initial events that trigger expectancy about reward, punishment, novelty, or information. Regions including (but not limited to) the hippocampus, nucleus accumbens, and prefrontal cortex may also contribute to coding this behaviorally relevant cueing event. During expectancy, either uncertainty or engagement in goal-directed behavior has the capacity to generate sustained, ramping dopamine firing in the VTA. This signal may or may not appear during other states of expectancy. At outcome, events that are entirely expected produce a very weak or non-existent dopamine response: there will be no additional dopamine release in downstream brain regions, including the hippocampus. However, events that violate expectations may then trigger a phasic response in the VTA, a signal that is consistent with learning from and updating expectations. The network activated by expectancy violation may or may not look different from the initial unexpected event. Physiological evidence suggests that phasic dopamine release from the VTA to the hippocampus benefits LTP and LTD, mechanisms by which long-term memory formation is enhanced. Extrasynaptic dopamine present as a result of sustained dopamine release in the hippocampus may benefit memory as well. Thus, expectancy modulates the amount of dopamine transmitted from the VTA to the hippocampus as a mechanism for impacting memory formation. Events that generate expectancy, occur during uncertainty, or

violate expectations should provide a dopaminergic neuromodulatory presence in the hippocampus beneficial for the formation of long-term episodic memories.

In the next section, remaining open questions will lead to experimental aims that can shed additional light on this model, which will then be revisited in the discussion of this dissertation.

1.6 Open questions and experimental aims

Many open questions still remain. In this section I will discuss some of the open questions examining the modulation of memory by expectancy via dopaminergic mechanisms. Each set of experimental aims (1, 2 and 3) represents an empirical chapter in this document.

Open question 1.1: Rewarded expectancy violations (e.g. reward prediction errors) drive the VTA (D'Ardenne, McClure, Nystrom, & Cohen, 2008). Novel expectancy violations also drive the VTA (Bunzeck & Duzel, 2006; Kafkas & Montaldi, 2014). Do non-novel, unvalenced expectancy violations (surprise) drive the VTA? Does novelty drive the VTA if it is expected? Neither of these has yet been demonstrated in human functional neuroimaging. In addition, what is the response in supporting mesolimbic circuitry? Does activation in the rest of the brain give us insight into the potential functional significance of the mesolimbic signal?

Aim 1.1: Characterize the independent contributions of expectancy violation (surprise) and novelty in driving the VTA and supporting mesolimbic circuitry.

Open question 1.2: There is a case in the literature that novelty is enhanced in memory via dopaminergic modulation (for a review, see Lisman & Grace, 2005). Does the expectancy of novelty influence the strength of memory formation?

Aim 1.2: Examine whether expectation of novelty influences episodic memory formation.

Open question 2.1: Expectation of reward drives mesolimbic circuitry at cue (Adcock et al., 2006). It also enhances memory for high reward events relative to low reward events (Adcock et al., 2006). Work in monkeys has demonstrated dopaminergic VTA activity for advance information about rewards (Bromberg-Martin & Hikosaka, 2009). To what extent does expectation of high-curiosity information resemble expectation of high reward? Similarities have been demonstrated at cue (Gruber et al., 2014). What does this signal look like at cue, during anticipation, and at outcome?

Aim 2.1: Characterize the mesolimbic neural response associated with the expectancy generation, anticipation, and resolution of high- versus low-curiosity information.

Open question 2.2: We know that memory is enhanced for information encoded in a high curiosity state (Gruber et al., 2014). How might mesolimbic circuitry be involved in enhancing memory formation for high-curiosity information?

Aim 2.2: Describe how a high-curiosity expectancy state supports episodic memory formation.

Open question 3: It has been demonstrated that phasic VTA dopamine firing and release tracks expected reward value whereas sustained VTA dopamine firing and release scales with reward uncertainty (Fiorillo et al., 2003). The dissimilar timing of phasic and sustained dopamine activity suggests that there may be distinct temporal windows during reward anticipation in which expected reward value versus reward uncertainty may benefit encoding. Can we tease apart the downstream repercussions of these dopaminergic patterns on memory formation?

Aim 3: Dissociate the influences of expected reward value and reward uncertainty, putatively driving phasic and sustained VTA dopamine release respectively, on episodic memory formation.

2. Mesolimbic and cortical networks differentiate novelty versus surprise

This chapter examines the role of the mesolimbic dopamine system in signaling expectancy violations (surprise) versus novelty. It also examines how expectancy influences episodic memory formation for novel imagery. This work helps illuminate the role of expectation in the human dopamine system and in memory formation, independent of both novelty and valence.

2.1 Introduction

Unpredicted events capture attention, influence behavior, and enhance perception and memory (Bubic, Cramon, & Schubotz, 2010; Corbetta, Patel, & Shulman, 2008; Kishiyama, Yonelinas, & Knight, 2009; Pearce & Hall, 1980). While recent research has examined how motivation and memory systems in the brain encode unpredictable events (Baldassarre & Mirolli, 2013; Duzel et al., 2010; Henson & Gagnepain, 2010; Kafkas & Montaldi, 2014; Li et al., 2003; Lisman et al., 2011; Ranganath & Rainer, 2003; Shohamy & Adcock, 2010), much of the extant work conflates two possible aspects: novelty and surprise (Barto et al., 2013). Novelty and surprise frequently co-occur in natural settings. For instance, a fire-bellied toad hopping through a lecture hall likely offers both—the novelty of seeing a fire-bellied toad for the first time, and the surprise generated by encountering one during a lecture. At the zoo, a fire-bellied toad, though novel, could be expected. Conversely, in the context of a lecture, one's mother, though

familiar, could be surprising. While both are failures of prediction, novelty requires previously unexperienced sensory input, whereas surprise requires violated expectations.

The pervasive conflation of novelty and surprise in the literature demands resolution, with particular implications for how mesolimbic circuits influence novelty encoding (Duzel et al., 2010; Lisman & Grace, 2005; Shohamy & Adcock, 2010). According to one influential model (Lisman & Grace, 2005), hippocampal novelty signals travel a multi-synaptic pathway via the nucleus accumbens (NAcc) to disinhibit dopamine neurons in the ventral tegmental area (VTA); disinhibition renders these neurons more responsive to excitatory inputs, potentially increasing dopamine release in the hippocampus. In humans, previous studies have demonstrated intrinsic connectivity among these regions (Kahn & Shohamy, 2012; Murty et al., 2014). However, studies in humans have either confounded novelty and surprise (Axmacher et al., 2010) or implemented designs that precluded clean dissociations (Bunzeck & Duzel, 2006; Wittmann et al., 2007; Zaehle et al., 2013), for reasons discussed previously (Murty et al., 2013) and in detail in the discussion. In animals, early research demonstrated involvement of dopaminergic circuitry during non-novel, surprising events (Horvitz et al., 1997). Thus, although surprising novel events purportedly activate mesolimbic systems, extant evidence leaves unspecified whether the critical driver is novelty or

surprise. To understand mesolimbic contributions to memory formation, observed neural responses in the hippocampus, NAcc, and VTA require disambiguation.

The distinct properties of surprising and novel events also imply differences in distributed brain representations. Encoding novelty creates new sensory representations (Blumenfeld, Preminger, Sagi, & Tsodyks, 2006). Encoding surprise involves updating expectations, and thus retrieving memories that underlie the prediction (Berlyne, 1960). We hypothesized that novel events would activate sensory regions, while surprising events would activate systems for memory retrieval and probabilistic learning. Thus, we predicted common activation of *a priori* mesolimbic regions in the setting of dissociable whole-brain networks.

In the current whole-brain fMRI study, we sought to characterize neural processing of behaviorally relevant unpredicted events, crossing novelty (novel vs. familiar stimuli) with surprise (surprising vs. expected stimuli) in a factorial design permitting identification of common versus selective networks for novelty and surprise.

2.2 Methods

2.2.1 Subjects

Participants were 23 healthy, right-handed volunteers. All participants provided informed consent, as approved by the Duke University Institutional Review Board. Four participants were excluded – one due to scanning artifacts, two for not meeting neurological/psychiatric history inclusion criteria (vertigo; taking antidepressants), and

one for falling asleep in the scanner. The 19 remaining participants (15 female; age range 21-35, mean age 26) were included in analysis.

2.2.2 Stimuli

Stimuli were 315 images of everyday objects drawn from the set previously assembled by Stark and colleagues (Bakker, Kirwan, Miller, & Stark, 2008). Fifteen were “familiarized” prior to encoding and the remaining 300 were presented as unique novel objects seen only once during the task. The stimuli were considered novel during this task since participants were screened to ensure that they had no previous experience with the stimulus set.

2.2.3 Tasks

2.2.3.1 Familiarization

Participants completed a one-back working memory task to become familiarized with the stimuli, which were five sequences of three everyday objects. A three-object sequence was briefly displayed on the screen (500 ms per object, 200 ms inter-stimulus interval) followed by a fixation cross (1 s) and then another sequence. Participants pressed the keyboard space bar to indicate a match for entire consecutive sequences. Five sequences were pseudorandomly presented 12 times each, such that no sequences occurred three times consecutively. These sequences were thus “familiar” during the subsequent encoding phase. In a separate pilot group ($n = 12$), participants completed

this task and were then given the 15 individual object stimuli to put into the five correctly ordered sequences. All participants performed at ceiling, 100% correct.

2.2.3.2 Incidental encoding

During the encoding phase, a lexical cue was presented at trial onset (2 s). This cue indicated to participants which categorical decision they would be making on that trial: natural/ manmade, edible/inedible, smaller/larger than a shoebox, or hard/soft.

Three trains of sequences followed. Participants made button press responses to the final stimulus in each sequence.

The cue and the first two sequences were part of the expectancy generation period, and the third sequence was the outcome. Each object in the sequence was presented for 500 ms with a 200 ms blank screen between objects. A 1 s fixation cross was presented after the first sequence, and a fixation cross jitter of 2.5 – 6.5 s occurred following the second sequence to isolate the neural systems associated with the third sequence (**Figure 2**) (Adcock et al., 2006). After the third sequence a jittered fixation cross was presented (1s – 20s). Timing of the cue onset and trial type was optimized by OptSeq2 (Dale, 1999).

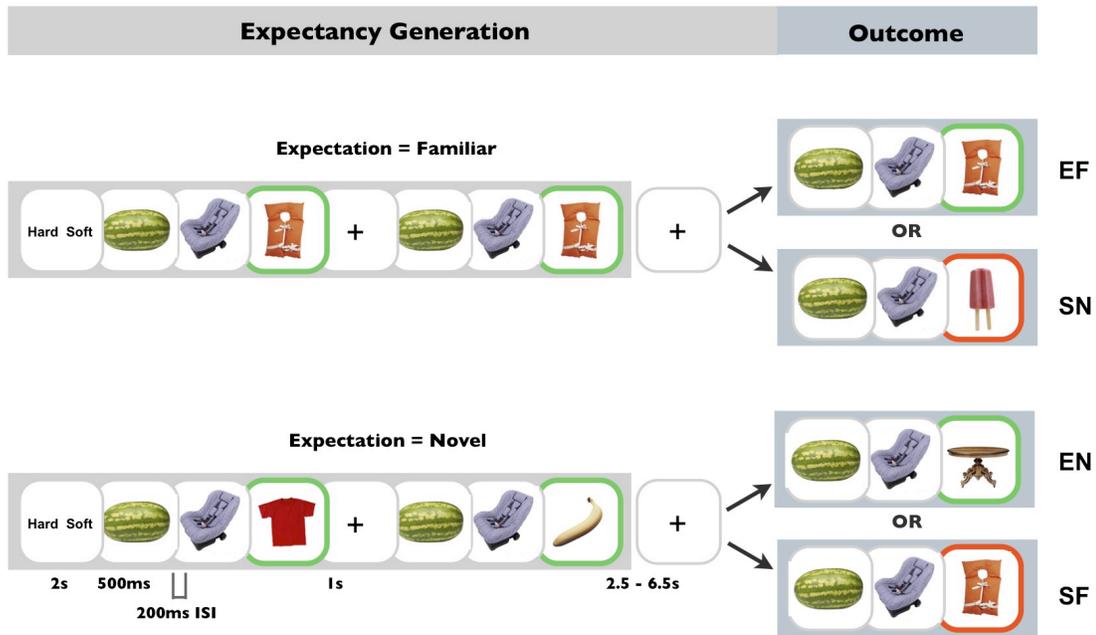


Figure 2: Experimental design of the incidental encoding task. A categorical decision cue indicated the response to be made on each trial. Color pictures of objects were subsequently presented in sequences of three objects, with a train of three pre-familiarized sequences per trial. Responses were made to the final object in each sequence. The first two sequences comprised the expectancy generation period whereas the last sequence was the outcome. Expectancy was generated for the novelty or familiarity of the final object in each sequence as follows: In half of all trials the final object of the first two sequences was part of the pre-familiarized sequence; this pattern predicted the same object in the final sequence eighty percent of the time. In the other half of trials, the final object of the first two sequences was a unique novel object; this pattern predicted a novel object in the final sequence eighty percent of the time. On twenty percent of trials, the final object violated expectations and was thus surprising. There were four possible conditions at outcome: expected familiar (EF), surprising novel (SN), expected novel (EN), and surprising familiar (SF).

Of the 200 trials, half (100 trials) had intact first and second familiar sequences.

Eighty percent of those were followed by an intact third sequence, such that the final object presented in the entire trial was both expected and familiar (EF condition).

Twenty percent instead had a novel object in the final position of the third sequence, so

that the object was both surprising and novel (SN condition). The remaining half (100 trials) had a novel object in the third position of the first and second sequences. Eighty percent of those were followed by a third sequence with a novel final object, such that it was expected but novel (EN condition). Twenty percent instead had the original familiar object in the final position of the third sequence, so that the object was surprising but familiar (SF condition, **Figure 2**). There were 10 runs of 20 trials each – eight EF, two SN, eight EN, and two SF per run.

2.2.3.3 Recognition memory test

In order to compare memory encoding for expected versus surprising novel events, participants viewed 40 completely new stimuli and 40 stimuli from encoding - 20 SN and 20 EN (we pseudorandomly selected two EN stimuli per run to match the SN condition). Each object was presented on the screen until the participant made a response, with a minimum time of 1 s per object, and was replaced by the next stimulus. Participants made a button press to indicate whether each item was old or new, and to rate their confidence as high, low, or guess. Memory in these conditions could not be compared to either the EF or SF condition, since familiar objects were repeated many times throughout the experiment (and thus memory was expected to be perfect for these objects).

2.2.4 Procedure

Participants completed the familiarization task on a computer outside of the scanner. During their anatomical scans, participants practiced the encoding task to learn the rules, using an independent set of stick-figure objects to avoid interference. During practice, participants received feedback about whether their responses were fast enough to be counted (1.5 s from cue onset). The encoding task was conducted over 10 functional scanning runs, with 20 counterbalanced trials per run. A surprise recognition memory test was completed outside of the scanner immediately following encoding.

2.2.5 Behavioral analysis

Reaction times for the final object of each trial were analyzed using a two-way repeated measures ANOVA with novelty (novel vs. familiar) and surprise (expected vs. surprising) as the within-subject factors. Upon reaching a significance level of $P < 0.05$ for the interaction, we followed with pairwise comparisons tests to clarify the pattern of reaction times. Categorization accuracy was measured by counting the most commonly submitted response for each object as correct. A linear regression was run to check for a correlation between accuracy and memory performance across subjects. Recognition memory for surprising versus expected novel stimuli (with surprise as a within-subjects factor) was tested by submitting the corrected percent correct (% hits - % false alarms) to a Student's t-test with a significance level of $P < 0.05$. Analyses included only high and low confidence responses; guesses were excluded.

2.2.6 MRI data preprocessing and acquisition

fMRI data were acquired on a 3.0-T GE Signa MRI scanner using a standard echo-planar imaging (EPI) sequence (TE = 27 ms, flip = 77 degrees, TR = 2 s, 34 contiguous slices, size = 3.75 mm x 3.75 mm x 3.8 mm) with coverage across the whole brain. Each of the 10 functional runs consisted of 198 volumes. Prior to the functional runs, we collected a whole-brain, inversion recovery, spoiled gradient high-resolution anatomical image (voxel size = 1mm isotropic) for use in spatial normalization. Data from a minimum of eight runs was analyzed per subject; five participants did not have enough time to complete all runs, and two participants had a run excluded due to excessive motion. fMRI preprocessing was performed using fMRI Expert Analysis Tool (FEAT) Version 5.98, as implemented in FSL 4.1.5 (www.fmrib.ox.ac.uk/fsl). The first six volumes were discarded to allow for signal saturation. BOLD images were skull stripped using the Brain Extraction Tool (S. M. Smith, 2002). Images were then realigned within-run, intensity normalized by a single multiplicative factor, spatially smoothed with a 4 mm full width half maximum (FWHM) kernel, and subjected to a high-pass filter (100 s). This 4mm smoothing kernel was chosen to optimize differentiation of midbrain, hippocampal, and ventral striatal activations (Adcock et al., 2006; Sacchet & Knutson, 2012). Spatial normalization was performed using a two-step procedure on fMRIB Linear Registration Tool. First, mean EPIs from each run were co-registered to the high-resolution anatomical image. Then the high-resolution anatomical image was

normalized to the high-resolution standard space image in Montreal Neurological Institute (MNI) space using a nonlinear transformation with a 10-mm warp resolution, as implemented by fMRI Nonlinear Registration Tool. All coordinates are reported in MNI space.

2.2.7 Definition of a priori ROIs

Circuit models have implicated the hippocampus, VTA, and NAcc in encoding novel events. We were interested in whether these regions respond to novel or surprising events and therefore anatomically defined each as *a priori* regions of interest (ROIs). The bilateral hippocampus ROI used for small volume correction in the voxelwise analyses was defined anatomically by the WFU PickAtlas (Maldjian, Laurienti, Kraft, & Burdette, 2003), NAcc was anatomically defined by the Harvard-Oxford Subcortical Structural Atlas (available with FSL). Although our primary region of interest in the midbrain was the VTA, we also examined the substantia nigra (SubN) since they are historically intertwined in the human neuroimaging literature (Duzel et al., 2009). The VTA and SubN regions were defined using a probabilistic anatomical atlas (Murty et al., 2014) and segmented by thresholding each region at 50% and assigning overlapping voxels to the statistically more likely region.

2.2.8 Functional MRI data analysis

Functional MRI data were analyzed using FEAT version 5.98 as implemented in FSL 4.1.5. Time-series statistical analyses used FILM with local auto-correlation correction (Woolrich, Ripley, Brady, & Smith, 2001).

2.2.8.1 General Linear Model: Task-Related Activations

To investigate task-related activations, first-level (i.e. within-run) general linear models (GLMs) included eight regressors of interest. These modeled expectancy generation (cue and first two stimulus sequences) and outcome (final stimulus sequence) events for each of the four conditions: EF, SF, EN, and SN. The events were modeled with a standard amplitude of one, and event durations of 7.1 s for the expectancy generation period and 1.9 s for outcome. They were then convolved with a double-gamma hemodynamic response function. Using this GLM, individual maps of parameter estimates were generated for nine contrasts of interest, all at outcome: Each condition against resting baseline (B), $EF > B$, $EN > B$, $SF > B$, and $SN > B$; conditions isolating novelty, $EN > EF$ and $SN > SF$; conditions isolating surprise, $SF > EF$ and $SN > EN$; and the condition describing surprising novelty, $SN > EF$. Second level analyses for each of these contrasts (across runs, but within-subject) were modeled using a fixed effects analysis.

2.2.8.2 Group-Level Analysis

Third-level analyses (across participants) were modeled using FSL's mixed effects analyses (FLAME 1), which accounts for within-session/subject variance calculated at the first and second levels, on the parameter estimates for contrasts of interest derived from the second-level analysis. To investigate the main effects, two-way repeated measures ANOVAs were run in FSL with novelty and surprise as within-subject factors.

Statistical tests for all fMRI analyses were set to an overall alpha = 0.05 family-wise error rate as calculated within the AlphaSim tool in AFNI (<http://afni.nimh.nih.gov/afni/doc/manual/AlphaSim>) which uses actual data structure to determine the number of independent statistical tests and thus balance Type 1 and Type 2 errors. With 1000 Monte Carlo simulations and a voxelwise cluster-forming significance of $P < 0.001$ (Bubic et al., 2010; Corbetta et al., 2008; Kishiyama et al., 2009; Pearce & Hall, 1980; Woo, Krishnan, & Wager, 2014), a smoothing kernel of 4.46 mm FWHM (the estimated intrinsic smoothness of the data, derived from the first level residual noise), an overall alpha of 0.05 corresponded to a cluster extent minimum of 21 voxels for the whole brain. Follow-up conjunction analyses were run to determine which voxels were significantly activated by both categories of novel events (EN > EF and SN > SF) and both categories of surprising events (SF > EF and SN > EN). In this way, we

could be sure the main effects were not driven only by one instance of novelty or surprise.

2.2.8.3 A Priori ROIs Analysis

Given the known heterogeneity across the long-axis of the hippocampus (Barto et al., 2013; Poppenk, Evensmoen, & Moscovitch, 2013), we took a voxel-wise approach to identifying responses distributed throughout this region. Like the whole-brain analyses, statistical tests were set to an overall $\alpha = 0.05$ family-wise error rate as calculated within the AlphaSim tool in AFNI. A voxelwise significance of $P < 0.01$ corresponded to a cluster extent minimum of 19 voxels in the hippocampus. This threshold was used in this anatomically limited region because there is often lower signal to noise and less signal change in the hippocampus than in cortical areas (Duzel et al., 2010; Greicius et al., 2003; Lisman & Grace, 2005; Shohamy & Adcock, 2010; Stark & Squire, 2001). (Of note, changing the primary threshold did not affect the overall multiple comparisons threshold.) Correcting for the volume of the bilateral hippocampus, we identified clusters activated by the two-way repeated measures ANOVA for the main effects of novelty and surprise. In addition, planned pairwise comparisons were run to separately look at the contrasts for novelty, EN > EF and SN > SF; surprise, SF > EF and SN > EN; and surprising novelty, SN > EF.

For VTA, NAcc, and SubN ROI analyses, the Featquery tool in FSL was used to extract mean activation across each anatomical region. Two-way repeated measures

ANOVAs were run for each region to determine the main effects of novelty and surprise. Upon finding potentially different patterns of activation in the VTA and NAcc, we decided to test whether these regions were demonstrating significantly different responses from one another by running a three-way repeated measures ANOVA with region, surprise, and novelty as within-subject factors.

2.2.9 Memory Correlation

Our ability to examine subsequent memory activations within subjects was limited because of the low number of surprise-condition trials. Therefore, we investigated the relationship between recognition memory performance and task-related activations across subjects. To determine which brain regions correlated with memory performance for each condition, we included EN and SN memory performance across subjects as regressors of interest in two separate GLM contrasts, EN > EF and SN > EF. As above, the primary threshold for the whole brain was $P < 0.001$, the overall threshold was $P < 0.05$ and the minimum cluster extent throughout the brain was 21 voxels.

2.3 Results

2.3.1 Behavior: Reaction Times and Accuracy

To confirm that participants generated expectancies, we ran a two-way ANOVA to examine the reaction times for the final objects within the critical sequences in each condition. There were significant main effects of novelty and surprise (novelty: $F(1,18) = 100.664$, $p < 0.0001$; surprise: $F(1,18) = 32.064$, $p < 0.0001$), which were driven by a

significant interaction such that subjects benefited from familiarity when it was expected but not when it was surprising (interaction of novelty by surprise: $F(1,18) = 54.641$, $p < 0.0001$); **Table 1**). Faster reaction times for the EF condition confirmed that participants successfully anticipated the exact object and response when possible, indicating our design of this condition was effective.

Table 1: Behavioral performance. Reaction times and corrected recognition memory.

	EF	SF	EN	SN
Reaction Times at Outcome	0.510***	0.683	0.729	0.705
(SEM)	(0.0282)	(0.0294)	(0.0286)	(0.0253)
Corrected Recognition Memory	---	---	0.690	0.668
(SEM)			(0.0333)	(0.0358)

*** $p < 0.0001$

As intended, overall categorization accuracy was high (mean = 86.27, SEM = 1.902) and had no relationship with recognition memory performance across participants ($F(1,17) = 0.0153$, $p = 0.903$, $r = 0.0300$).

2.3.2 fMRI: A Priori ROI Analyses

2.3.2.1 Hippocampus: Main Effects

The two-way repeated measures ANOVA revealed significant bilateral clusters of activation in the hippocampus for the main effect of surprise, but not novelty (MNI coordinates: -30, -24, -10 and 24, -26, -10). There was no novelty by surprise interaction.

2.3.2.2 Hippocampus: Paired Contrasts

We completed planned comparisons to examine the possible anatomical heterogeneity of the hippocampus response to novelty, surprise, and surprising novelty. These contrasts revealed significant bilateral clusters of activation in the hippocampus for all unpredicted events relative to expected familiar ones: novelty, EN > EF; surprise, SF > EF; and surprising novelty, SN > EF (**Figure 3**). However, surprising novel events did not increase activation relative to those that were either surprising or novel (SN > SF or SN > EN), suggesting that the hippocampal response to unpredicted stimuli may be more discrete than graded. In sum, the hippocampus was engaged by both surprising and novel events relative to expected familiar events.

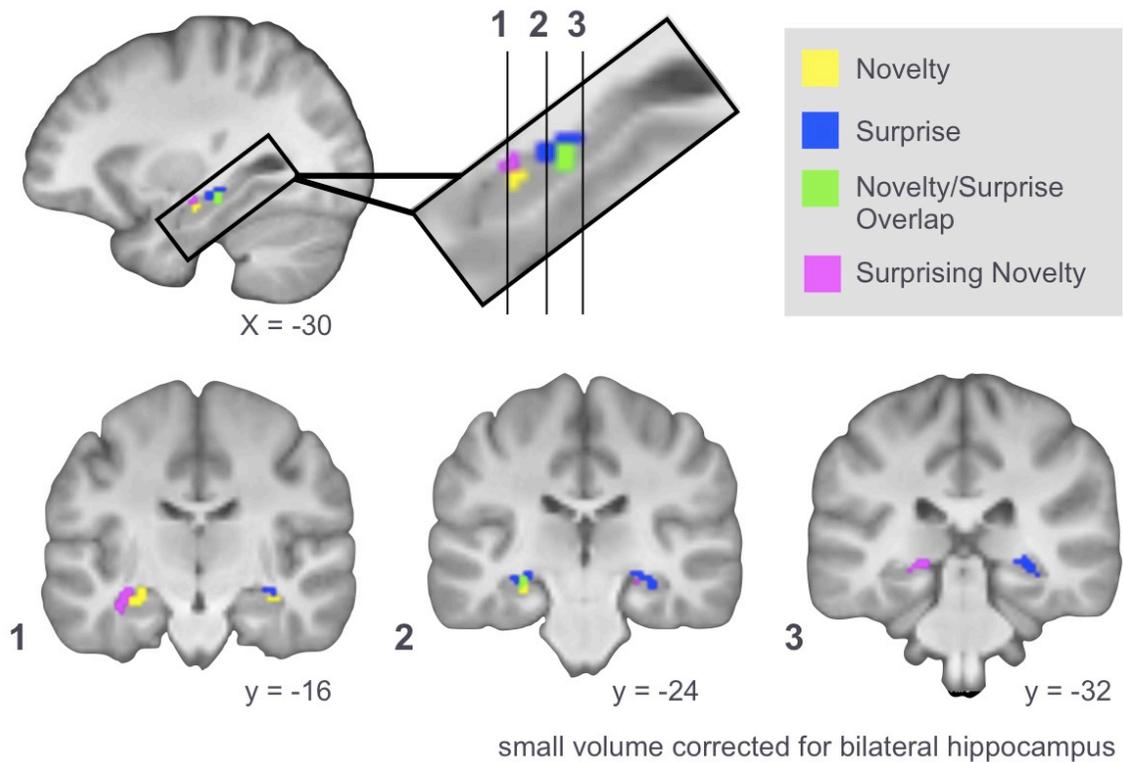


Figure 3: Novelty and surprise are distributed across the hippocampus. Clusters of activation were identified in the hippocampus for all unpredicted events relative to those that were both expected and familiar (novelty, $EN > EF$; surprise, $SF > EF$; and surprising novelty, $SN > EF$). All activations were small volume cluster corrected for the anatomical bilateral hippocampus at a primary threshold of $p < 0.01$, such that 19 contiguous voxels resulted in an overall threshold of $p < 0.05$.

2.3.2.3 VTA and NAcc: Main Effects

In our anatomically defined *a priori* ROIs the VTA and NAcc, mean activation levels were extracted for all conditions and two-way repeated measures ANOVAs were run within each region. The contrast of novel to familiar stimuli revealed a significant main effect of novelty only in the VTA ($F(1,18) = 8.557, p = 0.009$), with no difference in NAcc ($F(1,18) = 0.839, p = 0.372$; **Figure 4**). However, both ROIs revealed a significant

main effect of surprise (contrast of surprising to expected stimuli: VTA ($F(1,18) = 8.747$, $p = 0.008$), NAcc ($F(1,18) = 11.814$, $p = 0.003$; **Figure 4**). There was no novelty by surprise interaction in either region: VTA ($F(1,18) = 0.038$, $p = 0.848$), NAcc ($F(1,18) = 0.050$, $p = 0.825$).

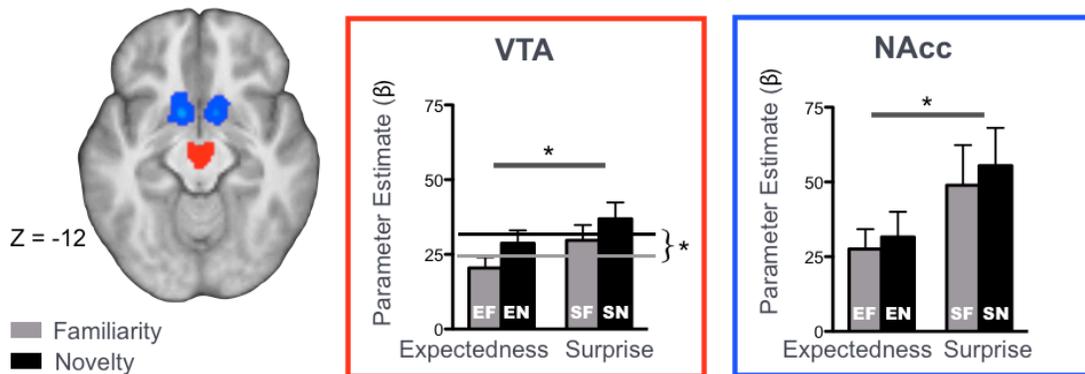


Figure 4: VTA signals both surprise and novelty while NAcc selectively signals surprise. Mean activations were extracted across our anatomical *a priori* ROIs. A two-way repeated measures ANOVA with novelty and surprise as within-subjects factors revealed main effects of novelty and surprise in the VTA, but only a main effect of surprise in NAcc. A three-way ANOVA with novelty, surprise, and region as within-subject factors revealed an interaction of region by surprise, confirming a larger surprise effect in NAcc than VTA (* $p < 0.05$).

We were intrigued by potentially different patterns of activation in the VTA and NAcc since these regions that are strongly linked anatomically and functionally. To test for differences in activation patterns between regions, we ran a three-way ANOVA with region, novelty, and surprise as within-subject factors. We found a significant main effect of surprise ($F(1,18) = 14.365$, $p = 0.001$), and the surprise effect was stronger in the NAcc than the VTA, as evidenced by a significant region by surprise interaction ($F(1,18) = 5.416$, $p = 0.032$). There was no main effect of novelty across regions ($F(1,18) = 2.667$, $p =$

0.120) and no region by novelty interaction ($F(1,18) = 0.334, p = 0.571$). In sum, the VTA responded to both surprise and novelty, while the NAcc responded to surprise alone—significantly more so than did the VTA.

2.3.2.4 SubN: Main Effects

The SubN is often combined with the VTA into a single ROI in the human neuroimaging literature; therefore, we additionally queried its modulation by novelty and surprise. We found a main effect of surprise ($F(1,18) = 6.246, p = 0.022$), but a non-significant main effect of novelty ($F(1,18) = 2.839, p = 0.109$). There was no novelty by surprise interaction ($F(1,18) = 0.029, p = 0.866$). In sum, the SubN was only significantly engaged by surprise, but demonstrated an overall pattern of activation that was similar to that of the VTA.

2.3.3 fMRI: Whole Brain Analyses

2.3.3.1 Main Effects

To identify brain regions that were modulated by novelty, we ran a two-way repeated-measures ANOVA and identified the main effects of novelty. We found clusters showing greater activation for novel than familiar outcomes in orbital frontal cortex, inferior frontal, precentral, and paracingulate gyri, lateral occipital cortex, and ventral visual stream (**Figure 5A**).

To identify brain regions that were modulated by surprise, we ran a two-way repeated-measures ANOVA and identified the main effects of surprise. We found

greater activation for surprising than expected outcomes in orbital frontal cortex, inferior frontal, middle frontal, precentral and paracingulate gyri, frontal pole, anterior insula, medial and lateral parietal cortex, striatum, thalamus, cerebellum, lateral occipital cortex and occipital pole (**Figure 5A**).

In the inverse contrasts, there were no significant main effects of expectedness. Familiar objects engaged the precuneus and cuneal cortex, posterior cingulate cortex, lateral occipital cortex, and frontal pole. The interaction of novelty by surprise revealed a region in lateral occipital cortex.

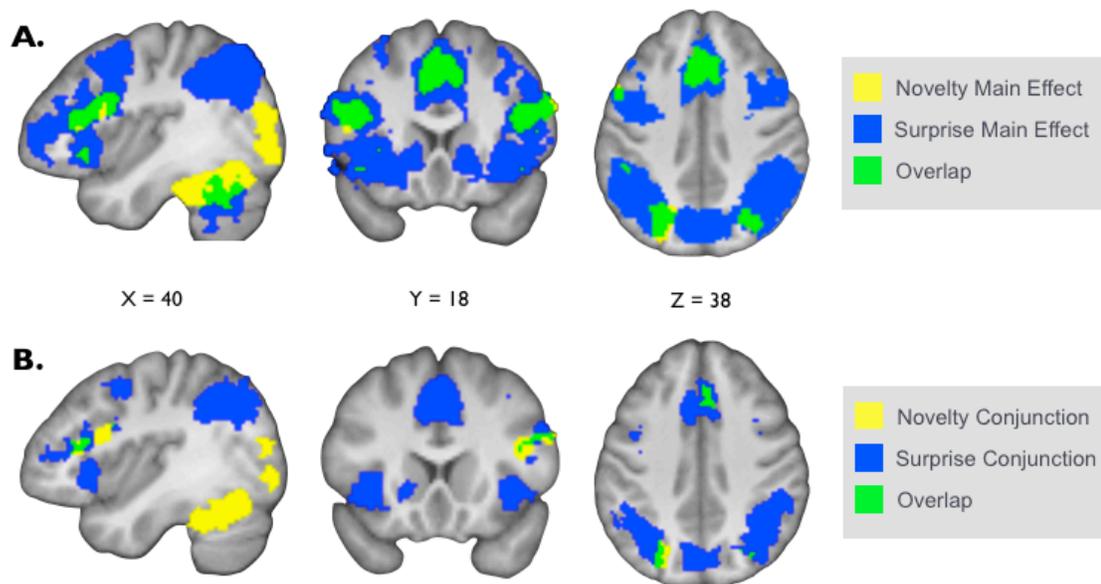


Figure 5: Brain regions engaged by novelty (yellow), surprise (blue), and their overlap (green). A, main effects maps from a whole brain two-way repeated measures ANOVA with surprise and novelty as the within-subject factors. B, conjunction maps for novelty (combining contrasts EN > EF and SN > SF) and surprise (combining contrasts SF > EF and SN > EN). All activations were cluster corrected at a primary threshold of $p < 0.001$, such that 21 contiguous voxels resulted in an overall threshold of $p < 0.05$.

2.3.3.2 Conjunctions

In order to find the network of regions that was significantly responsive to all instances of novelty, we performed a conjunction analysis that revealed overlapping significant voxels for EN > EF and SN > SF. To find the network associated with surprise, we performed a conjunction analysis of the contrasts SF > EF and SN > EN.

The novelty conjunction located activation in inferior frontal and precentral gyri, lateral occipital cortex, and the ventral visual stream (**Figure 5B**). The surprise conjunction found activation in orbital frontal cortex, anterior insula, inferior frontal, middle frontal, precentral and paracingulate gyri, medial and lateral parietal cortex, lateral occipital cortex, and occipital pole (**Figure 5B**). The only region selective for novelty was the ventral visual stream. In contrast, anterior insula, some areas of medial and lateral PFC, medial and lateral posterior parietal cortex, and visual cortex were selective for surprise.

2.3.4 Behavior and fMRI: Recognition Memory

If novelty and surprise rely on distinct networks, they may use that circuitry to differentially support memory. To test this, we examined recognition memory for novel objects from encoding (conducted after exiting the scanner) and found that hits were significantly greater than false alarms (mean +/- SEM: expected novel = 87.16 +/- 2.56, $t(18) = 30.29$, $p < 0.0001$), surprising novel = 84.99 +/- 2.81 ($t(18) = 34.02$, $p < 0.0001$), false

alarms = 18.17 +/- 2.88; **Table 1**). However, the expectedness of novel items did not influence encoding performance. Corrected recognition rates (hits – false alarms, excluding guessing responses) were not significantly different from one another ($t(18) = 1.031$, $p = 0.3162$), even when all responses were included ($t(18) = 0.6467$, $p = 0.5260$).

Linear regression identified regions correlated with memory for either expected or surprising novelty. Across participants, memory for SN events was not correlated with any brain activation. However, memory for EN events was correlated with activation in left precentral gyrus ($F(1,17) = 11.36$, $p = 0.0036$, $r = 0.633$).

2.4 Discussion

2.4.1 Summary

In this study, we investigated neural systems engaged by behaviorally relevant unpredicted events using a 2x2 factorial design crossing novelty with surprise. Notably, we found dissociations among regions in our *a priori* mesolimbic network: The hippocampus and VTA signaled both types of unpredicted events, while the NAcc signaled surprise but not novelty. Beyond mesolimbic regions, the distributed networks engaged by both categories of events included partially overlapping prefrontal and occipital regions. However, surprise drove a broader network, including additional striatal, prefrontal, and parietal regions, whereas novelty selectively activated the ventral visual stream. Further dissociating novelty and surprise, we found that memory for expected but not surprising novelty correlated with left precentral gyrus activation.

2.4.2 Notable Findings: A Priori Regions

In addition to differential contributions of dopaminergic midbrain, hippocampus, and NAcc to encoding novelty and surprise, we identified novel findings within each region. In the midbrain, the VTA responded both to novelty and surprise in isolation; this pattern comprised two notable findings. First, we disambiguated the role of the VTA in novelty processing, showing that, even when it was expected, novelty engaged the VTA. Second, we demonstrated non-valenced, non-novel surprise signals in the human VTA. Responses to surprise had previously been shown in the human VTA only for surprisingly valenced events, namely, surprising rewards, punishments, or reward omissions (D'Ardenne et al., 2008; D'Ardenne, Lohrenz, & Bartley, 2013; Lisman & Grace, 2005; Menon et al., 2007; Murray et al., 2008). Importantly, the current work is the first human neuroimaging study to demonstrate VTA responses to neutral surprise, providing support for this region's role in signaling surprising events that motivate behavior or learning, independent of reward value.

As predicted, we found that the hippocampus responded to both novelty and surprise. Planned comparisons revealed activation of the hippocampus by both novelty and surprise relative to events that were expected and familiar, consistent with previous literature (Axmacher et al., 2010; Chen, Olsen, Preston, Glover, & Wagner, 2011; Dickerson, Li, & Delgado, 2011; Köhler, Danckert, Gati, & Menon, 2005; Kumaran & Maguire, 2006; Stern, Sherman, Kirchoff, & Hasselmo, 2001). Novelty and surprise

responses were differentially distributed over hippocampal anatomy, suggesting that dissociable cortical networks differentially engage hippocampal machinery for encoding novel versus surprising events. However, the surprise signal was statistically predominant, as evidenced a significant main effect of surprise but not of novelty.

Unexpectedly, we did not detect NAcc activation for novelty unless the event was also surprising. Furthermore, although there were no differences in novelty responses, the NAcc was more strongly engaged by surprise than was the VTA. Previous work has implicated the NAcc in novelty encoding (Bunzeck & Duzel, 2006; Ihalaenen, Riekkinen, & Feenstra, 1999; Legault & Wise, 2001; Wittmann et al., 2007; Zaehle et al., 2013), although interestingly, the response in those studies diminished over the duration of novelty exposure. This pattern is consistent with findings that unexpected novelty but not prolonged exposure drives activity in the NAcc shell (Rebec et al., 1997b). We propose that in previous work, the critical driver of the NAcc responses was the surprising onset of novelty, reconciling prior reports with our finding that NAcc predominantly signaled surprise.

2.4.3 Dopaminergic Circuits in Encoding Novelty and Surprise: Integrating with Prior Human Imaging

Interestingly, previous fMRI studies of novelty and surprise in humans (Bunzeck & Duzel, 2006; Horvitz et al., 1997; Wittmann et al., 2007; Zaehle et al., 2013) have reported neither VTA nor NAcc activation by surprise. Several design choices

differentiate these prior studies from the current study and may explain the conflicting results.

First, we fully dissociated novelty and surprise in our current paradigm.

Previous studies confounded these two classes of stimuli in their paradigms by only including novel events that were also surprising (Blumenfeld et al., 2006; Bunzeck & Duzel, 2006; Zaehle et al., 2013). Second, we equated behavioral relevance across all of our trial types. Previous studies required no response for the novel or surprising events (Berlyne, 1960; Bunzeck & Duzel, 2006; Wittmann et al., 2007; Zaehle et al., 2013).

Behavioral relevance of events has been shown to gate responses in the striatum (Bakker et al., 2008; Guitart-Masip et al., 2011; Tricomi, Delgado, & Fiez, 2004; Zink, Pagnoni, Martin-Skurski, Chappelow, & Berns, 2004). Third, in the current study, predictions on each trial were necessarily dependent on learning from previous trials. Surprise had the ability to prompt learning: that is, to update predictions about behaviorally relevant events. In previous studies, explicit instructions about cue validity made outcomes uninformative for future trials (Adcock et al., 2006; Wittmann et al., 2007) so that surprise signals could not contribute to adaptive learning. Consistent with theories that an event's relationship to attentional reorienting and learning goals is critical in driving a dopamine signal (Dale, 1999; Redgrave & Gurney, 2006; Steinfels et al., 1983), we propose that VTA and NAcc responses to all events are well accounted for by the need

to update predictions following a behaviorally relevant surprise, whether or not it is novel.

2.4.4 Theoretical implications

A mechanistic loop from hippocampus to VTA via NAcc, implicated by previous work in both awake behaving animals and rodent slice physiology, has been proposed to control the entry of novel events into long-term memory (Lisman & Grace, 2005; Smith, 2002). Here, novelty activated the hippocampus and the VTA, but did not significantly activate the NAcc. Thus, novelty did not engage the entire circuit predicted by the hippocampus-NAcc-VTA model. Surprise, however, engaged all three regions, including NAcc. Our findings suggest that the pathway connecting hippocampus through the NAcc may in fact be more relevant for surprising than for novel events.

Where else might VTA novelty signals come from? Lateral septum has been shown to transmit context-reward associations between the hippocampus and VTA (Adcock et al., 2006b; Luo, Tahsili-Fahadan, Wise, Lupica, & Aston-Jones, 2011; Sacchet & Knutson, 2012); this pathway could theoretically relay signals about behaviorally relevant task contexts. Additionally, the hippocampus projects to prefrontal cortex (Barbas & Blatt, 1995; Goldman-Rakic, Selemon, & Schwartz, 1984; Maldjian et al., 2003), which has direct (Frankle et al., 2006; Murty et al., 2014), indirect (Carr & Sesack, 2000) and functional (Au-Young, Shen, & Yang, 1999; Woolrich et al., 2001) excitatory projections to the VTA. The prefrontal cortex has been demonstrated to drive VTA

activation in response to task demands relevant to the current study (Ballard et al., 2011). Finally, a route for processing basic sensory surprise in anesthetized animals transmits signal directly from superior colliculus to the VTA (Dommett, 2005). Future work utilizing network connectivity approaches will be required to investigate how novelty and surprise signals reach the VTA, and how they are then disseminated to downstream targets.

In summary, co-activation of hippocampus and VTA supports proposed frameworks in which behaviorally relevant unpredicted events are detected by the hippocampus and transmitted to the VTA. The stronger signal in the hippocampus together with stronger NAcc responses is consistent with disinhibition of the VTA via this route for surprise, but not novelty – although novelty is nonetheless communicated to the VTA from some source. Previous work implicating the NAcc as a critical component of this loop may chiefly reflect responses to surprising novelty, since these aspects of unpredicted events often co-occur, especially at onset. Taken together, our data suggest that future models should consider surprise and behavioral relevance as determinants of anatomical substrates for novelty encoding.

2.4.5 Novelty and Surprise in Distributed Brain Networks

Beyond the *a priori* mesolimbic circuit, our results clarify common and selective cortical networks for novelty versus surprise. We found prefrontal and occipital networks common to both, consistent with prior work broadly implicating these regions

in processing unpredicted events (Corbetta et al., 2008; Halgren, Marinkovic, & Chauvel, 1998; Kiehl, Laurens, & Duty, 2001; Rossion, Schiltz, & Crommelinck, 2003).

The ventral visual stream, a region associated with repetition suppression as visual novelty decreases (Grill-Spector, Henson, & Martin, 2006), was the only region selectively activated by novelty but not surprise. Prior authors have debated whether repetition suppression represents release from novelty or rareness (Egner, Monti, & Summerfield, 2010; Larsson & Smith, 2012; Summerfield, Trittschuh, Monti, Mesulam, & Egner, 2008). While we cannot speak directly to that debate, our data at least suggest that it is novelty, not surprise, that elicits increased activation in the ventral visual stream.

In contrast, surprise alone engaged a much more widespread network, including parietal regions associated with memory retrieval (Cabeza, Ciaramelli, & Olson, 2008; Wagner, Shannon, Kahn, & Buckner, 2005), medial prefrontal regions associated with updating behavior following surprise (O'Reilly et al., 2013) and striatal regions associated with probabilistic learning (Shohamy et al., 2008). These patterns are consistent with hypothesized adaptations following surprise: activating memories of recent prior events, updating representations to reflect actual outcomes, and evaluating appropriate behaviors.

2.4.6 Limitations

While our study provides novel insights into engagement of mesolimbic dopamine networks during surprise and novelty, fMRI cannot measure neurotransmitter release. BOLD activation in the VTA/SubN has been correlated with striatal dopamine release (Schott et al., 2008). However, additional studies are needed to relate fMRI observations to dopaminergic activity; novelty and surprise signals could be conveyed by different populations, including non-dopaminergic neurons.

2.4.7 Conclusions

Both novelty and surprise represent failures of prediction. Here, teasing the two apart permitted us to expand known functions of the human VTA and NAcc to include signaling surprises irrespective of novelty and with no associated valence. The selectivity of NAcc activation identifies surprise as a potential driver of novelty encoding via the hippocampus-NAcc-VTA circuit, and highlights questions about how novelty signals arise in the VTA. Amid similar recruitment of attention and memory systems, experiences of novelty and surprise call for different kinds of adaptation and have appropriately distinct neural representations.

3. The Neural Characterization of Curiosity at Cue, Anticipation, and Outcome

This chapter aims to characterize the mesolimbic neural response associated with the expectancy generation, anticipation, and resolution of high- versus low-curiosity information. It also seeks to describe how a high-curiosity expectancy state supports declarative memory formation.

3.1 Introduction

In his seminal characterization of epistemic curiosity in 1960, Berlyne noted that a question serves to induce a “quest for knowledge,” acting as both a cue for associative knowledge and motivation to resolve uncertainty (Berlyne, 1960). Over a half century later, following considerable discussion about the conceptual nature of curiosity, surprisingly little is known about the neural characterization of curiosity.

Curiosity has motivational salience; students who are more curious remember more information in school (Maw & Maw, 1966; Mittman & Terrell, 1964). Likewise, curiosity-inducing questions bring online neural networks allied with motivation, including the dopaminergic midbrain and nucleus accumbens (Gruber et al., 2014), neural regions associated with representing subjective value, and these regions together with the hippocampus support memory for high-curiosity information (Gruber et al., 2014). However, according to many accounts, curiosity is associated with distinct stages. These include the following: when individuals become aware of an information gap (we

designate this the cue period), experience uncertainty, conflict, deprivation, or arousal caused by the information gap (designated anticipation), and attain the desired information (designated outcome) (Berlyne, 1960; Litman, 2005; Loewenstein, 1994). It remains to be seen how the motivational signal in the brain is carried throughout these cue, anticipation, and outcome phases of curiosity and how these signals relate to memory outcomes.

Additionally, curiosity may interact with related motivational signals during these epochs, in particular when goal attainment is contingent on behavioral action (Guitart-Masip et al., 2011; Guitart-Masip et al., 2012; Howe et al., 2013). This relationship has interesting implications for classroom learning, where curiosity and action are tightly intertwined in pursuit of learning. Thus, in addition to examining curiosity over time, we were also interested in curiosity's relationship with action at each epoch.

Although there is a dearth of studies examining curiosity in neural systems (only (Gruber et al., 2014; Kang et al., 2009), curiosity is thought to share properties with reward (Gruber et al., 2014), which has been studied in the mesolimbic dopamine pathway at cue (Adcock et al., 2006; Schultz et al., 1997; Wittmann et al., 2005), anticipation (Fiorillo et al., 2003; Howe et al., 2013; Totah et al., 2013), and outcome (Schultz et al., 1997). Thus, studies of reward allowed us to make predictions about

neural activation; divergence would instead suggest differences between reward and curiosity.

Throughout each epoch, we predicted that curiosity would bring online the same motivational networks associated with reward: dopaminergic midbrain, nucleus accumbens, hippocampus, and prefrontal cortex. We also predicted that neural regions would be more sensitive to differences in curiosity without competing motivational signals (when the outcome was not instrumental, or contingent on an action). Finally, we thought that greater curiosity would increase the sensitivity of subcortical and cortical memory networks at each stage of processing.

In this study, we strove to characterize the neural activation associated with curiosity at cue, anticipation, and outcome. We examined how curiosity is influenced by the presence or absence of an action contingency. And we examined how curiosity influenced memory through neural activation. To look at curiosity in this way, we designed an fMRI study in which we queried memory for trivia answers. Critically, these answers followed questions for which we crossed curiosity state (high/low) with action contingency (present/absent). Analyses separated event-related cue activation, event-related outcome activation, and background anticipatory activation. By looking at curiosity and its interactions with action and memory at each stage of processing, we were able to create the most complete neural characterization of curiosity to date.

3.2 Methods

3.2.1 Subjects

Participants were 25 healthy, right-handed volunteers. All participants provided informed consent, as approved by the Duke University Institutional Review Board. Two participants were excluded – one participant was excluded for falling asleep in the scanner and one discontinued the scan due to a headache. The 23 remaining participants (10 female; age range 19-35, mean age 26.39) were included in analysis.

3.2.2 Stimuli

Stimuli consisted of 360 trivia questions and answers, a subset of the 375 trivia questions and answers used by Gruber and colleagues (Gruber et al., 2014).

3.2.3 Tasks

Participants completed four phases – screening, incidental encoding, surprise recall, and final ratings. The first three phases are included in analysis and discussed in this study.

3.2.3.1 Screening

Participant-specific curiosity ratings were used to sort trivia questions in high- and low-curiosity categories. Trivia questions were organized into three counterbalanced lists, so that exposure to trivia questions differed across participants. After the presentation of a trivia question, participants gave two self-paced ratings, marking their responses to the following questions on a continuous scale from “Least”

to “Most”: “How likely is it that you know the answer?” and “How curious are you about the answer?” (Figure 6A) Questions in which participants were likely to know the answer (> 90% on the scale) were excluded from the rest of the study. Participants responded until 216 trivia questions were eligible for inclusion. These questions were then binned into thirds according to curiosity ratings – the 72 lowest rated became the low curiosity questions during encoding; the 72 highest rated became the high curiosity questions during encoding. Twelve of the 72 middle-rated questions were included for catch trials during encoding but were not included in analysis.

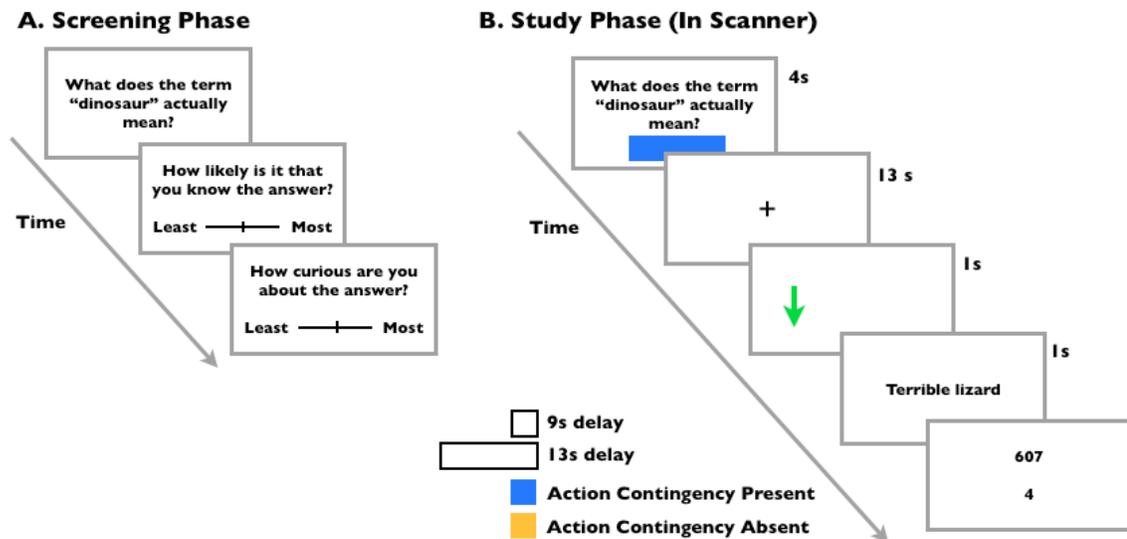


Figure 6: Experimental design screening and incidental encoding. Participants completed a screening phase (A) from which trivia questions were sorted into high- and low-curiosity designations. In the MRI scanner, participants completed a study phase (B) in which they saw a trivia question from screening, experienced an anticipatory delay, and saw the trivia answer. Receipt of the trivia answer was contingent on a button press for half of all trials.

3.2.3.2 Incidental Encoding

During the incidental encoding phase in the MRI scanner, the 156 selected trivia questions were presented along with the associated answers (**Figure 6B**). A trivia question and colored rectangle were presented at trial onset (4 s). The trivia question served as a cue to anticipate the upcoming trivia answer. The colored rectangle was a symbolic cue indicating the duration and action contingency on that trial. The rectangle could be one of two lengths, short or long. On short trials, the anticipation period, during which time a fixation cross appeared on the screen, would be 9 seconds in duration. On long trials, the anticipation period would be 13 seconds in duration. The rectangle could also be one of two colors. An orange rectangle indicated that there was no action contingency on that trial; the trivia answer (1 s) always appeared at the end of the anticipation period. A blue rectangle indicated that participants would have to successfully make a speeded button press (action contingency present) at the end of the anticipation period to see the trivia answer. On these trials, a large green arrow randomly appeared on either the left or the right side of the screen (1s). Participants had 1 second to make a correct left or right button press on the button box. If they were incorrect or too slow, they saw the string 'XXXXX' instead of the trivia answer (1 s). If they responded correctly and fast enough, they saw the trivia answer (1 s). After seeing the trivia answer, participants performed a jittered active baseline task in which they counted backward from unique starting numbers (1s – 20s). On 12 catch trials,

participants were presented with a counting probe (1 s) which asked whether they counted under or over a given number to encourage completion of the counting task. Trivia questions shown after the counting probe were drawn from the middle curiosity questions and were not included in analysis. There were 6 runs (10 min each), with 12 high curiosity trials, 12 low curiosity trials, and 2 catch trials in each. Within each curiosity condition, half of the trials had an action contingency and half did not. Among those, half had 9s anticipation periods and half had 13s anticipation periods. Timing of the cue onset and trial type was optimized by OptSeq2 (Dale, 1999).

3.2.3.3 Surprise Recall

Outside of the scanner, participants were given a list in Microsoft Excel with the 144 high and low curiosity trivia questions in a random order. Participants were encouraged to take approximately 25 min to type out the correct answers without guessing any answers.

3.2.3.4 Final Ratings

After the surprise recall phase, participants were shown both the trivia question and the answer and asked to rate how surprised they were by the answer (when they saw it for the first time in the scanner). They responded to the question "How surprised were you by the answer?" on a 4-point scale from "Least" to "Most." They also responded to the question "How confident are you about your response?" with either "Sure," "Pretty Sure," or "Just Guessing."

3.2.4 Procedure

Participants first completed demographics questionnaires and the Curiosity and Exploration Inventory II personality questionnaire, measuring self-reported trait tendencies toward curiosity and exploration in everyday life (Kashdan et al., 2009). Participants completed the screening task on a computer outside of the scanner. They then did a practice of the incidental encoding task outside of the scanner, using middle curiosity questions from screening that were not included in the encoding task. In this way participants were familiar with the timing and motor contingency cues prior to entering the MRI scanner. During their anatomical scans, participants again practiced the encoding task to become familiarized with the button box. The encoding task was conducted over 6 functional scanning runs, with 26 counterbalanced trials per run. The surprise recall phase and the final ratings phase were completed outside of the scanner immediately following encoding.

3.2.5 Analysis

3.2.5.1 Behavioral Analysis

Memory was analyzed as a function of curiosity, anticipatory delay, and motor contingency condition using a 2x2x2 repeated measures ANOVA.

3.2.5.2 MRI Data Preprocessing and Acquisition

FMRI data were acquired on a 3.0-T GE Signa MRI scanner using a standard echo-planar imaging (EPI) sequence (TE = 27 ms, flip = 77 degrees, TR = 2 s, 34

contiguous slices, size = 3.75 mm x 3.75 mm x 3.8 mm) with coverage across the whole brain. Each of the 6 functional runs consisted of 298 volumes. Prior to the functional runs, we collected a whole-brain, inversion recovery, spoiled gradient high-resolution anatomical image (voxel size = 1mm isotropic) for use in spatial normalization. Cardiac and respiratory physiological data was collected during functional scans using BioPac hardware. Functional MRI preprocessing was performed using fMRI Expert Analysis Tool (FEAT) Version 6.00, as implemented in FSL 5.0.8 (www.fmrib.ox.ac.uk/fsl). The first six volumes were discarded to allow for signal saturation. BOLD images were skull stripped using the Brain Extraction Tool (S. M. Smith, 2002). Physiological noise was removed using the Physiological Noise Modelling toolbox in FSL (Brooks et al., 2008). Images were then realigned within-run, intensity normalized by a single multiplicative factor, spatially smoothed with a 4 mm full width half maximum (FWHM) kernel, and subjected to a high-pass filter (80 s). This 4mm smoothing kernel was chosen to optimize differentiation of midbrain, hippocampal, and ventral striatal activations (Adcock et al., 2006; Sacchet & Knutson, 2012). Spatial normalization was performed using a two-step procedure on fMRIB Linear Registration Tool. First, mean EPIs from each run were co-registered to the high-resolution anatomical image. Then the high-resolution anatomical image was normalized to the high-resolution standard space image in Montreal Neurological Institute (MNI) space using a nonlinear transformation with a 10-mm warp

resolution, as implemented by fMRI Nonlinear Registration Tool. All coordinates are reported in MNI space.

3.2.5.3 Definition of A Priori ROIs

Circuit models have implicated the hippocampus, VTA, and NAcc in the anticipation and encoding of reward and curiosity. We therefore defined each as *a priori* regions of interest (ROIs). The bilateral hippocampus ROI used for small volume correction in the voxelwise analyses was defined anatomically by the WFU PickAtlas (Maldjian et al., 2003). NAcc was anatomically defined by the Harvard-Oxford Subcortical Structural Atlas (available with FSL). The VTA was defined using a probabilistic anatomical atlas (Murty et al., 2014), thresholded at 50% and excluding voxels statistically more likely to belong to the substantia nigra.

3.2.5.4 fMRI Data Analysis

fMRI data were analyzed using FEAT version 6.00 as implemented in FSL 5.0.8. Time-series statistical analyses used FILM with local auto-correlation correction (Woolrich et al., 2001).

3.2.5.4.1 General Linear Model: Task-Related Activations

To investigate task-related activations, first-level (i.e. within-run) general linear models (GLMs) included 16 regressors of interest. These modeled cue and outcome events for each 2x2x2 condition: high/low curiosity, action contingency present/absent, and remembered/forgotten. The events were modeled with a standard amplitude of one, and event durations of 4 s for the cue period and 1 s for contingency absent outcomes or

2 s for contingency present outcomes. They were then convolved with a double-gamma hemodynamic response function. Using this GLM, individual maps of parameter estimates were generated for 7 contrasts of interest, at both cue and outcome: high > low curiosity, present > absent action contingency, remembered > forgotten, high curiosity only: remembered > forgotten, low curiosity only: remembered > forgotten, action contingency present: high > low curiosity, action contingency absent: high > low curiosity. Second level analyses for each of these contrasts (across runs, but within-subject) were modeled using a fixed effects analysis.

3.2.5.4.2 *Group-Level Analysis*

Third-level analyses (across participants) were modeled using FSL's mixed effects analyses (FLAME 1), which accounts for within-session/subject variance calculated at the first and second levels, on the parameter estimates for contrasts of interest derived from the second-level analysis.

To test for differences in high curiosity: remembered > forgotten versus low curiosity: remembered > forgotten, as well as action contingency present: high > low curiosity versus action contingency absent: high > low curiosity, third level analyses were run contrasting each pair of contrasts. Twenty-three regressors of no interest removed subject specific variance in each case.

Statistical tests for all fMRI analyses were set to an overall alpha = 0.05 family-wise error rate as calculated within the AlphaSim tool in AFNI (<http://afni.nimh.nih.gov/afni/doc/manual/AlphaSim>) which uses actual data structure

to determine the number of independent statistical tests and thus balance Type 1 and Type 2 errors. With 1000 Monte Carlo simulations and a voxelwise cluster-forming significance of $P < 0.01$, a smoothing kernel of 5.00 mm FWHM (the estimated intrinsic smoothness of the data, derived from the first level residual noise), an overall alpha of 0.05 corresponded to a cluster extent minimum of 90 voxels for the whole brain.

3.2.5.4.3 *A Priori ROIs Analysis*

We took a voxel-wise approach to identifying responses distributed throughout the *a priori* regions of interest. Like the whole-brain analyses, statistical tests were set to an overall alpha = 0.05 family-wise error rate as calculated within the AlphaSim tool in AFNI. A voxelwise significance of $P < 0.01$ corresponded to a cluster extent minimum of 21 voxels in the hippocampus, 18 voxels in the VTA, and 19 in NAcc.

3.2.5.4.4 *Anticipatory Activations*

Following preprocessing, another analysis pipeline was run to examine whole-brain activation during anticipation, independent of cue- and outcome-related effects. In order to isolate the effects of anticipation independent of catch-, cue- or outcome-related activity, a first level GLM was created including all cue, outcome, and catch events. The residuals from this analysis were 4D data with event-related activity regressed out. GLMs group-level analyses, and ROI analyses were then conducted as they were for task-related activations.

3.2.5.5 Correlations between ROI activation and behavioral personality scores

Participants were scored on their responses to the Curiosity and Exploration Inventory II (CEI II) questionnaire, with a higher score indicating the participant was more motivated by curiosity in everyday life. We wondered whether curiosity-related activation at cue and outcome would be modulated by individual differences in CEI II score. Therefore, we included CEI II score as an across subjects regressor in the group level map for high > low curiosity (at cue, anticipation, and outcome) to see which brain regions' curiosity activation scaled with individual personality differences in trait curiosity.

3.3 Results

3.3.1 Behavior: Memory

We examined memory recall performance for the trivia answers that were incidentally encoded. Participants demonstrated significant memory, with an overall hit rate of $51.26\% \pm 3.18\%$. Because we were interested in how curiosity and action motivated memory encoding, and to ensure that the length of the anticipatory delay did not influence encoding, we ran a 2x2x2 ANOVA with curiosity, motor contingency, and delay as within-subject effects. We found an extremely significant main effect of curiosity, with participants remembering the answer to high curiosity trivia questions more than low curiosity trivia questions ($F(1,22) = 86.08, p < 0.0000001$). There was also a main effect of action contingency, such that participants recalled the trivia answer better

when they made a button press than when they did not ($F(1,22) = 15.16, p = 0.0008$). As anticipated, there was no main effect of anticipatory delay length ($F(1,22) = 0.81, p = 0.38$). There was a significant interaction was between curiosity and motor contingency ($F(1,22) = 5.16, p = 0.033$). Curiosity and motor contingency benefits interfered with one another, such that the boost for high > low curiosity was less when participants made a button press than when they did not.

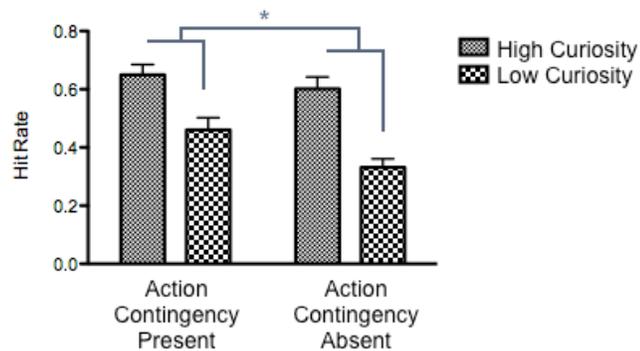


Figure 7: Curiosity and action contingency impact memory. Not only were there main effects of curiosity and action contingency on memory performance, there was also a significant curiosity by action contingency interaction such that curiosity differences were greater when participants were not required to make a motor action. (* $p < 0.05$)

3.3.2 Behavior: Reaction Times and Accuracy

As we hoped, participants were largely successful on the motor task, with a mean reaction time of 0.56 seconds \pm .02 seconds, and an accuracy of 98.58% \pm 0.42%. Thus, there were very few trials on which participants did not see the trivia answer.

3.3.3 fMRI: Cue-Related Activation

3.3.3.1 Main Effects of Curiosity

We were interested in how curiosity activates the mesolimbic pathway as well as the rest of the brain. For the contrast of high curiosity greater than low curiosity, we found significant activation in the hippocampus at whole brain correction and VTA at small volume correction, but not in the NAcc (Figure 8A).

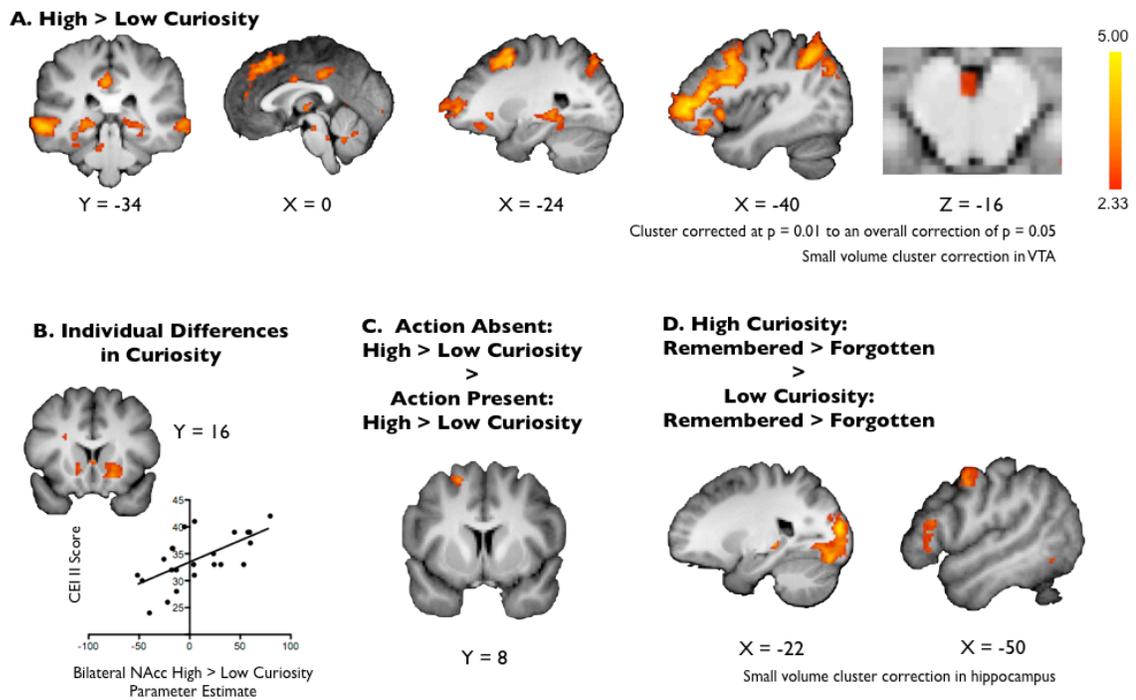


Figure 8: Cue-related activation. Main effects of curiosity (A), individual differences in curiosity (B), action contingency by curiosity interaction (C), and memory by curiosity interaction (D).

Additional activation throughout the brain included the following regions: medial and lateral prefrontal cortex, orbital frontal cortex, anterior insula, dorsal

striatum, parahippocampal cortex, cerebellum, posterior cingulate cortex, posterior parietal lobe, and primary visual cortex (**Figure 8A**).

3.3.3.2 Individual Differences in Curiosity

We wondered whether activation for high > low curiosity would differ with individual differences in personality measures of curiosity. Of our regions of interest, bilateral clusters in the NAcc showed a linear relationship between CEI II score and activation for high > low curiosity (NAcc: $R^2 = 0.41$, $p = 0.001$; **Figure 8B**).

Additional regions throughout the brain scaling with personality measures of curiosity were the inferior frontal gyrus, caudate, anterior cingulate cortex and posterior cingulate cortex.

3.3.3.3 Main Effects of Motor Contingency

For the contrast of contingency present > contingency absent, there were no activations in any of our regions of interest. Instead, we found effects in thalamus, posterior pons, cerebellum, medial and lateral prefrontal cortex, anterior insula, posterior parietal lobe, and primary and lateral occipital lobe.

There was no significant activation for the contrast of contingency absent > contingency present.

3.3.3.4 Curiosity by Motor Contingency Interaction

We examined whether curiosity effects in the brain were different across action contingencies. A region in dorsal prefrontal cortex was more responsive to high > low

curiosity for the contingency absent condition than the contingency present condition (**Figure 8C**). There were no other differences. Thus, activation in prefrontal cortex at cue mirrored the behavioral interaction in our memory findings, with greater curiosity differences in the absence of an action contingency.

3.3.3.5 Main Effects of Subsequent Memory

For the contrast of subsequently remembered > forgotten, there were no cue-related activations in any of our regions of interest. Instead, we found a difference in anterior insula, prefrontal cortex, and medial occipital lobe.

3.3.3.6 Memory by Curiosity Interaction

We thought that greater curiosity might enhance differences in activation for subsequently remembered > forgotten. Thus, we compared the memory response across curiosity conditions. We found that the hippocampus, prefrontal cortex, and ventral visual stream were more responsive to remembered > forgotten for high curiosity than for low curiosity (**Figure 8D**). Medial prefrontal cortex and posterior insula were more activated for remembered > forgotten in the low curiosity state than during the high curiosity state.

3.3.4 fMRI: Anticipatory Activation

3.3.4.1 Main Effects of Curiosity

We were interested in how the curiosity signal is held during anticipation, independent of cue-related activations. For the contrast of high curiosity greater than

low curiosity, we found significant activation in the hippocampus at small volume correction, but none in the VTA or NAcc (**Figure 9A**).

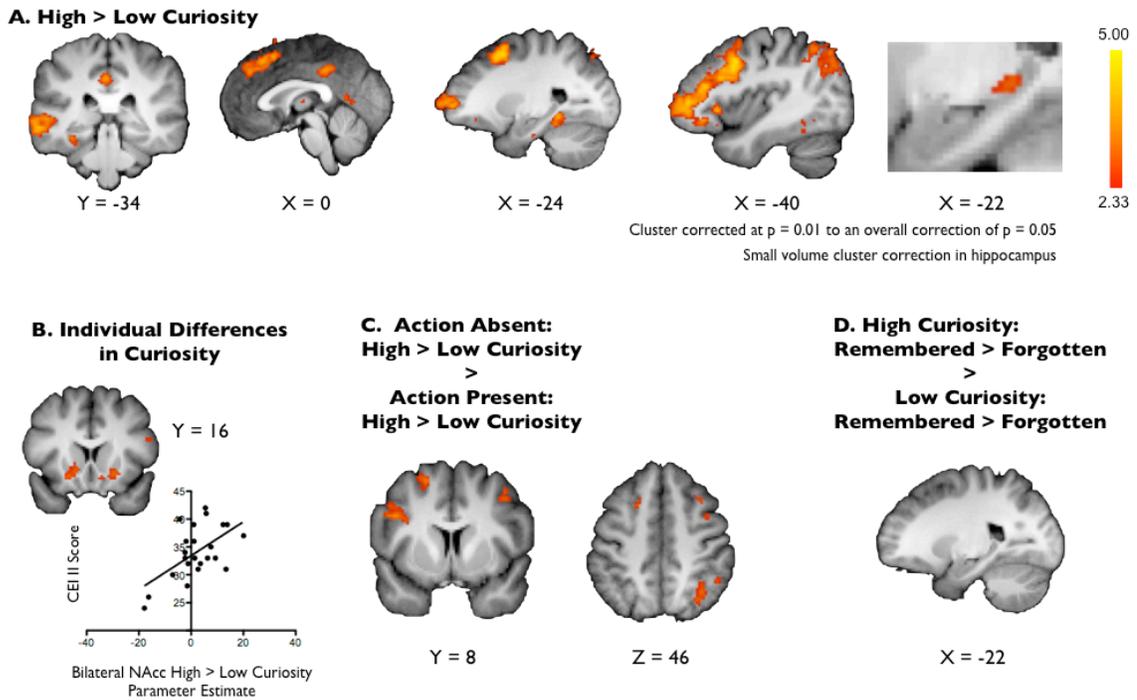


Figure 9: Anticipation-related activation. Main effects of curiosity (A), individual differences in curiosity (B), action contingency by curiosity interaction (C), and memory by curiosity interaction (D).

Additional activation throughout the brain included the following regions: medial and lateral prefrontal cortex, orbital frontal cortex, anterior insula, dorsal striatum, parahippocampal cortex, cerebellum, posterior cingulate cortex, posterior parietal lobe, and primary visual cortex (**Figure 9A**).

3.3.4.2 Individual Differences in Curiosity

We wondered whether activation for high > low curiosity would differ with individual differences in personality measures of curiosity during anticipation. Of our

regions of interest, bilateral clusters in the NAcc showed a linear relationship between CEI II score and activation for high > low curiosity (NAcc: $R^2 = 0.33$, $p = 0.004$; **Figure 9B**). Additional regions throughout the brain scaling with personality measures of curiosity were the inferior frontal gyrus, caudate, anterior cingulate cortex and posterior cingulate cortex.

3.3.4.3 Main Effects of Motor Contingency

For the contrast of contingency present > contingency absent, there were no activations in any of our regions on interest. Instead, we found effects in dorsomedial PFC, precentral gyrus, inferior parietal lobe, precuneous, and cerebellum.

There was no significant activation for the contrast of contingency absent > contingency present.

3.3.4.4 Curiosity by Motor Contingency Interaction

We examined whether curiosity effects in the brain were different across action contingencies at anticipation. Inferior and middle frontal gyrus in prefrontal cortex and inferior parietal lobe were more responsive to high > low curiosity for the contingency absent condition than the contingency present condition (**Figure 8C**). Thus, activation in prefrontal and parietal cortex at cue mirrored the behavioral interaction in our memory findings, with greater curiosity differences in the absence of an action contingency.

3.3.4.5 Main Effects of Subsequent Memory

For the contrast of subsequently remembered > forgotten, there was greater activation in left inferior/middle frontal gyrus. This region overlapped with that seen for high > low curiosity.

In the inverse direction, subsequently forgotten > remembered, there was widespread activation including hippocampus and NAcc in our ROIs. In the whole brain, we found activation in superior frontal gyrus, ventromedial frontal gyrus, middle insula, parahippocampal cortex, dorsal striatum, temporoparietal junction, anterior and posterior cingulate gyrus, cuneus, precuneus, and cerebellum.

3.3.4.6 Memory by Curiosity Interaction

There was no memory by curiosity interaction during anticipation (**Figure 9D**).

3.3.5 fMRI: Outcome-Related Activation

3.3.5.1 Main Effects of Curiosity

We were interested in how curiosity activates the mesolimbic pathway as well as the rest of the brain. For the contrast of high curiosity greater than low curiosity, we found significant activation in the VTA at small volume correction, but not in the NAcc or hippocampus (**Figure 10A**). At whole brain correction, we found activation in the following regions: medial and ventrolateral PFC, anterior insula, and posterior parietal lobe (**Figure 10A**).

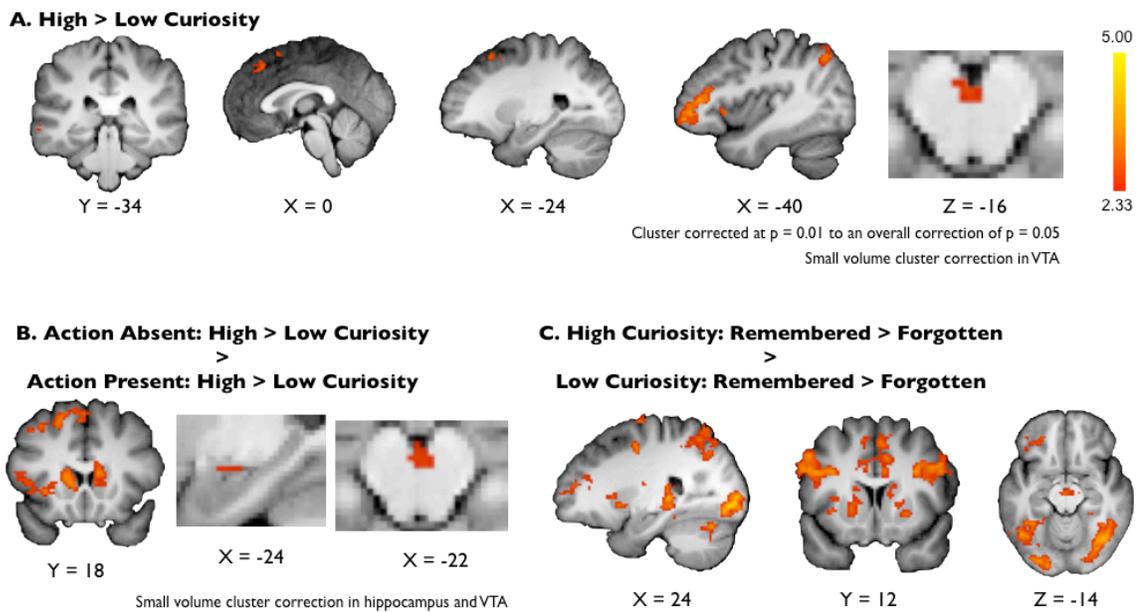


Figure 10: Outcome-related activation. Main effects of curiosity (A), action contingency by curiosity interaction (B), and memory by curiosity interaction (C).

3.3.5.2 Individual Differences in Curiosity

No regions demonstrated a relationship between individual differences in personality measures of curiosity and activation for high > low curiosity.

3.3.5.3 Main Effects of Action Contingency

For the contrast of action contingency present > action contingency absent, there were no activations in any of our regions of interest. Instead, we found effects in VLPFC, dorsal striatum, posterior insula, left motor and parietal cortex, and ventral visual stream.

There was stronger activation for the contrast of action contingency absent > contingency present in right hippocampus, lateral prefrontal cortex, lateral temporal cortex, and lateral occipital cortex.

3.3.5.4 Curiosity by Action Contingency Interaction

As at cue, we found a significantly greater curiosity effect in the action contingency absent than in the action contingency present conditions. These effects at outcome were, however, fairly widespread (**Figure 10B**). Small volume cluster correction revealed a greater curiosity effect for action contingency absent than action contingency present in the VTA and the left hippocampus. Throughout the whole brain, this interaction was also seen in medial and lateral prefrontal cortex, anterior insula, and dorsal striatum.

3.3.5.5 Main Effects of Subsequent Memory

For the contrast of subsequently remembered > forgotten, the only ROI that showed activation was the left hippocampus. We also found a difference in orbitofrontal cortex, DMPFC, anterior insula, entorhinal cortex, parahippocampal cortex, lateral temporal lobe, thalamus, and posterior cingulate cortex.

3.3.5.6 Memory by Curiosity Interaction

We thought that greater curiosity might enhance differences in activation for subsequently remembered > forgotten. Thus, we compared the memory response across curiosity conditions. We found that the VTA, hippocampus, prefrontal cortex, lateral occipital lobe, dorsal striatum, anterior insula, and ventral visual stream were more responsive to remembered > forgotten for high curiosity than for low curiosity (**Figure**

10C). There were no greater activations for remembered > forgotten in the low curiosity condition than during high curiosity.

3.4 Discussion

3.4.1 Summary

Cue-induced epistemic curiosity activated mesolimbic reward circuitry, even in the absence of an action contingency. In fact, activation in PFC at cue was slightly increased for high curiosity specifically in the contingency-absent condition, suggesting that if anything, curiosity resonated more strongly without movement to anticipate. Hippocampus and ventral visual stream activation at cue for high curiosity was predictive of subsequent memory success for the trivia answer at outcome.

During anticipation, the curiosity signal was held in many of the same regions as at cue such as PFC, but was absent in the VTA and weaker in the hippocampus. As at cue, the curiosity effect was even stronger in PFC when there was no action contingency. Individual differences in curiosity were still present in NAcc. There were no curiosity by memory interactions.

At outcome, the curiosity effect was more muted than at cue, but still activated VTA and prefrontal cortex. However, high curiosity additionally brought online hippocampus and dorsal caudate specifically if there was no action contingency. And while we showed that the typical remembered > forgotten networks were present at outcome, we also found significantly more memory-related activation for high curiosity

outcomes, including enhanced activation in the VTA, hippocampus, striatum, prefrontal cortex, and visual cortex.

3.4.2 Novel findings

The competition between action and valence in the dopaminergic system has received considerable attention in recent years (Guitart-Masip et al., 2011; Guitart-Masip et al., 2012; Tricomi et al., 2004), with some results suggesting that action is a critical feature of motivation. In this study, we demonstrated that both curiosity and action improve memory recall. However, the mnemonic benefit of curiosity is stronger in the absence of action, a competing motivational signal. Mechanistically, high curiosity states drive mesolimbic circuitry even in the absence of movement, providing support for the idea that reward anticipation and receipt, even without a corresponding action, is sufficient to activate dopaminergic regions. In fact, dopaminergic regions important for learning are more sensitive to information in a high curiosity state when the motor contingency is absent.

This finding, that action only improves memory outcomes when curiosity is lacking, has important implications for learning in the classroom. It has already been shown that monetary incentives only benefit learning for low-interest information (Murayama & Kuhbandner, 2011). Integrating our study with years of research showing that curiosity improves memory (Berlyne, 1954; Gruber et al., 2014; Kang et al., 2009; Maw & Maw, 1966), research demonstrating a delay-of-feedback benefit (Carpenter &

Vul, 2011; Mullaney, Carpenter, Grotenhuis, & Burianek, 2014), and results showing an interaction between interest and external rewards ((Murayama, Matsumoto, Izuma, Matsumoto, & Smith, 2010); for meta-analysis and review, see (Cameron, 2001; Deci, Koestner, & Ryan, 1999), we can further refine our model for helping students learn. Interest, or curiosity, is key, and although waiting for feedback can boost learning when students are interested, making learning action contingent does not. However, when interest or curiosity is low, both external rewards and action activate motivational systems that overcome the lack of intrinsic motivation in order to help students learn.

Not only can we update our behavioral model linking curiosity and action, our results expand on the known neural relationship between curiosity and memory. We found that the hippocampus, prefrontal cortex, and ventral visual streams are more predictive of memory in high curiosity states than low ones. At outcome, the dopaminergic system robustly supported memory when participants were highly curious. Thus, like reward (Adcock et al., 2006; Wittmann et al., 2005), curiosity can turn up the gain for memory.

Our study dissociated activation due to cue, anticipation, and outcome. While these phases have been dissociated in the rodent and non-human primate literatures (Cohen et al., 2012; Fiorillo et al., 2003), it is unique in the fMRI literature to separately analyze event-related cue activation and background anticipatory activation. Thus, we were able to show that the curiosity signal is not held throughout anticipation in the

mesolimbic system, but is present in prefrontal cortex. Whereas a sustained dopamine signal may be present when approaching a reward (Fiorillo et al., 2003; Howe et al., 2013; Totah et al., 2013), passive anticipation of a motivationally salient outcome may instead be held in working memory.

3.4.3 Caveats

Although we were interested in how curiosity and action interacted in support of memory, we had too few trials to perform an effective 3-way ANOVA. Thus, our neuroimaging results only address the curiosity-action and curiosity-memory relationships separately. In addition, although all of our discussion is about the importance of curiosity for learning, McGillivray and colleagues demonstrated that post-answer interest is more predictive of memory than curiosity (McGillivray, Murayama, & Castel, 2015). However, in that study, curiosity alone did correlate with memory. Because we were interested in the cue-related and anticipatory effects of curiosity, it was more relevant to characterize the motivational state of curiosity present prior to outcome.

3.4.4 Conclusion

Curiosity is a powerful motivator, one that is extremely important for learning and memory. Motivational circuitry in the dopaminergic system both codes and supports enhanced memory for this high curiosity state, although the signal is distinct at cue, anticipation and outcome. Furthermore, in the absence of additional motivational

signals such as action contingency, curiosity is even more effective. Thus, curiosity, supported by dopaminergic circuitry in the brain, is an important tool in the classroom and in everyday life for determining the contents of our memory.

4. Expected reward value and uncertainty have dissociable effects on memory formation

The goal of this chapter was to dissociate the influences of expected reward value and reward uncertainty, putatively driving phasic and sustained VTA dopamine release respectively, on episodic memory formation. This will provide insight into distinct mechanisms of memory modulation during different states of reward expectancy.

4.1 Introduction

Memory formation is enhanced during reward anticipation (Shohamy & Adcock, 2010). Just as the desire to get an 'A' can motivate students to remember information, research has likewise shown that the promise of money can motivate individuals to form new memories (Adcock et al., 2006; Gruber & Otten, 2010; Wittmann et al., 2005). Even incidental events are more strongly encoded during anticipation of upcoming rewards (Murty & Adcock, 2013). However, the mechanisms of memory enhancement during reward anticipation remain undetermined.

One possible mechanism is that during reward anticipation dopamine release augments long-term memory formation. Anatomically, VTA afferents project to the hippocampus (Gasbarri 1994; Gasbarri 1997). Applying dopamine receptor antagonists in the hippocampus blocks memory formation for new, rewarding events (Bethus et al. 2010). Extant evidence has also shown that activation of the dopaminergic midbrain (Adcock et al., 2006; Wittmann et al., 2005) and increased midbrain connectivity with the

hippocampus (Adcock et al., 2006) during reward anticipation benefits memory.

However, this mechanism must be elaborated to incorporate knowledge about multiple temporal profiles of dopamine neuronal response. A rapid phasic burst response scales with the expected value of a reward or a cue predicting reward, whereas a slower, sustained response has been reported associated with reward uncertainty (Fiorillo et al., 2003).

We proposed that within several seconds of reward anticipation, phasic and ramping dopamine neuronal excitation differentially modulate memory formation. Specifically, we hypothesized that when expected reward value is high, dopamine release is greater early and any effects on memory would occur for items presented temporally close to cues. On the other hand, when reward uncertainty is high, dopamine release is greater later during reward anticipation, so any effects on memory would occur closer to reward outcome.

Beyond the timing of action within reward anticipation, phasic versus sustained dopamine may influence memory via discrete mechanisms. Action may occur either synaptically or extrasynaptically, depending on the transient or sustained nature of release (Floresco et al., 2003). Importantly, effects may occur immediately or may be apparent only after a delay. Thus, we also sought to tease apart whether phasic versus sustained dopaminergic influences on memory would be present immediately or after a 24-hour period that would allow for consolidation.

We set out to dissociate the putative influence of these two distinct dopaminergic responses on memory formation during reward anticipation. To parse these effects, we designed a study in which we used overlearned abstract cues to indicate reward probability, establishing expected value independently from uncertainty. We further manipulated the epoch of encoding during reward anticipation: we presented items either early, to capture a rapid dopamine response associated with high expected reward value, or late, to capture a sustained dopamine response associated with high reward uncertainty. Finally, we manipulated retrieval time, either 15 minutes- or 24 hours-post encoding, to examine dependence on consolidation. With this paradigm, we were able to establish whether and in what way expected value and reward uncertainty each influenced memory formation.

4.2 Methods

4.2.1 Subjects

Participants were 40 healthy young adult volunteers. All participants provided informed consent, as approved by Duke University Institutional Review Board. Data from additional subjects were excluded due to failure to follow the instructions (n = 1), poor cue-outcome learning (n = 2) or computer error (n = 3). Subjects were assigned to participate in one of two experiments run in succession: Experiment 1 (n = 20, 12 female, mean age = 27.45 ± 3.82 years) or Experiment 2 (n = 20, 12 female, mean age = 21.90 ± 3.23 years).

4.2.2 Design and Procedure

4.2.2.1 Reward Learning

Participants were presented with four abstract cues, all Tibetan characters, which were 100%, 50% constant, 50% variable, or 0% predictive of subsequent monetary reward. They were instructed to try to learn the relationship between the cues and reward. They were presented with the cue (1s), a unique image of an everyday object (2s), then an image of either a dollar bill or a scrambled dollar bill (400ms), indicating a reward or no reward respectively. A jittered fixation-cross separated trials. No motor contingency was required to earn the reward. Participants saw 40 trials per condition, distributed evenly over 5 blocks. Prior to the first block and following each of the 5, participants were asked to rate their certainty of receiving reward following each cue along a sliding scale from “Certain: No Reward” to “Certain: Reward.” Independent of performance, participants were paid a monetary bonus equal to the amount earned in one block of the task. To be included in the analysis, participants during learning had to at a minimum consistently identify the 100% rewarded cues to be more certainly associated with reward than the 0% rewarded cues.

4.2.2.2 Incidental Encoding

The 100%, 50% constant, and 0% predictive cues from the reward-learning phase were present during incidental encoding (**Figure 11**). Participants saw a cue (400ms), followed by a unique novel object (1s) either immediately after the cue (400ms post-cue

onset and 3.2s pre-reward delivery) or just prior to reward delivery (3-3.6s post-cue onset and 0-0.6s pre-reward delivery). A fixation cross was displayed during the anticipatory delay when no image was present. The dollar bill or scrambled dollar bill (400ms), indicating a reward or no reward respectively, appeared 3.6 s after cue-onset for all trials. After reward feedback, participants were probed with the question, "Did you receive a reward?" (1s). Participants were instructed to quickly and accurately make a "yes" or "no" button press. The motor component could not be anticipated since the yes/no, right/left location was random from trial-to-trial. All together, there were 6 conditions in a 3 x 2 factorial design: probability of reward (100%, 50%, and 0%) at the early or late encoding epochs. There were 20 trials per condition, evenly dispersed among 5 blocks.

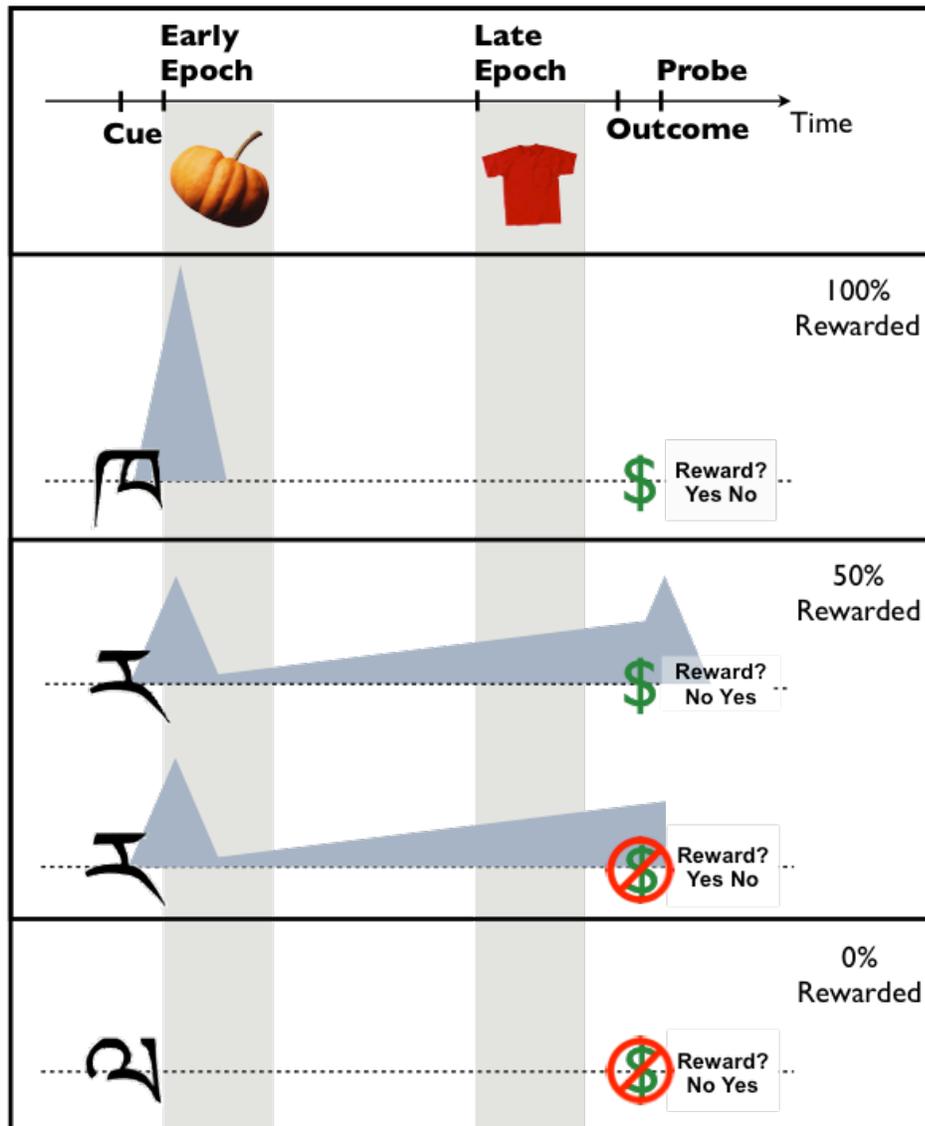


Figure 11: Figure 1. Experimental design: incidental encoding task. The task was designed to dissociate two physiological profiles of a putative dopamine response during reward anticipation - a phasic response that occurs rapidly and scales with expected value and a sustained response that increases with uncertainty (Fiorillo et al. 2003). Shaded triangles model these dopamine profiles relative to the cue, encoding, outcome, and probe events in each trial. Cues associated with 100%, 50%, or 0% reward probability were presented for 400ms. Incidental encoding objects were presented for 1s either immediately following the cue or shortly before anticipated reward outcome (400ms). Finally, a probe asking participants whether or not they received a reward, with the yes/no location counterbalanced, followed for 1s.

4.2.2.3 Recognition Memory Test

Participants saw 280 “new” objects and 280 “old” objects. Only the old objects from the incidental encoding phase, not from reward learning, were included in analyses to calculate memory performance. Participants performed a recognition memory test indicating whether they thought an item was old or new. They then indicated their confidence by saying “Definitely Sure,” “Pretty Sure,” or “Just Guessing.”

4.2.2.4 Experiment 1 – 24 hour

In Experiment 1, participants returned at the same time the next day to complete the recognition memory test, approximately 24 hours after encoding.

4.2.2.5 Experiment 2 – Immediate

In Experiment 2, participants completed the recognition memory test 15 minutes after the end of encoding.

4.2.3 Analysis

4.2.3.1 Within Experiments

A 3 x 2 repeated-measures ANOVA was used to examine the effects of reward probability and encoding epoch on subsequent memory performance. A significant interaction between reward probability and encoding epoch in each experiment prompted additional analyses. One-way repeated-measures ANOVAs at early encoding and late encoding were used to examine how reward probability related to memory

formation at each encoding epoch during anticipation. Significant one-way ANOVAs prompted follow-up analyses. Specifically, a test for a linear trend increasing with probability was used to examine how expected reward value related to memory, and paired Student's t-tests were used to compare memory on certain versus uncertain trials.

Participants indicated their confidence for each recognition response. One-sample t-tests revealed whether memory for guesses was above chance in each condition. Significant memory for guesses resulted in trials of all confidence being included in the analysis.

To determine whether memory for items presented following the 50% predictive cue was influenced by reward outcome, we completed two-tailed paired Student's t-tests to see if there were differences in memory for rewarded versus unrewarded trials. These were completed separately at the early and late encoding epochs since the effects could differ by time.

To examine whether general attentional differences across conditions resembled memory performance, we performed 1-way ANOVAs and follow-up pairwise Student's t-tests and tests for a linear trend to determine whether reaction time or accuracy for the probe varied by reward condition within each encoding epoch.

4.2.3.2 Across Experiments

Since both Experiment 1 and Experiment 2 revealed differences in the pattern of memory formation for early versus late encoding epochs, we decided to test whether the

patterns at each encoding epoch significantly differed according to retrieval time. We thus performed 3×2 ANOVAs with reward probability as a within-subjects factor and retrieval time as an across-subjects factor. We did this at both early encoding and late encoding. A significant interaction between reward probability and retrieval time prompted follow-up pairwise ANOVAs to see if the deltas between 15-minute and 24-hour retrieval were significantly different across reward probability conditions.

4.3 Results

4.3.1 Reward Learning

Participants in both groups learned the meaning of the cues during the reward-learning phase. In the 24-hour memory group, participants in the final block reported the 100% probable cue as 99.46% (± 0.20) likely to predict reward, the 50% cue as 52.65% (± 3.39) likely to predict reward and the 0% cue as 2.29% (± 1.79) likely to predict reward. In the immediate memory group, participants in the final block reported the 100% probable cue as 99.37% (± 0.43) likely to predict reward, the 50% cue as 55.26% (± 2.68) likely to predict reward and the 0% cue as 1.44% (± 1.07) likely to predict reward.

4.3.2 Experiment 1 – 24 hour retrieval

Because the aim of the study was to manipulate distinct temporal components of reward anticipation and relate those components to determinants of dopamine physiology, we completed a 3×2 ANOVA looking at memory performance as a function of reward probability and encoding epoch. We found a main effect of reward

probability ($F(2,18) = 5.56, p = 0.01$) and no main effect of encoding epoch ($F(1,19) = 2.57, p = 0.13$). Importantly, however, there was a strong interaction between encoding epoch and reward probability ($F(2,18) = 7.50, p = 0.004$), suggesting that further analyses were needed to interpret the effects. Indeed, one-way ANOVAs within early and late encoding epochs revealed a significant difference in memory late ($F(2,19) = 13.25, p < 0.0001$) and a trend for a difference early ($F(2,19) = 2.41, p = 0.10$).

Follow-up pairwise t-tests to examine memory performance during late encoding revealed greater memory following 50% cues compared to either 100% or 0% cues, with no difference between the latter two (100% vs. 50%: $t(19) = 4.34, p = 0.0004$; 50% vs. 0%: $t(19) = 4.20, p = 0.0005$; 100% vs. 0%: $t(19) = 0.31, p = 0.76$). Increased reward uncertainty benefitted memory late during reward anticipation (**Figure 12A**).

Follow-up tests revealed a significant linear trend such that memory scaled with increasing reward probability (Linear trend: $R^2 = 0.03, p = 0.04$). Thus, early during anticipation, memory performance linearly tracked expected reward value (**Figure 12A**).

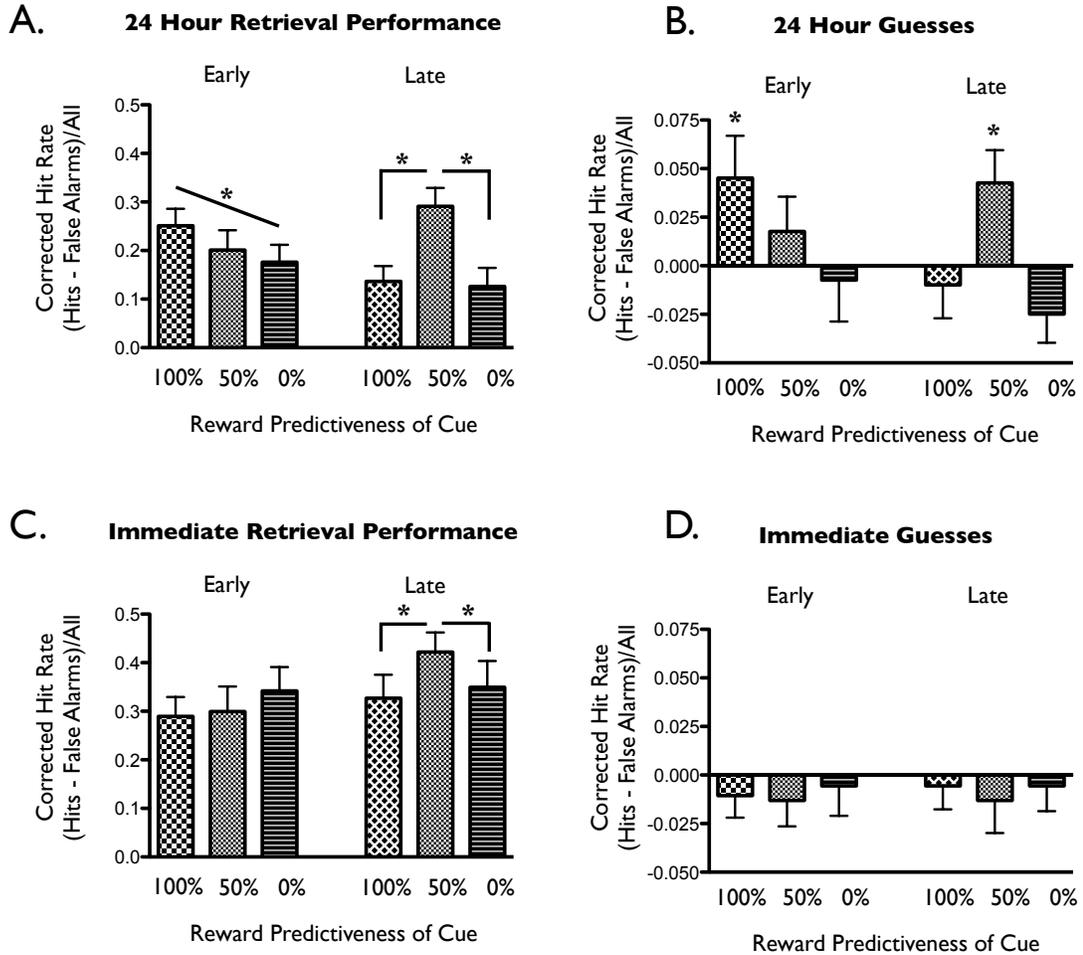


Figure 12: Figure 2. Memory performance for early and late encoding items at 24-hour and immediate retrieval. A. In the 24-hour retrieval group, early memory linearly increased with expected value; late memory was greatest for items encoded during reward uncertainty. This pattern was supported by significant memory for guesses only at the highest expected value early and during reward uncertainty late. B. In the immediate retrieval group, early memory did not differ by expected value; late memory was greatest for items encoded during reward uncertainty. There was no significant memory for guesses at this retrieval period. (* $p < 0.05$)

To determine whether there was significant memory in guesses, we examined whether memory was above chance across encoding epochs and reward probabilities.

We found significant memory in guesses at 100% early and at 50% late (100% early: $t(19)$

= 2.08, $p = 0.05$; 50% late: $t(19) = 2.54$, $p = 0.02$; all others $t(19) < 1.68$, $p > 0.11$). Thus, there was significant memory in guesses for items encoded under high uncertainty late, and high probability early, paralleling the overall memory findings (**Figure 12B**).

It was also possible that the uncertainty memory benefit we attributed to the anticipatory context could be explained instead by associations with reward outcomes. To investigate this alternative explanation, we performed t-tests between the rewarded and unrewarded uncertain trials, during both early and late epochs. We found no differences in either epoch (early, rewarded vs. unrewarded: $t(19) = 0.10$, $p = 0.92$; late, rewarded vs. unrewarded: $t(19) = 1.09$, $p = 0.29$).

4.3.3 Experiment 2 – Immediate retrieval

The 24-hour retrieval test did not allow us to distinguish between effects acting at encoding versus consolidation. Thus, in Experiment 2, participants completed an immediate retrieval test, 15 minutes after encoding. All analyses for Experiment 1 were repeated for Experiment 2. Analyses of immediate retrieval performance replicated effects of reward uncertainty on items presented late in the anticipation epoch, but differentially showed no effects of reward probability on items presented early in the epoch, as follows:

A 3 x 2 ANOVA revealed a trend for a main effect of reward probability ($F(18) = 3.33$, $p = 0.06$) and a main effect of encoding epoch, with memory greater at late than early encoding epochs ($F(19) = 8.008$, $p = 0.01$). Importantly, there was again an

interaction between reward probability and encoding epoch ($F(18) = 3.711$, $p = 0.04$).

One-way ANOVAs within early and late encoding epochs revealed a significant difference in memory late ($F(2,19) = 4.95$, $p = 0.01$) but no difference early ($F(2,19) = 1.31$, $p = 0.28$).

Replicating the memory benefit for uncertainty in the late encoding epoch seen at 24-hours, follow-up t-tests late again revealed a difference following 50% cues relative to both 100% and 0% cues, with no difference between the latter (100% vs. 50%: $t(19) = 2.97$, $p = 0.008$; 50% vs. 0%: $t(19) = 2.57$, $p = 0.02$; 100% vs. 0%: $t(19) = 0.66$, $p = 0.52$). Thus, the uncertainty effect was not dependent on consolidation (**Figure 12C**).

By contrast, the effect of reward probability seen at 24-hour retrieval for items in the early encoding epoch was not present at immediate retrieval. Although the ANOVA demonstrated no significant difference by reward probability early, the test for a linear trend was an a priori analyses for Experiment 2. We found no significant linear trend ($R^2 = 0.01$, $p = 0.14$). The influence of reward probability on memory early during anticipation was not present during immediate retrieval, and only appeared after 24 hours (**Figure 12C**).

At immediate retrieval, unlike 24-hour retrieval, there was no significant memory in guesses for any condition (all contrasts: $t(19) < 0.97$, $p > 0.34$; **Figure 12D**).

Finally, as was the case at 24-hour retrieval, analyses for immediate retrieval revealed no effects of reward outcome on memory during uncertain trials (early,

rewarded vs. unrewarded $t(19) < 0.0001$, $p = 1.00$; late, rewarded vs. unrewarded $t(19) = 1.33$, $p = 0.20$).

4.3.4 Experiments 1 & 2: Contrasting Delayed and Immediate memory results

To quantify whether memory patterns within early and late encoding epochs changed over a 24-hour period of consolidation, we ran 3 x 2 ANOVAs with the factors reward probability and retrieval time and looked for an interaction between the two.

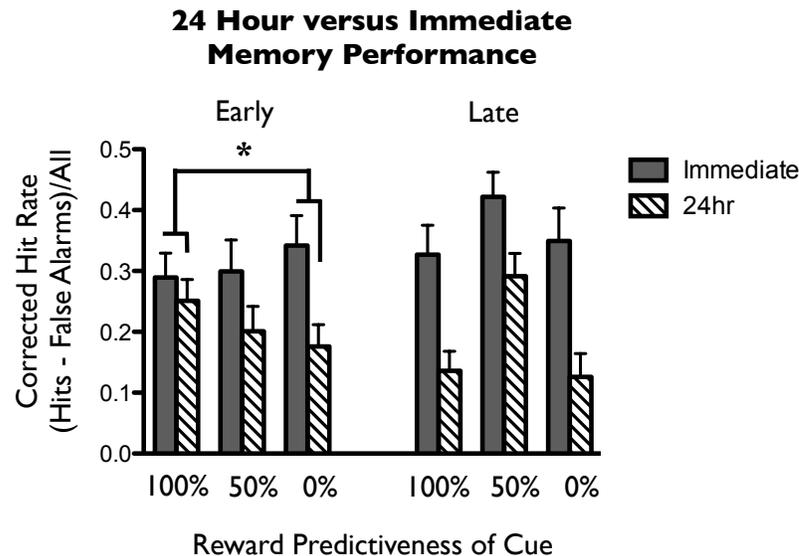


Figure 13: Retrieval group by expected value interaction. Within the early encoding period, there was a significant interaction between reward probability and retrieval group. There was a smaller difference between the 24-hour and immediate retrieval groups at 100% reward probability than at 0% reward probability, consistent with increased consolidation for higher probability items encoded early during anticipation. (* $p < 0.05$)

We found a significant interaction early ($F(2,37) = 4.281$, $p = 0.021$) but not late ($F(2,37) = 1.826$, $p = 0.175$). Follow-up pairwise ANOVAs revealed a significant

difference between 24-hour and immediate retrieval that was greater for 0% than 100% (0% Immediate-24hr > 100% Immediate-24hr: $F(1,38) = 8.76, p = 0.005$), with no other significant differences (all other ANOVAs: $F(1,38) < 1.75, p > 0.19$). After consolidation, early-encoded memory decreased more for 0% than it did for 100%. Thus, the relationship between memory and reward anticipation was consistent at the late encoding epoch, but changed significantly across retrieval periods in the early encoding epoch (**Figure 13**).

4.3.5 Experiments 1 & 2: Attentional Performance during encoding

To help rule out general effects of attention as the cause of the relationships between reward anticipation and memory, we asked whether the patterns of accuracy or reaction time resembled that of memory across conditions. In both retrieval groups, there were no accuracy differences by reward probability at early encoding (all contrasts: $F(2,19) < 2.37, p > 0.11$); however accuracy at late encoding in both groups revealed significant linear trends such that people were more accurate as reward probability increased (both retrieval groups: $R_{\text{square}} > 0.04, p < 0.004$; **Figure 14A and 14B**). Reaction time at late encoding in both groups revealed slower RTs for 50% rewarded trials relative to 0% rewarded trials (both retrieval groups: $t(1,19) > 2.22, p < 0.04$), but no other significant differences (all others: $t(1,19) < 1.45, p > 0.16$; **Figure 14C and 14D**). Thus, attention could not account for the above memory effects.

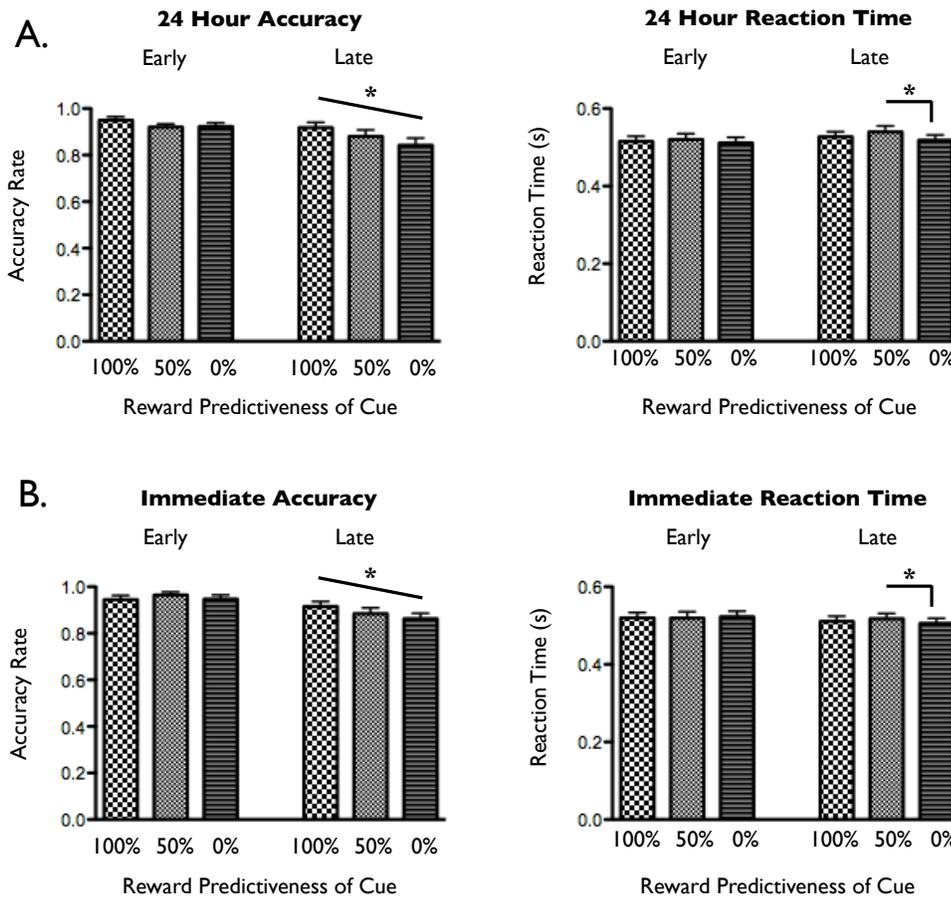


Figure 14: Figure 4. Attention for early and late encoding items at 24-hour and immediate retrieval. A. In the 24-hour retrieval group, early accuracy and early reaction time did not differ by expected value. Late accuracy linearly increased with increasing expected value. Late reaction times were slower for the 50% trials than the 0% trials B. In the 15-minute retrieval group, the patterns of accuracy and reaction time were the same as for the 24-hour retrieval group. (* $p < 0.05$)

4.4 Discussion

4.4.1 Summary

Our findings demonstrate temporally distinct reward anticipation influences on memory formation. During the early encoding epoch, 400ms after the presentation of the reward cue, memory scaled with expected value. During the late encoding epoch,

approximately three seconds after the presentation of the reward cue and just prior to the reward outcome, memory was greatest during high uncertainty. Interestingly, the uncertainty benefit was present both immediately and 24 hours after encoding, whereas the expected value benefit only emerged after 24 hours, suggesting a differential relationship with consolidation. Moreover, the pattern of correct guesses at 24 hours suggests that greater expected value during encoding resulted in items retaining a weak memory representation when they would otherwise be forgotten.

4.4.2 Integration with existing literature

While this was the first behavioral demonstration of dissociable temporal contexts for encoding within reward anticipation, the results build on expectations generated from prior neuroimaging and physiological studies. Previous fMRI work has demonstrated dissociable neural responses within the dopaminergic system for expected value and uncertainty (Preuschoff, Bossaerts, & Quartz, 2006; Tobler, O'Doherty, Dolan, & Schultz, 2006), with one study making a temporal distinction by demonstrating striatal activation in the first second after a reward cue for expected value versus the following seconds leading up to reward outcome during uncertainty (Preuschoff et al., 2006). Physiologically, cues associated with greater expected value elicit greater phasic dopamine firing in the midbrain at latencies under approximately 400ms (Fiorillo et al., 2003; Tobler et al., 2005), whereas a sustained dopaminergic ramp has been shown to increase with greater reward uncertainty over a two second period of reward

anticipation (Fiorillo et al., 2003). Importantly, increased dopamine has been theorized to enhance memory formation and retention (Duzel et al., 2010; Shohamy & Adcock, 2010). This study suggests new, specific relationships between memory and phasic or sustained dopamine: phasic dopamine may benefit memory early in the reward anticipation period and sustained anticipatory dopamine may have an effect close to the reward outcome.

The differential relationship with consolidation could help explain why some reward anticipation effects on memory are not consolidation-dependent (Gruber et al., 2014) even though rodent work has demonstrated the importance of consolidation for dopamine-dependent memory formation (Bethus et al., 2010; McNamara et al., 2014) and resulting theoretical mechanisms of dopamine synaptic activity emphasize consolidation (Lisman et al., 2011; Redondo & Morris, 2011). Our data integrates across previous studies, suggesting that phasic dopamine effects, potentially acting via synaptic release, may be reliant on consolidation, whereas sustained dopamine effects, potentially acting via extrasynaptic release, may occur at encoding.

4.4.3 Alternative Accounts

Multiple alternative accounts have been considered as potential explanations of our memory findings. One intuitive possibility is that enhanced encoding was a result of greater attention (Maunsell, 2004). However, attention, in the form of reaction time or accuracy, could not account for our mnemonic differences here. Another alternative we

considered was that the late uncertainty benefit was due to a phasic dopaminergic response to reward delivery. However, there was no evidence to support this interpretation since there was no memory difference for items presented prior to rewarded versus unrewarded outcomes on the uncertain trials.

4.4.4 Caveats

An important caveat for this work is that we did not measure dopamine directly, thus the mnemonic benefits may not have been dopaminergic in nature. In particular, the literature is not in consensus that there is a sustained dopamine signal during uncertainty (Niv et al., 2005). Other neurotransmitters, such as acetylcholine or noradrenaline, are likely to be involved. In fact, acetylcholine has been discussed as important for expected uncertainty (Yu & Dayan, 2005), which may be similar to the cued uncertainty in this study. Alternatively, noradrenergic fibers may be the primary source of dopamine release in the hippocampus (Smith & Greene, 2012), thus noradrenaline may play an important role as well. Further work remains to explore these alternate possibilities.

Finally, although we discuss these findings as they relate to consolidation, the across-subject design for immediate and 24-hour retrievals did not allow us to determine which encoded items were retained versus forgotten. A test-retest strategy would have introduced interference, and we did not have enough trials in each condition to split items across immediate and 24-hour retrievals. Thus, we cannot say with certainty that

fewer items were forgotten for the high expected value items presented early in encoding. However, we can conjecture with support from previous literature (Lisman et al., 2011; Redondo & Morris, 2011) that based on the memory performance for guesses, weakly encoded items with more dopamine present were retained at lower confidence rather than outright forgotten.

4.4.5 Conclusions

Anticipating reward shapes the way we learn about the world around us. However, a simple story, that reward anticipation uniformly benefits encoding, is insufficient to explain reward effects on memory. The present study shows that the temporal context within reward anticipation modulates how the expected value and uncertainty of reward cues impact memory outcomes. By mapping reward value and uncertainty onto putative physiological profiles, this work suggests a novel and testable model of dopaminergic influence on memory formation. We propose that phasic synaptic dopamine release acts during encoding to benefit memory, whereas sustained extrasynaptic dopamine release facilitates enhanced consolidation to benefit memory. This model is capable of integrating disparate findings across neuroscience on how rewards influence memory formation and paves the way for future research examining the contextually-regulated mechanisms responsible for reward-enhanced memory formation.

5. Discussion

5.1 Summary of findings

In Chapter 2, we had two aims. The first aim was to characterize the independent contributions of expectancy violation (surprise) and novelty in driving the VTA and supporting mesolimbic circuitry. Both novelty and surprise are unpredicted events. However, one results in an unpredicted sensory experience whereas the other violates expectations. In order to tease these phenomena apart, we developed a paradigm in which we crossed novelty and surprise, enabling us to isolate novelty independent of surprise, and surprise independent of novelty.

We demonstrated for the first time in humans non-novel, non-valenced surprise signals in the VTA. Expectancy violations drove the VTA regardless of novelty status. This is consistent with animal physiology studies, but an important demonstration in humans. We also showed that the VTA was responsive to novelty, even if it was not surprising. This suggests that the VTA is sensitive to all behaviorally relevant unpredicted events. These results are consistent with the idea that the VTA signals an orienting response when events are unpredicted, regardless of valence.

Activation was also present in the hippocampus and NAcc. The hippocampus was sensitive to novelty and surprise, in distributed clusters throughout the region. As a whole, the hippocampus demonstrated a main effect of surprise but not novelty. Interestingly, NAcc was sensitive to surprise but not novelty. It also demonstrated a

stronger response to surprise than the VTA. The above results suggest that a route from the hippocampus to the NAcc to the VTA might be more relevant for surprise than novelty.

Differential whole brain networks suggest that relationship of the dopaminergic system to behavior may differ for novelty and surprise. Novelty selectively activated the visual ventral stream, consistent with the possibility that the dopamine response may orient the visual system to new sensory information. Surprise activated a more widespread, whole-brain network, including regions involved in retrieval and updating, consistent with a role for the dopamine system in initiating learning from violations of expectation.

The second aim of Chapter 2 was to examine whether expectation of novelty influences declarative memory formation. We found no difference in memory for expected novel events versus surprising novel events. However, a region in prefrontal cortex correlated with memory for expected novelty but not surprising novelty, suggesting possible top down cognitive control influences resulting from successful expectancy generation.

In Chapter 3, we also had 2 aims. First, we wanted to characterize the mesolimbic neural response associated with the expectancy generation, anticipation, and resolution of high- versus low-curiosity information. To do this, we presented participants with trivia questions associated with either high or low curiosity for the answers. These trivia

questions were paired with symbolic cues representing the length of the upcoming anticipatory delay, and whether or not there would be an action contingency to see the trivia answer. Participants then waited 9 or 13 seconds, and either saw the trivia answer, or made a button press and then saw the answer.

Both the VTA and the hippocampus were more activated by high curiosity than low curiosity at cue. NAcc activation for curiosity scaled with individual differences in personality measures of curiosity. This would suggest that VTA and hippocampus were activated by unexpected, behaviorally relevant events, regardless of individual differences in personality measures of curiosity, but NAcc was only activated as much as people valued curiosity in their everyday lives. A region in PFC showed this curiosity effect more strongly when participants were not expecting an action contingency.

During anticipation, expectancy of high-curiosity information was held in PFC and hippocampus, but not the VTA. Again, this was stronger in PFC when there was no action contingency to anticipate. NAcc was still sensitive to individual differences, suggesting trait-level differences in curiosity impacted persistent anticipatory activation.

The VTA was also activated by high curiosity trivia answers more than low curiosity trivia answers. Thus, the satisfaction of high-curiosity expectation was signaled in the VTA. Notably, while the reward component of this information was predictable, the content was still uncertain. It was also present in PFC among other regions. The curiosity effect in the VTA and PFC was even stronger in the absence of an action

contingency, and in fact, only emerged in the hippocampus if participants were not also making a button press. It appears that making a movement may have suppressed the difference in these regions between high and low curiosity states, and activation was more robust when curiosity was the only motivational factor.

The second aim of Chapter 3 was to describe how induction of a high-curiosity expectancy state supports declarative memory formation. When participants generated an expectation of high-curiosity information, they had remarkably better recall for subsequent trivia answers. This was supported by increased brain activation for remembered > forgotten in high curiosity relative to low curiosity trials. This interaction was demonstrated at cue in the hippocampus, PFC, and ventral visual stream. At outcome, VTA, hippocampus, and PFC among others regions showed this effect. Thus, activation at cue and outcome was more predictive of future memory success when participants generated expectancy for high-curiosity information rather than low-curiosity information

In Chapter 4, our final aim was to dissociate the influences of expected reward value and reward uncertainty, putatively driving phasic and sustained VTA dopamine release respectively, on declarative memory formation. We used overlearned reward cues to indicate reward probability (100%, 50%, or 0%), establishing expected reward value independently from reward uncertainty. We presented items either early after the reward cue, to capture the effects of a phasic dopamine response, or several seconds

later just prior to reward feedback to capture a sustained dopamine response. These items were not relevant to the participants during reward anticipation, but were later tested for incidental memory. We also manipulated retrieval time across groups to examine whether memory was impacted at the time of encoding or after a period of consolidation.

In the 24-hour retrieval group, participants' memory for early epoch items scaled with expected reward value. However, memory for late epoch items was greatest during reward uncertainty. This was supported by significant memory for guesses at the highest expected reward value during the early epoch and the uncertain reward during the late epoch. Interestingly, the uncertainty effect was present and not significantly different in the immediate retrieval group, suggesting that the uncertainty benefit for memory occurred at encoding and was not altered during consolidation. The effect of expected reward value, on the other hand, was not present at immediate retrieval, and was in fact significantly different than the 24-hour group. It emerged only after a period of consolidation. Thus, it appears that expected reward value benefits memory for a brief window during reward anticipation just after the cue, whereas uncertainty is more beneficial during reward anticipation closer to the outcome.

These results are consistent with physiology studies showing a rapid, phasic dopamine response to reward cues scaling with expected reward value (Fiorillo et al., 2003; Tobler et al., 2005) and a sustained, ramping dopamine response scaling with

reward uncertainty (Fiorillo et al., 2003). Thus, given the demonstrated capacity of dopamine release in the hippocampus to modulate memory formation, a possible mechanism of memory enhancement is via dopaminergic modulation. These results integrate across many studies examining the relationship between reward anticipation and memory, suggesting that phasic dopamine may be critical for consolidation whereas sustained dopamine may benefit encoding.

5.2 Synthesizing findings into larger model

When examined independently, each of the studies included in this dissertation invites inquisition into distinct psychological phenomena, such as novelty and expectancy violation in Chapter 2, curiosity in Chapter 3, and reward anticipation in Chapter 4. They also examine the dopamine system at different timescales: cue, early and late anticipation, and outcome. It is only when grounding these works in the broader literature that they become part of a coherent story. Each of the studies here addresses targeted questions in the broad realm of expectation, memory, and the dopamine system. By synthesizing across studies in a common framework, we can update our model of how expectations influence memory via the dopaminergic system.

What do the studies together tell us about the following series of events? An animal experiences an unexpected, behaviorally relevant event. This event may trigger expectations about a future event. There is some period of sustained expectancy, for which active maintenance of the potential outcome may be more or less valued by the

organism, and may include more or less uncertainty about the outcome. The final event resolves expectations by either meeting or violating them. How is the dopaminergic system responding to each circumstance and what are the downstream effects on episodic memory formation?

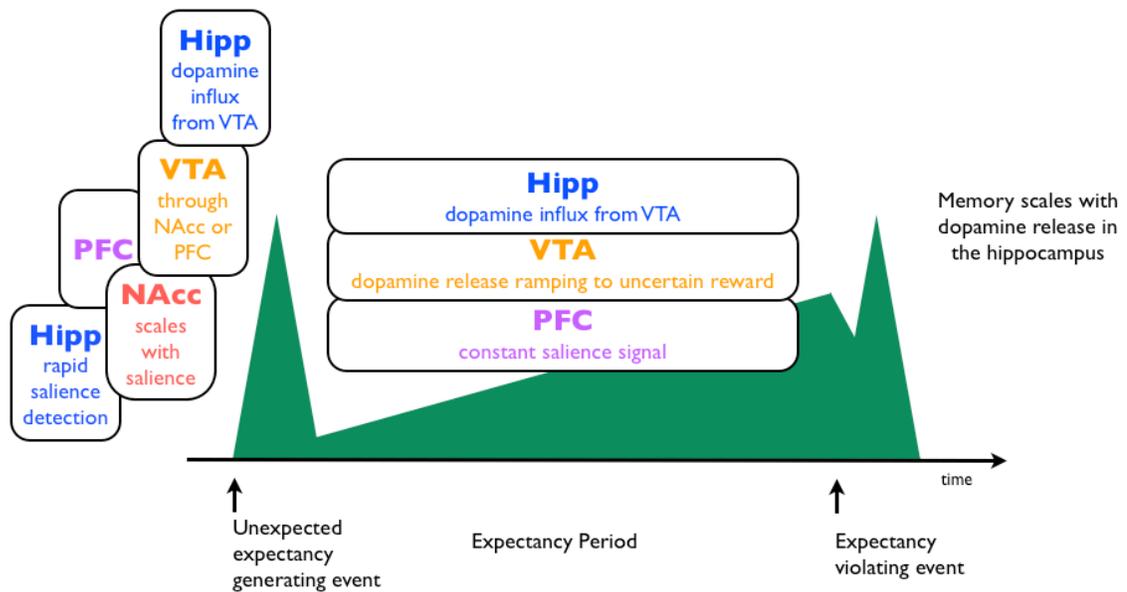


Figure 15: An updated model integrating expectation, the dopamine system and memory. Our data support claims from the model in the introduction. It additionally suggests NAcc scales with the salience or value of an event to an individual. During a sustained expectancy period, PFC may hold the expectation in working memory. Memory may scale with uncertainty or salience, which both influence dopamine release in the hippocampus.

Chapters 2 and 3 suggest that the dopaminergic system signals unpredicted, behaviorally relevant events, consistent with the proposal by Howitz and colleagues (Horvitz, 2000). This was shown via a dopaminergic midbrain response to novelty, which did not violate expectations, but whose content was not fully predictable. It was also demonstrated by a VTA response to trivia questions, the timing and content of

which were unpredictable, and trivia answers, to which the content was unpredictable. NAcc activation was distinct from VTA activation, seemingly moving with the salience of the event. This was evidenced by NAcc activation for curiosity scaling with individual differences in personality measures of curiosity as well as greater NAcc activation for surprise but not novelty. The PFC and hippocampus however were activated fairly consistently with the VTA, consistent with a potential path of information flow from the VTA to the PFC to the hippocampus.

As shown in Chapter 4, better memory for items rapidly following unexpected events thought to elicit more phasic dopamine firing, such as those that predict higher expected reward value (Fiorillo et al., 2003; Tobler et al., 2005), suggests that increased activation in the dopaminergic midbrain has the capacity to facilitate memory formation. Interestingly, our finding of enhanced memory at 24-hour retrieval but not immediate retrieval is consistent with a mechanism of dopaminergic modulation of LTP, manifesting over a period of consolidation (Redondo & Morris, 2011).

The studies here can start to address a period of expectancy on the order of seconds. A ramping, sustained dopamine signal has been demonstrated during expectancy scaling with reward uncertainty (Fiorillo et al., 2003), with signal in the striatum also following this pattern (Preuschoff et al., 2006). We showed that memory formation improves with greater uncertainty late, but not early in reward anticipation. This would seem to suggest that more dopamine release in the hippocampus late in

reward anticipation leads to better memory. We can only speak to uncertainty here, but there may be other contexts in which reward magnitude (Howe et al., 2013), or anticipatory success (Totah et al., 2013) may facilitate memory formation via the mesolimbic dopamine pathway during anticipation.

Interestingly, we did not find evidence of VTA activation during expectancy of high curiosity information. However, this period was long (9-13 seconds in Chapter 3, versus 3 seconds for Chapter 4) and there was no active goal pursuit, which may be important for keeping the system online. Instead, we found that the expectation of high curiosity information was held in PFC and hippocampus. Thus, it may be that top-down cognitive control and working memory holds the expectation more than the dopamine system itself does.

Once expectations have been generated, the outcome may meet or violate them. If the expectations are fuzzy, then the event may be unexpected (or unpredicted) without explicitly violating expectations. We found increased VTA, NAcc, hippocampus, and PFC activation for expectancy violations relative to expected events. In addition, we also found VTA activation for unexpected events (both novelty and high curiosity information), although weaker in NAcc and hippocampus. This would suggest that VTA signals a mismatch between expectations and outcome, potentially providing a learning signal for the rest of the brain. However, NAcc may scale with salience and may be more activated when there is a stronger initiation of goal-directed behavior

(Murty et al., 2013). The hippocampus may more discretely come online when detecting and remembering any unpredicted event. We found no difference in memory formation for expected versus surprising novelty, perhaps because memory may scale with increasing dopamine for uncertainty or salience, both of which are present for the two types of novel events.

5.3 Other considerations

5.3.1 Limitations

Although our findings integrate with existing literature, there are a number of caveats to consider. In this section, I will discuss the following: dopamine transmission versus dopaminergic midbrain activation; memory formation in the hippocampus versus the surrounding medial temporal lobe (MTL) cortex; behaviorally relevant versus non-behaviorally relevant events; and other brains regions that could be playing an important role such as the amygdala, anterior cingulate cortex (ACC) and orbitofrontal cortex (OFC).

The methods used in these studies, in particular functional neuroimaging, limit our capacity to make claims about the exact nature of activation in the VTA. Functional MRI measures blood oxygenation level dependent (BOLD) signal, which is the amount of local blood oxygen in a region and is thought to increase with increasing neural activity (Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001). However, because we could not isolate activity in the VTA due to dopaminergic signals alone, our activation in

the VTA almost assuredly captures activity by GABAergic and glutamatergic neurons as well. In fact, GABA may contribute to the sustained signal during reward anticipation, as shown by Cohen and colleagues (Cohen et al., 2012). GABAergic neurons often demonstrate similar patterns of activity as dopamine neurons (e.g. Totah et al., 2013) and likely modulate local dopamine activity. Cell labeling has shown that of all neurons in the VTA, approximately 70% are dopaminergic, 35% are GABAergic, and 2-3% are glutamatergic (Nair-Roberts et al., 2008), with some neurons co-expressing both dopamine and GABA or glutamate. However, in a study utilizing both fMRI and positron emission tomography (PET), researchers demonstrated that the BOLD signal in the dopaminergic midbrain correlates with dopamine release in the ventral striatum (Schott et al., 2008), providing support for the interpretation that VTA BOLD activation is at least partially capturing neuronal dopamine activity. Thus, although we cannot say for certain that we are measuring dopamine activity, our signal in the VTA likely correlates with downstream dopamine release.

Additionally, our model only speaks to behaviorally relevant events. None of the included studies examine the dopaminergic implications of irrelevant events (although we do show improved memory for irrelevant items during expectancy of reward). However, based on the literature showing dopaminergic habituation to irrelevant salient events (Ljungberg et al., 1992), we would make the prediction that only irrelevant events

that are unexpected and for which the relevance must be still be determined would activate the VTA.

Another important oversimplification of our model is that all dopaminergic modulation of episodic memory formation occurs in the hippocampus. Empirical data show that regions in the surrounding MTL cortex are sufficient for item-level episodic memory formation (Davachi, Mitchell, & Wagner, 2003). They also receive afferents from the VTA (Insausti et al., 1998; Room & Groenewegen, 1986). Thus, some modulation may take place in surrounding MTL areas. We cannot address this issue in our behavioral work. However, our neuroimaging findings from Chapters 2 did not demonstrate a response in MTL cortex to novelty or surprise. We also did not see activation in the MTL for the interaction of curiosity and memory in Chapter 3, so we have no evidence for expectancy-modulated memory contributions by the MTL. It may be possible that small volume corrections would have allowed us to detect activation in these regions. However, given our results in the hippocampus and the demonstrated dopaminergic modulation of the hippocampus, our model remains simple in this work.

The circuit discussed in this work is also limited in scope. We emphasize the mesolimbic dopamine pathway, which includes the VTA, NAcc and hippocampus. However, the mesolimbic pathway also includes the amygdala, which will be discussed more here. The data also show signs of activation in the dACC/mPFC and OFC, thus those regions will be briefly considered as well.

Although the amygdala has primarily been associated with expectation and memory during aversive states (Ablner, Erk, Herwig, & Walter, 2007; Johansen, Treppey, LeDoux, & Blair, 2010; Murty et al., 2012; Sarinopoulos et al., 2009), neurons in the amygdala have been shown to be responsive to expectancy violating stimuli both pleasant and aversive (Belova, Paton, Morrison, & Salzman, 2007). The amygdala projects indirectly to the VTA via the NAcc and receives direct afferents from the VTA (Shohamy & Adcock, 2010). In the present studies, activation in the amygdala was only noted for remembered > forgotten trivia answers, with no relationship at cue and no relationship with curiosity state. There was also no detectable amygdala activation for non-valenced novelty or surprise. Given distinct populations of neurons in the VTA that fire in response to valenced or neutral events (Bromberg-Martin et al., 2010), it may be that dopamine more strongly modulates the amygdala when there is an associated rewarding or punishing component. It may also be that the amygdala response to expectancy violations scales with an associative learning signal that we did not directly measure (Holland & Gallagher, 1999). Future work that closely examines amygdala in these and other contexts may illuminate the functional role of the amygdala in the modulation of memory by expectation independent of valence.

One recent theory of dACC/mPFC function is that it encodes surprise, specifically the non-occurrence of an expected event (Alexander & Brown, 2011; Egner, 2011). Others include conflict monitoring (Weissman, Giesbrecht, Song, Mangun, &

Woldorff, 2003) or error detection (Holroyd, Nieuwenhuis, & Yeung, 2004). In any of these models, ACC signal could then contribute to learning from errors of expected action-outcome contingencies. In our studies, as in many others, we do in fact see dACC/mPFC activation. The region is responsive to both novelty and surprise, as well as curiosity at both cue and outcome. The region is also more predictive of memory success for high curiosity information. Thus, the dACC/mPFC is active in many instances that the VTA is active in our data. Its response to expected novelty > expected familiar, in which there is no non-occurrence of an expected event is more consistent with ACC being responsive to all unpredicted events. It may therefore be an important part of the pathway including hippocampus, NAcc, and VTA and should be further investigated in future work.

OFC has also been tightly associated with expectation. OFC neurons are active when animals initiate expectation-driven behavior and reward-seeking (Moorman & Aston-Jones, 2014; Nobre, Coull, Frith, & Mesulam, 1999; Tremblay & Schultz, 2000). There is a connection between OFC and VTA (Lodge, 2011). In our studies, although we see considerable PFC activation in other places, OFC is absent for both novelty and surprise, perhaps because behavior does not change substantially as a result of these events. We do see OFC activation for high curiosity cues > low curiosity cues, and OFC is also more predictive of memory success for high curiosity outcomes than low

curiosity outcomes. OFC may then be more involved in the circuit when behavior is changed as a result of expected value.

In addition, other neurotransmitter systems are likely to be involved, such as acetylcholine and noradrenaline. Acetylcholine has been discussed as important for expected uncertainty (Yu & Dayan, 2005). Thus, while we might not presume acetylcholine signals all expectancy cuing or violating events, contexts in which uncertainty is consistent might elevate acetylcholine modulation of the hippocampus, a region which is densely populated with acetylcholine receptors (Alkondon & Albuquerque, 1993). In fact, uncertainty can be expected to varying extents in all three studies, thus it is likely that acetylcholine is influencing the hippocampus simultaneously with dopamine. This remains to be further investigated in future work.

Alternatively, norepinephrine is associated with sudden increases in arousal (Foote, Aston-Jones, & Bloom, 1980) and its source nucleus, the locus coeruleus, shares reciprocal connections with the VTA (Sara, 2009). Noradrenergic fibers may even be the primary source of dopamine release in the hippocampus (Smith & Greene, 2012). Thus, it is quite possible norepinephrine is playing an important role in the modulation of memory. However, locus coeruleus is very small and thus very challenging to examine via fMRI. Similar studies in non-human systems may better be able to address the functional role of norepinephrine.

5.4 Future Directions

The work presented in this dissertation can be used to inform a model for how expectations modulate memory formation. This model is both testable and expandable, and future work can contribute greatly to clarifying the current unknowns.

While we were interested in circuit dynamics, we did not examine connectivity in these studies. Thus, future work may examine functional connectivity via fMRI, or more direct circuit dynamics using neurobiology systems-level tools such as optogenetic manipulation. For instance, we propose here that expectancy information may be conveyed from the hippocampus to the VTA via PFC, rather than through the NAcc, but this remains to be empirically tested. Directed connectivity could also help inform whether the NAcc signal we see in our fMRI studies is being conveyed to the VTA or is a downstream result of VTA activation.

We also propose that memory scales with dopamine release in the hippocampus, but this has yet to be demonstrated directly. An interesting future study would use optogenetic manipulation to not only look at the impact of phasic versus sustained dopamine release on hippocampal synapses (Rosen et al., 2015), but also on subsequent memory outcomes. It would also be interesting to test our hypothesis that synaptic modulation of hippocampal LTP by phasic dopamine enhances consolidation, whereas extrasynaptic modulation by sustained dopamine enhances encoding.

As noted above, there are many other regions and neurotransmitter systems that may be involved in shaping memory as a result of expectation. The contributions of the amygdala, dACC, OFC, and MTL cortex should be further investigated; likewise, the influences of acetylcholine and norepinephrine could be targets for future study.

In addition, while the present model attempts to integrate expectancy across motivational valence (e.g. reward versus punishment versus non-valenced events), there may be differences in dopaminergic engagement due to valence, above and beyond expectancy. Clarifying when motivational valence changes expectancy circuits is still an important step, and work in this domain is ongoing (Murty, LaBar, & Adcock, 2016b).

Finally, all of the present work has been completed in healthy young adults. Expectations change over the course of aging, and future work could consider how expectancy modulation of memory evolves over the course of development. Examination of patients, such as those with Parkinson's disease or MTL lesions could provide insight into mnemonic outcomes when the system is disrupted. This could also be accomplished by knocking out potentially critical regions in rodents or monkeys to determine the necessity of these regions for the changing of memory formation by expectation. Thus, studies in different developmental groups, in patients, or in animals, could extend our knowledge of the system beyond healthy young adults.

5.5 Conclusions

We demonstrated in the current studies that non-novel, non-valenced expectancy violations activate the VTA, as does non-surprising novelty. This was the first demonstration of these effects in humans, since other studies typically use appetitive or aversive events or conflate novelty and expectancy violation. We did not find a difference in memory for surprising versus expected novel events, suggesting that behaviorally relevant novel events (which are necessarily unpredictable) induce the same amount of dopaminergic modulation regardless of expectation. We also showed that the induction and resolution of high-curiosity expectancy activates the VTA, although during anticipation this signal was held in PFC and hippocampus. The VTA and hippocampus were more predictive of memory for information following a high-curiosity expectancy state, providing support for a dopaminergic mechanism of memory enhancement. We also demonstrated that during reward anticipation, memory for items immediately following cues for higher expected reward values were better remembered; however, uncertainty benefitted memory just prior to the outcome. This was the first demonstrated behavioral consequence for increased phasic dopamine in response to higher expected value but increased sustained dopamine for greater uncertainty; in both cases, memory scaled with the putative amount of dopamine release.

Together, these results provide support for a dopaminergic mechanism of expectation-induced memory modulation. Integrating our work and existing literature

allows us to propose the following testable model. Expectations modulate memory formation via the mesolimbic dopaminergic pathway. Motivationally significant events that induce expectancy, are unexpected, or violate expectations are detected by the hippocampus and activate the VTA, via either the NAcc or PFC. Downstream dopamine release in the hippocampus modulates LTP, synaptically for events that cause phasic dopamine release, and extrasynaptically for events that cause sustained dopamine release. Hippocampal LTP is critically important for episodic memory formation. Thus, expectations (that may scale with salience, value, or certainty) elicit activity in the dopamine system, which acts on the hippocampus to modulate memory formation.

References

- Abler, B., Erk, S., Herwig, U., & Walter, H. (2007). Anticipation of aversive stimuli activates extended amygdala in unipolar depression. *Journal of Psychiatric Research, 41*(6), 511–522.
- Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. D. E. (2006). Reward-Motivated Learning: Mesolimbic Activation Precedes Memory Formation. *Neuron, 50*(3), 507–517.
- Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience, 14*(10), 1338–1344.
- Alkondon, M., & Albuquerque, E. X. (1993). Diversity of nicotinic acetylcholine receptors in rat hippocampal neurons. I. Pharmacological and functional evidence for distinct structural subtypes. *Journal of Pharmacology and Experimental Therapeutics, 265*(3), 1455–1473.
- Au-Young, S. M., Shen, H., & Yang, C. R. (1999). Medial prefrontal cortical output neurons to the ventral tegmental area (VTA) and their responses to burst-patterned stimulation of the VTA: neuroanatomical and in vivo electrophysiological analyses. *Synapse (New York, N.Y.), 34*(4), 245–255.
- Axmacher, N., Cohen, M. X., Fell, J., Haupt, S., Dumpelmann, M., Elger, C. E., et al. (2010). Intracranial EEG Correlates of Expectancy and Memory Formation in the Human Hippocampus and Nucleus Accumbens. *Neuron, 65*(4), 541–549.
- Bakker, A., Kirwan, C. B., Miller, M., & Stark, C. E. L. (2008). Pattern Separation in the Human Hippocampal CA3 and Dentate Gyrus. *Science, 319*(5870), 1640–1642.
- Baldassarre, G., & Mirolli, M. (2013). Intrinsically Motivated Learning in Natural and Artificial Systems. *Intrinsically Motivated Learning in Natural and Artificial Systems, 49*–72.
- Balkwell, J. W. (1991). From Expectations to Behavior: An Improved Postulate for Expectation States Theory. *American Sociological Review, 56*(3), 355.

- Ballard, I. C., Murty, V. P., Carter, R. M., MacInnes, J. J., Huettel, S. A., & Adcock, R. A. (2011). Dorsolateral Prefrontal Cortex Drives Mesolimbic Dopaminergic Regions to Initiate Motivated Behavior. *Journal of Neuroscience*, 31(28), 10340–10346.
- Barbas, H., & Blatt, G. J. (1995). Topographically specific hippocampal projections target functionally distinct prefrontal areas in the rhesus monkey. *Hippocampus*, 5(6), 511–533.
- Barto, A., Mirolli, M., & Baldassarre, G. (2013). Novelty or Surprise? *Frontiers in Psychology*, 4.
- Beckstead, R. M., Domesick, V. B., & Nauta, W. J. H. (1979). Efferent connections of the substantia nigra and ventral tegmental area in the rat. *Brain Research*, 175(2), 191–217.
- Belova, M. A., Paton, J. J., Morrison, S. E., & Salzman, C. D. (2007). Expectation Modulates Neural Responses to Pleasant and Aversive Stimuli in Primate Amygdala. *Neuron*, 55(6), 970–984.
- Berlyne, D. E. (1954). An Experimental Study of Human Curiosity. *British Journal of Psychology General Section*, 45(4), 256–265.
- Berlyne, D. E. (1960). Conflict, Arousal, and Curiosity.
- Bethus, I., Tse, D., & Morris, R. G. M. (2010). Dopamine and Memory: Modulation of the Persistence of Memory for Novel Hippocampal NMDA Receptor-Dependent Paired Associates. *Journal of Neuroscience*, 30(5), 1610–1618.
- Blumenfeld, B., Preminger, S., Sagi, D., & Tsodyks, M. (2006). Dynamics of memory representations in networks with novelty-facilitated synaptic plasticity. *Neuron*, 52(2), 383–394.
- Bollinger, J., Rubens, M. T., Zanto, T. P., & Gazzaley, A. (2010). Expectation-Driven Changes in Cortical Functional Connectivity Influence Working Memory and Long-Term Memory Performance. *The Journal of Neuroscience*, 30(43), 14399–14410.

- Bromberg-Martin, E. S., & Hikosaka, O. (2009). Midbrain Dopamine Neurons Signal Preference for Advance Information about Upcoming Rewards. *Neuron*, 63(1), 119–126.
- Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in Motivational Control: Rewarding, Aversive, and Alerting. *Neuron*, 68(5), 815–834.
- Bubic, A., von Cramon, D. Y., & Schubotz, R. I. (2010). Prediction, Cognition and the Brain. *Frontiers in Human Neuroscience*, 4.
- Bunzeck, N., & Duzel, E. (2006). Absolute Coding of Stimulus Novelty in the Human Substantia Nigra/VTA. *Neuron*, 51(3), 369–379.
- Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, 35(4), 625–642.
- Cabeza, R., Ciaramelli, E., & Olson, I. R. (2008). The parietal cortex and episodic memory: an attentional account. *Nature Reviews Neuroscience*, 9(8), 613–625.
- Cameron, J. (2001). Negative Effects of Reward on Intrinsic Motivation: A Limited Phenomenon: Comment on Deci, Koestner, and Ryan (2001). *Review of Educational Research*, 71(1), 29–42.
- Carpenter, S. K., & Vul, E. (2011). Delaying feedback by three seconds benefits retention of face–name pairs: the role of active anticipatory processing. *Memory & Cognition*, 39(7), 1211–1221.
- Carr, D. B., & Sesack, S. R. (1996). Hippocampal afferents to the rat prefrontal cortex: Synaptic targets and relation to dopamine terminals. *The Journal of Comparative Neurology*, 369(1), 1–15.
- Carr, D. B., & Sesack, S. R. (2000). Projections from the rat prefrontal cortex to the ventral tegmental area: target specificity in the synaptic associations with mesoaccumbens and mesocortical neurons. *Journal of Neuroscience*, 20(10), 3864–3873.

- Chen, J., Olsen, R. K., Preston, A. R., Glover, G. H., & Wagner, A. D. (2011). Associative retrieval processes in the human medial temporal lobe: Hippocampal retrieval success and CA1 mismatch detection. *Learning & Memory, 18*(8), 523–528.
- Chun, M. M., & Turk-Browne, N. B. (2007). Interactions between attention and memory. *Current Opinion in Neurobiology, 17*(2), 177–184.
- Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B., & Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature, 482*(7383), 85–88.
- Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *NeuroImage, 35*(2), 968–978.
- Colombo, M. (2014). Deep and beautiful. The reward prediction error hypothesis of dopamine. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 45*, 57–67.
- Corbetta, M., Patel, G., & Shulman, G. L. (2008). The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron, 58*(3), 306–324.
- Curtis, M. E., & Bharucha, J. J. (2009). Memory and Musical Expectation for Tones in Cultural Context. *Music Perception: an Interdisciplinary Journal, 26*(4), 365–375.
- D'Ardenne, K., Lohrenz, T., & Bartley, K. A. (2013). Computational heterogeneity in the human mesencephalic dopamine system. *Cognitive, Affective, & Behavioral Neuroscience, 13*(4), 747–756.
- D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD Responses Reflecting Dopaminergic Signals in the Human Ventral Tegmental Area. *Science, 319*(5867), 1264–1267.
- Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Human Brain Mapping, 8*(2-3), 109–114.

- Davachi, L., Mitchell, J. P., & Wagner, A. D. (2003). Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences*, 100(4), 2157–2162.
- Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6), 627–668.
- DeLong, M. R., Crutcher, M. D., & Georgopoulos, A. P. (1983). Relations between movement and single cell discharge in the substantia nigra of the behaving monkey. *Journal of Neuroscience*, 3(8), 1599–606.
- Dickerson, K. C., Li, J., & Delgado, M. R. (2011). Parallel contributions of distinct human memory systems during probabilistic learning. *NeuroImage*, 55(1), 266–276.
- Dommett, E. (2005). How Visual Stimuli Activate Dopaminergic Neurons at Short Latency. *Science*, 307(5714), 1476–1479.
- Duzel, E., Bunzeck, N., Guitart-Masip, M., & Duzel, S. (2010). NOvelty-related Motivation of Anticipation and exploration by Dopamine (NOMAD): Implications for healthy aging. *Neuroscience and Biobehavioral Reviews*, 34(5), 660–669.
- Duzel, E., Bunzeck, N., Guitart-Masip, M., Wittmann, B., Schott, B. H., & Tobler, P. N. (2009). Functional imaging of the human dopaminergic midbrain. *Trends in Neurosciences*, 32(6), 321–328.
- Hirsch, E., Graybiel, A. M., & Agid, Y. A. (1988). Melanized dopaminergic neurons are differentially susceptible to degeneration in Parkinson's disease. *Nature*, 334(6180), 345–348.
- Egner, T. (2011). Surprise! A unifying model of dorsal anterior cingulate function? *Nature Neuroscience*, 14(10), 1219.
- Egner, T., Monti, J. M., & Summerfield, C. (2010). Expectation and surprise determine neural population responses in the ventral visual stream. *Journal of Neuroscience*, 30(49), 16601–16608.

- Eichenbaum, H. (2000). A cortical-hippocampal system for declarative memory. *Nature Reviews Neuroscience*, 1(1), 41–50.
- Eshel, N., Tian, J., Bukwich, M., & Uchida, N. (2016). Dopamine neurons share common response function for reward prediction error. *Nature Neuroscience*.
- Esterman, M., & Yantis, S. (2009). Perceptual Expectation Evokes Category-Selective Cortical Activity. *Cerebral Cortex*, 20(5), 188–1253.
- Fazio, L. K., & Marsh, E. J. (2009). Surprising feedback improves later memory. *Psychonomic Bulletin & Review*, 16(1), 88–92.
- Federmeier, K. D., & Kutas, M. (1999). A Rose by Any Other Name: Long-Term Memory Structure and Sentence Processing. *Journal of Memory and Language*, 41(4), 469–495.
- Fenker, D. B., Frey, J. U., Schuetze, H., Heipertz, D., Heinze, H.-J., & Duzel, E. (2008). Putting the Pieces Together: The Role of Dorsolateral Prefrontal Cortex in Relational Memory Encoding. *Journal of Cognitive Neuroscience*, 20(7), 1250–1265.
- Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete Coding of Reward Probability and Uncertainty by Dopamine Neurons. *Science*, 299(5614), 1898–1902.
- Fiorillo, C. D., Newsome, W. T., & Schultz, W. (2008). The temporal precision of reward prediction in dopamine neurons. *Nature Neuroscience*, 11(8), 966–973.
- Floresco, S. B., Todd, C. L., & Grace, A. A. (2001). Glutamatergic afferents from the hippocampus to the nucleus accumbens regulate activity of ventral tegmental area dopamine neurons. *The Journal of Neuroscience*, 21(13), 4915–4922.
- Floresco, S. B., West, A. R., Ash, B., Moore, H., & Grace, A. A. (2003). Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nature Neuroscience*, 6(9), 968–973.
- Foote, S. L., Aston-Jones, G., & Bloom, F. E. (1980). Impulse activity of locus coeruleus neurons in awake rats and monkeys is a function of sensory stimulation and arousal.

Proceedings of the National Academy of Sciences, 77(5), 3033–3037.

- Frankle, W. G., Laruelle, M., & Haber, S. N. (2006). Prefrontal cortical projections to the midbrain in primates: evidence for a sparse connection. *Neuropsychopharmacology*, 31(8), 1627–1636.
- Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, 108(3), 316–355.
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138.
- Friston, K. J., & Stephan, K. E. (2007). Free-energy and the brain. *Synthese*, 159(3), 417–458.
- Gasbarri, A., Packard, M. G., Campana, E., & Pacitti, C. (1994). Anterograde and retrograde tracing of projections from the ventral tegmental area to the hippocampal formation in the rat. *Brain Research Bulletin*, 33(4), 445–452.
- Gasbarri, A., Sulli, A., & Packard, M. G. (1997). The dopaminergic mesencephalic projections to the hippocampal formation in the rat. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 21(1), 1–22.
- Goldman-Rakic, P. S., Selemon, L. D., & Schwartz, M. L. (1984). Dual pathways connecting the dorsolateral prefrontal cortex with the hippocampal formation and parahippocampal cortex in the rhesus monkey. *Neuroscience*, 12(3), 719–743.
- Goto, Y., & Grace, A. A. (2005). Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nature Neuroscience*, 8(6), 805–812.
- Green, R. T. (1956). Surprise as a factor in the von Restorff effect. *Journal of Experimental Psychology*, 52(5), 340–344.
- Greicius, M. D., Krasnow, B., Boyett-Anderson, J. M., Eliez, S., Schatzberg, A. F., Reiss, A. L., & Menon, V. (2003). Regional analysis of hippocampal activation during

- memory encoding and retrieval: fMRI study. *Hippocampus*, 13(1), 164–174.
- Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, 10(1), 14–23.
- Gruber, M. J., & Otten, L. J. (2010). Voluntary control over prestimulus activity related to encoding. *The Journal of Neuroscience*, 30(29), 9793–9800.
- Gruber, M. J., Gelman, B. D., & Ranganath, C. (2014). States of Curiosity Modulate Hippocampus-Dependent Learning via the Dopaminergic Circuit. *Neuron*, 84(2), 486–496.
- Guarraci, F. A., & Kapp, B. S. (1999). An electrophysiological characterization of ventral tegmental area dopaminergic neurons during differential pavlovian fear conditioning in the awake rabbit. *Behavioural Brain Research*, 99(2), 169–179.
- Guitart-Masip, M., Chowdhury, R., Sharot, T., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Action controls dopaminergic enhancement of reward representations. *Proceedings of the National Academy of Sciences of the United States of America*, 109(19), 7511–7516.
- Guitart-Masip, M., Fuentemilla, L., Bach, D. R., Huys, Q. J. M., Dayan, P., Dolan, R. J., & Duzel, E. (2011). Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. *Journal of Neuroscience*, 31(21), 7867–7875.
- Halgren, E., Marinkovic, K., & Chauvel, P. (1998). Generators of the late cognitive potentials in auditory and visual oddball tasks. *Electroencephalography and Clinical Neurophysiology*, 106(2), 156–164.
- Henson, R. N., & Gagnepain, P. (2010). Predictive, interactive multiple memory systems. *Hippocampus*, 20(11), 1315–1326.
- Hirshman, E. (1988). The expectation-violation effect: Paradoxical effects of semantic relatedness. *Journal of Memory and Language*, 27(1), 40–58.

- Holland, P. C., & Gallagher, M. (1999). Amygdala circuitry in attentional and representational processes. *Trends in Cognitive Sciences*, 3(2), 65–73.
- Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R. B., Coles, M. G., & Cohen, J. D. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nature neuroscience*, 7(5), 497–498.
- Horvitz, J. (2000). Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience*, 96(4), 651–656.
- Horvitz, J. C., Stewart, T., & Jacobs, B. L. (1997). Burst activity of ventral tegmental dopamine neurons is elicited by sensory stimuli in the awake cat. *Brain Research*, 759(2), 251–258.
- Howe, M. W., Tierney, P. L., Sandberg, S. G., Phillips, P. E. M., & Graybiel, A. M. (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature*, 500(7464), 575–579.
- Ihalainen, J. A., Riekkinen, P., Jr, & Feenstra, M. G. P. (1999). Comparison of dopamine and noradrenaline release in mouse prefrontal cortex, striatum and hippocampus using microdialysis. *Neuroscience Letters*, 277(2), 71–74.
- Insausti, R., Juottonen, K., Soininen, H., Insausti, A. M., Partanen, K., Vainio, P., et al. (1998). MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *American Journal of Neuroradiology*, 19(4), 659–671.
- Jay, T. M., & Witter, M. P. (1991). Distribution of hippocampal CA1 and subicular efferents in the prefrontal cortex of the rat studied by means of anterograde transport of Phaseolus vulgaris-leucoagglutinin. *The Journal of Comparative Neurology*, 313(4), 574–586.
- Johansen, J. P., Tarpley, J. W., LeDoux, J. E., & Blair, H. T. (2010). Neural substrates for expectation-modulated fear learning in the amygdala and periaqueductal gray. *Nature Neuroscience*, 13(8), 979–986.
- Kafkas, A., & Montaldi, D. (2014). Two separate, but interacting, neural systems for

- familiarity and novelty detection: A dual-route mechanism. *Hippocampus*, 24(5), 516–527.
- Kahn, I., & Shohamy, D. (2012). Intrinsic connectivity between the hippocampus, nucleus accumbens, and ventral tegmental area in humans. *Hippocampus*, 23(3), 187–192.
- Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T.-Y., & Camerer, C. F. (2009). The Wick in the Candle of Learning: Epistemic Curiosity Activates Reward Circuitry and Enhances Memory. *Psychological Science*, 20(8), 963–973.
- Kashdan, T. B., Gallagher, M. W., Silvia, P. J., Winterstein, B. P., Breen, W. E., Terhar, D., & Steger, M. F. (2009). The curiosity and exploration inventory-II: Development, factor structure, and psychometrics. *Journal of Research in Personality*, 43(6), 987–998.
- Kiehl, K. A., Laurens, K. R., & Duty, T. L. (2001). An event-related fMRI study of visual and auditory oddball tasks. *Journal of Psychophysiology*, 15(4), 221.
- Kish, S. J., Shannak, K., & Hornykiewicz, O. (1988). Uneven Pattern of Dopamine Loss in the Striatum of Patients with Idiopathic Parkinson's Disease. *The New England Journal of Medicine*, 318(14), 876–880.
- Kishiyama, M. M., & Yonelinas, A. P. (2003). Novelty effects on recollection and familiarity in recognition memory. *Memory & Cognition*, 31(7), 1045–1051.
- Kishiyama, M. M., Yonelinas, A. P., & Knight, R. T. (2009). Novelty enhancements in memory are dependent on lateral prefrontal cortex. *The Journal of Neuroscience*, 29(25), 8114–8118.
- Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *Journal of Neuroscience*, 21(16), RC159.
- Köhler, S., Danckert, S., Gati, J. S., & Menon, R. S. (2005). Novelty responses to relational and non-relational information in the hippocampus and the parahippocampal

- region: A comparison based on event-related fMRI. *Hippocampus*, 15(6), 763–774.
- Krebs, R. M., Heipertz, D., Schuetze, H., & Duzel, E. (2011). Novelty increases the mesolimbic functional connectivity of the substantia nigra/ventral tegmental area (SN/VTA) during reward anticipation: Evidence from high-resolution fMRI. *NeuroImage*, 58(2), 647–655.
- Kumaran, D., & Maguire, E. A. (2006). An Unexpected Sequence of Events: Mismatch Detection in the Human Hippocampus. *PLoS Biology*, 4(12), e424.
- Lak, A., Stauffer, W. R., & Schultz, W. (2014). Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proceedings of the National Academy of Sciences of the United States of America*, 111(6), 2343–2348.
- Lammel, S., Ion, D. I., Roeper, J., & Malenka, R. C. (2011). Projection-Specific Modulation of Dopamine Neuron Synapses by Aversive and Rewarding Stimuli. *Neuron*, 70(5), 855–862.
- Larsson, J., & Smith, A. T. (2012). fMRI Repetition Suppression: Neuronal Adaptation or Stimulus Expectation? *Cerebral Cortex*, 22(3), 567–576.
- Legault, M., & Wise, R. A. (2001). Novelty-evoked elevations of nucleus accumbens dopamine: dependence on impulse flow from the ventral subiculum and glutamatergic neurotransmission in the ventral tegmental area. *European Journal of Neuroscience*, 13(4), 819–828.
- Lemon, N. (2006). Dopamine D1/D5 Receptors Gate the Acquisition of Novel Information through Hippocampal Long-Term Potentiation and Long-Term Depression. *Journal of Neuroscience*, 26(29), 7723–7729.
- Li, S., Cullen, W. K., Anwyl, R., & Rowan, M. J. (2003). Dopamine-dependent facilitation of LTP induction in hippocampal CA1 by exposure to spatial novelty. *Nature Neuroscience*, 6(5), 526–531.
- Light, L. L., Kayra-Stuart, F., & Hollander, S. (1979). Recognition memory for typical and unusual faces. *Journal of Experimental Psychology: Human Learning and Memory*, 5(3),

212–228.

- Lisman, J. E., & Grace, A. A. (2005). The Hippocampal-VTA Loop: Controlling the Entry of Information into Long-Term Memory. *Neuron*, *46*(5), 703–713.
- Lisman, J. E., & Otmakhova, N. A. (2001). Storage, recall, and novelty detection of sequences by the hippocampus: Elaborating on the SOCRATIC model to account for normal and aberrant effects of dopamine. *Hippocampus*, *11*(5), 551–568.
- Lisman, J., Grace, A. A., & Duzel, E. (2011). A neoHebbian framework for episodic memory; role of dopamine-dependent late LTP. *Trends in Neurosciences*, *34*(10), 536–547.
- Litman, J. (2005). Curiosity and the pleasures of learning: Wanting and liking new information. *Cognition & Emotion*, *19*(6), 793–814.
- Ljungberg, T., Apicella, P., & Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *Journal of Neurophysiology*, *67*(1), 145–163.
- Lodge, D. J. (2011). The Medial Prefrontal and Orbitofrontal Cortices Differentially Regulate Dopamine System Function. *Neuropsychopharmacology*, *36*(6), 1227–1236.
- Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*(1), 75–98.
- Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150–157.
- Loh, E., Kumaran, D., Koster, R., Berron, D., Dolan, R., & Duzel, E. (2015). Context-specific activation of hippocampus and SN/VTA by reward is related to enhanced long-term memory for embedded objects. *Neurobiology of Learning and Memory*.
- Luo, A. H., Tahsili-Fahadan, P., Wise, R. A., Lupica, C. R., & Aston-Jones, G. (2011).

- Linking Context with Reward: A Functional Circuit from Hippocampal CA3 to Ventral Tegmental Area. *Science*, 333(6040), 353–357.
- Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, 19(3), 1233–1239.
- Martig, A. K., & Mizumori, S. J. Y. (2011). Ventral tegmental area disruption selectively affects CA1/CA2 but not CA3 place fields during a differential reward working memory task. *Hippocampus*, 21(2), 172–184.
- Matsumoto, Masayuki, & Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*, 459(7248), 837–841.
- Matsumoto, Masayuki, & Takada, M. (2013). Distinct Representations of Cognitive and Motivational Signals in Midbrain Dopamine Neurons. *Neuron*, 79(5), 1011–1024.
- Maunsell, J. H. R. (2004). Neuronal representations of cognitive state: reward or attention? *Trends in Cognitive Sciences*, 8(6), 261–265.
- Maw, W. H., & Maw, E. W. (1966). An Attempt to Measure Curiosity in Elementary School Children. *American Educational Research Journal*, 3(2), 147–156.
- McGillivray, S., Murayama, K., & Castel, A. D. (2015). Thirst for knowledge: The effects of curiosity and interest on memory in younger and older adults. *Psychology and Aging*, 30(4), 835–841.
- McNamara, C. G., Tejero-Cantero, Á., Trouche, S., Campo-Urriza, N., & Dupret, D. (2014). Dopaminergic neurons promote hippocampal reactivation and spatial memory persistence. *Nature Neuroscience*, 17(12), 1658–1660.
- Menon, M., Jensen, J., Vitcu, I., Graff-Guerrero, A., Crawley, A., Smith, M. A., & Kapur, S. (2007). Temporal difference modeling of the blood-oxygen level dependent response during aversive conditioning in humans: effects of dopaminergic modulation. *Biological Psychiatry*, 62(7), 765–772.

- Mittman, L. R., & Terrell, G. (1964). An Experimental Study of Curiosity in Children. *Child Development, 35*(3), 851–855.
- Moncada, D., & Viola, H. (2007). Induction of Long-Term Memory by Exposure to Novelty Requires Protein Synthesis: Evidence for a Behavioral Tagging. *Journal of Neuroscience, 27*(28), 7476–7481.
- Moorman, D. E., & Aston-Jones, G. (2014). Orbitofrontal Cortical Neurons Encode Expectation-Driven Initiation of Reward-Seeking. *The Journal of Neuroscience, 34*(31), 10234–10246.
- Moraga-Amaro, R., González, H., Ugalde, V., Donoso-Ramos, J. P., Quintana-Donoso, D., Lara, M., et al. (2016). Dopamine receptor D5 deficiency results in a selective reduction of hippocampal NMDA receptor subunit NR2B expression and impaired memory. *Neuropharmacology, 103*, 222–235.
- Mullaney, K. M., Carpenter, S. K., Grotenhuis, C., & Burianek, S. (2014). Waiting for feedback helps if you want to know the answer: the role of curiosity in the delay-of-feedback benefit. *Memory & Cognition, 42*(8), 1273–1284.
- Murayama, K., & Kuhbandner, C. (2011). Money enhances memory consolidation—But only for boring material. *Cognition, 119*(1), 120–124.
- Murayama, K., Matsumoto, M., Izuma, K., Matsumoto, K., & Smith, E. E. (2010). Neural basis of the undermining effect of monetary reward on intrinsic motivation. *Proceedings of the National Academy of Sciences of the United States of America, 107*(49), 20911–20916.
- Murray, G. K., Corlett, P. R., Clark, L., Pessiglione, M., Blackwell, A. D., Honey, G., et al. (2008). Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Molecular Psychiatry, 13*(3), 239–267–76.
- Murty, V. P., & Adcock, R. A. (2013). Enriched encoding: Reward motivation organizes cortical networks for hippocampal detection of unexpected events. *Cerebral Cortex, bht063*.

- Murty, V. P., Ballard, I. C., & Adcock, R. A. (2016a). Hippocampus and Prefrontal Cortex Predict Distinct Timescales of Activation in the Human Ventral Tegmental Area. *Cerebral Cortex*, bhw005.
- Murty, V. P., LaBar, K. S., & Adcock, R. A. (2012). Threat of Punishment Motivates Memory Encoding via Amygdala, Not Midbrain, Interactions with the Medial Temporal Lobe. *The Journal of Neuroscience*, 32(26), 8969–8976.
- Murty, V. P., LaBar, K. S., & Adcock, R. A. (2016b). Distinct medial temporal networks encode surprise during motivation by reward versus punishment. *Neurobiology of Learning and Memory*.
- Murty, V. P., Shermohammed, M., Smith, D. V., Carter, R. M., Huettel, S. A., & Adcock, R. A. (2014). Resting state networks distinguish human ventral tegmental area from substantia nigra. *NeuroImage*, 100, 580–589.
- Murty, V. P., Stanek, J. K., & Heusser, A. C. (2013). Representations of Distinct Salience Signals in the Nucleus Accumbens. *The Journal of Neuroscience* 33(39), 15319–15320.
- Nair-Roberts, R. G., Chatelain-Badie, S. D., Benson, E., White-Cooper, H., Bolam, J. P., & Ungless, M. A. (2008). Stereological estimates of dopaminergic, GABAergic and glutamatergic neurons in the ventral tegmental area, substantia nigra and retrorubral field in the rat. *Neuroscience*, 152(4), 1024–1031.
- Naus, M. J., Ornstein, P. A., & Kreshtool, K. (1977). Developmental differences in recall and recognition: The relationship between rehearsal and memory as test expectation changes. *Journal of Experimental Child Psychology*, 23(2), 252–265.
- Niv, Y., Duff, M. O., & Dayan, P. (2005). Dopamine, uncertainty and TD learning. *Behavioral and Brain Functions*, 1(1), 1.
- Nobre, A. C., Coull, J. T., Frith, C. D., & Mesulam, M. M. (1999). Orbitofrontal cortex is activated during breaches of expectation in tasks of visual attention. *Nature Neuroscience*, 2(1), 11–12.
- O'Carroll, C. M., Martin, S. J., Sandin, J., Frenguelli, B., & Morris, R. G. M. (2006).

- Dopaminergic modulation of the persistence of one-trial hippocampus-dependent memory. *Learning & Memory*, 13(6), 760–769.
- O'Reilly, J. X., Schüffelgen, U., Cuell, S. F., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (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*, 110(38), E3660–9.
- Otmakhova, N. A., & Lisman, J. E. (1996). D1/D5 dopamine receptor activation increases the magnitude of early long-term potentiation at CA1 hippocampal synapses. *The Journal of Neuroscience*, 16(23), 7478–7486.
- Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, 87(6), 532–552.
- Pedzek, K., Whetstone, T., Reynolds, K., & Askari, N. (1989). Memory for real-world scenes: The role of consistency with schema expectations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 15, 587-595.
- Poppenk, J., Evensmoen, H. R., & Moscovitch, M. (2013). Long-axis specialization of the human hippocampus. *Trends in Cognitive Sciences*, 12(5), 230-240.
- Porubanova, M., Shaw, D. J., McKay, R., & Xygalatas, D. (2014). Memory for Expectation-Violating Concepts: The Effects of Agents and Cultural Familiarity. *PLoS ONE*, 9(4), e90684.
- Preuschoff, K., Bossaerts, P., & Quartz, S. R. (2006). Neural Differentiation of Expected Reward and Risk in Human Subcortical Structures. *Neuron*, 51(3), 381–390.
- Ranganath, C., & Rainer, G. (2003). Neural mechanisms for detecting and remembering novel events. *Nature Reviews Neuroscience*, 4(3), 193–202.
- Rebec, G. V., Christensen, J. R. C., Guerra, C., & Bardo, M. T. (1997a). Regional and temporal differences in real-time dopamine efflux in the nucleus accumbens during free-choice novelty. *Brain Research*, 776(1-2), 61–67.

- Rebec, G. V., Grabner, C. P., Johnson, M., Pierce, R. C., & Bardo, M. T. (1997b). Transient increases in catecholaminergic activity in medial prefrontal cortex and nucleus accumbens shell during novelty. *Neuroscience*, 76(3), 707–714.
- Redgrave, P., & Gurney, K. (2006). The short-latency dopamine signal: a role in discovering novel actions? *Nature Reviews Neuroscience*, 7(12), 967–975.
- Redgrave, P., Prescott, T. J., & Gurney, K. (1999). Is the short-latency dopamine response too short to signal reward error? *Trends in Neurosciences*, 22(4), 146–151.
- Redondo, R. L., & Morris, R. G. M. (2011). Making memories last: the synaptic tagging and capture hypothesis. *Nature Reviews Neuroscience*, 12(17), 1–14.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory*, 2, 64–99.
- Restorff, Von, H. (1933). Über die wirkung von bereichsbildungen im spurenfeld. *Psychological Research*, 18(1), 299–342.
- Richter, F. R., & Yeung, N. (2014). Corresponding influences of top-down control on task switching and long-term memory. *The Quarterly Journal of Experimental Psychology*, 68(6), 1124–1147.
- Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(4), 803–814.
- Roesch, M. R., Esber, G. R., Li, J., Daw, N. D., & Schoenbaum, G. (2012). Surprise! Neural correlates of Pearce–Hall and Rescorla–Wagner coexist within the brain. *European Journal of Neuroscience*, 35(7), 1190–1200.
- Roggenhofer, E., Fidzinski, P., Bartsch, J., Kurz, F., Shor, O., & Behr, J. (2010). Activation of dopamine D1/D5 receptors facilitates the induction of presynaptic long-term potentiation at hippocampal output synapses. *European Journal of Neuroscience*, 32(4), 598–605.

- Roitman, M. F., Stuber, G. D., Phillips, P. E., Wightman, R. M., & Carelli, R. M. (2004). Dopamine operates as a subsecond modulator of food seeking. *The Journal of Neuroscience*, *24*(6), 1265-1271.
- Rojahn, K., & Pettigrew, T. F. (2011). Memory for schema-relevant information: A meta-analytic resolution. *British Journal of Social Psychology*, *31*(2), 81–109.
- Rolls, E. T. (2010). A computational theory of episodic memory formation in the hippocampus. *Behavioural Brain Research*, *215*(2), 180–196.
- Room, P., & Groenewegen, H. J. (1986). Connections of the parahippocampal cortex in the cat. II. Subcortical afferents. *The Journal of Comparative Neurology*, *251*(4), 451–473.
- Rosen, Z. B., Cheung, S., & Siegelbaum, S. A. (2015). Midbrain dopamine neurons bidirectionally regulate CA3-CA1 synaptic drive. *Nature Neuroscience*, *18*(12), 1763–1771.
- Rossion, B., Schiltz, C., & Crommelinck, M. (2003). The functionally defined right occipital and fusiform “face areas” discriminate novel from visually familiar faces. *NeuroImage*, *19*(3), 877–883.
- Sacchet, M. D., & Knutson, B. (2012). Spatial smoothing systematically biases the localization of reward-related brain activity. *NeuroImage*, *66C*, 270–277.
- Sajikumar, S., & Frey, J. U. (2004). Late-associativity, synaptic tagging, and the role of dopamine during LTP and LTD. *Neurobiology of Learning and Memory*, *82*(1), 12–25.
- Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience*, *10*(3), 211–223.
- Sarinopoulos, I., Grupe, D. W., Mackiewicz, K. L., Herrington, J. D., Lor, M., Steege, E. E., & Nitschke, J. B. (2009). Uncertainty during Anticipation Modulates Neural Responses to Aversion in Human Insula and Amygdala. *Cerebral Cortex*, *20*(4), bhp155–940.

- Schott, B. H. (2004). Activation of Midbrain Structures by Associative Novelty and the Formation of Explicit Memory in Humans. *Learning & Memory*, 11(4), 383–387.
- Schott, B. H., Minuzzi, L., Krebs, R. M., Elmenhorst, D., Lang, M., Winz, O. H., et al. (2008). Mesolimbic Functional Magnetic Resonance Imaging Activations during Reward Anticipation Correlate with Reward-Related Ventral Striatal Dopamine Release. *The Journal of Neuroscience*, 28(52), 14311–14319.
- Schultz, W. (2002). Getting Formal with Dopamine and Reward. *Neuron*, 36(2), 241–263.
- Schultz, W. (2016). Dopamine reward prediction-error signalling: a two-component response. *Nature Reviews Neuroscience*, 17(3), 183–195.
- Schultz, W., & Hollerman, J. R. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience*, 1(4), 304–309.
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, 275(5306), 1593–1599.
- Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery & Psychiatry*, 20(1), 11–21.
- Seeman, P. (1980). Brain dopamine receptors. *Pharmacological Reviews*, 32(3), 229–313.
- Shohamy, D., & Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends in Cognitive Sciences*, 14(10), 464–472.
- Shohamy, D., & Wagner, A. D. (2008). Integrating Memories in the Human Brain: Hippocampal-Midbrain Encoding of Overlapping Events. *Neuron*, 60(2), 378–389.
- Shohamy, D., Myers, C. E., Kalanithi, J., & Gluck, M. A. (2008). Basal ganglia and dopamine contributions to probabilistic category learning. *Neuroscience and Biobehavioral Reviews*, 32(2), 219–236.
- Smith, C. C., & Greene, R. W. (2012). CNS Dopamine Transmission Mediated by

- Noradrenergic Innervation. *The Journal of Neuroscience*, 32(18), 6072–6080.
- Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, 17(3), 143-155.
- Somers, L. A., Beyene, M., Carelli, R. M., & Mark Wightman, R. (2009). Synaptic Overflow of Dopamine in the Nucleus Accumbens Arises from Neuronal Activity in the Ventral Tegmental Area. *The Journal of Neuroscience*, 29(6), 1735–1742.
- Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science*, 348(6230), 91–94.
- Staresina, B. P., & Davachi, L. (2009). Mind the Gap: Binding Experiences across Space and Time in the Human Hippocampus. *Neuron*, 63(2), 267–276.
- Stark, C. E., & Squire, L. R. (2001). When zero is not zero: the problem of ambiguous baseline conditions in fMRI. *Proceedings of the National Academy of Sciences*, 98(22), 12760–12766.
- Steinfels, G. F., Heym, J., Strecker, R. E., & Jacobs, B. L. (1983). Behavioral correlates of dopaminergic unit activity in freely moving cats. *Brain Research*, 258(2), 217–228.
- Stern, C. E., Sherman, S. J., Kirchoff, B. A., & Hasselmo, M. E. (2001). Medial temporal and prefrontal contributions to working memory tasks with novel and familiar stimuli. *Hippocampus*, 11(4), 337–346.
- Stuber, G. D., Roitman, M. F., Phillips, P. E. M., Carelli, R. M., & Wightman, R. M. (2005). Rapid Dopamine Signaling in the Nucleus Accumbens during Contingent and Noncontingent Cocaine Administration. *Neuropsychopharmacology*, 30(5), 853–863.
- Summerfield, C., & Egnér, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, 13(9), 403–409.
- Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M. M., & Egnér, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature*

Neuroscience, 11(9), 1004–1006.

Swanson, L. W. (1982). The projections of the ventral tegmental area and adjacent regions: a combined fluorescent retrograde tracer and immunofluorescence study in the rat. *Brain Research Bulletin*, 9(1), 321–353.

Szpunar, K. K., McDermott, K. B., & Roediger, H. L. (2007). Expectation of a final cumulative test enhances long-term retention. *Memory & Cognition*, 35(5), 1007–1013.

Takeuchi, T., Duzskiewicz, A. J., & Morris, R. G. M. (2013). The synaptic plasticity and memory hypothesis: encoding, storage and persistence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1633), 20130288.

Tobler, P. N. (2005). Adaptive Coding of Reward Value by Dopamine Neurons. *Science*, 307(5715), 1642–1645.

Tobler, P. N., O'Doherty, J. P., Dolan, R. J., & Schultz, W. (2006). Reward Value Coding Distinct From Risk Attitude-Related Uncertainty Coding in Human Reward Systems. *Journal of Neurophysiology*, 97(2), 1621–1632.

Total, N. K. B., Kim, Y., & Moghaddam, B. (2013). Distinct prestimulus and poststimulus activation of VTA neurons correlates with stimulus detection. *Journal of Neurophysiology*, 110(1), 75–85.

Tremblay, L., & Schultz, W. (2000). Reward-Related Neuronal Activity During Go-Nogo Task Performance in Primate Orbitofrontal Cortex. *Journal of Neurophysiology*, 83(4), 1864–1876.

Tricomi, E. M., Delgado, M. R., & Fiez, J. A. (2004). Modulation of caudate activity by action contingency. *Neuron*, 41(2), 281–292.

Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: Role of the hippocampus. *Hippocampus*, 8(3), 198–204.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The Automaticity of Visual

- Statistical Learning. *Journal of Experimental Psychology: General*, 134(4), 552–564.
- Tversky, B. (1973). Encoding processes in recognition and recall. *Cognitive Psychology*, 5(3), 275–287.
- Ungless, M. A. (2004). Dopamine: the salient issue. *Trends in Neurosciences*, 27(12), 702–706.
- Vago, D. R., Bevan, A., & Kesner, R. P. (2007). The role of the direct perforant path input to the CA1 subregion of the dorsal hippocampus in memory retention and retrieval. *Hippocampus*, 17(10), 977–987.
- Waelti, P., Dickinson, A., & Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, 412(6842), 43–48.
- Wagner, A. D., Shannon, B. J., Kahn, I., & Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends in Cognitive Sciences*, 9(9), 445–453.
- Wang, S.-H., & Morris, R. G. M. (2009). Hippocampal-Neocortical Interactions in Memory Formation, Consolidation, and Reconsolidation. *Dx.Doi.org*, 61(1), 49–79.
- Weissman, D. H., Giesbrecht, B., Song, A. W., Mangun, G. R., & Woldorff, M. G. (2003). Conflict monitoring in the human anterior cingulate cortex during selective attention to global and local object features. *NeuroImage*, 19(4), 1361–1368.
- Wittmann, B. C., Bunzeck, N., Dolan, R. J., & Duzel, E. (2007). Anticipation of novelty recruits reward system and hippocampus while promoting recollection. *NeuroImage*, 38(1), 194–202.
- Wittmann, B. C., Schott, B. H., Guderian, S., Frey, J. U., Heinze, H.-J., & Duzel, E. (2005). Reward-Related fMRI Activation of Dopaminergic Midbrain Is Associated with Enhanced Hippocampus-Dependent Long-Term Memory Formation. *Neuron*, 45(3), 459–467.
- Wolosin, S. M., Zeithamova, D., & Preston, A. R. (2012). Reward Modulation of

Hippocampal Subfield Activation during Successful Associative Encoding and Retrieval. *Journal of Cognitive Neuroscience*, 24(7), 1532–1547.

Woo, C.-W., Krishnan, A., & Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. *NeuroImage*, 91, 412–419.

Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage*, 14(6), 1370–1386.

Yu, A. J., & Dayan, P. (2005). Uncertainty, Neuromodulation, and Attention. *Neuron*, 46(4), 681–692.

Zaehle, T., Bauch, E. M., Hinrichs, H., Schmitt, F. C., Voges, J., Heinze, H.-J., & Bunzeck, N. (2013). Nucleus Accumbens Activity Dissociates Different Forms of Saliency: Evidence from Human Intracranial Recordings. *The Journal of Neuroscience*, 33(20), 8764–8771.

Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., Chappelow, J. C., & Berns, G. S. (2004). Human striatal responses to monetary reward depend on saliency. *Neuron*, 42(3), 509–517.

Biography

Jessica K. Stanek was born to Michael and Diane Wilson in San Francisco, California on September 29th, 1986. During her undergraduate career at University of California, Berkeley, she studied sleep and memory in the laboratory of Dr. Richard Ivry. She graduated with degrees in Psychology and Molecular and Cellular Biology, earning highest honors for the former, in May 2008. During her graduate career at Duke University, she was awarded a National Science Foundation Graduate Research Fellowship from 2012-2015. She completed her dissertation research on expectation and memory in the dopaminergic system under the mentorship of Dr. R. Alison Adcock, earning her Master of Arts in Psychology and Neuroscience in 2013 and Doctor of Philosophy in Psychology and Neuroscience in 2016. She looks forward to continuing her academic and professional career as a postdoctoral fellow for the National Exposure Research Laboratory in the Environmental Protection Agency.

Publications

- Murty, V. P., **Stanek, J. K.**, Heusser A. (2013). Representations of distinct salience signals in the nucleus accumbens. *Journal of Neuroscience*, 33(39), 15319-15320.
- Addicott, M. A., Pearson, J. M., **Wilson, J.**, Platt M. L., McClernon F. J. (2013). Smoking and the bandit: a preliminary study of smoker and nonsmoker differences in exploratory behavior measured with a multiarmed bandit task. *Experimental and Clinical Psychopharmacology*, 21(1), 66-73.
- Wilson, J.**, Baran. B., Pace-Schott, E., Ivry, R., & Spencer R. M. (2012). Offline processing in aging: Sleep modulates declarative but not procedural learning in healthy older adults. *Neurobiology of Aging*, 33, 991-1000.
- Baran, B., **Wilson, J.**, & Spencer, R. M. (2010). REM-dependent repair of competitive memory suppression. *Experimental Brain Research*, 203(2), 471-477.