MAKING DECISIONS ON THE FLY - UNRAVELLING THE COMPUTATIONAL PRINCIPLES GOVERNING CHOICE BEHAVIOR IN THE *DROSOPHILA MELANOGASTER* BRAIN

by

Adithya E. Rajagopalan

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

September 2023

Abstract

Foraging animals adopt decision-making strategies that successfully adapt to the dynamically changing availability of rewards in the environment, as well as account for the complexity of the sensory world around them. Understanding the neural principles that underlie these complex decision strategies has been an area of interest in neuroscience for a long time. While a lot of progress has been made on these fronts, a complete picture of these neural algorithms has eluded us for two reasons. In studies of larger vertebrates that have used tasks that account for richness of the natural world, the lack of manipulability of these complex brains has restricted our ability to map the underlying neural algorithms. On the other hand, studies in more accessible small brains have been limited to simple Pavlovian learning tasks for the most part. In this dissertation, my colleagues and I develop a novel foraging task for *Drosophila melanogaster*, the details of which are described in chapter 2, and leverage this task to provide insight into the neural algorithms underlying decision-making.

In chapter 3, we deal with dynamic probabilistic environments, where several animals are known to distribute their choices in proportion to the rewards received from available options - Herrnstein's operant matching law. Theoretical work suggests an elegant mechanistic explanation for this ubiquitous behavior, as operant matching follows automatically from simple expectation-based synaptic plasticity rules acting within behaviorally relevant neural circuits. However, no past work has mapped operant matching onto plasticity mechanisms in the brain, leaving the biological relevance of the theory unclear. Here we discovered operant matching in *Drosophila* and using a combination of behavior, computational modeling and optogenetics showed that it requires reward expectation based synaptic plasticity that acts in the mushroom body. Our results reveal the first synapse-level mechanisms of operant matching in the fly brain.

In chapter 4, we provide the first biological test of Marr and Albus' expansion layer theory about sensory discrimination. In particular, the ratio between sensory channels and expansion layer neurons and the number of sensory inputs that individual expansion layer neurons receive are theorized to be key parameters. Leveraging the development of tools that manipulate these parameters in the context of the fly mushroom body, we show that fly behavior agrees with many theoretical predictions. An increase in expansion layer neuron number improves discrimination of odors while an increase in input connectivity causes worsened discrimination. There are however some key differences, suggesting that theoretical models can learn from the experiments and be modified to better explain sensory discrimination.

In chapter 5, we move beyond just sensory discrimination and tackle the question of flexible decision-making that must depend on the available options. We find that *Drosopihla melanogaster* can successfully learn to both distinguish between very similar stimuli and generalize across cues. Rather than forming memories that strike a balance between specificity and generality, we find that flies flexibly categorize a given stimulus into different groups by performing a side-by-side comparison over time of the available options. Together the work in this thesis combines multiple important avenues of neuroscience, providing insight into the neural principles underlying decision-making in dynamic and sensory rich environments.

Advisory committee:

Dr. Glenn C. Turner (Thesis Advisor)

Dr. Christopher J. Potter (Second Reader, Chair)

Dr. Yoshinori Aso

Dr. Jeremiah Y. Cohen

Dr. Ann M. Hermundstad

Dr. Vivek Jayaraman

Dr. Erik L. Snapp

Acknowledgments

Finishing a Ph.D. and writing a thesis is not unique or special in its own right. Thousands of people across the globe do this every year. What makes a Ph.D. special and unique, particularly one that was hampered by a global pandemic, is the people that carry you over the finish line. I'd therefore like to highlight all of the people who have helped me over the finish line of this important chapter of my life and career. Buckle up. This is going to be long.

First, my mentor Dr. Glenn Turner. Thank you for all of your help, support and guidance throughout. From when I reached out to you out of the blue to go over a project proposal, I had written so that I could apply to the Janelia grad program, to your thoughtful advice about future career options and post-doc labs, you have provided me with exactly the mentorship I needed. Working with you has refined not only my scientific thinking, but my writing and communication and my football skills (I can only be hope be playing as well as you when/if I become a PI)! So many pieces of advice you have given me over the course of this journey have stayed with me. I am writing this soon after getting back form a Neuroethology conference and so the one that is most prominent in my mind is your advice to not temper my science to appease "cortical chauvinists" that dominate the decision-making field, but rather to do the best science I can. I think that advice speaks to the kind of science you do and it has been my pleasure to work with and train under such an amazing scientist. Thank you and I hope we continue to collaborate in the future.

On that note, I'd like to thank my many collaborators.

Dr. James Fitzgerald, thank you for your amazing mentorship surrounding our operant matching project. It has been a privilege to work with you on this and I have learnt so much about the kind of computational neuroscience I like to and would like to continue working on from our collaboration. Adopting your schedule management system and use of calendar has been *the* tool that has gotten me to the end of this Ph.D. in one piece and cannot thank you enough. I look forward to what comes out of your lab in the future and wish you the best of luck on your move to Northwestern and I hope to continue collaborating with you (I'm already thinking about ways to do this for my post-doc project!).

Dr. Ran Darshan, thank you for introducing me to the ideas of operant matching and the work from Loewenstein and Seung. These ideas and so many others that you have exposed me to have been so central to my Ph.D. and are already shaping the kind of science I hope to continue pursuing in my career. I wish you the best of luck with you own new lab at Tel Aviv University and I look forward to a continued collaboration.

Dr. Maria Ahmed, thank you for introducing me to the world of developmental neuroscience and how it can and should intersect with the world of system neuroscience. When I was working on my master project of modeling the effect of PN-KC connectivity on discrimination the only question that came up in my defense that I couldn't answer was about how I would test my results experimentally. When you came to Glenn and me with the developmental manipulations you and the Clowney lab had been working on to actually change PN-KC connectivity in vivo it truly felt like black magic to me. I hope to stay in touch and look forward to your post-doc work and here's to hoping you never again have to do an experiment that involves shipping chemically ablated flies 1000s of miles for someone else to run behavior experiments.

Dr. Josie Clowney, thank you so much the developmental + genetic circuit hacking you and your lab have begun working on. These tools are going to be so incredibly important for neuroscience going forward and I look forward to what's going to come out next and hopefully I can continue to collaborate and discuss science with you and your lab (I'm already excited to see what Yijie is going to be working on, the project sounds incredibly cool).

Dr. Mehrab Modi, thank you so much for welcoming me with such openness to the Turner lab. From when I reached out to you when I was still in Baltimore to discuss me potentially joining the lab, you have been an important source of insight and advice. Thank you for letting me work on the incredible project about flexible decision-making that you had going. It helped me get started in the lab and begin thinking about the ideas I wanted to explore in my Ph.D. Working with you has taught me how to balance workload during a project as well as how to write a paper. Most importantly thank you for advising me to not get into a postdoc with my eyes closed and look into the options available for me before making a decision. This has had and will continue to have a tremendous impact on my career choices.

Dr. Bilal Bari, thank you for being an inspiration of a grad student and for getting me started in thinking about foraging and decision-making. I am continuously amazed by your productivity and look forward to borrowing many more inspirational ideas from your future work.

To present and past members of the Turner and Aso labs - Yichun Shuai, Mehrab Modi, Kevin Chen, Martin Brill, Anu Bhukel, Karen Hibbard, Rishika Mohanta, He Yang and Jasper Simon plus our extended lab meeting/outing group of Dan Bushey, Lisha Shao and Thomas Ravenscroft, thank you for being a wonderful, hospitable, kind and intellectually stimulating group. My science and I have grown leaps and bounds with your help. In particular, thank you Dr. Karen Hibbard for all the help with/teaching about flies, fly genetics and all your amazing work with the tissue specific

CRISPR. The operant matching project would have been incomplete without you and your work and insights. Also, to Rishika Mohanta, thank you so much for joining the lab and sharing in my excitement about flies doing choice behavior. You are a powerhouse and are going to do great doing our Ph.D. Thank you for all the modeling discussion and more than anything for the work on putting the high-throughput Y-arena together. I look forward to seeing what you are going to get up to during your Ph.D. See you in NYC.

To Yash Mehta, Jan Funke, Danil Tyulmankov, Chuntao Dan, Kevin Miller, Maria Eckstein and Tzuhsuan Ma, our collaborations may have been brief or just beginning but the ideas we have discussed have opened up interesting directions in which I hope to take my science and i looking forward to our continued collaboration.

Next, I'd like to thank the members of my thesis committee and other mentors. Thank you all so much for all your insightful and helpful comments during our many meetings. You have shaped the projects in this dissertation.

Dr. Jeremiah Cohen, thank you for introducing me to the wonderful world of value-based decision-making. The questions that inspired this thesis may have existed in some nascent state from early in my undergrad, but working with you and Bilal really fleshed them out for me and has significantly shaped the direction of my scientific interests. In another life I can see myself doing a Ph.D. in your lab and enjoying it tremendously.

Dr. Vivek Jayaraman, in addition to your inspirational work on the central complex that has played a significant role in guiding how i think a scientific project should be structure, thank you for showing me how to lead an institution well while still having time to sit in on all my thesis committee meeting, meetings with the DeepMind team amongst so many others. I remember one

conversation where you pulled out a random fact about another Janelian that I would have just assumed someone who has to interact with as many people as you do on the daily would have forgotten (sorry for the vagueness did not want to mention any specifics here), that has stayed with me. Thank you also for the presentation opportunities you have put me up for.

Dr. Ann Hermundstad, thank you for the conversations both early on in my project where I was trying to understand how best to model this behavior and later on with Maz and Rishika that helped me feel like the work I had started was meaningful and going to be pursued by others both in Glenn's lab and outside. Also, for your amazing talks, they are an inspiration and even made me finally decide to cave to Janelia pressure and switch to black slides (though that took ages to actually implement!)

Dr. Chris Potter, thank you for being the second person to read this thesis after Glenn and for your support both in the committee meetings and as director of the Neuroscience program. Your prompt responses and help made it easy for me to identify and pursue an internship outside of academia which eventually helped me decide to stay and do a postdoc.

Dr. Erik Snapp, thank you for being the leading supporter of all things Ph.D. student related at Janelia. I have learnt so much about a life in science and science communication through the various mentorship groups, journal clubs and events that you have organized.

Dr. Yoshi Aso, thank you for being the first person to introduce me to Janelia science and the world of the insect MB all the way back in 2015. You were an inspirational mentor then and to this day will be the person I turn to with any questions about the fly olfactory system and learning and memory. Your encyclopedic knowledge has always astounded me, and I aspire to be able to know

half of what you know. Thank you for all the ideas you have so kindly shared both in committee meetings and in our joint lab meetings. I can't wait to see what you do with the bumblebees!

Dr. Shreesh Mysore, I learnt a lot during my rotation in your lab and the ideas about stimulus selection that I was introduced to there have stayed with me. Thank you for being a great mentor and again I can see another life in which I did my Ph.D. in the lab and thoroughly enjoyed it.

Rita Ragan, Beth Wood-Roig and Dr. Hita Adwanikar, thank you for being such amazing organizers of the Hopkins neuroscience training program. Without you getting through the requirements for the Ph.D. would not have been as seamless as they were and I am sure the whole experience would have been so much more frustrating.

Dr. Amy Braun, thank you for leading the internship initiative at Professional Development and Careers Office (PDCO) at Hopkins and for all your help identifying, applying for and getting started on a non-academic internship during my Ph.D. This experience that was instrumental in helping me decide my future career path would not have been even close to as easy as it was without you. I will spread your praises whenever I can and hope that many students reach out to you for future internship opportunities.

Next, I'd like to thank my friends and family. The brief words below cannot express just how thankful I am to all of you. I can't spell out all of your contributions to my life and science quite as tangibly as I have with my colleagues and mentors above (who I also consider friends and who this statement applies to as well!) but know that you have been incredibly important parts of this journey.

To my core board game group - Drs. Amrita Singh, Arthur Zhao, Judith Hoeller and Chirag Gupta, thank you for helping get me through the isolation of the second half of the pandemic. You are all great friends as well as great scientists and I look forward to many more game nights online and whenever we meet in the future! Thank you also for all the dinners and hikes and for getting the “Gentle soccer” group started with me, I really needed football back in my life by the end of 2021 and it would not have happened without you all.

So on to the gentle soccer group, though I know it got not so gentle for a while there. Dr. Emmanuel Marquez Legorreta, Dr. Pratik Kumar, Dr. Miguel Nunez, Dr. Angel Stanoev, Dr. Daniel Bushey, Saul, Oscar, Dr. Zhengchang Lei, Dr. Davis Bennett, Gregory Hernandez, Dr. Glenn Turner, Dr. Amrita Singh, Dr. Arthur Zhao, Dr. Judith Hoeller, Dr. Chirag Gupta and so many others, thank you for being part of the thing I looked forward to most during my week. Football (I refuse to call it soccer except when it relates to the name of our whatsapp group) has always been my stress buster and you all made it possible and extremely fun. I’m dreading having to find a new football group once I move.

I lived in Arlington for a large part of the Ph.D. and so the Janelia shuttle and amazing folks who took it with me were another big part of my time at Janelia. Drs. Emmanuel Marquez Legorreta, Emily Behrman, Christina Gladkova, Mark Eddison, Mehrab Modi, Katie Schreiter, John Bogovic and so many others including our wonderful drivers thank you so much for making my commute feel both product and fun. Thanks for the after shuttle drinks, I wish we had had time for more of those. Emily, Christina, Mehrab and Emmanuel, I really enjoyed our kayaking trip and the various comedy shows!

Once I did move on to campus the time I spent on the shuttle was replaced in part by playing more Badminton. So thanks to the whole group including Dr. Pratik Kumar, Dr. Laura Grima, Dr.

Davis Bennett, Dr. Jinyang Liu, Dr. Andrian Gutu and Dr. Caiying Guo for all of the fun I had playing and for improving my game. Maybe when I visit my cousins in Chennai next they won't beat me as bad as they did the last time we played.

Thanks also to my group of Alamo Drafthouse/ movie friends, Dr. Sanna Koskela, Dr. Laura Grima, Tom Hennigan, Ben Foster, Sarah Plutkis, Michalis Michaelos for helping me use my alamo pass wisely and for all the sage commentary and not-so-sage reddit post sharing following the movies we watched. And also for introducing me to so truly weird movies that were supposedly for kids though I can't image they would be rated that way if they were made today.

So many amazing people at Janelia, not listed as part of any of the groups above made my time so special and I hope that despite leaving academia keeps us all in touch! Dr. Gabriela Michel, Dr. Brian Lustig, Dr. Hannah Haberkern, and soon to be Drs., Shivam Chitnis, Atika Syeda, thank you all so much for both the scientific discussions and the fun random conversation at lunches, dinners, at conferences, on the football field or on hikes!

This Ph.D. would have also not been as enjoyable without the grad student community I was a part of. First, my Hopkins Cohort – Drs. or soon-to-be Drs. Abel Corver, Lim Jing Xuan, Emily Han, Gabrielle Ewall, Matt Brown, Anthony Ramnauth, Lionel Rodriguez, Emma Spikol, Linghua Zhang, Calvin Kersbergen and Jared Hinkle, thank you for making that first year in the program and my first year living in a country other than India so easy. I miss us all meeting together in Neurocog class but am glad that in the last year of my Ph.D. I have had more opportunities to come to Baltimore and meet up with you all. Next, the community of grad students at Janelia, we may be a small group but you all are awesome! Thanks Drs or soon to be Drs. Lim Jing Xuan, Misha Proskurin, Virginia Rutten, Alex Chen, Shivam Chitnis, Atika Syeda, Farah Du, Maanasa Natrajan and other for being such supportive people and for all of our lunch journal clubs.

In addition to the people at Janelia and Hopkins, so many friends and family both in the US and in India have provided me with a strong support system. To my large family disributed all over the US, Suchi athai, Suku chitthappa, Girija perimma, Nalini perimma, Rajiv, Chammu chitthappa, Viji mama, Geetha athai and Raju, Vinod and so many others thank you for all the times I've just dropped by out of the blue to stay at your places and for also being there to help even when I did my usual thing of not contacting you all for months before suddently asking for favors.

To my amazing friends from Chennai and Vidya Mandir, Krishna, Shrinath, Madhu, Avinaash, Pathu, Vishal, Veena. Thank you for always just being the steady support and the same friendly presence in my life whenever I call or message or come back to India. Knowing I can turn to you guys after long breaks in communication and we can pick things off where we left off as if no time passed is so important for my mental wellbeing. Krishna, thank you for such an amazing long friendship and with regards to the Ph.D. thank you for all the music. The amount of music you have introduced me to and how much I rely on it to get through difficult periods of projects cannot be competed with. Shrinath, thank you for all the flamingo feathers I have turned into fly pushing equipment and you have turned into bookmarks :P. So glad that you all will always be in my life

To my amazing family and extended family in Chennai and elsewhere, Regina, Sandeep mama, Gayathri mami, Nishanth, Drishya, Kavita, Meghana, Sashi aunty, Shyam uncle, and so so so many more people that I don't want to name anymore because I will feel bad if I try to make and exhaustive list and forget people, you have all helped me become the person I am today and I am so so thankful to have you all in my life.

Dr. (yes I am being premature but its going to happen soon) Shruthi Ravindranath. What can I say? Don't forget me when you become a hot shot PI. Thank you for sharing a career in science

with me from VM to IISER to Janelia as undergrads and to our Ph.Ds and onwards, it isn't going to stop. Thank you for all the perspective you've given me about science but also about life and myself. I have grown as a person so much from having you as a part of my life. Thank you for all the trips and holidays and adventures and books and music and movies we've shared. I really just don't have enough words to list everything you've given me. Thank you for being the biggest piece of my support system for so long and thanks in advance for all the things we will continue to do together!

Last and most definitely not the least, my parents. Thank you both so much for raising me to be who I am. I would not be here without you and your support. You made life easy for me and I know I always have a place to go back to in case anything doesn't work out and that makes taking risks in my life and science possible.Thank you Appa for all the Richard Dawkins and Stephen Jay Gould books you suggested I read during high school that really pushed me along my chosen scientific path. Thank you for all the conversations about medicine, science, statistics, data interpretation and more recently ML that have inspired a lot of my own thinking. Thank you also for cleaning my apartment when you visit :P. Amma, thank you for finding out about IISER for me when I was thinking about what I would do after school. The choice to go there has been so so important to my life and career and I cannot imagine where I would be if you hadn't shared that opportunity with me. Also for those of you who have made it all the way and don't know, my mom decided around the time I was starting grad school that in addition to being a doctor, she would out-publish me. I know your decision was purely work related, but I can't help but think there was a little bit of "I should do this so that I can help Adi through things if he runs into publishing issues" going on there. It has been so nice to see your work getting recognized in this way as well as being able to talk about the publication process with you. It will likely take my career to catch up to your publication list now but as always you and Appa both set inspiring goals for me to meet. Thank you!

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

Chapter 2. A novel olfactory dynamic foraging task for *Drosophila*

Chapter 3. Covariance-based learning rules that incorporate reward expectation direct learning and underlie operant matching in the *Drosophila* Mushroom body

Chapter 5. Flexible discrimination or generalization between similar odors in *Drosophila* depends on the presented alternative

List of Abbreviations

OFC	orbitofrontal cortex	GABA	gamma-aminobutyric acid
PFC	prefrontal cortex	RNAi	RNA interference
LIP	lateral intraparietal cortex	*mud*	MB defect
MB	mushroom body	*tao*	thousand and one kinase
KC	Kenyon cell	HU	hydroxyurea
PN	projection neuron	ROI	region of interest
MBON	MB output neuron	PI	preference index
DAN	dopaminergic neuron	SEM	standard error of the mean
OCT	3-octanol	TH	tyrosine hydroxylase
MCH	4-methyl-3-cyclo-hexanol	IR	infrared
PA	pentyl acetate	DDM	drift diffusion model
BA	butyl acetate	RPE	reward prediction error
EL	ethyl lactate	2AFC	two alternative forced choice
VS	ventral striatum	DA	dopamine

UAS	Upstream activation sequence
DTH	*Drosophila* tyrosine hydroxylase
PPL1	posterior inferiorlateral protocerebrum
PAM	protocerebral anterior medial
DScam	Down Syndrome cell adhesion molecule
CRISPR	clustered regularly interspaced short palindromic repeats

Chapter 1. Introduction

I remember first thinking about the core question that has guided this dissertation during my undergraduate studies. Sitting in a classroom late at night waiting for the quiz club's trivia night to get started, I wondered to myself why I chose to attend this event rather than the music event that was happening across campus at the same time. Indeed, such questions regarding how and why humans, and other animals, choose one out of multiple options available to them has been of great interest to philosophers and scientists from as far back as the days of Aristotle and Plato (Ober 2022).

More recently, these questions have fallen within the realm of study of neuroscientists, ecologists and ethologists (Hayden and Walton 2014; Hayden 2018; Stephens and Krebs 1986; El Hady, Davidson, and Gordon 2019; Juavinett, Erlich, and Churchland 2018; Doya and Shadlen 2012). Through the use of the modern scientific approaches utilized in these fields, we have begun to understand a great deal about the behavioral strategies used whilst making decisions and the computations that are being performed in our brains. The burgeoning field of neuroeconomics, in particular, has contributed significantly to how we think about the decision-making process (Glimcher and Fehr 2013). Researchers have identified neural signatures that correspond to economically important concepts such as value - the worth of an option (Platt and Glimcher 1999; Schoenbaum et al. 2003; Sugrue, Corrado, and Newsome 2004a; Padoa-Schioppa and Assad 2006; Lau and Glimcher 2008; Bari et al. 2019; Hayden and Niv 2021) - amongst others, and have constructed models that use these representations to determine choice (Loewenstein and Seung 2006; Soltani and Wang 2006; Song, Yang, and Wang 2017; Rustichini and Padoa-Schioppa 2015). Together the results from this field propose a framework in which decisions are

made by combining subjective value, risk etc. together for a given option and comparing these between options to determine a winner.

Despite these detailed advances in understanding elements of the neural algorithms underlying decision-making, such as the value representations mentioned above, a complete mapping of the decision process to neural circuitry has not yet been arrived at in a single brain. A large reason for this is the complexity of the mammalian brains that have classically been used to study this process. Value signals for example have been found in many cortical and subcortical regions making it an incredibly painstaking task to understand how these signals are updated and used to eventually guide choices (Platt and Glimcher 1999; Schoenbaum et al. 2003; Sugrue, Corrado, and Newsome 2004a; Padoa-Schioppa and Assad 2006; Lau and Glimcher 2008; Bari et al. 2019; Hayden and Niv 2021). Second, the tasks that have classically been used to study decision-making are often too simple (Krakauer et al. 2017; Juavinett, Erlich, and Churchland 2018; Dennis et al. 2021). As a result, the solutions we arrive at are under-constrained; multiple neural algorithms could serve as equally good explanations of the decision-making process in these tasks. Important aspects of the natural environment such as sensory overlap, reward uncertainty (where cues that predict reward may only do so probabilistically and not everytime a cue is responded to) and dynamic changes in cue-reward relationships must be added into the repertoire of tasks being used to study behavior to help us identify the actual computations and principles being used in our brains.

In the work detailed in this dissertation, my colleagues and I use the “fruit fly”, *Drosophila melanogaster*, with its small, well-characterized brain as a means to bypass the roadblocks to the understanding of decision-making algorithms put in place by simplistic task structures and the complexity and lack of manipulability of traditionally studied brains. We develop a task for fruit flies (henceforth referred to as fly/flies) that allows us to study decision-making in a manner that

is similar to mammalian tasks, while also incorporating complex features of natural environments such as sensory overlap and dynamic uncertain cue-reward associations. We use this task in conjunction with the well-mapped and genetically tractable neural architecture of the fly brain to understand the neural computations and principles that guide decision-making in flies. In this first chapter, I will provide an introductory overview of neuroscientists' thinking on how brains make decisions, link these ideas to what is known about the *Drosophila* neural circuitry and motivate the work described in the later chapters.

1.1 Decision-making: A bridge linking economics and neuroscience

From as early as the 19th century, economists including William Stanley Jevons and Vilfredo Pareto proposed utility, or subjective value, of an option as the most important factor used by a rational individual to make a decision (Jevons 1879; Blaug 1997; Pareto 2014).

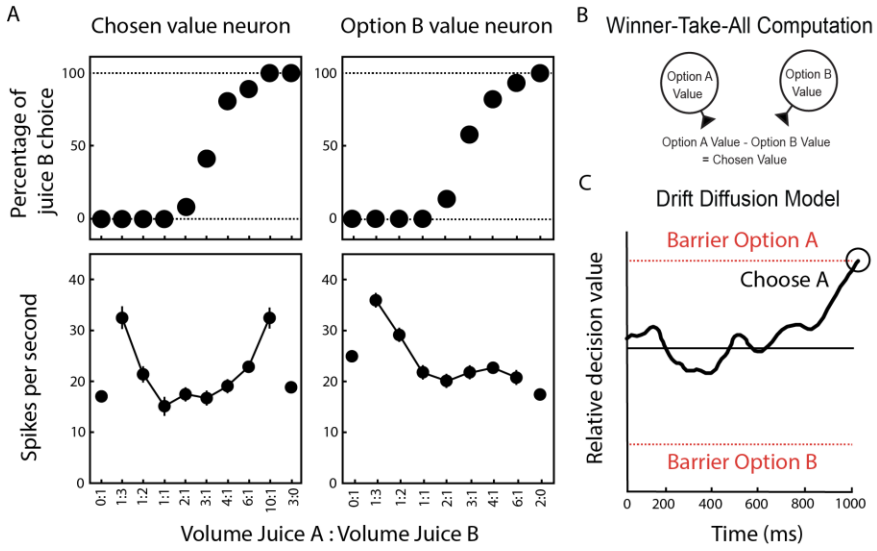

Figure 1.1 Neural representations of value and their role in decision-making. (A) An illustration of choice behavior and correlated neural activity in a two-option task adapted from Padoa-Schioppa et al. 2006. *Top* Percentage of choices made towards option B as a function of ratio of volume of each option shows a monkey's preference for the two options, with an inflection point suggesting that 3 units of juice A is equally valuable as 1 unit of juice B. *Bottom* Neural activity recording from the OFC during the behavior in *Top* shows two neurons, *left* responds as a function of the chosen option, *right* responds as a function of the value/amount of juice B. (B) A

schematic representing a simple winner-take-all computation in the context of the neural responses shown in (A). (C) A schematic of a generic drift diffusion model adapted from Krajbich et al. 2010, showing how a relative decision value can be estimated over time when the sensory information available to make a decision is noisy.

Value in their minds was a measure of the satisfaction that the consumer obtained from a given good or option. Accordingly, this framework implied that a rational individual should make decisions so as to maximize value (Aleskerov, Bouyssou, and Monjardet 2007). This utility theory was then paired with the work of Samuelson who proposed that these highly subjective and intrinsic measures (values or utility functions) assigned by an individual to different options could be indicated via the readily observable choices made by that individual (Samuelson 1938, 1948). This “Revealed Preference” theory of choice allowed observers to estimate future choices of individuals based on their revealed values and allowed economists to make global predictions about choice behavior in a variety of settings.

Such a behavioral approach to economic choices allowed neuroeconomists, beginning in the 1990s, to treat decisions made by animals in the same way (Glimcher and Fehr 2013). In early experiments, researchers working with monkeys had their subjects choose between two options presented on a computer screen that were rewarded with different amounts or types of juice reward (Platt and Glimcher 1999; Padoa-Schioppa and Assad 2006). They found that the behavioral estimates of value from choices correlated with neural responses in both the lateral intra-parietal (LIP) (Platt and Glimcher 1999) and orbitofrontal (OFC) (Padoa-Schioppa and Assad 2006) cortices, with the neural signatures indicating a variety of value-related signatures including chosen value and relative value (Fig. 1.1A). Inspired by these early results, experiments measuring human brain activity using fMRI identified signatures of values in regions including the OFC, medial prefrontal cortex (mPFC) and ventral striatum (VS), in tasks where the subject had to indicate how much they were willing to pay rather than choose between options (Plassmann, O’Doherty, and Rangel 2007; Grabenhorst et al. 2010). Similar work in a go/no-go task in rats has

implicated a role for the basolateral amygdala in the representation of value, particularly in the OFC (Schoenbaum et al. 2003).

This ubiquity of value-related neural signatures across brain regions and species prompted several theorists to propose explanations for how these signals could be used to produce a decision when choosing between options. The most prominent set of proposed computations assume that the value comparison and action production step is independent from the value representation process, and is calculated in different downstream brain regions (Glimcher 2011a; Padoa-Schioppa and Conen 2017). This set includes ideas that range from simple winner-take-all models, where competing output neurons representing each option's values were compared and the higher one was chosen (Loewenstein and Seung 2006; Soltani and Wang 2006), to drift diffusions models (DDMs) where option-value information is accumulated over time till a choice threshold for either option is crossed (Gold and Shadlen 2007; Ratcliff and McKoon 2008) (Fig. 1.1B-C). This second flavor of models, while performing the same winner-take-all comparison at its heart, better accounts for factors such as sensory noise and shifting attention(Krajbich, Armel, and Rangel 2010), and accurately captures reaction times (Milosavljevic et al. 2010). Based on these computations, biophysically realistic neural network models have been developed that accurately capture value representation and choices (Soltani and Wang 2006; Rustichini and Padoa-Schioppa 2015). While this idea of action and value being independent is predominant, another set of models suggest that action production and value representations are intertwined concepts, with value related neural signals directly guiding decisions and actions. This hypothesis was initially discounted due to a paucity of experimental evidence but has recently gained more support (Rushworth et al. 2012; Hunt and Hayden 2017; Hayden and Niv 2021). It is clear from these results and several others not cited here that the importance given to value in economic decisions has a basis in biological reality, linking together economics and neuroscience in an important and incredibly fruitful way.

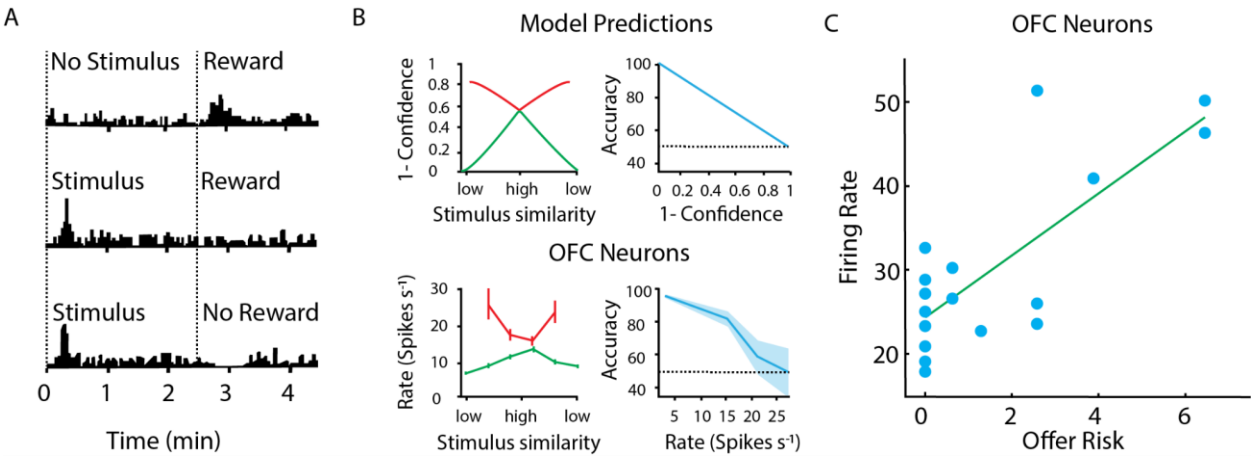

Figure 1.2 Neural representation of other economic variables (A) An illustration of reward prediction error signals observed in dopaminergic neurons present in the mammalian midbrain adapted from Schultz et al 1997. (B) Data illustrating that the spike rate of neurons in the OFC of rats *(bottom)* accurately reflect confidence of choice as predicted by a computation model of confidence *(top),* adapted from Kepecs et al 2008. (C). Data adapted from Raghuraman and Padoa-Schioppa 2014, indicating that firing rate of neurons in the OFC covary with the amount of risk associated with a given offer.

1.2 Beyond simple value: The representation of other economic variables in the brain

Importantly though, a simple value isn't the only economic variable that has been shown to be critical for decision-making, nor is it the only one that is represented in the brain. Soon after identifying its usefulness, economists recognized a limitation to their theory of value. It does not appropriately deal with uncertain or probabilistic outcomes. To account for this von Neumann and Morgenstern turned to the mathematical idea of expected value, i.e. the mean value of an option (von Neumann and Morgenstern 2007). If such an expectation were calculated over several interactions with an option, an average value for that option could then be arrived at that accounts for both the presence and absence of a rewarding outcome. These expected values could then be compared to aid in making decisions.

This economic idea of expected value has been widely adopted into neuroscience and is the most prominent way in which the brain has been shown to deal with probabilities. This is done implicitly

via the concept of reward prediction errors (RPE) that brings together the original concept of value with its expectation (Schultz, Dayan, and Montague 1997; Schultz 2004; Glimcher 2011b; Sutton and Barto 2018). First inspired by computational models of value learning (Bush and Mosteller 1951a, [b] 1951; Rescorla and Wagner 1972; Sutton and Barto 2018), RPEs estimate the value of options over time by comparing the currently experienced value/outcome of a decision with the chosen option's expected value. In seminal work, Wolfram Schultz and colleagues observed that the activity of midbrain dopaminergic neurons in the monkey represented such RPE signals (Schultz, Dayan, and Montague 1997) (Fig. 1.2A). This finding has since been replicated in a variety of species (Bayer and Glimcher 2005; Cohen et al. 2012; Hart et al. 2014) and has inspired a significant body of work aimed at understanding how RPE is calculated (Keiflin and Janak 2015) and utilized (Schultz 1997; Glimcher 2011b).

RPE is however not the only way in which probabilities are represented in the brain. One more direct representation of probability that has been identified both in the context of uncertain rewards as well as uncertain stimulus categorization is confidence (Pouget, Drugowitsch, and Kepecs 2016). Confidence is defined as the probability that the choice made is correct. If one considers a simple choice between two options that have the same expected value, where one provides a reliable small reward, while the other provides a reward that's twice as large but only half as often, the confidence in one's choice is simply the probability that the second option provides a reward (here, 0.5). In more difficult tasks, confidence can take more complicated forms. A large body of work has shown that the OFC, PFC as well as visual regions such as the pulvinar nuclei of the thalamus have neurons whose activity correlates with the subjects confidence in their choices (Kepecs et al. 2008; De Martino et al. 2013; Komura et al. 2013), thus more directly representing a probability value in neural activity (Fig. 1.2B).

A third concept closely related to uncertainty in economics is that of risk and risk-appetite. Consider the same example with one safe and one risky alternative introduced above while defining confidence. The reason confidence is half in this case is related to the fact that comparing expected rewards to make a decision in this case is not useful as they are both the same. Animals that compare expected values to make decisions in this case should choose both options equally. However it has been observed in such scenarios that monkeys and humans do show biases (Kahneman and Tversky 1979; Stauffer et al. 2015), preferring the risky alternative when rewards are small on average (So and Stuphorn 2010; Raghuraman and Padoa-Schioppa 2014) and the safe alternative when rewards are large (H. Yamada et al. 2013; Stauffer, Lak, and Schultz 2014). In each case these behavior observations were supported by neural observations that showed that neural activity in regions including the Supplementary Eye Field, OFC and midbrain dopaminergic neurons reflect the risk-appetite of the animal (Fig. 1.2C).

1.3 An incomplete picture of the decision-making process

This rich history of neuroeconomics, primarily involving the identification of decision-variables, provides tremendous explanatory power when predicting decisions and behavioral variability. However, our knowledge falls short when attempting to put together a complete picture of the architecture of (and computations performed by) the neural circuitry underlying decision-making. This is particularly problematic when moving downstream of the decision-variable (eg. value) representations to understand the computations that govern their use (eg. value updating over time, decision-making). Existing models of decision-making that incorporate biophysical realities and appropriate cortical architectures to capture decision-variable representations make simplifying assumptions about the how these are used to eventually make the decision (Rustichini and Padoa-Schioppa 2015), while models that do make novel predictions about the computations

that give rise to behavioral outputs are hard to test (Loewenstein and Seung 2006; Soltani and Wang 2006)

There appear to be two primary reasons for this. First, a large number of hypothetical frameworks can accurately describe behavior in the kinds of simple decision-making tasks that have primarily been used in the literature cited earlier on in this introduction (for example read (Pereira-Obilinovic et al. 2022)). This is because the tasks that have been primarily used to study decision-making under-constrain the space of potential solutions (Krakauer et al. 2017). Animal brains and their underlying computational algorithms have evolved to deal with more complex environments, therefore more naturalistic and complex tasks that can distinguish between alternate computational hypotheses are essential for further understanding of the system. This approach has gained significant traction lately (Gomez-Marin et al. 2014; Krakauer et al. 2017; Juavinett, Erlich, and Churchland 2018; Dennis et al. 2021) and two approaches to introduce naturalistic complexity into decision-making tasks will be the primary focus of sections 1.4 and 1.5 of this introduction. Second, as referenced earlier in this introduction, a large variety of brain regions have been implicated in the production of choice behavior. Understanding how all these different players interact is a herculean task, particularly in large mammalian brains. Understanding how the different elements such as value, confidence, expected value etc come together in a simpler more genetically tractable brain might provide greater insight into how larger brains solve the problem. In section 1.6 of this introduction I will provide an overview of a particular brain-circuit my colleagues and I feel is well suited to answering questions about the neural underpinnings of decision-making, the Mushroom Body (MB) of the fruit fly *Drosophila melanogaster.*

1.4 Making decisions in dynamic environments.

One approach for complex task-design that has produced significant insight into decision-making computations involves a dynamic foraging paradigm (Sugrue, Corrado, and Newsome 2004a). This paradigm expands on a simple two-option choice task in two ways. The first is to incorporate probabilistic rewards associated with the two options. This has already been discussed earlier in this introduction and doesn't provide novel avenues of narrowing down the neural computations compared to the previously cited literature. The second change however is significant, the probabilities of reward associated with options change over time. By observing how animals adapt their behavior in the face of such a dynamic environment, we can distinguish between previously indistinguishable behavioral strategies and their neuronal basis. For example, a simple strategy that animals could use when faced with repeated choices between options is to stay with the same option if it was rewarding on the immediately previous trial or switch to another option if the previous choice was unrewarded. Such a strategy has the advantages of being computationally simple but is not as optimal at accumulating reward as say Herrenstein's operant matching law, where choices between options are divided in proportion to the rewards received (income) (Fig. 1.3A). These two strategies (amongst others) show significantly different choice dynamics when the reward probabilities in the world change and require a dynamic foraging paradigm to be distinguished between.

Sugrue and colleagues, using a version of this task with visual cues, found that rhesus macaques follow Herrenstein's operant matching law when faced with this task (Sugrue, Corrado, and Newsome 2004a). This is a strategy that equalizes the return on investment across options (Fig. 1.3A). Richard Herrnstein first formulated this long-standing empirical law in 1961 based on observations made on pigeons in a simpler and non-probabilistic context (R. J. Herrnstein 1961; Richard J. Herrnstein 1997). Sugrue and colleagues simultaneously found neurons in the LIP

whose activity tracked the local fractional income, suggesting that these value-like representations could be guiding behavior in agreement with the previously noted economic theories.

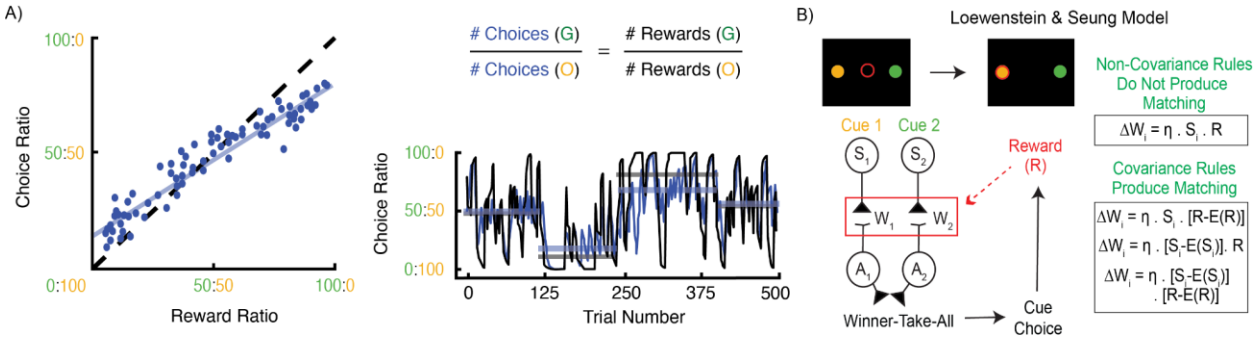

Figure 1.3 Operant matching in dynamic foraging tasks and its proposed theoretical underpinnings. (A) Data adapted from Sugrue et al. 2004, showing that the choice behavior of monkeys in a dynamic foraging task follow Herrnstein's operant matching law. The law states that choices between options are divided in proportion to the rewards received (income) and is defined on the top right. Choice ratio from blocks of monkey behavior where reward probabilities are held constant are plotted against reward rations from these blocks on the left. Both choice and reward ratios calculated overtime using a moving window average are plotted on the bottom right (B) An illustration summarizing the theoretical work of Loewenstein and Seung 2006, which shows mathematically that any network that stores and learns values of options available for choice as synaptic weights must update these weights using covariance-based rules to produce operant matching.

However, the most important advance this approach contributed, was that it suggested that these value representations need to be updated over time and updated in a specific way that allows animals to follow the operant matching strategy rather than the economists rational value maximization strategy that had been observed in tasks with stationary environments (Padoa-Schioppa and Assad 2006). Following this original finding multiple groups replicated this observation of operant matching, using similar tasks in both monkeys and mice (Lau and Glimcher 2005, 2008; Tsutsui et al. 2016; Bari et al. 2019; Iigaya et al. 2019). This body of work made progress in identifying the local decision-making rules that could give rise to this global phenomenon of the matching-law. Based on the observed behavior a simple win-stay; lose-switch

strategy was ruled out in favor of a strategy in which an option's value was updated in time by integrating over multiple past rewards and choices (Lau and Glimcher 2005; Bari et al. 2019; Beron et al. 2022). This was paired with the identification of neural signatures in a variety of brain regions including the PFC and the striatum, that tracked this change in value over time, either through putative changes in the strength of synapses onto the neurons of interest (Sugrue, Corrado, and Newsome 2004a; Tsutsui et al. 2016), or through work-memory-like persistent changes in neural activity (Bari et al. 2019).

Together, this body of literature allowed theorists to propose specific hypotheses regarding the way in which value representation must change over time to allow for matching. One particularly elegant hypothesis put forward by Loewenstein and Seung (Loewenstein and Seung 2006), makes a very general claim regarding the learning rules that must be used to update values over time if values are stored as synaptic weights. They prove that operant matching is the inevitable outcome of plasticity rules that modify synaptic weights according to the covariation of neural activity signaling reward and sensory input (Fig. 1.3B). Mathematically, covariance is the product of two variables with at least one being subtracted by its expectation. Covariance in this form not only nicely aligns with the economic idea of expected value and the neuroscientific idea of RPE, it expands this to say that the expectation needs to be incorporated in a very specific mathematical form to give rise to matching. Importantly, only an animal that follows operant matching would receive rewards at a rate equal to the reward expectation for both options, which leads weights to stabilize. Loewenstein & Seung mathematically formalized this intuitive link between expectation and matching and showed that when synaptic plasticity is the basis for operant matching, only a covariance-based plasticity rule can account for matching.

It is here unfortunately that we reach the limits of working with large mammalian brains. While Loewenstein and Seungs theory is specific, it is difficult to test experimentally in a mammal brain.

There are too many potential locations where such a plasticity rule could be implemented and the specificity of the tools necessary to manipulate the nature of plasticity rules remain too broad. As a result, there has been no mapping of these rules onto particular synapses or plasticity mechanisms in any animal and a deep investigation of this hypothesis by manipulating and testing the nature of plasticity rules underlying operant matching has been intractable. In Chapter 3 of this thesis, I will describe the approach we have taken to test Loewenstein and Seung's hypothesis in the MB of the fruit fly, using a combination of a novel dynamic foraging task for flies (described in Chapter 2), computational modeling approaches and optogenetic manipulation.

1.5 Making decisions when faced with overlapping or nearly indistinguishable options.

A second approach for adding complexity to decision-making tasks to better understand the underlying principles borrows from the rich field of sensory representation and categorization. Assigning values to options is not always as simple as in the previously described tasks where the alternatives provided to the subjects are clearly distinct. Distinct sensory stimuli can be very similar to one another leading values to be misassigned. As a result, neural circuits have evolved to appropriately discriminate between inputs, aiming to group multiple noisy experiences of the same stimulus together, while separating distinct yet similar stimuli. The neural computations that must be performed to correctly link options to rewards and value are upstream of the kinds of learning rules we discussed in the immediately preceding section. However, they still play an important role in the decision-making process and in the construction of value representations. A complete picture of the neural circuitry underlying decisions would be incomplete without an understanding of option discrimination.

One seminal theory developed by Marr and Albus' in the context of cerebellar-like networks suggests that such neural circuits are able to separate and uniquely represent different inputs

with the help of an “expansion layer” (Marr 1969; Albus 1971). An ‘‘expansion layer’’ is a large set of postsynaptic cells that contain a sparse representation of stimuli based upon combinatorial information received from multiple different sensory input channels. Theoretical models of expansion layer circuits (that include the insect mushroom-body (Modi, Shuai, and Turner 2020)) have found that two important architectural parameters affect the circuit’s discrimination capabilities: 1) the number of cells in the expansion layer (Babadi and Sompolinsky 2014; Cayco-Gajic and Silver 2019), and 2) the density of connection between the input and expansion layers (Jortner, Farivar, and Laurent 2007; Litwin-Kumar et al. 2017; Rajagopalan and Assisi 2020).

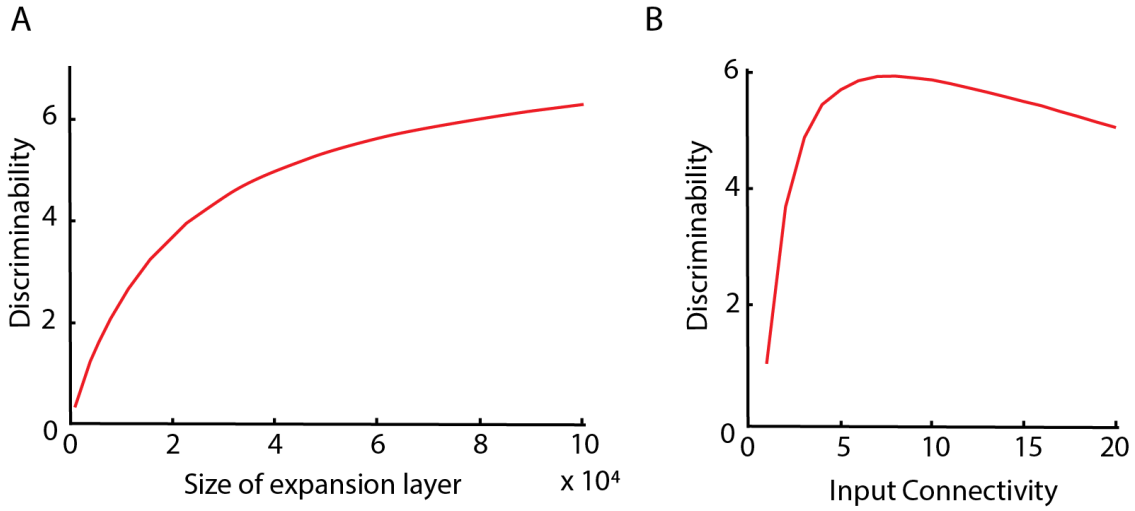

Figure 1.4 Theoretically identified relationships between circuit architecture and discrimination in “expansion layer” networks. (A) A plot adapted from Babadi and Sompolinsky 2014, shows that the ability of a “expansion layer” network to discriminate between inputs increases as a function of expansion layer size. (B) A plot adapted from Litwin-Kumar et al. 2017, shows how the ability of a “expansion layer” network to discriminate between inputs varies as a function of input connectivity, with optimal discrimination obtained at an intermediate value of connectivity.

As the number of cells in the expansion layer increase, the ability to separate similar inputs increases suggesting that more expansion layer neurons will allow circuits to more easily separate inputs and assign values appropriately to them (Babadi and Sompolinsky 2014; Cayco-Gajic and Silver 2019) (Fig. 1.4A). Energy and animal-size constraints likely limit the number of neurons

found in any given expansion layer circuit. The density of connection between the two layers on the other-hand seems to have arrived at an optimal value in both the cerebellum and the MB (Litwin-Kumar et al. 2017; Jortner 2013). Both increases and decreases in connectivity tend to make these circuits worse at discrimination (Fig. 1.4B). However here too, as was the case with Loewenstein and Seung's theory, testing these hypotheses in the cerebellum has been impossible because of a lack of tools to manipulate cell number or connectivity. Ground-breaking work from our collaborators in Dr. E. Josephine Clowney's lab at the University of Michigan has produced developmental tools in the fruit fly that allow for such circuit manipulations. In Chapter 4 we combine this work with our novel behavioral assay described in Chapter 2 to test these theories of cell number and connectivity.

While discrimination is an important step for appropriate value assignment, in many cases similar options do have similar values and so storing and updating values for every potential option separately would be computationally taxing (Seger and Miller 2010). Instead, it is beneficial for the brain to maintain slightly overlapping sensory representations, which would allow options to be grouped downstream into categories based on sensory similarity and assigned with common values (Seger 2008). Such a coding-scheme would allow animals to distinguish between options in different categories and choose the highest value stimulus, taking into account the values of available alternatives (Glimcher and Fehr 2013; Padoa-Schioppa and Conen 2017; Hayden 2018; Kudryavitskaya et al. 2021). However, this scheme would not readily allow for distinguishing between two options from the same category. It is therefore essential in such a coding-scheme that category boundaries be flexible. One prominent hypothesis suggests that animals perform a side-by-side comparison of the value representation activated by the available stimuli and make use of a relative value signal to guide flexible categorization (Itti and Koch 2001; Carello and Krauzlis 2004; Mysore and Knudsen 2011; Mysore, Asadollahi, and Knudsen 2011). In Chapter

5, in collaboration with Dr. Mehrab Modi and colleagues, we ask how this flexible categorization arises and identify neural correlates of stimulus comparison in the *Drosophila* mushroom body.

1.6 The form and function of the *Drosophila* Mushroom Body

The mushroom body is a key center for learning and memory in the fly brain that serves as a junction where sensory representations are assigned value and converted into behavioral outputs (Martin Heisenberg 2003; Davis 2005; Modi, Shuai, and Turner 2020; Adel and Griffith 2021). In this final section of the introduction, I will describe the structure and function of the MB circuit.

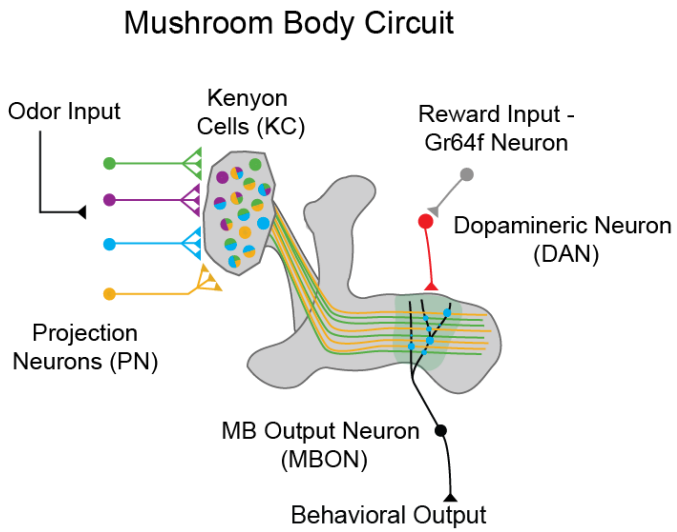

Figure 1.5 A schematic of the Drosophila MB circuit. Inputs to the primary cells of the MB (KCs) arrive via PNs. KC axons form compartments in the MB lobes. Each compartment is innervated by compartment specific MBONs and DANs. A subset of rewarding DANs receive input from reward sensory neurons including Gr64f sugar sensory neurons.

Structurally, the MB consists of primary cells called Kenyon cells (KCs) whose cell bodies are located in a structure called the calyx that forms the head of the mushroom-shaped structure (Fig. 1.5). These are third odor neurons along the olfactory sensory pathway (Caron et al. 2013; Wilson 2014) and represent odors in a sparse spatiotemporal pattern of activity (Turner, Bazhenov, and Laurent 2008; Honegger, Campbell, and Turner 2011; Robert A. A. Campbell et al. 2013). KCs

form the “expansion layer” in the MB circuit and receive their olfactory inputs from second order neurons called projection neurons (PNs) (Caron et al. 2013). KC axons project into a set of horizontal and vertical lobes that form the stalk of the mushroom and are made up of 15 distinct subdomains called compartments. The mushroom body output neurons (MBONs) that receive input from the KCs typically innervate only one of these compartments (Aso, Hattori, et al. 2014; Li et al. 2020). The activity of these MBONs serve as a signal that guides action selection (Aso, Sitaraman, et al. 2014; Owald et al. 2015). This conversion from the sensory representation of the KC to the action guiding activity of MBONs is mediated by the KC-MBON synapses that are thought to store the value of a given olfactory input (Villar et al. 2022). These synapses are plastic and modified by a rule that depends on the coincident activity of odor-representing KCs and release of dopamine by reward-signaling DANs (Séjourné et al. 2011; Hige et al. 2015; Cohn, Morantte, and Ruta 2015; Owald et al. 2015; Aso and Rubin 2016; Berry, Phan, and Davis 2018; Handler et al. 2019; Jiang and Litwin-Kumar 2021). Recent theoretical work has also attempted to link the features of learning rules identified with the help of these behavioral experiments to the plasticity mechanisms identified using imaging and electrophysiology (Jiang and Litwin-Kumar 2021; Springer and Nawrot 2021; Bennett, Philippides, and Nowotny 2021; Adel and Griffith 2021; Gkanias et al. 2022).

On the functional front, over the last half century, researchers have shown that flies can learn a wide variety of Pavlovian and operant associations between cues and rewards in an MB dependent manner(Quinn, Harris, and Benzer 1974; Tully and Quinn 1985; de Belle and Heisenberg 1994; M. Heisenberg et al. 1985; Brembs and Heisenberg 2000; Rohrsen et al. 2021). They exhibit features common in vertebrate memory, such as forgetting, re-evaluation, second-order condition and consolidation, amongst others (Quinn, Harris, and Benzer 1974; Tully and Quinn 1985; Aso, Sitaraman, et al. 2014; Ichinose et al. 2015; Aso and Rubin 2016; Sayin et al. 2019; König et al. 2019; Sabandal, Berry, and Davis 2021; D. Yamada et al. 2023). With the help

of advances in functional and anatomical tools(Barret D. Pfeiffer et al. 2010; Jenett et al. 2012; Klapoetke et al. 2014; Riabinina et al. 2015; Zheng et al. 2018; Li et al. 2020), specific features of learning have been linked to specific parts of the MB circuit. For example, learning can take place on multiple time-scales depending on the MB compartment in which synaptic plasticity takes place (Shuai et al. 2011; Aso and Rubin 2016).

Despite this progress, there remains several unanswered questions regarding the extent of the decision-making capabilities of flies. There is a dearth of understanding about how flies learn in natural environments (but see (Seidenbecher et al. 2020)). For example, animals in naturalistic scenarios have to be able to form associations between multiple different options and rewards, yet evidence in flies suggests that some associations are labile and easily overwritten(Aso and Rubin 2016). Second, choice behavior has rarely been investigated at the individual fly level(Claridge-Chang et al. 2009; Seidenbecher et al. 2020; Honegger et al. 2020; Lesar et al. 2021), and never in the context of flies making repeated choices between two probabilistically rewarding options. It is therefore unclear whether flies can learn associations between options and probabilistic rewards. It is also unknown whether they can integrate probabilistic reward events over multiple past experiences to form analog expectations. Even if such analog expectations can be formed, it is unclear if they lead to matching behavior through covariance-based plasticity in the fly brain. Further, evidence has been mixed as to whether this learning rule makes use of value or reward expectation(Dylla et al. 2017; Riemensperger et al. 2005; Felsenberg et al. 2017, 2018; Eschbach et al. 2020).

One factor that has contributed to this lack of knowledge is that the tasks predominantly used to study learning in flies focus on groups of 10s of flies making single choices. This is related to the ideas of more complex tasks mentioned in section 1.4 and 1.5. The adoption of the “dynamic foraging task” in mammalian research for example allowed for the deeper understanding of the

neural principles underlying learning and decision-making. We therefore designed a novel olfactory “dynamic foraging” two-alternative forced choice (2AFC) task for individual *Drosophila* inspired by earlier behavior assays for flies(Simonnet, Berthelot-Grosjean, and Grosjean 2014; Buchanan, Kain, and de Bivort 2015; Lewis et al. 2017; Mohandasan et al. 2021; Wiggin et al. 2021) and foraging related 2AFC tasks in vertebrates(Sugrue, Corrado, and Newsome 2004b; Lau and Glimcher 2008; Tsutsui et al. 2016; Bari et al. 2019). The assay allows us to measure hundreds of sequential choices from individual flies as we vary the probability of reward associated with different odor cues.

1.7 Combining a novel decision-making task, genetic tools and the *Drosophila* MB circuit to elucidate the neural principles underlying decision-making.

We propose that the fruit fly and its mushroom body circuitry, in combination with our novel foraging inspired decision-making task can serve as a versatile testing ground for the hypotheses described in sections 1.3 to 1.5. The MB circuit contains architectural and functional elements necessary to implement the hypothesized computations and the *Drosophila melanogaster* model system has the genetic tools necessary to specifically manipulate circuit structure and function with unparalleled specificity (Barret D. Pfeiffer et al. 2010; Jenett et al. 2012; Klapoetke et al. 2014; Riabinina et al. 2015; Zheng et al. 2018; Li et al. 2020). Over the course of the next four chapters I will describe the efforts taken to test these important hypotheses and map elements of the neural algorithms of decision-making onto biological realities in the MB circuitry.

Chapter 2. A novel olfactory dynamic foraging task for *Drosophila*

Adapted from Rajagopalan et al. 2022, bioRxiv 2022.05.24.493252

The one missing piece needed to address the challenges and test the hypotheses detailed in Chapter 1 is a behavioral task that allows us to assess decisions of individual flies in a more complex and natural context. For this purpose, we designed an olfactory 2AFC Y-arena (Fig. 2.1A). The Y-arena plays an important role in Chapters 3 and 4 and the materials and methods are common for all uses of this arena. Therefore, the experimental procedures related to the design, testing and function of this arena are detailed first in this chapter.

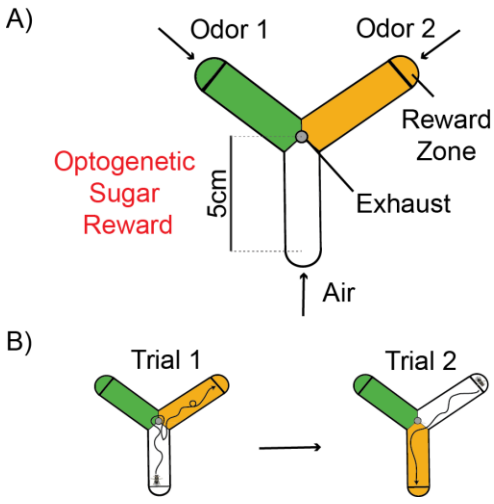

Figure 2.1 A schematic of the novel olfactory Y-arena. (A) Air flows from tips of each arm to an outlet in the center. Reward zones are demarcated by lines at the tips of the arms. A choice is registered when a fly crosses into the reward zone of an odorized arm, triggering Gr64f sugar sensory neuron or PAM DAN optogenetic activation with a 500 ms pulse of red light. (B) An example of the trial structure of this task. Once the fly completes one trial, the next trial commences with the chosen arm switching to air. The two odors (green/orange) are randomly reassigned to the other two arms.

In the Y-arena, a single fly begins a trial in an arm filled with clean air and can choose between two odor cues that are randomly assigned to the other two arms (Fig. 2.1A). The fly can freely move between arms, with a choice defined as the fly crossing into the reward zone at the end of the arm (Fig. 2.1A). Once a choice is made, we provide reward by optogenetically activating either sugar-sensing neurons using a Gr64f driver (Jiao et al. 2008; Haberkern et al. 2019) or the reward-related protocerebral anterior medial dopaminergic neurons (PAM-DANs). The Y-arena then resets, with the arm chosen by the fly filled with clean air and the other two arms randomly filled

with the two odors (Fig. 2.1B). This task design permits us to precisely control reward delivery without satiating the fly and enables us to monitor the choices of a single fly over hundreds of trials.

A task such as this allows us to distinguish between different foraging strategies, such as the matching law versus a simpler win-stay, lose-switch strategy. Importantly, our assay breaks the hard dichotomy between Pavlovian and operant conditioning. Unlike purely Pavlovian tasks(Aso and Rubin 2016; Quinn, Harris, and Benzer 1974), flies in our task do not passively experience olfactory cues and rewards. Rather the choices made by the fly dictate the odors and rewards experienced, a hallmark of operant learning tasks. However, unlike purely operant tasks, where animals learn that specific actions lead to rewards or punishment(Brembs and Heisenberg 2000; Rohrsen et al. 2021), flies in our task have to learn to perform stimulus-dependent actions. This relationship between stimulus, action and reward is very similar to the dynamic foraging tasks where operant matching has been observed in other species(Sugrue, Corrado, and Newsome 2004b; Iigaya et al. 2019; Bari et al. 2019; Tsutsui et al. 2016). The dynamic foraging task structure thereby allows us to readily translate past theoretical work into the context of the *Drosophila* brain to seek a mechanistic understanding of decision-making behaviors that could apply across animals**.**

2.1 Apparatus design

A detailed description of the material composition of the apparatus is provided in Figure 2.2. The Y chamber consists of two layers of white opaque plastic. The bottom is a single continuous circular layer and serves as the floor of the Y that flies navigate. The top is a circular layer with a Y shaped hole in the middle. The length of each arm from center to tip is 5 cm and the width of

each arm is 1 cm. These two layers are placed underneath an annulus of black aluminum. A transparent glass disk is located in the center of this annulus and acts as the ceiling of the Y - allowing for video recording of experiments. This transparent disk is rotatable and contains a small hole used to load flies. The black annulus houses three clamps that hold the circular disk in place. All three layers are held together and made airtight with the help of 12 screws that connect the layers.

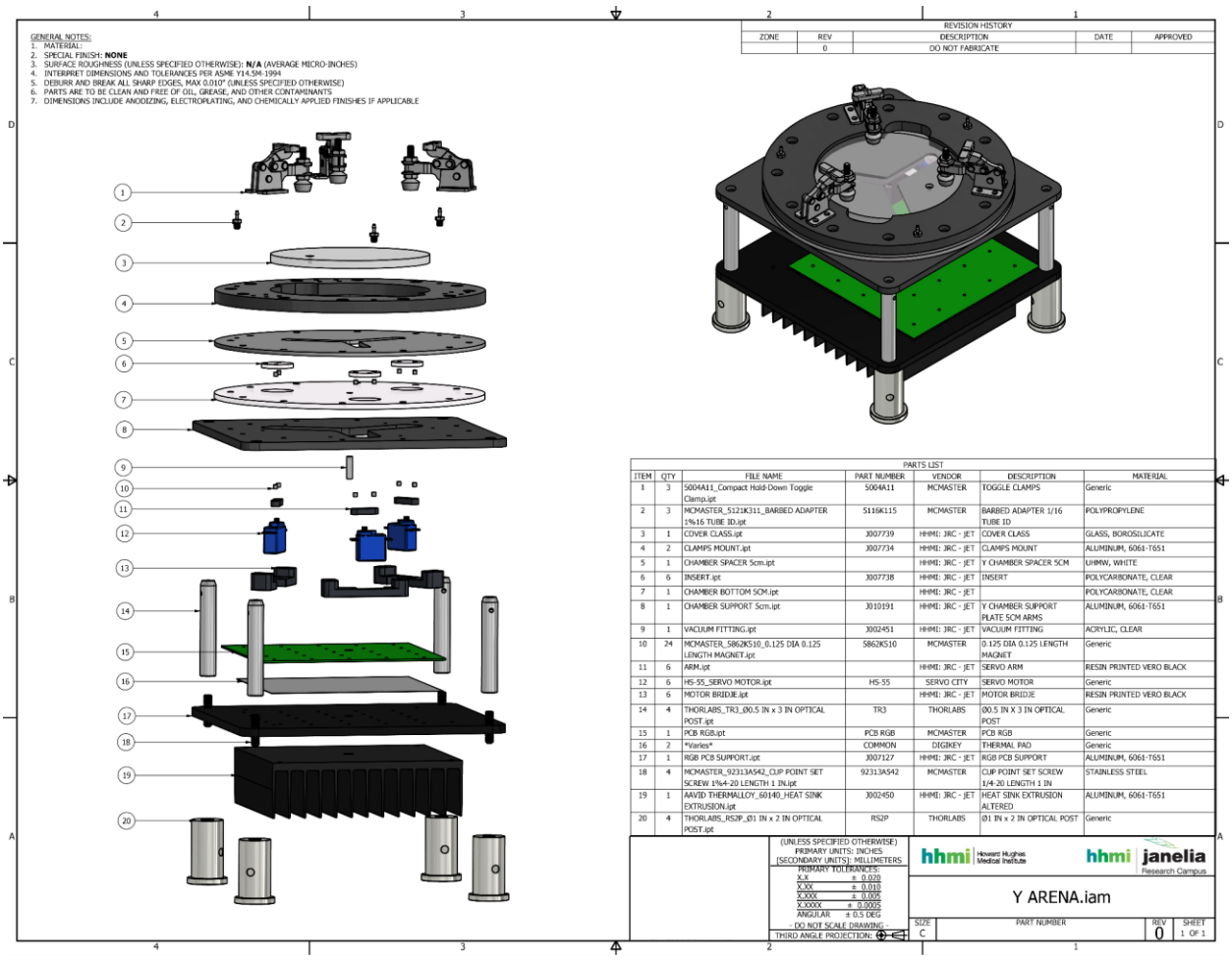

Figure 2.2 A detailed structural drawing of the novel olfactory Y-arena. This structural drawing of the Y-arena lists each of the parts used to build the arena including the part number and vendor from whom each part was purchased for ease of replication.

The Y chamber is mounted above an LED board that provides infrared illumination to monitor the fly's movements, and red (617 nm) light for optogenetic activation. The LED board consists of a

square array of red (617 nm peak emission, Red-Orange LUXEON Rebel LED, 122 lm at 700mA, 1.9mW/cm^2) and infrared (IR) LEDs that shine through an acrylic diffuser to illuminate flies. Experiments were recorded at ~5Hz from above the Y using a single USB3 camera (Flea3, model: FL3-U3-13E4M-C: 1.3 MP, 60 FPS, e2v EV76C560, Mono; Teledyne FLIR, with longpass filter of 800 nm).

Figure 2.3 A measurement of odor boundaries in the Y-arena. Moist blue litmus paper was placed in the Y-arena while the arm at the bottom was filled with carbon dioxide. This caused the color of the litmus paper to change, providing an estimate of how odor boundaries are formed at the center of the Y.

Each arm of the Y has a corresponding odor delivery system, capable of delivering up to 5 odors. For our experiments, olfactometers injected air/odor streams into each arm at a flow rate of 100 ml/min. A crisp boundary between odors and air is formed at the center of the Y. This was measured by passing CO_2 through one of the two arms and taping damp blue litmus paper on the glass disk that serves as the top of the arena to observe the associated color change (Fig. 2.3). Odors and concentrations used for each experiment are detailed in the behavioral task structure and design section of the methods. The center of the Y contains an exhaust port connected to a vacuum, which was set at 300 ml/min using a flow meter (Dwyer, Series VF Visi-Float® acrylic flowmeter) - matching total input flow in our experiments.

2.2 Fly tracking and operation

We wrote custom MATLAB code (MATLAB 2018b, Mathworks) to control the Y-arena and run experiments. The data collected by the USB3 camera was loaded into MATLAB in real time and the fly's location was identified using the MATLAB image processing toolbox as follows. A background image was calculated just before beginning the experiment by averaging multiple frames as the fly moved around in the Y. This background was subtracted from the frame being processed and the resulting image was thresholded, leaving the fly as a white shape on a black background. The location of the centroid of the fly was estimated using the MATLAB's *bwconncomp* and *regionprops* functions and then assigned to 1 of 6 regions in the Y. If the fly was located in one of the reward zones, a trial was deemed complete, and rewards were provided by switching on the red LEDs as defined by the reward contingencies of the task. The arena was then reset with air being pumped into the chosen arm and odors randomly reassigned to the two other arms (Fig. 2.1B). The location of the fly along with other information, such as reward presence and odor-arm assignments, were saved as a .mat file for further analysis. All subsequent analysis was based on this saved information.

2.3 Behavioral protocol structure and design

A variety of protocols were used over the course of this project. They involved different odors, different ways of associating probability to cues etc. Each protocol is provided with a name and detailed below. These named protocols will be referenced in later chapters. For all protocols, two or three of the following odorants were used to form cue-reward relationships:

1. 3-Octanol (OCT) [Sigma-Aldrich 218405] diluted in paraffin oil [Sigma-Aldrich 18512] at a 1:500 concentration and then air-diluted to a fourth of this concentration in air.

2. 4-Methyl-cyclo-hexanol (MCH) [Sigma-Aldrich 153095] diluted in paraffin oil at a 1:500 concentration and then air-diluted to a fourth of this concentration in air.

3. Pentyl Acetate (PA) [Sigma-Aldrich 109584] diluted in paraffin oil at a 1:5000 concentration and then air-diluted to a fourth of this concentration in air.

The simplest protocol used was the 100:0 protocol. Here flies were inserted randomly into one of the three arms. This arm was injected with a clean airstream and OCT and MCH were randomly assigned to the other two arms. For a given fly, one of OCT or MCH was paired with reward on 100% of the trials it was chosen and the other a 0% chance (hence 100:0). Once a fly reached the choice zone of either odor arm a choice was considered to have been made. If the rewarded odor was chosen, the fly was rewarded with a 500ms flash of red LED (617 nm, 1.9mW/cm^2) to activate the appropriate reward-related neurons. The arena was then reset with the arm chosen by the fly injected with clean air and OCT and MCH randomly assigned to the other two arms. This was repeated for many trials. In different versions of this protocol different amounts of trials were used. In version A flies were allowed to choose between OCT and MCH for 30 minutes when neither option was rewarding and then for 30 minutes when one of the options was consistently rewarding. In version B flies made 60 naive choices where neither option was rewarding and 60 training choices where one option was consistently rewarded. In version C flies made 60 rewarded choices where one option was consistently rewarded, and in version D flies made 120 rewarded choices where one option was consistently rewarded.

In a probabilistic version of this protocol (named - 80:0, 40:0 or 80:20 protocols), rather than one of the options being consistently rewarded and the other not, on every trial none, one or both options were rewarded probabilistically. At the beginning of each trial, the code used to run the

experiment determined whether either or both odors would provide a reward upon being chosen in that trial. Whenever probabilistic rewards were included in our tasks, reward baiting was incorporated as follows. If an option was unbaited (as it always is on the first trial of an experiment), the baiting probability assigned to the odor (e.g.80% for the rewarded odor in an 80:0 protocol) was comparing a uniform random number in the range [0,100] to determine if that option was baited on a given trial. If an option was baited on the previous trial (would have been rewarded on that trial) but was not chosen by the fly, the probability that that option produces a reward when chosen increases to 1 (a 100% certain reward) until the next trial on which that alternative was chosen. This means that the likelihood an odor cue holds a reward increases over time if it is unchosen for many trials. Importantly, the fly performing the task is never informed as to whether the unchosen option would have been rewarding on any given trial. The sum of baiting probabilities for both options for this version of the task was either 1 or 0.5 for the 80:20 protocol, if the sum of baiting probabilities of 0.5 then 80:20 could alternatively be called 40:10 thought we never refer to it this way.

In a third version of this task aimed to test if flies can store multiple simultaneous cue-reward relationships, two different odors were rewarded - one with high probability (p=0.8) and one with low probability (p=0.4). For this task we used three different odors, OCT, MCH and PA. On each trial, flies were presented with a choice of either OCT versus PA or MCH versus PA. These choices were delivered in alternation for a total of 80 trials to assess naive odor preference. We then assessed choices during an 80-trial reward pairing block in alternation, where one of OCT or MCH was paired with high probability of baiting reward (0.8) and the other with low probability of baiting reward (0.4), always keeping PA unrewarded. To give an example: a fly experiences alternating OCT vs. PA and MCH vs. PA choices to assess naive choice probability and then receives the same set of choices, but with OCT baited with 80% probability and MCH baited with

40% probability. For another fly the high/low baiting probability odor is swapped to balance out the dataset. A diagram describing this protocol is shown in Fig. 2.4.

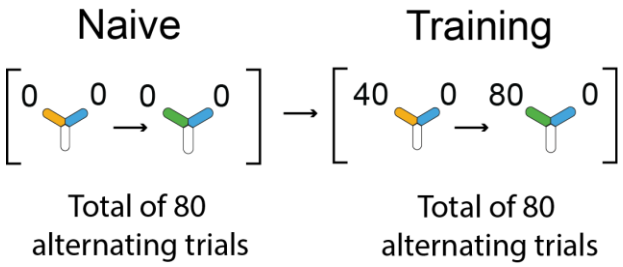

Figure 2.4 A description of the protocol used to show that flies can store multiple associations simultaneously. Two different odor choices are alternated throughout an unrewarded naive block and a reward block of 80 trials each where options were rewarded with baiting probability of 0.4 or 0.8.

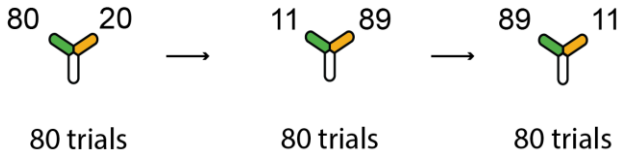

Figure 2.5 An example of the dynamic foraging protocol. Schematics indicate the reward baiting probabilities for each odor in the three 80-trial blocks.

Finally dynamic foraging protocol was adapted from monkey and mouse versions (Sugrue, Corrado, and Newsome 2004b; Lau and Glimcher 2008; Tsutsui et al. 2016; Bari et al. 2019) and incorporated the baiting and probabilistic reward features described earlier. OCT and MCH were the two odors used in this task. An example protocol of this task with details about the three-block structure and number of trials per block can be seen in Fig. 2.5. This example protocol is representative of all of the protocols experienced by flies performing this particular task except that the exact probabilities assigned to odors changed between individual examples. In this task flies experienced 3 blocks of 80 trials each. Within a block the baiting probabilities associated with the two odors were fixed (apart from baiting related changes) and changed across blocks. The relative reward ratios between the two odors for a given block were drawn from the following set [1:1, 1:2, 1:4, 1:8]. Similar to the probabilistic reward task detailed earlier, the code used to run the experiment determined at the beginning of a trial whether either or both of the odors would provide a reward upon being chosen. Here again the sum of baiting probabilities for both options for this task was either 1 or 0.5.

2.4 Quantitative analysis of fly movement

All analysis and modeling was performed using MATLAB 2020b **(**Mathworks). We used nonparametric statistical tests when quantifying statistical significance in our data in all cases as we could not definitively say that all data were from normal distributions. The Mann-Whitney test was used when testing hypotheses related to unpaired samples. The Wilcoxon signed-rank test was used when testing hypotheses related to paired samples. Specific descriptions of the hypotheses being tested are provided in the test describing the relevant results and figure legend in each case.

The *(x,y)* coordinates of the fly were analyzed to calculate: i) the distance of the fly from the center of the Y; ii) when the fly entered and exited a given odor arm; and iii) the time taken per trial to enter into the reward zone at the end of an odorized arm. These quantities were then used to produce the plots.

Distance from center was calculated by projecting the (*x,y*) location of the fly (p_t) onto a skeleton of the Y and this metric was used when plotting location over time plots. Here the subscript t denotes the time point at which the (*x,y*) location was observed. The skeleton consisted of three lines running down the middle of each arm to the center of the Y (v_0). Based on which arm the fly was located in, its (*x,y*) position was projected onto the appropriate (i^{th}) skeleton line using the following equations for projecting a point onto a line,

$$D_i = \frac{||b \cdot a||}{||a||} \ , \tag{1}$$

where $b = p_t - v_0$, $a = v_i - v_0$, and v_i is the (x,y) coordinates of the end of the i^{th} arm. The entries/exits of a fly into/from a particular odorant or air were estimated by tracking the region that the fly was located in at every time point and comparing it to the known odor-arm identity map (stored in the experiment .mat file). A turn (reversal) was considered to have been made whenever a fly entered an odor and then exited this odor without reaching the reward zone. An approach was considered to have been made whenever a fly entered an odor arm and then traveled all the way into the reward zone of that same arm without ever exiting it.

To calculate the time taken per trial, we made use of the timestamp vector that we saved along with the (*x,y*) vector. Time taken from the entire trial was calculated by subtracting the timestamp for the frame that the previous trial was completed by the timestamp of the frame when the current trial was completed. Time taken from first exit of the air arm was calculated by subtracting the timestamp of the frame that the fly first exited the air arm after a trial began by the timestamp of the frame when the current trial was completed.

Choices themselves were determined by identifying the arm in which the fly crossed into the reward zone and mapping that arm to its assigned odor on that trial. Once choices were determined we could calculate two important metrics. Choice ratio, defined as the ratio between the number of choices made towards option A to the number of choices made towards option B, and reward ratio, defined as the ratio between the number of rewards received upon choosing option A to the number of rewards received upon choosing option B. These ratios were calculated on one of two time-scales, i) the ratio over an entire block of 80 trials where baiting probabilities were constant, or ii) the ratio in a ten-trial moving window over the entire 240 trials of the experiment. The undermatching index used in some cases is defined as the mean square error between the instantaneous choice ratio and reward ratio curves produced for each fly.

2.5 Conclusion

Prior to the apparatus and decision-making paradigms described in this chapter, most existing associative learning tasks for flies relied on studying the place preference behavior of groups of flies, such as the commonly used T-maze (Tully and Quinn 1985) or circular arena (Aso, Sitaraman, et al. 2014). Such tasks limit the ability of the experimenter to provide reward contingent on the choices of any one fly, making it hard to study behavior in response to probabilistic reward or measure choice distributions over time (but see (Claridge-Chang et al. 2009; Buchanan, Kain, and de Bivort 2015; Seidenbecher et al. 2020; Wiggin et al. 2021)). In the work described in this dissertation we develop a novel olfactory decision-making task for individual *Drosophila melanogaster*, inspired by tasks in vertebrates (Sugrue, Corrado, and Newsome 2004a; Tsutsui et al. 2016; Bari et al. 2019; Iigaya et al. 2019). We leverage this apparatus to explore the behavior of flies in more complex environments. In Chapter 3, we study the decision-making strategy that the fly brain uses to cope with reward uncertainty and test a theory underlying the nature of learning rules that could give rise to the observed behaviors. In Chapter 4, we combine this apparatus with novel developmental manipulations produced by our collaborators in Dr. E. Josephine Clowney's lab to test theoretical hypothesis underlying decisions involving the discrimination of similar stimuli.

Chapter 3. Covariance-based learning rules that incorporate reward expectation direct learning and underlie operant matching in the *Drosophila* Mushroom body

Adapted from Rajagopalan et al. 2023, PNAS (in press)

3.1 Introduction

An animal's survival depends on its ability to adaptively forage between multiple potentially rewarding options(Stephens and Krebs 1986; Hayden and Walton 2014). To guide these foraging decisions appropriately, animals learn associations between options and rewards(Schultz 2004; Pearce 2008; Glimcher and Fehr 2013). Learning these associations in natural environments is complicated by the uncertainty of rewards that change dynamically over time. Both vertebrates and invertebrates employ decision-making strategies that account for this uncertainty (Beron et al. 2022; Pierce and Epling 1983; Greggers and Menzel 1993; Richard J. Herrnstein 1997; Lau and Glimcher 2008; Tsutsui et al. 2016; Bari et al. 2019; Iigaya et al. 2019; Sugrue, Corrado, and Newsome 2004b). A commonly observed strategy across the animal kingdom is to divide choices between options in proportion to the rewards received from each(Pierce and Epling 1983; Greggers and Menzel 1993; Richard J. Herrnstein 1997; Sugrue, Corrado, and Newsome 2004b; Lau and Glimcher 2005, 2008; Tsutsui et al. 2016; Bari et al. 2019; Iigaya et al. 2019). It has been hypothesized that animals that use this operant matching strategy make use of the expectation of reward – the recency-weighted rolling average over past rewards – to learn option-reward associations(Sugrue, Corrado, and Newsome 2004b; Lau and Glimcher 2005). Many studies further posit that this learning involves synaptic plasticity (Wickens, Reynolds, and Hyland 2003; Soltani and Wang 2006; Pereira-Obilinovic et al. 2022), and theoretical work has identified a characteristic relationship between operant matching and a specific form of expectation-based plasticity rule that incorporates the covariance between reward and neural activity(Loewenstein

and Seung 2006). Despite this strong link between plasticity rules and the matching strategy, there has been no mapping of these rules onto particular synapses or plasticity mechanisms in any animal. As a result, deeply investigating these theories by manipulating and testing the nature of plasticity rules underlying operant matching has been intractable. By using our novel behavioral assay described in Chapter 2 and leveraging this foraging framework in flies we aim to provide an insightful framework for testing the neural computations underlying decision-making strategies, such as matching.

3.2 Results

3.2.1 Flies can learn multiple probabilistic cue-reward associations

In our Y-arena, a single fly begins a trial in an arm filled with clean air and can choose between two odor cues that are randomly assigned to the other two arms (Fig. 3.1A; Chapter 2). The fly can freely move between arms, with a choice defined as the fly crossing into the reward zone at the end of the arm (Fig. 3.1A). Once a choice is made, we provide reward by optogenetically activating sugar-sensing neurons using a Gr64f driver(Jiao et al. 2008; Haberkern et al. 2019). The Y-arena then resets, with the arm chosen by the fly filled with clean air and the other two arms randomly filled with the two odors. This task design permits us to precisely control reward delivery without satiating the fly and enables us to monitor the choices of a single fly over hundreds of trials.

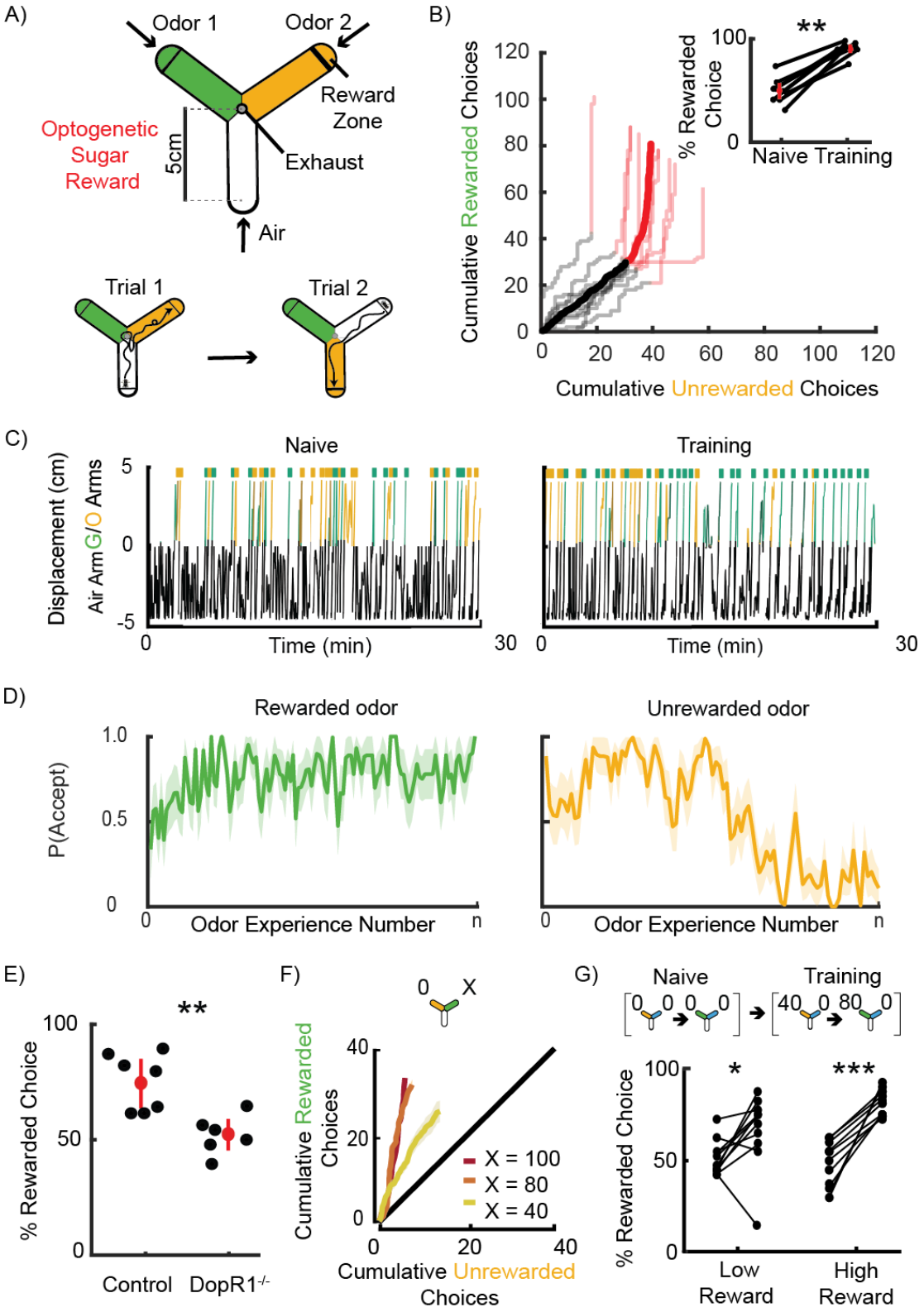
A)
Odor 1
Odor 2
Reward Zone
Optogenetic Sugar Reward
5cm
Exhaust
Air
Trial 1
Trial 2
B)
Cumulative Rewarded Choices
Cumulative Unrewarded Choices
% Rewarded Choice
Naive
Training
**
C)
Naive
Training
Displacement (cm)
Air Arm G/O Arms
Time (min)
D)
Rewarded odor
Unrewarded odor
P(Accept)
Odor Experience Number
E)
% Rewarded Choice
Control
DopR1-/-
F)
Cumulative Rewarded Choices
Cumulative Unrewarded Choices
X = 100
X = 80
X = 40
G)
Naive
Training
% Rewarded Choice
Low Reward
High Reward
*
***

Figure 3.1 Flies learn multiple probabilistic cue-reward associations (A) Schematic of Y-arena *(top)*. Air flows from tips of each arm to an outlet in the center. Reward zones are demarcated by lines. A choice is registered when a fly crosses into the reward zone of an odorized arm, triggering Gr64f sugar sensory neuron optogenetic activation with a 500 ms pulse of red light. The next trial commences as the chosen arm switches to air and the two odors (green/orange) are randomly reassigned to the other two arms *(bottom)*. (B) Cumulative choices made towards each option are shown (n = 9 flies, mean + individual flies). No rewards are available for the first 60 trials (Naive - black) and become available for the green option from the 61st trial onwards (Training - red). *Inset:* Percentage rewarded choices in naive and training blocks. Flies prefer the rewarded option in the training block compared to naive (Wilcoxon signed-rank test: p = 0.0039, n = 9).(C) Example trajectory of a fly in the Y before *(left)* and after *(right)* green odor is paired with reward. Distance in air arm is represented as negative values (black), while distances in odorized arms are represented as positive values (green/orange). Choices are represented by colored rasters. At choice points the arena resets and that arm switches to air, so the fly's position jumps to the tip of the air arm. (D) The probability of accept decisions are plotted as a function of time in the 100:0 protocol (n = 9 flies, mean - solid line, SE - shaded area). Flies show a high probability of accepting the rewarded odor *(left).* The probability of accepting the unrewarded odor drops over time *(right)*. (E) Controls (left) show higher percentage of choice made towards the rewarded option than DopR1 K.O. (right) flies in one 100:0 block of 60 trials (mean + SE - red point + line; individual fly scores - black; Mann-Whitney rank-sum test: p = 0.0022, control: n = 7, DopR1 -/- : n = 6). (F) Schematic describes the reward structure of the task *(top).* Cumulative rewarded and unrewarded choices plotted against each other, for three different protocols 100:0, 80:0, 40:0 *(bottom)*. Slope of all curves indicate that flies show a preference for the rewarded odor in all cases compared to a naive preference indicated by the black line (Mann-Whitney rank-sum test: 100:0, p = 4.4500*10-8, n = 18; 80:0, p = 5.8927*10-5, n = 10; 40:0, p = 0.0014, n = 10). (G) Schematic of the protocol for training flies with two simultaneous probabilistic cue-reward contingencies (*top*). Two different odor choices are alternated throughout an unrewarded naive block and a reward block where options were rewarded with baiting probability of 0.4 or 0.8. Performance (percentage of choices in which potentially rewarding option was chosen) on the low and high reward choices (*bottom*), indicates that flies learn both associations. An increased preference for the rewarded odors over unrewarded is observed (compared to naive preference) (Mann-Whitney rank-sum test: p = $2.3059*10^{-4}$ for high rewarding odor; n = 10; p = 0.01 for low rewarding odor, n = 10).

We first established that flies learn effectively in this apparatus by reliably rewarding flies only when they chose one of the odors - what we term a 100:0 protocol. Each fly first experienced the two odors (3-octanol; OCT and 4-methylcyclohexanol; MCH) unrewarded for a block of 60 trials, and then reward delivery was activated for the following block of 60 trials. As observed previously, although individual flies exhibited different odor biases in this naive phase(Claridge-Chang et al. 2009; Honegger et al. 2020; Smith et al. 2022), those biases averaged out over the population (Fig. 3.1B + inset). In this phase, flies spent a lot of time in the air arm and made variable choices, with little preference for either odor (Fig. 3.1C left, example fly). Once reward was made available, flies rapidly shifted to choosing the rewarded odor (Fig. 3.1B). This was accompanied by a faster interval between choices (Fig. 3.2A), and a decrease in meandering trajectories (Fig. 3.1C).

To analyze this choice behavior at a more elemental level, we adopted the common framework of considering foraging choices as a series of accept-reject decisions, where the animal decides whether or not to pursue an option(Hayden 2018). We defined reject decisions as when a fly enters an odorized arm but turns around and exits the arm before reaching the reward zone, while accept decisions reflect cases where the fly reaches the reward zone (see 3.4 Methods). Associating options with rewards changed the probability of accept decisions gradually over the course of a block. Acceptance probability increased for the rewarded odor and decreased for the unrewarded odor (Fig. 3.1D, Fig. 3.2D). On average, flies were around four times more likely to reject the unrewarded odor and seven times more likely to accept the rewarded odor (Fig. 3.2C). Interestingly, flies tended to reject odors quite close to the tip of the arm (Fig. 3.2E), suggesting that flies might accumulate evidence over time to make and commit to their decision – an aspect of fly behavior that has previously been studied(Groschner et al. 2018). These results indicate that fly choice behavior in this task can be thought of as a series of accept-reject decisions.

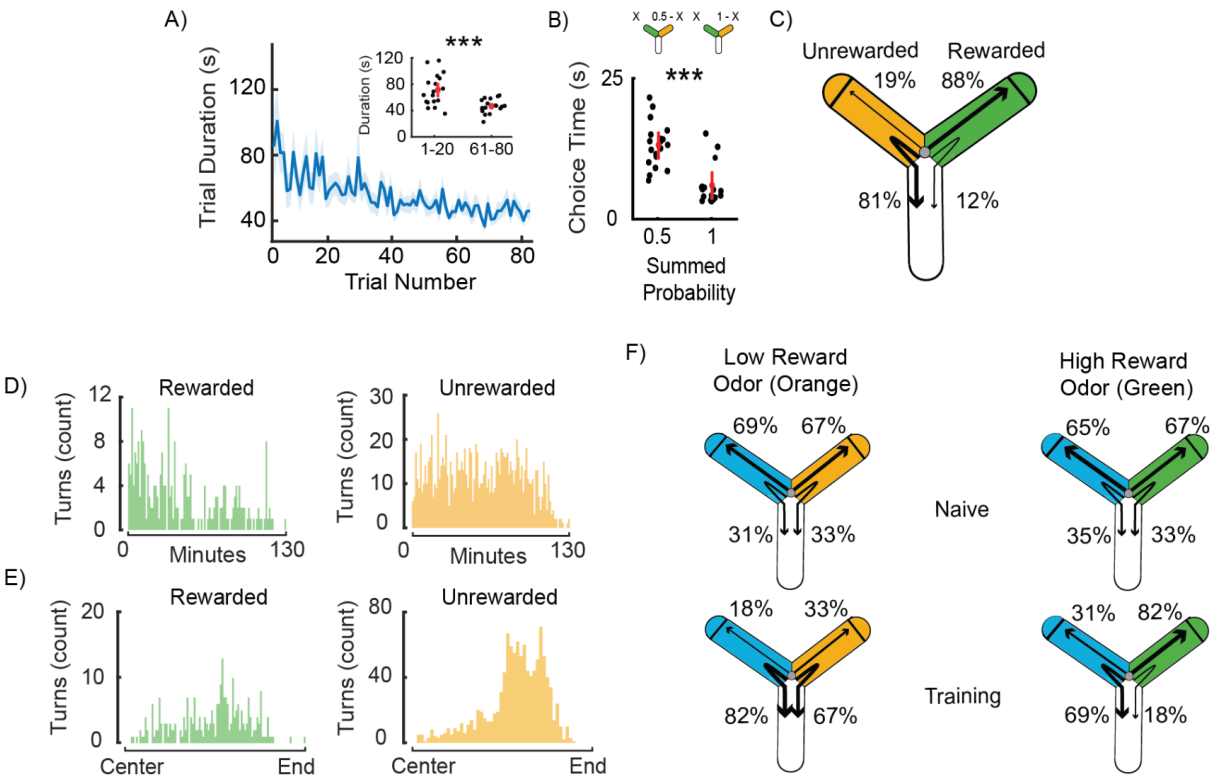

Figure 3.2 Quantification of learning in the novel Y-arena (A) The time taken to make a choice decreases once reward is made available to the fly (mean +/- SEM). *Inset:* Average trial time for the first 20 trials is longer than the average trial for the last 20 trials (Wilcoxon signed-rank test: $p = 6.292*10^{-4}$, n=18 flies). (B) The average choice time across 80 trials, measured from first exit of the air arm till entry into the reward zone, for two different summed probabilities of receiving reward (0.5 or 1). Average choice time decreases as the summed probability increases (Mann-Whitney rank-sum test: $p = 4.7653*10^{-4}$, n = 18 for summed probability = 1; n = 20 for summed probability = 0.5). (C) Percentage of odor arm entries that lead to accept or reject in a 100:0 protocol. Flies show increased rejection of the unrewarded odor (Mann-Whitney rank-sum test: $p = 3.2278 * 10^{-7}$, n = 18), and decreased rejection of the rewarded odor (Mann-Whitney rank-sum test: $p= 3.2166 * 10^{-7}$, n = 18). (D) Histogram showing the number of rejects over time for rewarded and unrewarded odor choices. Rejects decrease over time for the rewarded odor (Wilcoxon signed rank test: p =0.0171, n = 18, comparing bins 15-65 with 66-115 minutes - to exclude large time values when many flies had already finished the task), but not the unrewarded one (Wilcoxon signed rank test: p = 0.1839, n = 18). (E) Histogram showing the number of reversals as a function of distance along the odorized arm. The demarcation of "End" on the x-axis represents entry into the reward zone. (F) The average percentages of accepting and rejecting each odor - high-rewarding (green), low-rewarding (orange) and unrewarded (blue) are graphically represented in a schematic of the Y-arena (n=10 flies). Flies increasingly accepted the high-rewarding odor(Mann-Whitney rank-sum test: $p= 1.8267*10^{-4}$, n = 10), and displayed an increased probability of rejecting both low-rewarding and unrewarded odors, as compared to naive trials (Mann-Whitney rank-sum test: unrewarded odor: $p = 1.8165*10^{-4}$, n = 10; rewarded odor: $p = 7.6854*10^{-4}$, n = 10).

We found that the odor-reward associations learnt by flies in our assay were MB dependent. Learning-related plasticity in the MB circuit requires the activity of dopaminergic neurons (DANs) (Séjourné et al. 2011; Hige et al. 2015; Cohn, Morantte, and Ruta 2015; Owald et al. 2015; Aso and Rubin 2016; Berry, Phan, and Davis 2018; Handler et al. 2019). Dopamine is sensed by odor-representing Kenyon cells (KCs) and induces synaptic plasticity between these KCs and downstream mushroom body output neurons (MBONs)(Hige et al. 2015; Cohn, Morantte, and Ruta 2015). To interfere with this plasticity, we used a tissue-specific CRISPR knock-out strategy(Gratz et al. 2014) to knock out DopR1 receptors selectively in the KCs (Chapter 3.4 Methods), which are necessary for flies to associate odors with rewards in other paradigms(Handler et al. 2019). These flies showed no detectable learning in the 100:0 protocol, compared to control animals (Fig. 3.1E). These findings establish that odor-reward associations in our novel behavioral essay are mediated by MB plasticity.

We then asked whether flies could link odor cues with probabilistic rewards and distinguish between different reward probabilities, a key aspect of natural foraging. Importantly, we incorporated reward baiting into our probabilistic reward tasks(Sugrue, Corrado, and Newsome 2004b; Bari et al. 2019). This means that rewards probabilistically become available and then persist until the reward is collected (Chapter 3.4 Methods). Baiting is commonly used in mammalian 2AFC tasks, as it is thought to mimic the natural processes of resource depletion and replenishment over time. We began with experiments in which a single odor was rewarded with a range of baiting probabilities: 1 (100:0 task), 0.8 (80:0 task) or 0.4 (40:0 task). Flies showed a preference towards the rewarded odor in all cases compared to a naive lack of preference indicated by the black line (Fig. 3.1F). The extent of the preference varied with the probability of reward - a higher probability of reward led to a stronger preference. Interestingly, flies made faster choices when rewards were more probable (Fig. 3.2B).

These results show that flies can learn from probabilistic rewards but do not determine if they can store two associations simultaneously - another necessity for foraging. To test this, we designed a paradigm with a third odor, pentyl acetate (PA) included. This served as the unrewarded cue while we tested memory formation with the other two odors (Fig. 3.1G top). Flies first made 80 unrewarded choices consisting of 40 choices between OCT and PA and 40 choices between MCH and PA. In the next 80 (Training) trials, one of OCT or MCH was assigned a high baiting probability (0.8) and the other a low probability (0.4). We alternated the training trials for the two different odors to ensure both relationships would be learnt simultaneously (see Chapter 3.4 Methods). After pairing, flies preferred both rewarded odors over PA compared to their naive preference (Fig. 3.1G bottom). This choice preference was also reflected in their accept/reject behavior, with flies exhibiting a clear preference for accepting the high-rewarding odor (Fig. 3.2F right). Interestingly, in trials with the low-reward cue presented, there was an increased probability of rejecting both rewarded and unrewarded odors, as compared to naive trials (Fig. 3.2F left). This suggests the possibility that flies keep track of all the odor options potentially available in the environment and actually increase their rejection rate in the absence of the high-reward odor.

Overall, these experiments establish the fly as a capable animal model for studying foraging behaviors. Individual flies in the Y-arena can learn multiple odor-reward associations and can do so in the face of probabilistic reward. Importantly, these relationships are mediated by synaptic plasticity at the KC-MBON synapses in the MB. This establishes a foundation to test how these animals perform in dynamic foraging tasks and assess how they respond to reward baiting probabilities that change over time.

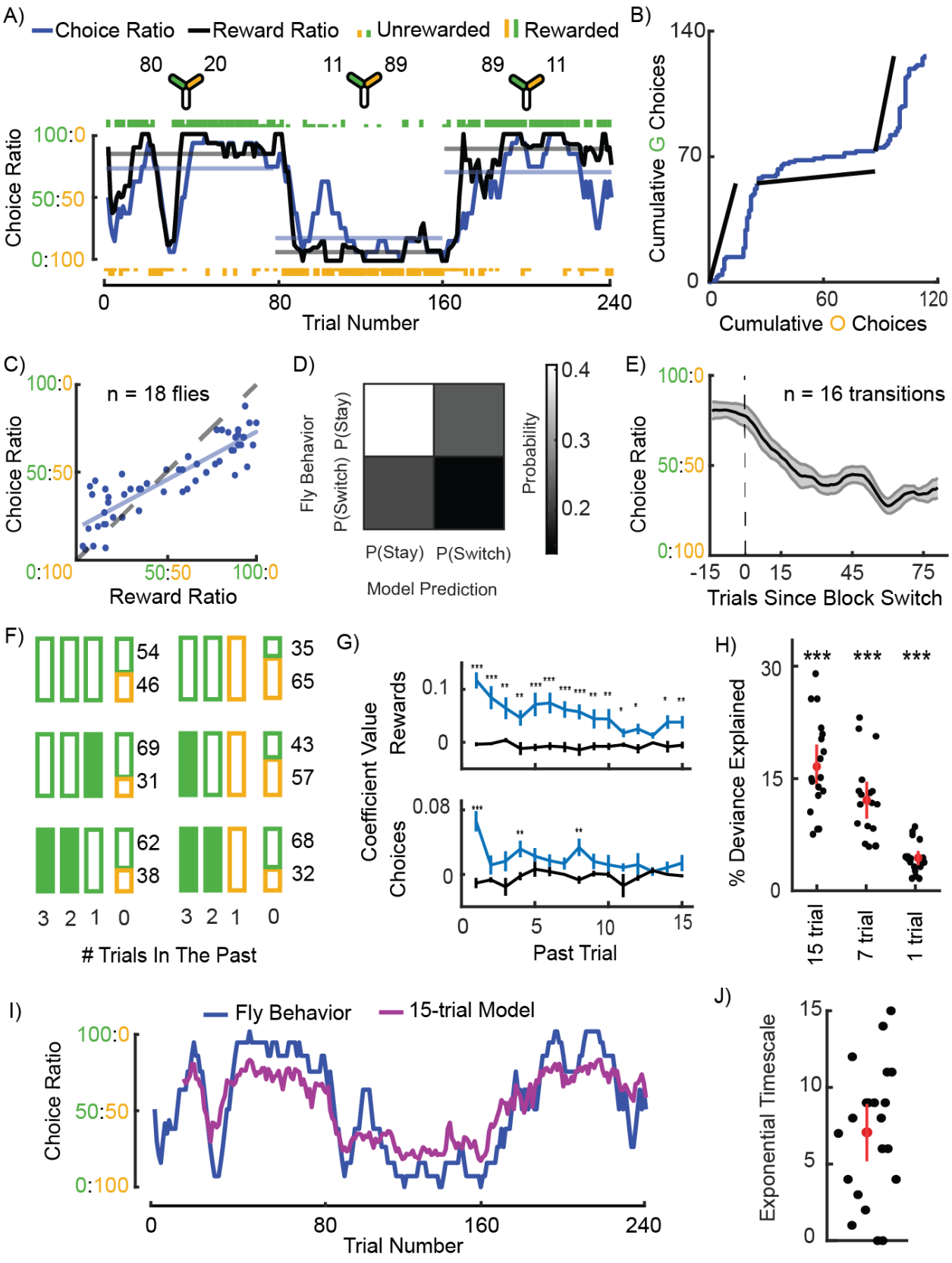
A)
Choice Ratio
Reward Ratio
Unrewarded
Rewarded
80
20
11
89
89
11
100:0
50:50
0:100
Choice Ratio
0
80
160
240
Trial Number
B)
140
70
0
Cumulative G Choices
0
60
120
Cumulative O Choices
C)
n = 18 flies
100:0
50:50
0:100
Choice Ratio
0:100
50:50
100:0
Reward Ratio
D)
Fly Behavior
P(Switch)
P(Stay)
P(Stay)
P(Switch)
Model Prediction
Probability
0.4
0.3
0.2
E)
n = 16 transitions
100:0
50:50
0:100
Choice Ratio
-15
0
15
45
75
Trials Since Block Switch
F)
54
46
35
65
69
31
43
57
62
38
68
32
3 2 1 0
3 2 1 0
Trials In The Past
G)
Coefficient Value
Rewards
0.1
0
Choices
0.08
0
0
5
10
15
Past Trial
H)
% Deviance Explained
30
15
0
***
***
***
15 trial
7 trial
1 trial
I)
Fly Behavior
15-trial Model
100:0
50:50
0:100
Choice Ratio
0
80
160
240
Trial Number
J)
Exponential Timescale
15
10
5
0

Figure 3.3. Flies follow Herrnstein's operant matching law (A) Matching of instantaneous choice ratio (blue) and reward ratio (black) in an example fly. Schematics indicate the reward baiting probabilities for each odor in the three 80-trial blocks (top). Individual odor choices are denoted by rasters, tall rasters - rewarded choices, short rasters - unrewarded choices. Curves show 10-trial averaged choice ratio and reward ratio, and horizontal lines the corresponding averages over the 80-trial blocks *(bottom)*. A description of how rewards are determined on any given trial in this task can be found in the 3.4 Methods section. (B) Cumulative choices of the same example fly. The slope of the black lines indicate the block-averaged reward ratio in the three successive blocks; the blue line indicates the cumulative choices with slope representing choice ratio. The parallel slopes of the two lines indicate matching. (C) Block-averaged choice ratio is approximately equal to reward ratio, following the matching law, but with some undermatching (n=54 blocks from n=18 flies). (D) A "win-stay; lose-switch" model does not accurately capture the trial-by-trial staying and switching probabilities of flies. A 2x2 probability table indicating the joint probability of the action predicted by the model and the action made by the fly (n=18 flies, 3.4 Methods). (E) Change in instantaneous choice ratio around block changes (n = 16 transitions with large changes in baiting probabilities between blocks). (F) Analysis of choices following particular histories of experience. Choices made by flies over three consecutive past trials are represented by boxes of different colors. Colors represent odors chosen, and rewarded choices are represented by filled boxes. Probabilities of choosing the green and orange odor on the current trial conditional on this history are illustrated with associated values. 6 out of 64 possible sequences are illustrated here. (G) Coefficients from logistic regression performed on fly choice behavior to determine the influence of 15 past rewards (top) and choices (bottom) on a fly's present choice (blue). These coefficients were compared to coefficients predicted on shuffled data (black) (Wilcoxon signed-rank test: *** - $p < 0.001$, ** - $p < 0.01$, * - $p<0.05$, n = 18 flies). (H) Model fit quality (percentage deviance explained) for 15-trial logistic regression, 7-trial logistic regression and 1-trial logistic regression models. Null model used to calculate the quality metric is a logistic regression with 0-trial history and only bias (Wilcoxon signed-rank test comparing the null model prediction with each test model prediction; shown here as test model prediction subtracted by null model: *** - $p < 0.001$, n = 18 flies). (I) 15-trial logistic regression fit (purple) on behavior (blue) from the example fly from panel A, plotted from the 15th trial onwards to avoid edge effects. (J) Exponential timescales (τ) for each fly shown in Supp Fig. 2, estimated from fitting the leaky-integrator model (see 3.4 Methods).

3.2.2 Flies follow Herrnstein's operant matching law

Foraging tasks are cognitively complex, involving two cues paired with different baiting probabilities that change with time. This requires animals to keep track of choice and reward history and form expectations to make adaptive choices. We designed our own dynamic foraging protocol to investigate how flies behave in such a scenario. The protocol consisted of three

consecutive blocks of 80 trials each. Flies made choices between two odors (OCT & MCH) that were paired with different baiting probabilities (Chapter 3.4 Methods). These probabilities remained fixed within a block and changed across blocks (Fig. 3.3A, example).

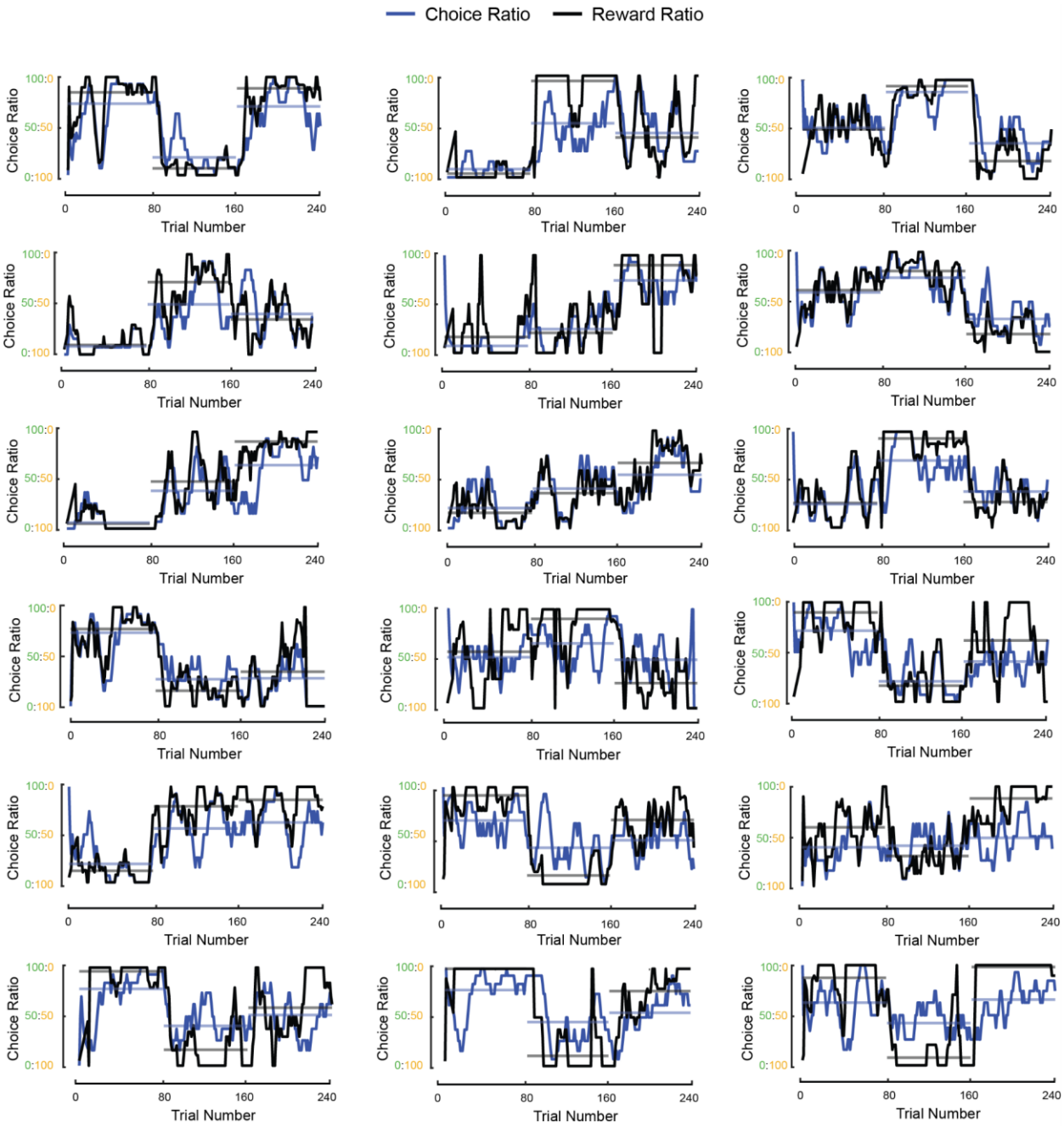

Figure 3.4 All example instantaneous choice ratio and reward ratio plots. Matching of instantaneous choice ratio (blue) and reward ratio (black) in all example flies.

We found that flies exhibit operant matching behavior, similar to observations in monkeys, mice and honeybees(Sugrue, Corrado, and Newsome 2004b; Tsutsui et al. 2016; Bari et al. 2019; Greggers and Menzel 1993). Individual flies exhibited a strong correlation between choice ratio (defined as the ratio between the number of choices made towards option A and option B), and reward ratio (defined as the ratio between the number of rewards received upon choosing option A and option B), either calculated over entire blocks or over a short (ten-trial) window to capture short-term dynamics/fluctuations (Fig. 3.3A,B - example fly; Fig. 3.4 - all 18 flies; Chapter 2). This holds true across flies, as seen in the relationship between block-averaged reward ratios and their choice ratios (Fig. 3.3C). In such a plot the matching law predicts that all points will fall along a line with slope equal to one (the unity line). Flies appear to approximately follow the matching law with a slight amount of undermatching, signified by a slope less than one. Undermatching is commonly observed across species(Sugrue, Corrado, and Newsome 2004b; Lau and Glimcher 2005; Tsutsui et al. 2016; Bari et al. 2019; Iigaya et al. 2019), and several reasons have been suggested for this tendency(Iigaya et al. 2019; Loewenstein and Seung 2006) (see Discussion).

Past work has suggested that animals form expectations of reward and use this to guide behavior in such dynamic foraging tasks(Sugrue, Corrado, and Newsome 2004b; Lau and Glimcher 2005; Loewenstein and Seung 2006; Iigaya et al. 2019). When rewards are delivered probabilistically, animals can only derive an expectation of reward by tallying information over multiple trials. However, such tallying could reflect a computation beyond the capabilities of flies. We wanted to explicitly address the alternative hypothesis that flies follow a simple win-stay/lose-switch strategy, which would suggest that their behavior is dictated by only the most recent reward/omission experience. Simulating choice sequences using this learning rule produced output that somewhat resembled that of the fly (example in Fig. 3.4A). However, it poorly captured the stay/switch probabilities actually observed in fly behavior data (Fig. 3.3D). In particular, switching occurred much more frequently than predicted. As further evidence that multiple past

outcomes affected behavior, choices of an individual fly at block transitions showed a lag between the choice ratio curve and the updated reward ratio at transition points (Fig. 3.3B), suggesting that the fly takes a few trials to adjust its behavior. Quantifying this across multiple transitions for all flies in the task showed flies require 15-20 trials to reach a new steady state choice behavior following block switches (Fig. 3.3E).

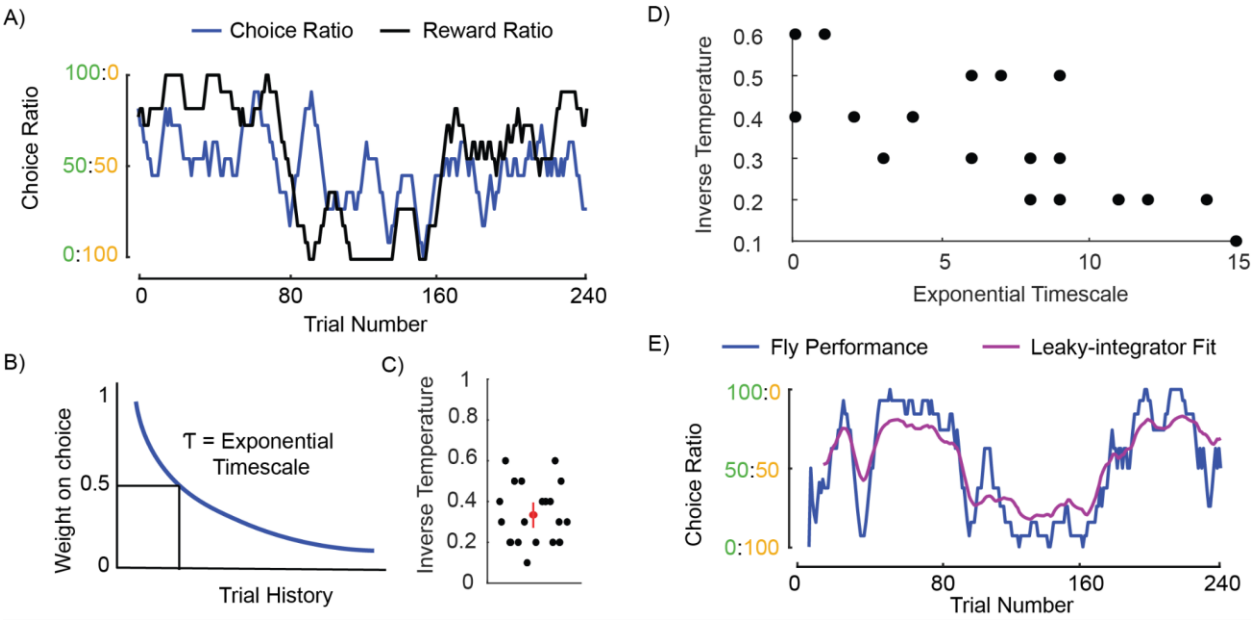

Figure 3.5 Analysis of the "win-stay lose-switch" and "leaky-integrator" models. (A) Example choice data generated by the "win-stay; lose-switch" model showing instantaneous choice ratio (blue) and reward ratio (black). (B) Schematic representing how the past trial history is weighted to calculate value in the "leaky-integrator" model. (C) Estimated inverse temperatures for each fly shown in Fig. 3.4. (D) Relationship between exponential timescale and inverse temperature. (E) Leaky-integrator model fit (purple) on behavior (blue) from the example fly in Fig. 3.3A, plotted from the 15th trial onwards to avoid edge effects.

It is possible that this lag could arise from averaging across multiple flies that switch at different trials after the transition. This could occur even if flies use just one past trial's worth of information to learn about the change in reward, consistent with observations in larvae(Lesar et al. 2021). To qualitatively illustrate that flies learn using multiple trials worth of past information, we first looked at the decisions made by flies following example triplets of choices and outcomes (Fig. 3.3F),

inspired by recent work in mice(Beron et al. 2022). For example, following three unrewarded choices of one particular odor, flies' next choice was roughly random (Fig. 3.3F top left). However, when an odor was rewarded on the most recent trial or more distant trials, choices were biased towards that option (Fig. 3.3F middle and bottom left). In another comparison, flies' tendency to switch back to an earlier choice (i.e. choose the green odor after an unrewarded choice of the orange odor) increased if that odor was rewarded in the recent past (Fig. 3.3F right).

To measure the relationship between current choice and past outcomes more systematically, we used logistic regression to determine how a fly's decisions depended on choice and reward history. Like other animals(Bari et al. 2019; Beron et al. 2022), flies showed a small amount of habitualness by choosing options that had been recently chosen more often; regression coefficients for a short history of recent choices were significantly positive compared to coefficients fit to shuffled data (Fig. 3.3G bottom). This approach also showed that the reward history was important for predicting choice, with many recent rewards weighted significantly (Fig. 3.3G top). We compared regression models that predicted behavior based on different lengths of outcome histories (15, 7 and 1 trial) and found that the percentage of deviance explained over a null model with a 0-trial history was greater for models that used longer outcome histories (Fig. 3.3H). An example fit from a regression model with a 15-trial history is shown in Fig. 3.3I. In alignment with the results of the regression model (Fig. 3.3G), we found that when fitting a leaky integrator model(Sugrue, Corrado, and Newsome 2004b), which assigns value to options using exponentially weighted reward histories (Fig. 3.4B-E), to the behavior of individual flies, an exponential timescale of 7 trials on average best predicted behavior (Fig. 3.3J). Together, these results show that flies' choices follow operant matching, with choices depending on the history of many past choices and outcomes.

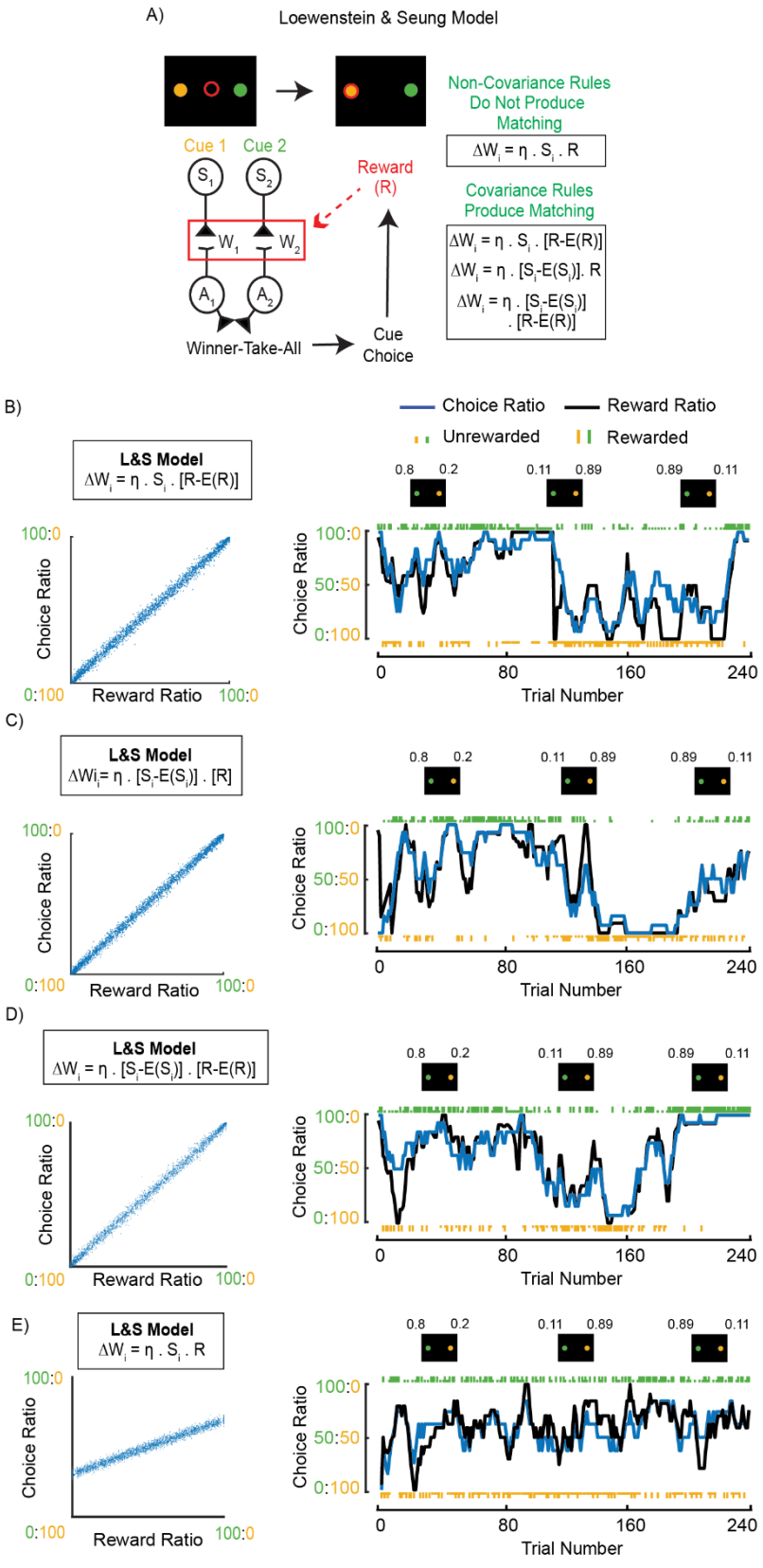
A)
Loewenstein & Seung Model
Non-Covariance Rules Do Not Produce Matching
ΔW_i = η . S_i . R
Covariance Rules Produce Matching
ΔW_i = η . S_i . [R-E(R)]
ΔW_i = η . [S_i-E(S_i)]. R
ΔW_i = η . [S_i-E(S_i)] . [R-E(R)]
Cue 1
Cue 2
Reward (R)
Winner-Take-All
Cue Choice
B)
L&S Model
ΔW_i = η . S_i . [R-E(R)]
C)
L&S Model
ΔWi= η . [S_i-E(S_i)] . [R]
D)
L&S Model
ΔW_i = η . [S_i-E(S_i)] . [R-E(R)]
E)
L&S Model
ΔW_i = η . S_i . R
Choice Ratio
Reward Ratio
Unrewarded
Rewarded
0.8 0.2
0.11 0.89
0.89 0.11
100:0
50:50
0:100
Choice Ratio
Reward Ratio
Trial Number
0
80
160
240

Figure 3.6 Covariance-based learning rules are necessary for operant matching (A) Description of the model developed by Loewenstein and Seung to study the requirements of matching behavior. The neural network consists of sensory neurons S_1 and S_2 that respond to one of the two simultaneously provided stimuli and synapse onto motor neurons A_1 and A_2 via synapses with weights W_1 and W_2. Choices are determined via a winner-take-all computation downstream of motor neurons. Upon choice, weights are updated according to one of the shown plasticity rules (*boxes-right*). Here, S_i is the activity of the i^{th} sensory neuron; R represents the presence or absence of reward; $E(S_i)$ is the mean or expectation of the activity of S_i and E(R) is the expectation of reward. (B) *Left:* Block-averaged choice ratio produced by the S_i*[R-E(R)] covariances-based rule (*box*) plotted against reward ratio. The model exhibits matching behavior (slope is 1). *Right:* An example simulation showing the performance in a 3-block task of a model incorporating a covariance-based rule S_i*[R-E(R)]. Task reward contingencies are the same as shown for the example fly in Fig. 2A. (C) Same as (B), but simulated with a $[S_i-E(S_i)]$*R rule. *Left:* The model exhibits matching behavior (slope is 1). *Right:* performance in a 3 block task where reward contingencies are the same as shown for the example fly in Fig. 2A. (D) Same as (B), but simulated with a $[S_i-E(S_i)]$*[R-E(R)] rule. *Left:* The model exhibits matching behavior (slope is 1). *Right:* performance in a 3-block task where reward contingencies are the same as shown for the example fly in Fig. 2A. (E) Same as (B), but simulated with a non-covariance (S_i*R) learning rule using the task and circuit structure of the Loewenstein and Seung model shown in Supp. Fig. 4. *Left:* The model produces behavior that does not show matching (slope < 1). *Right:* performance in a 3-block task does not accurately replicate fly behavior.

3.2.3 Covariance-based learning is required for matching behavior in a model of the MB

Theoretical work has placed strong, testable constraints on the nature of learning rules that could underlie operant matching. An elegant theory put forward by Loewenstein and Seung(Loewenstein and Seung 2006) proves that operant matching is the inevitable outcome of plasticity rules that modify synaptic weights according to the covariation of neural activity signaling reward and sensory input (3.4 Methods). Mathematically, covariance is the product of two variables with at least one being subtracted by its mean. The mean is simply the average reward and/or sensory input the animal experiences - an average that can also be expressed as the animal's expectation. Comparing the current value to its expectation ensures that weights can be adjusted up or down. Importantly, only an animal that follows operant matching would receive rewards at a rate equal to the reward expectation for both options, which leads weights to stabilize.

Loewenstein & Seung mathematically formalized this intuitive link between expectation and matching and showed that when synaptic plasticity is the basis for operant matching, only a covariance-based plasticity rule can account for matching.

They used a simple neural circuit model to illustrate their theory, with two different sensory inputs controlling different motor outputs and a decision determined by a winner-take-all interaction between those outputs (Fig. 3.6A). Interestingly, the structure of this model maps nicely onto the circuitry of the fly MB (Fig. 3.7A left). Sensory inputs are represented by the KCs, each odor activating a sparse subset of the KC population(Turner, Bazhenov, and Laurent 2008; Robert A. A. Campbell et al. 2013; Caron et al. 2013). KCs synapse onto MBONs, which guide motor output by signaling the valence of an odor i.e. its attractive/repulsive quality, rather than a specific action(Aso, Sitaraman, et al. 2014; Owald et al. 2015; Villar et al. 2022). KC-MBON synapses are modified by a plasticity rule that depends on the coincident activity of odor-representing KCs and release of dopamine by reward-signaling DANs (Séjourné et al. 2011; Hige et al. 2015; Cohn, Morantte, and Ruta 2015; Owald et al. 2015; Aso and Rubin 2016; Berry, Phan, and Davis 2018; Handler et al. 2019; Jiang and Litwin-Kumar 2021) (Fig. 3.7A - center, box). Current evidence indicates that postsynaptic activity of the MBON does not play a role in the plasticity (Modi, Shuai, and Turner 2020), so only the sensory and reward activity needs to be considered. Either or both of these terms could incorporate an expectation resulting in a covariance-based rule (Fig. 3.7A - center, box). DANs could incorporate reward expectation ([R - E(R)]) by subtracting a running average of reward activity (E(R)) from the current reward-related activity (R). Similarly, KCs could incorporate sensory expectation ([S_i - E(S_i)]) by calculating an average sensory experience, possibly by a mechanism that involves meta-plasticity and synaptic eligibility traces.

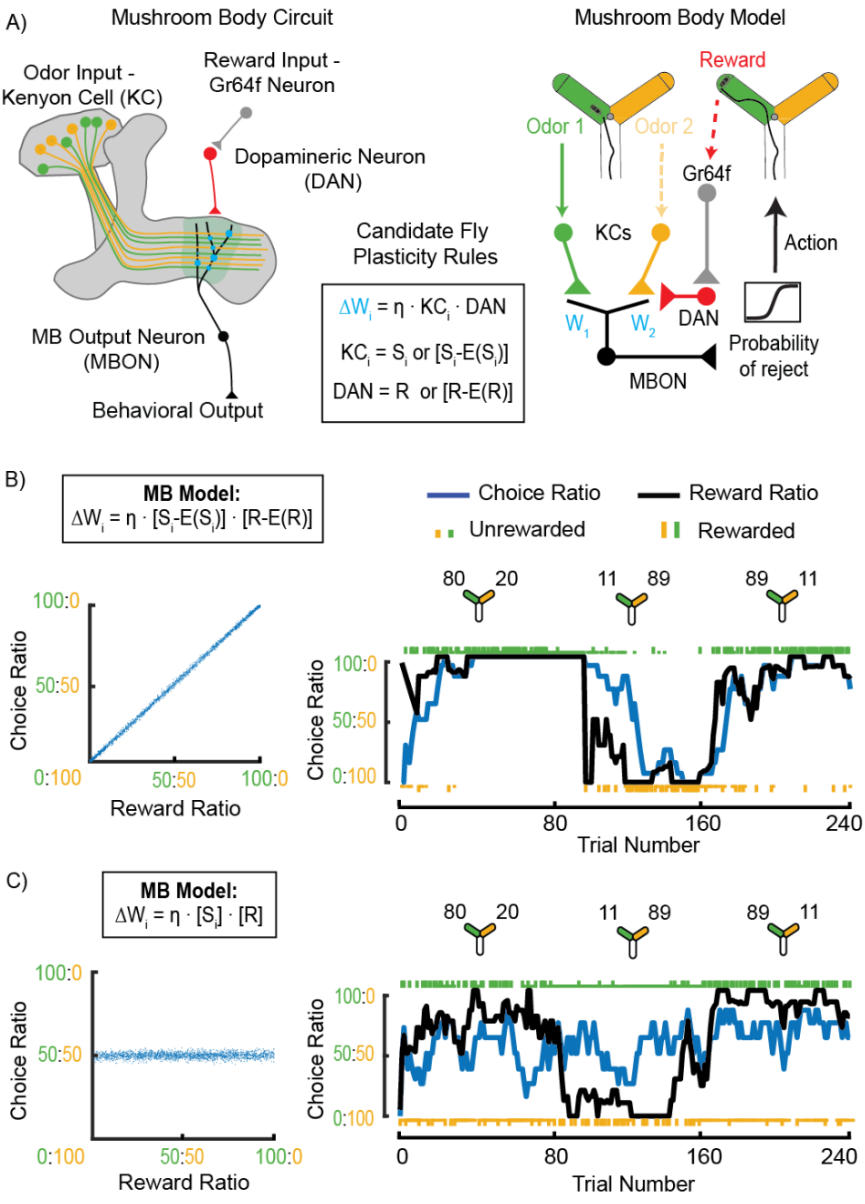

Figure 3.7 Covariance - based learning is required for matching behavior in a model of the MB. (A) *Left:* Schematic representing the MB with all relevant neurons shown in different colors (*key*). *Center:* Box containing candidate reward dependent synaptic plasticity rules at the KC-MBON synapse. *Right:* Schematic of our MB model developed by adapting Loewenstein & Seung's model to more closely resemble the MB and the features of our olfactory task. In the modified task, agents only experience one odor at a time. Reward information is provided to this circuit via DAN activity which either represents simply reward (R) or reward minus reward expectation (R-E[R]). Weights between inputs and MBON are modified according to plasticity rules shown in *Center*. MBON output determines probability of rejecting an odor and is passed through a sigmoidal nonlinearity to determine action. (B) *Left:* Block-averaged choice ratio produced by the $[S_i\text{-}E(S_i)]$*[R-E(R)] covariances-based rule (*box*) plotted against reward ratio. The model exhibits matching behavior (slope is 1). *Right:* An example simulation showing the performance in a 3 block task of a model incorporating a covariance-based rule $[S_i\text{-}E(S_i)]$*[R-E(R)]. Task reward contingencies are the same as shown for the example fly in Fig. 3.3A. (C) Same as (B) but simulated with a non-covariance learning rule. *Left:* The model produces behavior that does not show matching (slope < 1). *Right:* performance in a 3-block task does not show matching of choice and reward ratio.

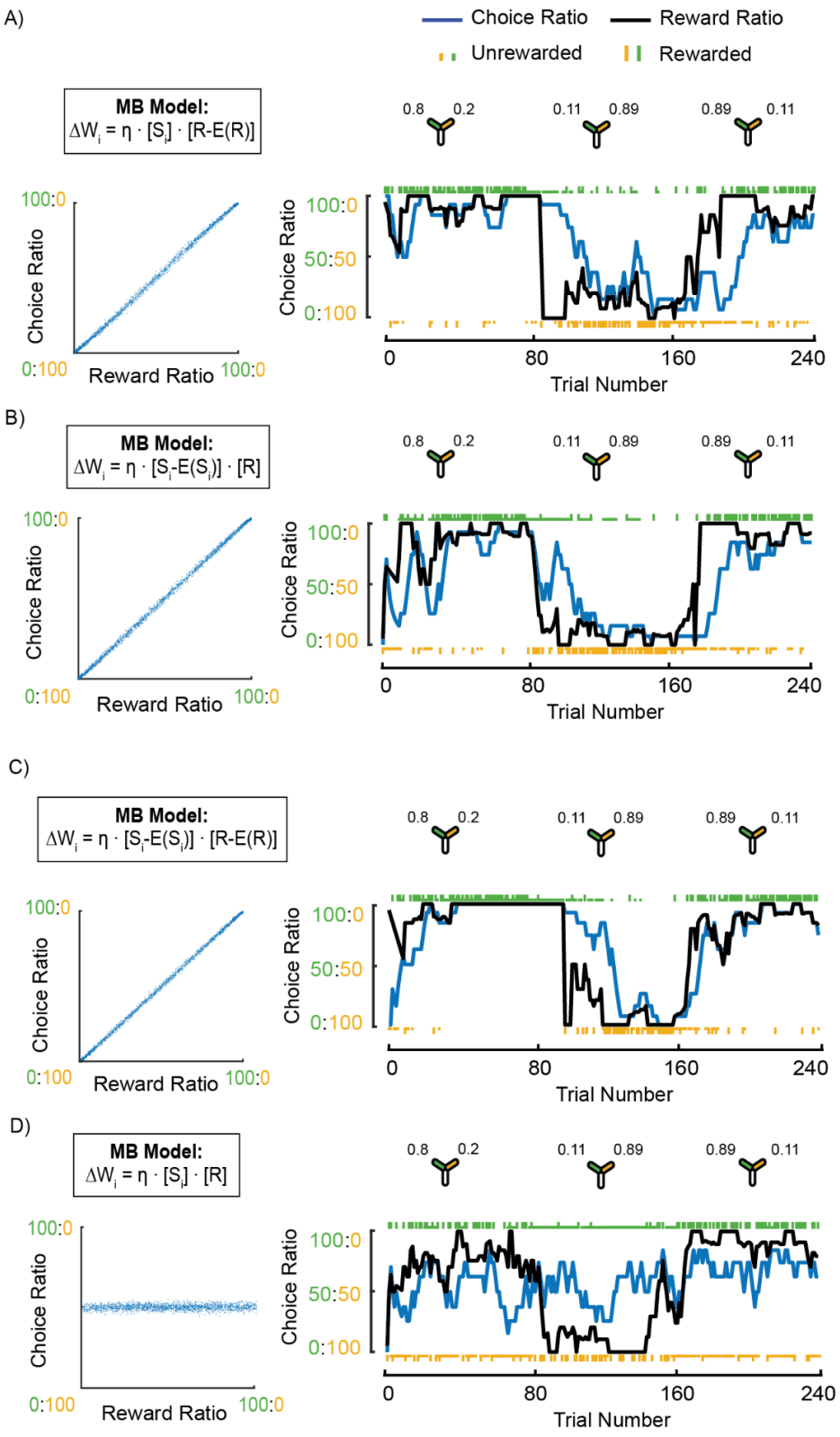
A)
Choice Ratio
Reward Ratio
Unrewarded
Rewarded
MB Model:
ΔW_i = η · [S_i] · [R-E(R)]
0.8 0.2
0.11 0.89
0.89 0.11
100:0
Choice Ratio
0:100
Reward Ratio
100:0
50:50
0:100
0
80
160
240
Trial Number
B)
MB Model:
ΔW_i = η · [S_i-E(S_i)] · [R]
C)
MB Model:
ΔW_i = η · [S_i-E(S_i)] · [R-E(R)]
D)
MB Model:
ΔW_i = η · [S_i] · [R]

Figure 3.8 Models using covariance-based learning rules produce behavior more similar to real fly behavior (A) *Left:* Block-averaged choice ratio produced by the S_i*[R-E(R)] covariances-based rule (*box*) plotted against reward ratio. The model exhibits matching behavior (slope is 1). *Right:* An example simulation showing the performance in a 3-block task of a model incorporating a covariance-based rule S_i*[R-E(R)]. Task reward contingencies are the same as shown for the example fly in Fig. 3.3A. (B) Same as (A), but simulated with a [S_i-E(S_i)]*R rule. *Left:* The model exhibits matching behavior (slope is 1). *Right:* performance in a 3 block task where reward contingencies are the same as shown for the example fly in Fig. 3.3A. (C) Same as (A), but simulated with a [S_i-E(S_i)]*[R-E(R)] rule. *Left:* The model exhibits matching behavior (slope is 1). *Right:* performance in a 3-block task where reward contingencies are the same as shown for the example fly in Fig. 3.3A. (D) Same as (A) but simulated with a non-covariance (S_i*R) learning rule. *Left:* The model produces behavior that does not show matching (slope < 1). *Right:* performance in a 3-block task does not show matching of choice and reward ratio.

To fully adapt the theoretical framework of Loewenstein & Seung to the biological network in the MB, we had to make a few changes (Fig. 3.7A right, 3.4 Methods). First, odors are represented by noisy populations of KCs (Turner, Bazhenov, and Laurent 2008; Caron et al. 2013; Robert A. A. Campbell et al. 2013). We thus parameterized input representations in the model to incorporate noise and overlap of KC subsets between options. Second, in our task flies only experience one odor at a time, so only one set of KCs is updated on a given trial. Although Loewenstein & Seung's original theory does not account for this possibility in its proof, we extended it to this case and found that a covariance-based plasticity rule is still necessary and sufficient to produce matching (3.4 Methods). Third, plasticity between MBONs and KCs are modified by a synaptic depression rule(Hige et al. 2015; Owald et al. 2015). We thus flipped the sign of the synaptic weight update rule. Finally, MBON activity determines whether flies accept or reject an odor rather than a winner-take-all decision mechanism (Aso, Sitaraman, et al. 2014; Owald et al. 2015) (Fig. 3.7A). We incorporated this into our model by having MBON activity encode the probability of rejecting an odor, with higher activity representing a greater probability to reject. This MBON activity was then passed through a sigmoidal nonlinearity to determine behavioral output.

We then evaluated whether these changes affect the relationship between covariance-based rules and matching. We used this MB-aligned model to simulate behavior arising from covariance rules that incorporated stimulus-expectation, reward-expectation or both (Fig. 3.7, Fig. 3.8). Consistent with the theory, all three covariance-based rules gave rise to a choice-ratio versus reward-ratio relationship that followed the matching law (Fig. 3.7B left, Fig. 3.8A,B,C left). In contrast, a rule that did not incorporate either reward or stimulus expectation did not follow the matching law and instead yielded a flat slope (Fig. 3.7 left, Fig. 3.8D left). For comparison, we also examined the behavior produced by the original model in a distinctly different task and observed similar results (Fig. 3.6A-E). Note that in the Loewenstein & Seung task, both options are always present when reward is delivered, which leads to a slope in between flat and unity when a non-covariance rule is used (Fig. 3.6E left). However, if only one option is present when an animal is rewarded, as in the fly task, synapses saturate and a non-covariance rule leads to a flat choice-ratio versus reward-ratio relationship (Fig. 3.7C left).

To get a more refined view of model performance, we examined the trial-by-trial behavior each plasticity rule generates. Models that incorporate covariance-based plasticity rules nicely replicate the trial-by-trial behavior of flies, tracking changes in the reward contingencies across blocks, with the resulting instantaneous choice ratio biased towards the more rewarded option in each block (Fig. 3.7B right; Fig. 3.6B-D; Fig. 3.8A-C right). On the other hand, both the MB-inspired model and Loewenstein & Seung's model do not capture trial-by-trial behavior well when a non-covariance rule is incorporated, with choices made roughly equally to both options throughout (Fig. 3.7C right, Fig. 3.6E right, Fig. 3.8D right). This reflects the fact that when value updates only depend on sensory input and reward, plasticity is unidirectional. Consequently, synapses representing the two options will both be driven to low levels, although at slightly different rates, so that ultimately both options are chosen roughly equally. Overall, these results show that a

model constrained by the network architecture of the MB more closely reproduces fly behavior when it operates according to a covariance-based plasticity rule.

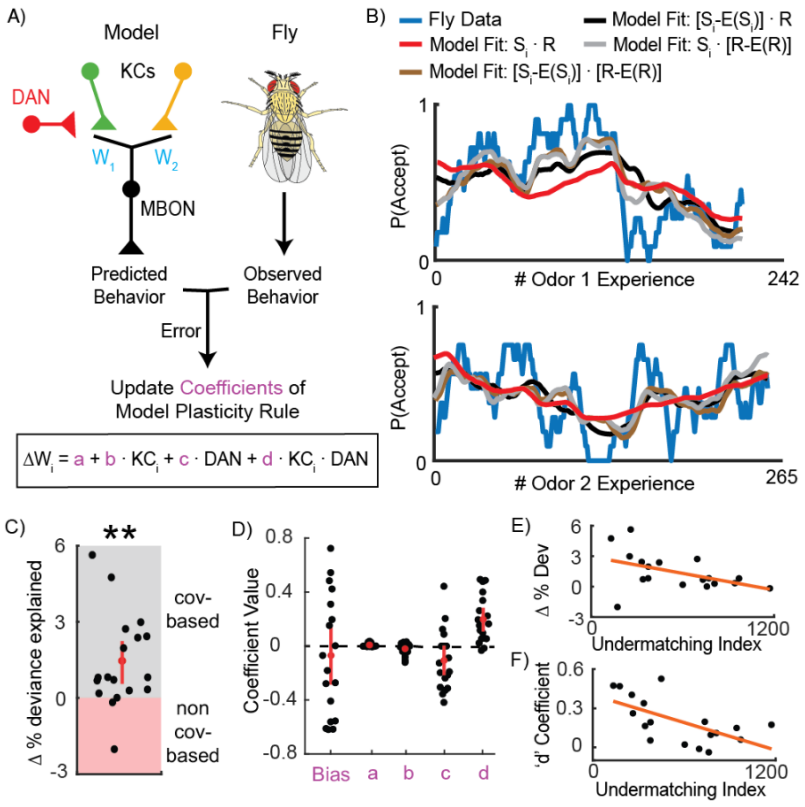

Figure 3.9 Identifying learning rules underlying dynamic foraging in the MB. (A) Schematic detailing the logic of the MB-inspired regression model. This model was used to predict the behavior of and learning rules used by each individual fly that experienced the task described in Fig. 3.3. (B) Example fly data (blue) showing the probability of accepting odor 1 *(top)* and odor 2 *(bottom)* calculated over a 6-trial window as a function of the number of times the fly experienced the given odor. This data was fit using an MB-inspired regression model (A) that incorporates either a covariance-based rule with sensory and reward expectations (brown), just sensory expectations (black), just reward expectation (grey), or a non-covariance rule (red). © Change in percentage deviance explained, computed by subtracting the percentage deviance explained of the non-covariance-based model from a covariance-based rule that incorporates reward expectation (n = 18 flies). On average, fly behavior was better predicted by the covariance-based model (Wilcoxon signed-rank test: p = 0.0018). Individual flies that were better fit by the covariance-based model have a positive value on this plot (gray region) while flies better fit by the non-covariance-based model have a negative value (red region). (D) Regression coefficients assigned to each term of the plasticity rule when the MB-inspired regression model using a covariance-based rule with reward-expectation was fit to the flies' behavior. As in C, the model was fit to each fly resulting in 18 different values for the coefficients. Largest coefficients were observed to have been assigned to the product term. (E) Change in percentage deviance explained (shown in C), plotted against a measure of undermatching (mean square error between instantaneous choice ratio and reward ratio lines) for each fly (n = 18). The best fit line of the scatter, calculated by a linear regression is shown in orange. (F) Coefficient value assigned to the product term (shown in D), plotted against a measure of undermatching for each fly (n = 18). The best fit line of the scatter, calculated by a linear regression is shown in orange.

3.2.4 Identifying learning rules underlying dynamic foraging in the MB

To test if our theoretical prediction of a covariance-based rule is supported by the observed behavior, we developed an approach that estimated the form of the plasticity rule being used in the fly MB. Our goal was to break the plasticity rule into components that span a space of possible rules and use optimization approaches to predict trial-by-trial behavior of each individual fly to assign coefficients to each of these components. In this way we would identify the form of the plasticity rule that best explained observed behavior and be able to conclude if this rule was a covariance-based rule.

We used the structure of the MB-inspired generative model (Fig. 3.7A) to build a predictive model and test how it fits the accept/rejection decisions made by the fly on each odor encounter. However, rather than utilizing a pre-defined plasticity rule, the predictive model used a rule composed of four terms that were candidate components of the MB learning rule (Fig. 3.9A). We then used logistic regression to assess which of these terms contributed the most when fitting fly behavioral data (3.4 Methods). The four terms were: a constant term governing overall learning rate, a KC term reflecting sensory input, a DAN term representing reward, and finally the product of KC and DAN activity. By definition, this product term becomes a covariance calculation when either of its elements are subtracted by their mean values, i.e., when either reward and/or sensory expectation are incorporated (Fig. 3.7A center box). We considered four model variants, a non-covariance one that lacked any expectation term and three different covariance-based rules where either KC or DAN or both were subtracted by their expectation. At every iteration of the logistic regression, the model prediction was compared to experimentally observed fly behavior, and regression coefficients were updated. Once the fit was optimized, we evaluated which term contributed the most to the fit by examining the weights of each coefficient.

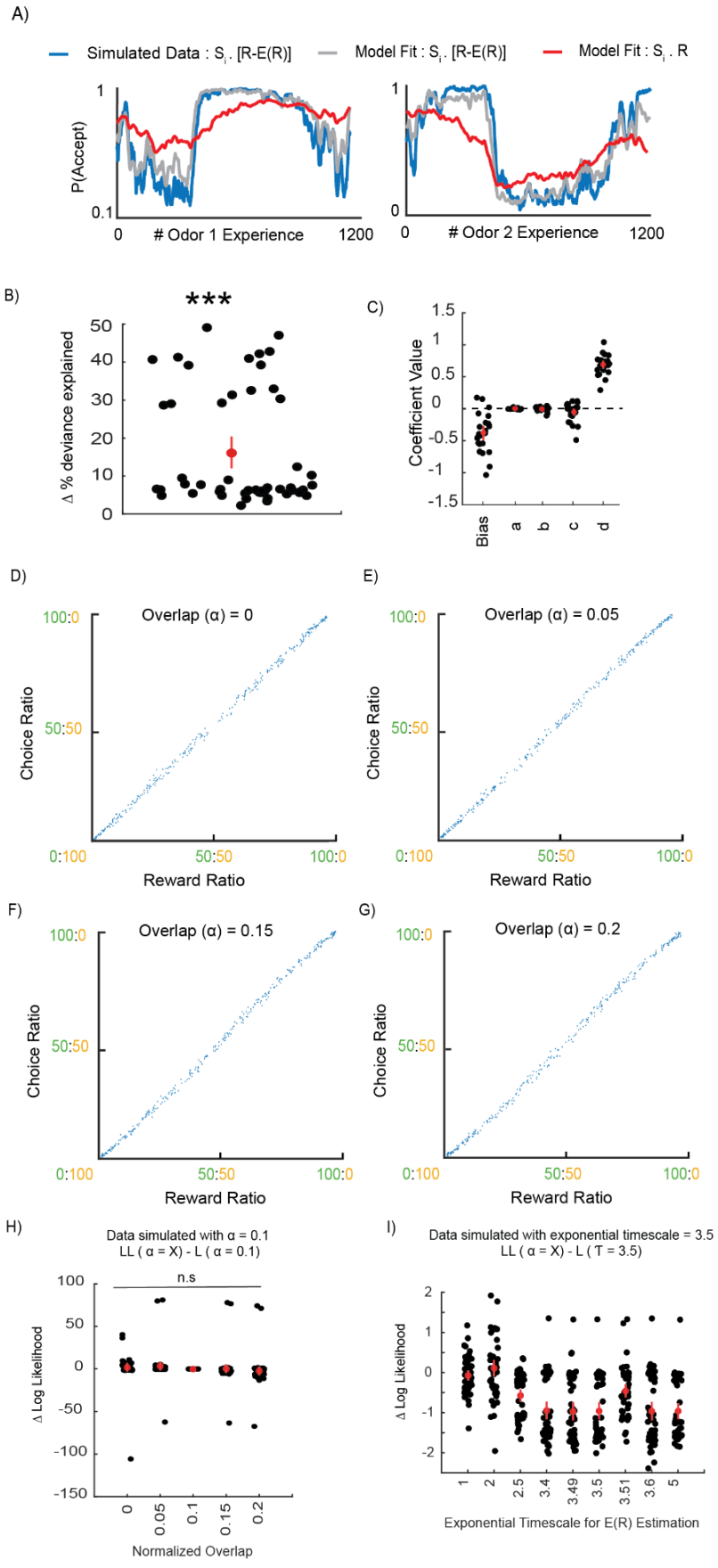
A)
Simulated Data : S_i . [R-E(R)]
Model Fit : S_i . [R-E(R)]
Model Fit : S_i . R
P(Accept)
1
0.1
0
Odor 1 Experience
1200
Odor 2 Experience
B)
***
Δ % deviance explained
C)
Coefficient Value
Bias
a
b
c
d
D)
Overlap (α) = 0
E)
Overlap (α) = 0.05
F)
Overlap (α) = 0.15
G)
Overlap (α) = 0.2
Choice Ratio
Reward Ratio
100:0
50:50
0:100
H)
Data simulated with α = 0.1
LL (α = X) - L (α = 0.1)
n.s
Δ Log Likelihood
Normalized Overlap
I)
Data simulated with exponential timescale = 3.5
LL (α = X) - L (T = 3.5)
Exponential Timescale for E(R) Estimation

Figure 3.10 Extent of sensory overlap does not affect the behavior of our model. (A) Example simulated data in an experiment that consisted of 3 blocks of 80 trials each with baiting probabilities changing between blocks, showing the probability of accepting odor 1 *(left)* and odor 2 *(right)* (blue), simulated using a covariance-based rule with reward expectation, and fit using an MB-inspired regression model (A) that incorporates either the same rule (gray), or a non-covariance rule (red). The predictions resulting from the model using the covariance-based rule is a better fit for the simulated data. Note: While the number of trials is 240, what we are plotting is the probability of accepting an odor upon experiencing it. Flies reject odors a lot and so the number of odor experiences is much more than 1 per trial leading to the x-axis of these plots having more than 1000 experiences. (B) Change in the percentage deviance explained between non-covariance and a covariance-based model. Both models are used to predict behavior simulated using a covariance-based rule (as in A). The change in goodness of fit is computed by subtracting the percentage deviance explained of the non-covariance-based regression models predictions from the predictions of the model with covariance-based rule. This is plotted for each simulation (n = 50). The covariance-based rule better fits the simulated behavior than the non-covariance-based rule (Wilcoxon signed-rank test: $p = 6.7595 * 10^{-9}$). (C) Regression coefficients assigned to each term of the learning rule when the MB-inspired regression model using a covariance-based rule with reward expectation was fit to the simulated behavior. Note the non-zero weight of the c and d terms. (D) Block-averaged choice ratios versus reward ratios (n = 300 simulations) from data simulated using a covariance-based rule with reward expectation are plotted against each other. Sensory overlap α = 0. (E) Block-averaged choice ratios versus reward ratios (n = 300 simulations) from data simulated using a covariance-based rule with reward expectation are plotted against each other. Sensory overlap α = 0.05. (F) Block-averaged choice ratios versus reward ratios (n = 300 simulations) from data simulated using a covariance-based rule with reward expectation are plotted against each other. Sensory overlap α = 0.15. (G) Block-averaged choice ratios versus reward ratios (n = 300 simulations) from data simulated using a covariance-based rule with reward expectation are plotted against each other. Sensory overlap α = 0.2. (H) Change in percentage deviance explained, computed by subtracting the percentage deviance explained of the covariance-based regression model with overlap = 0.1, from that of regression models with covariance-based rules and different amounts of overlap (0,0.05,0.15,0.2). Points are plotted for each simulation (n = 50). Simulations were run with overlap α = 0.1. (I) Change in percentage deviance explained, computed by subtracting the percentage deviance explained of the covariance-based regression model with exponential timescale (τ)= 3.4, from that of regression models with covariance-based rules and different exponential timescales. Points are plotted for each simulation (n = 50). Simulations were run with exponential timescale (τ)= 3.4.

Before applying this approach to fly data, we validated it by determining whether it correctly identifies the relevant term when tested with choice sequences that were simulated using a covariance-based learning rule that only incorporates reward expectation. Indeed, the fit quality was clearly better with a model that incorporated reward expectation (Fig. 3.10A,B). Moreover,

the largest weights were correctly assigned to the KC-DAN product term, the term that calculates the covariance between these two elements (Fig. 3.10C). Additionally, our simulations suggested that the extent of matching and the accuracy of learning rule fits were largely unaffected by either the degree of overlap in KC activity patterns or the timescale over which rewards were integrated (Fig. 3.10D-I). Consequently, for simplicity we then used overlap of zero and an exponential timescale of 3.4 trials in all future analysis.

We then applied our approach directly to the behavioral data from individual flies performing the dynamic foraging task. A representative example showing fly behavior and model predictions can be seen in Fig. 3.9B. This example suggests that models with covariance-based rules may better resemble the flies' behavior. To quantitatively compare fit quality of the different models, we calculated the percentage deviance explained for every individual fly. This metric showed that regressions that utilized rules with sensory expectation, reward expectation, or both were objectively better fits for fly behavior (Fig. 3.9C, Fig. 3.11A).

Overall, we found that learning rules that incorporated sensory and/or reward expectation yielded better fits to fly behavior than non-covariance rules. To distinguish between these different expectation-based learning rules, we examined the regression coefficients. When we fit a rule with only reward expectation, the regression assigned the KC-DAN product, i.e. the covariance term, the largest weight (Fig. 3.9D, Fig. 3.11B). On the other hand, fitting using either of the two covariance rules that incorporated sensory expectations yielded large coefficients to the non-covariance terms containing either KC or DAN activity alone (Fig. 3.11B). We observed a similar result when we fit simulated data from an agent using a reward expectation-based learning rule (Fig. 3.11C). However, when the behavior was simulated and fit using the same sensory expectation rule, the covariance term was given the most weight (Fig. 3.11E). These results suggest that flies use a reward-expectation based covariance rule to guide behavior.

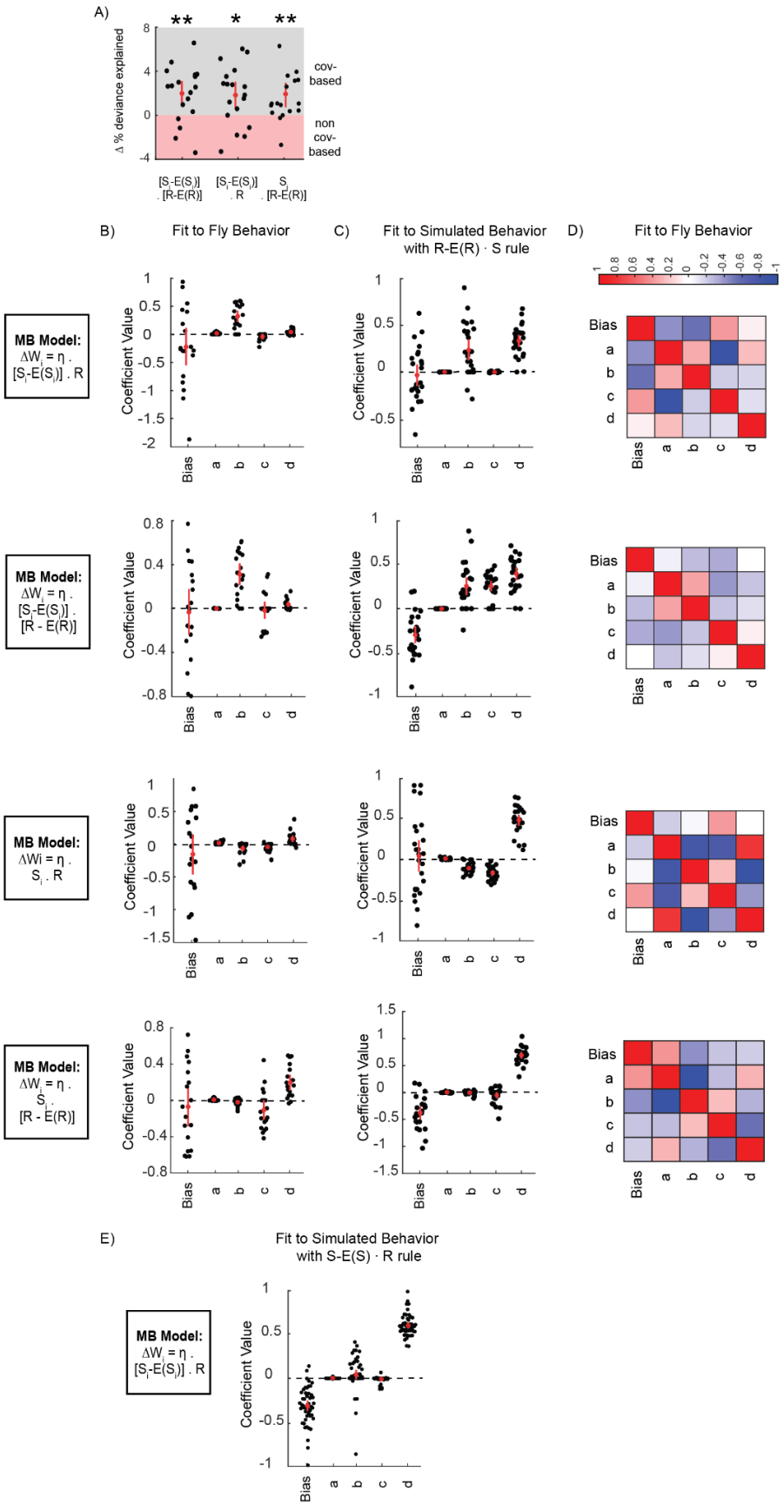
A)
**
*
**
Δ % deviance explained
cov-based
non cov-based
[S_i-E(S_i)] . [R-E(R)]
[S_i-E(S_i)] . R
S_i . [R-E(R)]
B)
Fit to Fly Behavior
C)
Fit to Simulated Behavior with R-E(R) · S rule
D)
Fit to Fly Behavior
MB Model: ΔW_i = η . [S_i-E(S_i)] . R
MB Model: ΔW_i = η . [S_i-E(S_i)] . [R - E(R)]
MB Model: ΔWi = η . S_i . R
MB Model: ΔW_i = η . S_i . [R - E(R)]
Coefficient Value
Bias
a
b
c
d
E)
Fit to Simulated Behavior with S-E(S) · R rule
MB Model: ΔW_i = η . [S_i-E(S_i)] . R

Figure 3.11 Covariance-based learning rules are better predictors on individual choice behavior. (A) Change in percentage deviance explained, computed by subtracting the percentage deviance explained of the non-covariance-based model from three models with covariance-based rules when fit to fly behavior (*left:* incorporating both stimulus and reward expectations; *center:* incorporating just stimulus expectation, *right:* incorporating just reward expectation - same as Fig. 3.9C) plotted for each fly (n = 18). Covariance-based rules were more predictive of fly behavior on average (Wilcoxon signed-rank test: *left*, p=0.0074; center, p = 0.0168, *right, p* = 0.0018). (B) Regression coefficients assigned to each term of the learning rule when the MB-inspired regression model was fit to the flies' behavior. All four flavors of model were fit and are indicated on the left. (C) Regression coefficients assigned to each term of the learning rule when the MB-inspired regression model was fit to data simulated with a reward-expectation-based covariance rule. All four flavors of model were fit and are indicated to the left of B. (D) Correlation between regression coefficients resulting from MB-inspired regression models fit to flies' behavior. All four flavors of model were fit and are indicated to the left of B. (E) Regression coefficients assigned to each term of a covariance-based learning rule with only stimulus expectation, when the MB-inspired regression model was fit to data simulated with the same rule.

Interestingly, we found that in some flies the simple expectation-free non-covariance rule was a better fit. One possible explanation for this result is that these flies showed operant matching to a lesser extent. We thus quantified matching by calculating the mean squared error between instantaneous choice and reward ratios and found that different strengths of matching across flies were correlated with how well an expectation-free plasticity rule fit the behavior data (3.4 Methods). Flies that were better fit by the expectation-free rule tended to show more undermatching, in line with our predictions (Fig. 3.9E). Consistent with this, the weight of the covariance term coefficient was greater in flies that exhibited stronger matching behavior (Fig. 3.9F). To examine whether some flies were better fit by a non-covariance rule because our approach might inaccurately assign weights to a combination of correlated terms in the learning rule, we examined the correlations between pairs of coefficients. However, we found no consistent statistical relationship (Fig. 3.11D; 3.4 Methods). Overall, this general approach allowed us to estimate the learning rule the fly uses directly from behavioral data, providing clear evidence that a reward-expectation-based covariance rule is important in the MB.

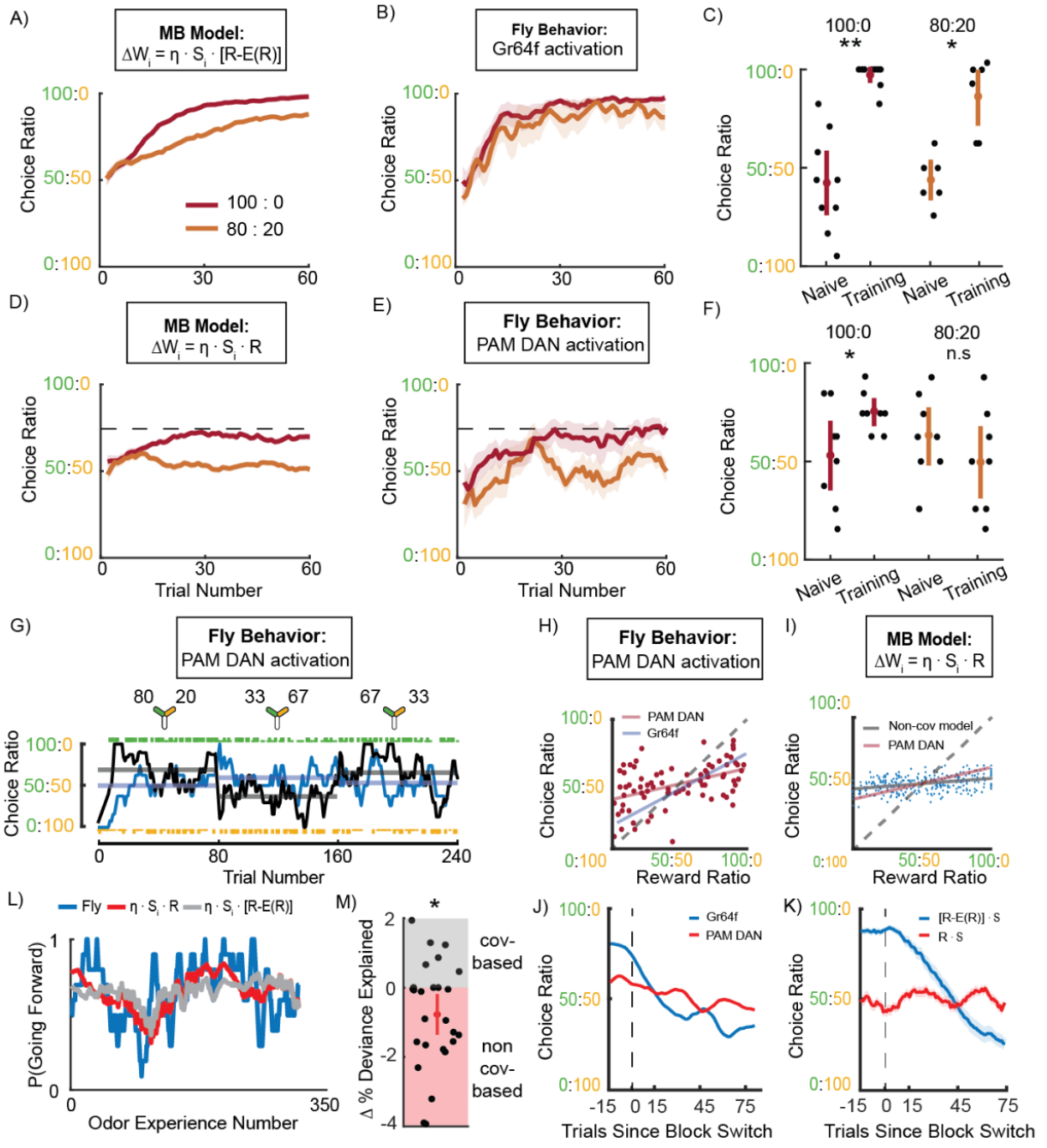

Figure 3.12 Behavioral evidence of reward expectation in DANs. (A) Instantaneous choice ratio over trial number, for a simulated agent using a covariance-based rule with reward expectation in 80:20 (orange) and 100:0 (red) tasks. (B) As (A) except it shows fly behavior when providing sugar sensory optogenetic reward instead of simulation (100:0, n=8 flies; 80:20, n = 6). (C) Average choice ratios of individual flies from C showing significant learning in both 100:0 and 80:20 protocols (Wilcoxon signed-rank test: 100:0, p = 0.0039; 80:20, p = 0.0312). (D) As (A), except for an agent using a non-covariance rule. (E) As (B) except reward provided via the PAM DANs using R58E02-Gal4 to drive UAS-CSChrimson (n=8 flies in both 80:20 and 100:0). Dashed

line in (D) and (E) indicates the maximum possible performance of agent in B in the 100:0 protocol. (F) Average choice ratios of individual flies from D. Flies showed a significant preference towards the rewarded odor in 100:0 but not 80:20 (100:0, p = 0.0391; 80:20, p = 0.1875). (G) The instantaneous choice ratio of an example fly receiving DAN optogenetic reward performing the dynamic foraging protocol plotted against trial number plotted as in Fig. 3.3A. (H) Block-averaged choice ratios against reward ratios for flies with DAN reward (n = 26 flies, 3 blocks each). Best fit lines: red - DAN reward, blue - Gr64f sugar sensory reward (Fig. 3.6C). (I) Block-averaged choice ratios against reward ratios (n = 50) from data simulated using a non-covariance-based rule. Best fit lines: black - simulated data, red - DAN reward. (J) Instantaneous choice ratio around block changes. Flies trained with Gr64f activation in blue, DAN activation in red. (K) As (J) but with simulated agents using either a covariance-based rule in blue or non-covariance rule in red. (L) Example fly data showing probability of accepting odors against experience number(blue) with DANs activated as reward. Fit using a model (see Fig. 3.9) that incorporates either a covariance-based rule (gray), or a non-covariance rule (red). (M) Change in percentage deviance explained, computed by subtracting the percentage deviance explained of the non-covariance-based model from a covariance-based rule, plotted for each fly (n = 26). On average, fly behavior was better predicted by the non-covariance-based model (Wilcoxon signed-rank test: p = 0.0164).

3.2.5 Behavioral evidence of reward expectation in DANs

We next wanted to experimentally verify that a reward-expectation-based covariance rule in particular guided learning and choice behavior in the fly MB. The mathematical differences between the three different covariance rules suggested a way forward (3.4 Methods). In particular, the rules differ in which terms - sensory input or reward - incorporate an expectation. Thus, to distinguish between the possible different covariance-based rules in the MB, we designed an experiment to manipulate the calculations of reward expectation using genetic tools that override the natural activity of the DANs. Specifically, we provided reward via optogenetic activation of the reward-related protocerebral anterior medial (PAM) DANs. This would bypass any upstream computation of reward expectation and simply provide a reward presence signal on every trial. Such a manipulation would change the learning rule from a covariance-based rule to a non-covariance rule if the following conditions were met: i) the animal's learning rule depended on the product of DAN and KC activities; ii) DAN activity incorporated reward expectation; iii) KC activity did not incorporate sensory expectation. This would in turn result in modified behavior. For this

test we initially focused on a task consisting of two blocks (naive and training) of 60 trials each, with a reward ratio of either 100:0 (one odor has a baiting probability of 100% and the other is never rewarded) or 80:20 (one odor has a baiting probability of 80% and the other 20%) in the second block (3.4 Methods).

We first predicted how behavior in these protocols would differ between covariance-based and non-covariance rules using simulations. As expected, covariance-based models learnt to choose the more rewarded option more often, with choice ratios reflecting reward ratios (Fig. 3.12A, Fig. 3.13A,B). The behavior of the model with any covariance-based rule was similar to the fly behavior when it was rewarded using the sugar neurons (Fig. 3.12B,C). On the other hand, non-covariance rules led to preferences saturated around 75% in 100:0 and 50% with the 80:20 reward ratio (Fig. 3.12D). These theoretically predicted preferences very closely match our observations of fly behavior in the DAN activation experiments (Fig. 3.12E,F). We observed low plateau performance in both tasks (Fig. 3.12E,F), with values strikingly similar to that predicted by the non-covariance rule (Fig. 3.12D).

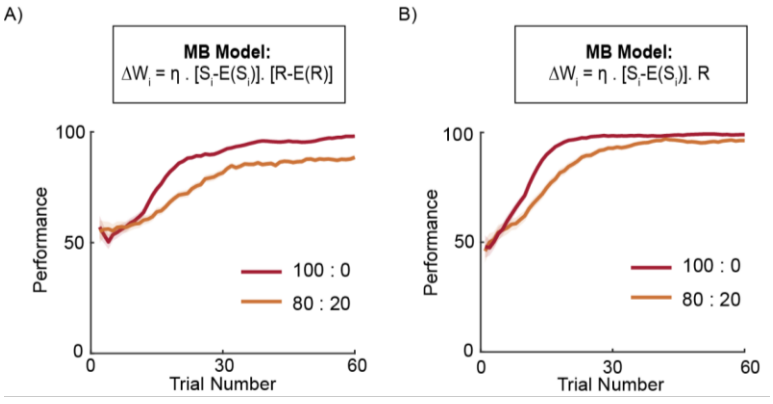

Figure 3.13 Covariance-based learning rules produce similar behavior in 100:0 and 80:20 tasks. (A)Simulated instantaneous performance plotted as a function of trial number (defined as the percentage of choices towards the option with higher pre-defined baiting probability in a 10-trial window) of an agent using a covariance-based rule with sensory and reward expectations in 80:20 (orange) and 100:0 (red) reward conditions. (B) Simulated instantaneous performance plotted as a function of trial number (defined as the percentage of choices towards the option with higher pre-defined baiting probability in a 10-trial window) of an agent using a covariance-based rule with sensory expectation.

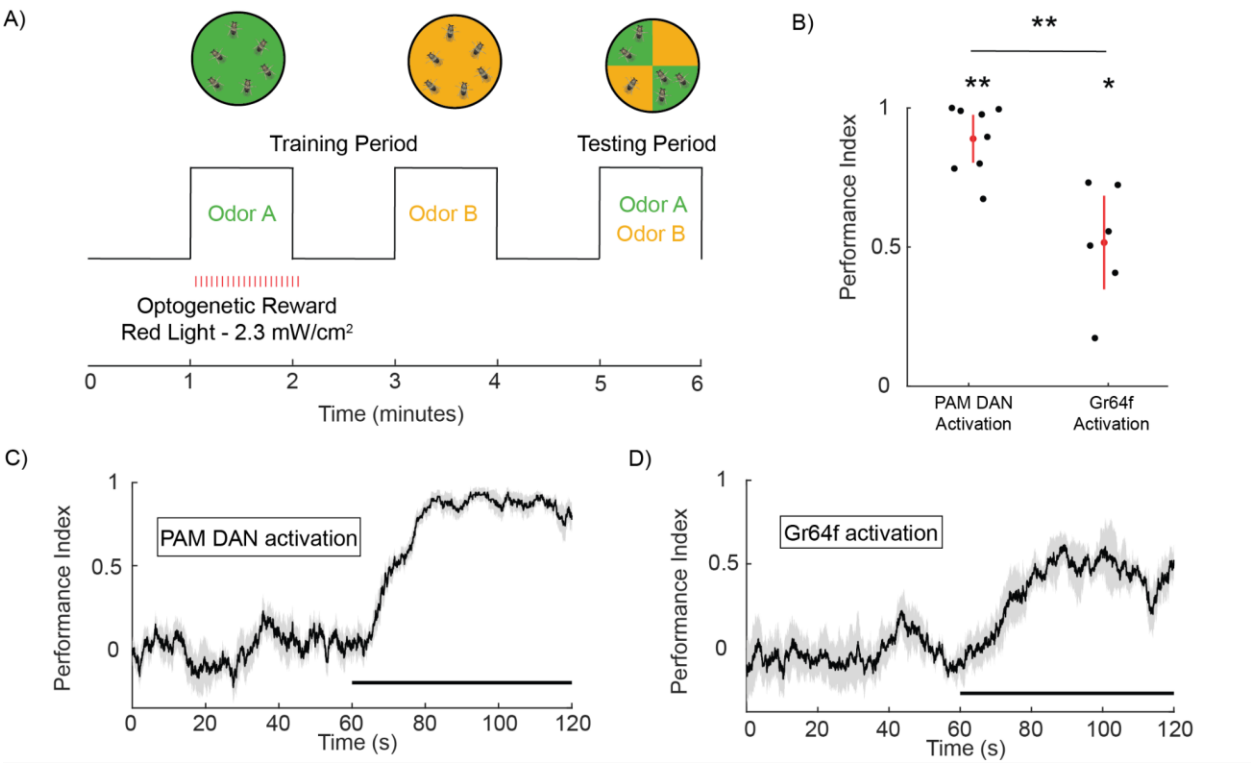

Figure 3.14 Circular arena experiments to control for rewarding red LED intensity. (A) Schematic of the experimental paradigm used to train flies in the circular arena. LED intensity chosen to be 2.3 mW/cm^2 to match intensity in the Y-arena. (B) Time averaged performance index (see 3.4 Methods) plotted for DAN trained and Gr64f trained flies show that both learn to prefer the reward-paired odor (Wilcoxon signed-rank test: Gr64f - n = 6 groups of ~20 flies each, p = 0.0312; DAN - n = 8 groups of ~20 flies each, p = 0.0078). (C) Performance index (mean in black, standard error in grey) plotted over timecourse of testing period for DAN trained flies. The black horizontal line represents the period during which odors were present in the arena. (D) Performance index (mean in black, standard error in grey) plotted over timecourse of testing period for Gr64f trained flies. The black horizontal line represents the period during which odors were present in the arena.

One potential concern with these experiments is that differences in the efficacy of optogenetic activation of the DANs deep in the central brain versus the peripherally located Gr64f neurons could contribute to these behavioral differences. However, when flies were instead made to choose between reward-associated or unrewarded odors in a circular arena previously used to assess learning in flies (Aso and Rubin 2016) we found that both PAM DAN and Gr64f sugar neuron activation drove similar learning (Fig. 3.14A-D). Since the LED intensity in the circular arena (2.3 mW/cm^2) was closely matched to that in the Y-arena (1.9 mW/cm^2), differences in

optogenetic efficacy cannot explain the range of behavioral patterns seen in the circular and Y arenas. All data are consistent with the interpretation that PAM activation bypasses the computation of reward expectation and converts a covariance rule into a non-covariance rule. In particular, learning via the non-covariance plasticity rule only modifies weights from Kenyon cells that respond to the rewarded odor, which increases the acceptance probability of the rewarded odor without changing behavior to the unrewarded odor. According to this model, performance saturates in the Y-arena because the fly repeatedly encounters the unrewarded odor by chance, and their initial tendencies for accepting the odor option never change; performance does not saturate in the circular arena because a fly that that has learned to accept the rewarded odor will stop exploring and cease to encounter the unrewarded option.

We next examined how bypassing reward expectation affects matching behavior. When tested with the same three-block matching design as earlier, but now providing a consistent reward signal via direct DAN stimulation, flies exhibited strongly diminished matching behavior (Fig. 3.12G,H). The slope of the choice-ratio versus reward-ratio plot was lower than that observed with Gr64f-driven reward and approached the flat line predicted by simulations of behavior with a non-covariance-based learning rule (Fig. 3.12I). The instantaneous choice ratio and reward ratio of an example fly (Fig. 3.12G) suggested that this flattening arises because choice ratios are never strongly biased to either odor. This is again explained by the uni-directional non-covariance rule. In agreement with this, changes in choice ratio at block transitions were much flatter with DAN reward than with Gr64f, as expected by the differences between the covariance-based and non-covariance models (Fig. 3.12J,K). To quantitatively evaluate whether providing reward with DAN activation changed the learning rule from covariance-based to a non-covariance rule, we fit our MB-inspired regression models (Fig. 3.9A) to fly data produced with DAN reward. We found that the non-covariance rule was the better fit (Fig. 3.12L,M). We found through these experiments that bypassing the computation of reward expectation changes fly choices from resembling

behavior produced by a covariance-based learning rule to behavior expected from a non-covariance rule. In particular, the results suggest this covariance-based rule is located in the fly MB and incorporates reward expectation but not sensory expectation.

Altogether, our results support the theory that covariance-based learning rules that incorporate reward expectation are necessary for operant matching. It suggests that reward expectation signal is calculated in the DANs of the fly MB and provides the first mapping of learning rules underlying operant matching onto plasticity mechanisms at specific synapses.

3.3 Discussion

The foraging strategies used by animals play a key role in their survival. Operant matching is one simple and ubiquitous behavioral strategy, utilized in dynamically changing and probabilistic environments. Despite the ubiquity of this strategy and its strong theoretical foundation, little is known about the underlying biological mechanisms. We leveraged the growing body of knowledge regarding learning in the fruit fly, and the plethora of available anatomical tools, to identify these mechanisms. We developed a foraging task that allowed us to monitor choices of individual fruitflies and showed, for the first time, that flies follow Herrnstein's operant matching law. Combining experimental results with computational modeling, we found that this behavior requires synaptic plasticity and uses a rule that incorporates expectation of reward via the rewarding PAM DANs. Our results provide the first mapping of the learning rule underlying operant matching onto the plasticity of specific synapses – the KC-MBON synapses in the MB.

3.3.1 Does the ubiquity of operant matching imply a common mechanistic framework?

When choosing between options that predict reward with different probabilities, mammals, birds and insects all follow Herrnstein's matching law(Greggers and Menzel 1993; Richard J. Herrnstein 1997; Sugrue, Corrado, and Newsome 2004b; Lau and Glimcher 2005; Tsutsui et al. 2016; Bari et al. 2019; Iigaya et al. 2019). This is clear at the global, trial-averaged level, where choice ratios are roughly equal to reward ratios, but is also true at the trial-by-trial level (Fig. 3.3A). In fact, we found that individual choices made by flies depended on choice and reward information received over multiple past trials (Fig. 3.3G-H). This is in agreement with what has been observed in mice and monkeys(Lau and Glimcher 2005; Bari et al. 2019) and suggests that these animals all make use of similar kinds of information to guide their behavior. Flies also show an increase in speed of choice when rewarded, another common signature of learnt behavior in mice and monkeys(Tsutsui et al. 2016; Bari et al. 2019) (Fig. 3.2B,C).

It is unclear whether these behavioral similarities result from underlying mechanisms that are shared across species. At its surface, mechanistic similarities seem likely. For example, neural signals that subtract reward expectation from reward – a key component of the plasticity rules underlying matching shown here – can be found in the form of a reward prediction error in many different animals(Schultz, Dayan, and Montague 1997; Berns et al. 2001). Nevertheless, such a signal on its own is not sufficient to produce matching; it needs to be incorporated into a covariance-based plasticity rule in a behaviorally relevant circuit. On the other hand, while learning values of options via synaptic plasticity is the traditional mechanistic framework thought to underlie such foraging decisions(Wickens, Reynolds, and Hyland 2003; Soltani and Wang 2006), recent work has found signatures of graded neural responses proportional to value during inter-trial-intervals, suggesting a persistent-activity-based mechanism for foraging decisions that may not require synaptic plasticity(Bari et al. 2019; Hattori et al. 2019). Associated modeling

efforts suggest matching can arise from models that don't incorporate synaptic learning(J. X. Wang et al. 2018; Loewenstein and Seung 2006).

While both synaptic plasticity and non-plasticity mechanisms can explain the observed behaviors, each makes different testable assumptions about the underlying neural architecture(Pereira-Obilinovic et al. 2022) and the effect of changing environmental conditions on the behavior. For example, if one eliminated reward baiting in our experiment, a circuit using a covariance-based plasticity rule would still give rise to behavior that follows Herrnstein's matching law. In this case, following such a law would lead the animal to always choose the option with higher reward probability. On the other hand, if matching behavior was produced using a different mechanism, the lack of reward-baiting might give rise to different strategies, such as the probability matching strategy commonly observed in mice under these conditions(Beron et al. 2022). Experiments to identify which mechanisms are used by different brains, and theoretical work to understand why, would therefore provide important insight into circuit function and the neural basis of operant matching.

3.3.2 Beyond covariance-based synaptic plasticity

Our behavioral evidence suggests that synaptic plasticity in the mushroom body depends on reward expectations through a simple covariance-based plasticity rule. We identified this plasticity rule by the process of elimination. First, we narrowed our focus to the three minimal covariance-based plasticity rules inspired by architecture of the MB. Importantly, Loewenstein and Seung showed that these rules produce matching. We then showed that only one of the three rules also explains the results of the DAN-activation experiment. It's important to recognize that more complex plasticity rules may be consistent with our data and necessary to explain future mechanistic and behavioral data. For instance, the plasticity rule could be augmented by adding

any term that averages to zero in the matching task. The plasticity rule could also be changed to involve a nonlinear function of the current synaptic weight, presynaptic KC activity, and postsynaptic MBON activity. The fundamental requirement of Loewenstein and Seung's theory is merely that the plasticity rule ultimately drives the covariance between neural activity and reward to zero.

Loewenstein and Seung's theory provides an impressively general link between operant matching and covariance-based plasticity, but it does make several assumptions that may be violated in the fly. For instance, the theory assumes that plasticity only occurs when the animal makes a choice, with weights fixed between decisions (3.4 Methods). In our current paradigm, this means that no plasticity occurs when the fly rejects an odor or otherwise explores and navigates through its environment. In contrast, DANs encode a variety of motor variables and are not locked to choice or reward (Cohn, Morantte, and Ruta 2015; Zolin et al. 2021). These motor-related DAN signals would presumably modify synaptic connections in the MB, and such off-task plasticity could generate important variability in synaptic weights and choice behavior. Interestingly, recent work has also found that these same DANs do not have a consistent effect on action-reward learning in a purely operant task (Rohrsen et al. 2021). This suggests that motor-related DAN signals are not the substrate for operant learning, and MB plasticity may specifically act to link sensory cues to rewarding actions. Further, the theory assumes that neural activity and reward are conditionally independent given choice. The MB represents reward via DAN activity, so this assumption could be violated if KC and DAN activity have correlated variability across trials that is not related to choice. Such correlations are feasible given indirect connections from KCs to DANs and the complexity of DAN activity(Zheng et al. 2018; Li et al. 2020; Zolin et al. 2021). Finally, the theory pertains to tasks where the animal decides between two options. Some animals have also been found to exhibit operant matching behavior when choosing between three or more options(Greggers and Menzel 1993; Harris and Carpenter 2011). In this setting, operant matching

still implies that the covariance between neural activity and reward vanishes, so there is hope that covariance-based plasticity rules would generate matching. However, other behavioral strategies can also lead to vanishing covariance (see 3.4 Methods - *Vanishing covariance does not imply matching between more than two alternatives*). It would be interesting to investigate whether modified learning rules can more reliably produce matching in naturalistic foraging scenarios or multi-option choice tasks.

3.3.3 Plasticity in multiple MB compartments could explain deviations from matching

One complication to the framework of expectation-based learning rules and matching is that flies, like several other animals, don't perfectly follow the matching law: rather they undermatch. Two hypotheses have been proposed to account for this deviation. The first proposes that animals that undermatch make use of a learning rule that deviates from a strictly covariance-based rule(Loewenstein and Seung 2006). Synaptic saturation and representation of motor variables in DAN response, as discussed in the previous section, offer particularly simple possibilities. Another important possibility for how this could occur is to have plasticity at multiple sites contributing to the overall learning, with different plasticity rules at each site. Indeed, the MB is divided into multiple compartments that contribute to behavior but exhibit important differences in learning(Aso and Rubin 2016; Aso, Sitaraman, et al. 2014). It is possible that some compartments make use of reward expectation in a covariance-based learning rule, while others do not. Alternatively, undermatching can also result if reward expectations are estimated over long timescales(Iigaya et al. 2019), even if all compartments made use of a covariance-based rule. This idea suggests that in a dynamic environment where baiting probabilities change quickly, the memory of past experiences acts as a bias that prevents the animal from correctly estimating the present cue-reward relationships. This is possible in the MB, as different compartments form and decay over different time scales(Aso and Rubin 2016). Whether either or both of these

hypotheses explain undermatching in flies can be studied in future experiments by manipulating different compartments of the MB circuitry and analyzing the effect of such a manipulation on undermatching. Relatedly, it would be interesting to check if animals could adapt the timescales used to estimate reward expectations to the dynamics of the behavior task.

3.3.4 An approach for inferring learning rules from behavior

Here we introduced a statistical method that uses logistic regression to infer learning rules from behavioral data. While we specifically applied our approach to infer learning rules for the fly mushroom body, the inference of learning rules is of importance to many areas of neuroscience(Lim et al. 2015; Ashwood et al. 2020; Confavreux et al. 2020). In fact, this method could be similarly applied to model other learnt behaviors in the fly and other animals. In the current work, we considered learning rules that only depended on the current sensory stimulus (KC response) and reward (DAN response), but our methodology would also generalize to the inference of learning rules that incorporated a longer time-scale history of sensory input and reward. For example, the framework would be able to estimate rules that incorporated the weighted average of recent sensory experience. The framework would also apply to the fitting of rules that depend on alternate nonlinear combinations of sensory input and reward.

However, it's important to realize that the logistic regression formalization would break down entirely for learning rules that depend on the magnitudes of synaptic weights or postsynaptic activity. Such terms would induce different nonlinear dependencies between the choice sequence and learning rule parameters, preventing us from converting these choice and reward histories into logistic regression inputs related to each component of the learning rule (see 3.4 Methods). Our approach was appropriate here because the plasticity rule in the mushroom body is not thought to involve these terms. However, many biological learning rules do depend on

postsynaptic activity and current synaptic weights, and future work should explore more flexible methodologies from modern machine learning to develop generally applicable approaches(Confavreux et al. 2020).

3.3.5 Circuit mechanisms for matching and reward expectation in *Drosophila*

We have shown that operant matching is mediated by synaptic plasticity in the fly mushroom body and involves the calculation of a reward expectation. However, the mechanisms underlying this calculation remain unclear.

The proposed mechanism underlying the calculation of reward prediction error (RPE) in mammals provides a hint at one option(Schultz, Dayan, and Montague 1997). Here dopaminergic neurons implicitly represent expectation by calculating the difference between the received reward and the reward expectation. This has been found to involve the summation of positive 'reward' inputs and negative GABA-ergic 'expected reward' inputs to the dopaminergic neurons(Keiflin and Janak 2015). MB DANs could represent reward expectation in a similar way. In fact, the recently released hemi-brain connectome(Li et al. 2020) has found many direct and indirect feedback connections from MBONs to DANs that theoretical work has shown could support such a computation(Bennett, Philippides, and Nowotny 2021; Jiang and Litwin-Kumar 2021). In the MB circuit, MBON activity is related to the expectation of reward associated with a given odor(Aso, Sitaraman, et al. 2014; Owald et al. 2015). An inhibitory feedback loop, via GABA-ergic interneuron(s) for example, could potentially carry reward expectation related information from MBONs to DANs. The negative expected reward signal from this interneuron could be combined in the DANs with a positive reward signal from sensory neurons, allowing DAN activity to represent the type of reward expectation signal needed by a covariance-based rule.

It is important to note that such a mechanism would have a major difference from mammalian RPEs. Since MBON activity is linked to the presence of odor, the reward expectation signal would vary across stimuli and only be present when the stimulus was too. Thus, this signal would not have the temporal features of mammalian RPEs. This difference in temporal structure of the reward expectation signal could explain the mixed observations from past studies aimed at identifying reward expectation in flies. For instance, a study that used temporally distinct cues and reinforcements suggested that DANs do not incorporate reward expectation(Dylla et al. 2017), while studies that used temporally overlapping cues and reinforcements did find signatures of reward expectation(Riemensperger et al. 2005; Eschbach et al. 2020), albeit with different temporal properties than the typical mammalian RPE.

It's also possible that reward expectations are incorporated into mushroom body plasticity by adjusting the levels of reward and punishment needed to achieve a given dopamine signal. In this scheme reward-related dopamine neurons could represent how much a reward exceeds expectations, and punishment-related dopamine neurons could respond when expectations are not met. This is reminiscent of the idea from Felsenberg and colleagues that interactions between reward and punishment-related compartments in the MB can guide bi-directional learning(Felsenberg et al. 2018; König et al. 2019; Adel and Griffith 2021; Gkanias et al. 2022). However, here we extend the idea by proposing that reward would not only modify KC-MBON synapses, but also modulate the baseline dopamine release or firing threshold of reward-related dopaminergic neurons. Similarly, upon missing an expected reward, learning would do the same for MBONs and DANs in punishment-related compartments. The resulting behavior would depend on balance between the activity of both reward and punishment compartments, and if the reward and punishment baselines were updated correctly, such a mechanism could produce a covariance-based rule and support operant matching. This mechanism would also tie into the

notion that phasic dopamine release (i.e. the difference of dopamine from its baseline level) mediates the RPE signal in mammals.

Future experiments can distinguish between these hypotheses. For instance, neural recordings can probe how DAN activity changes over the course of the task, and connectomics can identify other neurons in the system that may be important for the computing of reward expectation. These types of experiments are easily doable in the *Drosophila melanogaster* model. Paired with further modeling efforts and the foraging framework we developed; the fly MB promises to be a system in which we can understand decision making at a level of detail that is presently unparalleled in systems neuroscience.

3.4 Methods

3.4.1 Fly strains and rearing

Drosophila melanogaster used for behavior experiments were raised on standard cornmeal food supplemented with 0.2 mM all-trans-retinal at 25℃ [for Gr64f-Gal4 lines - see table] or 21℃ [for other lines - see following table] with 60% relative humidity and kept in dark throughout. The details of all flies used for experiments in this manuscript can be found in Table 3.1. Cross progeny (2-5 days old) were sorted on a cold plate at around 4℃ and females of the appropriate genotype were transferred to starvation vials. Starvation vials contained nutrient-free 1% agarose to prevent desiccation. Flies were starved between 28 - 42 hrs before being aspirated into the Y-arena for experiments.

Table 3.1 Fly genetic lines

Genotype	Expression target/reporter description	Bloomington stock number /Reference (if applicable)
w; Gr64f-Gal4/CyO; Gr64f-Gal4/TM3 X 20XUAS-CsChrimson-mVenus attp18	Optogenetic activation of Gr64f expressing sugar sensory neurons in behavior experiments	(Dahanukar et. al. 2007; Haberkern et. al. 2019) (Dahanukar et al. 2007; Haberkern et al. 2019) X BDSC:55134(Klapoetke et. al. 2014) (Klapoetke et al. 2014)
w; +; 58E02-Gal4 X 20XUAS-CsChrimson-mVenus attp18	Optogenetic activation of PAM cluster DAN in behavior experiments	BDSC:41347(Jenett et. al. 2012) (Jenett et al. 2012) X BDSC:55134(Klapoetke et. al. 2014) (Klapoetke et al. 2014)
w; 10XUAS-opCas9wt in ZH51C; DopR1-4gRNA JK65C/CyO::TM6B X w; +; 13F02-Gal4 attP2, Gr64LexAp65 JK73A, 13XLexAop-IVS-Syn 21-Chrimson88::tdT/TM3	Optogenetic activation of Gr64f expressing sugar sensory neurons and CRISPR mediated knockout of DopR1 receptors in KCs for behavior experiments	New Stocks - see Cloning sub-section below
w; 10XUAS-opCas9wt in ZH51C; DopR1-4gRNA JK65C/CyO::TM6B X w; +; Gr64LexAp65 JK73A, 13XLexAop-IVS-Syn 21-Chrimson88::tdT/TM3	Optogenetic activation of Gr64f expressing sugar sensory neurons without expression of cas9 protein needed for knockout of DopR1 in behavior experiments	New Stocks - see Cloning sub-section below

3.4.2 Cloning

The Gr64f promoter was PCR amplified using Q5 High-Fidelity 2× Master Mix (New England Biolabs) from the Gr64f-GAL4 (Dahanukar et al. 2007) and cloned into the FseI/ EcoRI digested backbone of pBPLexAp65 (Barret D. Pfeiffer et al. 2010) using NEBuilder HiFi DNA Assembly (New England Biolabs). Primer sequences were:

Table 3.2 Primer sequences for Gr64f

NEB_GR64f_fwd	GAGGCCCTTTCGTCTTCAAGAATTCCAGCGATTGTCTCTTAGCTGTTAAAATC
NEB_GR64f_rev	CCCCGGGCGAGCTCGGCCGGCCCCTAGGACCTGCTGGGGTAAAC

Four gRNA for the gene Dop1R1 were designed using https://flycrispr.org/target-finder/ (Gratz et al. 2014). The gRNA were then cloned into pCFD5_5 following the protocol published in Port and Bullock, 2016(Port and Bullock 2016) . Dop1R1 gRNA target sites (5'-3') were:

Table 3.3 Dop1R1 gRNA target sites

Dop1R1 gRNA 1	GACATCCAACTGCTGACAAA
Dop1R1 gRNA 2	GCTGCAGCTCACGACCGCCA
Dop1R1 gRNA 3	CGTGGAATTCGTGGAGAATC
Dop1R1 gRNA 4	ACTGGTGTGATTCCCGCCGA

Primer sequences were as in Table 3.4. Transgenic injections were performed by Genetivision using fC31 integrase mediated integration into attP dock sites. Gr64f-LexAp65 was integrated into P{CaryP}JK73A and Dop1R1 was integrated into P{CaryP}JK65C.

Table 3.4 Primer sequences for Dop1R1 4gRNA

f.PCR1-4gRNA-Dop1R1	GCGGCCCGGGTTCGATTCCCGGCCGATGCGACATCCAACTGCTG ACAAA GTTTTAGAGCTAGAAATAGCAAG
r.PCR1-4gRNA-Dop1R1	TGGCGGTCGTGAGCTGCAGCTGCACCAGCCGGGAATCGAACCC
f.PCR2-4gRNA-Dop1R1	GCTGCAGCTCACGACCGCCAGTTTTAGAGCTAGAAATAGCAAG
r.PCR2-4gRNA-Dop1R1	GATTCTCCACGAATTCCACGTGCACCAGCCGGGAATCGAACCC
f.PCR3-4gRNA-Dop1R1	CGTGGAATTCGTGGAGAATCGTTTTAGAGCTAGAAATAGCAAG
r.PCR3-4gRNA-Dop1R1	ATTTTAACTTGCTATTTCTAGCTCTAAAACTCGGCGGGAATCACAC CAGTTGCACCAGCCGGGAATCGAACCC

3.4.3 Circular olfactory arena

Group learning experiments in Fig. 3.14 were performed in a previously described circular olfactory arena(Aso, Sitaraman, et al. 2014). A schematic of the task performed in the circular arena is shown in Fig. 3.14A. OCT and MCH were used as odors for these experiments. Odors were presented sequentially (and separated in time) for one minute each, with one of the odors paired with reward. To mimic the relationship between odor time and reward time experienced by the fly in the Y-arena, 1 sec of reward (red light, 617nm, 2.3mW/cm^2), was provided after every 3 seconds of odor experience. Flies were finally tested by dividing the circular arena into four quadrants with two opposite quadrants receiving one odor and the other two quadrants receiving the other. Videos of a fly's movements in the circular arena were read into MATLAB frame by frame and the location of the fly's centroid was identified using the MATLAB image processing toolbox. Once identified, the number of flies in each quadrant was used to calculate the preference

index (PI) metric on a per frame basis. PI is defined as the difference between the number of flies in each pair of odor-matched quadrants divided by the total number of flies. Time-averaged PIs could then be calculated by taking the average of the PIs from each individual frame.

3.4.4 Logistic regression to estimate influence of past rewards and choices on behavior

To estimate the role of choice and reward histories in determining fly choices in the dynamic foraging task, we fit the following logistic regression for each fly as

$$\ln(\frac{P(C(t)=1)}{1-P(C(t)=1)}) = \sum_{i=1}^{T} \beta_i^R \cdot R(t-i) + \sum_{i=1}^{T} \beta_i^C \cdot C(t-i) + \beta_0 \ , \tag{2}$$

where t is the present trial and i is the variable used to iterate over the past T trials. $C(t) = 1$ if the chosen odor was OCT and -1 if the chosen odor was MCH. $R(t)$ = 1 if chosen OCT option produced reward, -1 if chosen MCH option produced reward, and 0 otherwise. β_0 represents the weight assigned to the bias term, β_i^C represents the weight assigned to the i^{th} past choice and β_i^R represents the weight assigned to the i^{th} past reward. We chose to look at the past T = 15 trials to align with previous studies (Lau and Glimcher 2005; Bari et al. 2019). The regression coefficients generated were 10-fold cross-validated, and the regression model included an elastic net regularization (MATLAB function - lassoglm). The weight of lasso versus ridge optimization was set to 0.1 as this value provided best behavioral fits. These fly-specific regression coefficients could be combined with the flies reward and choice histories to predict trial choice probability and estimate the log-likelihood (ℓ) and percent deviance explained(PD):

$$\ell = \sum_{i=1}^{T_{max}} \ln(\sum_{j=1}^{2} p(x_{ji}) \cdot x_{ji}) \ , \tag{3}$$

$$PD = \frac{-2 \cdot \ell_{null} + 2 \cdot \ell_{model}}{-2 \cdot \ell_{null}} \ , \tag{4}$$

where T_{max} is the total number of trials in the data being fit, i indexes trials, j indexes possible options, $p(x_{ji})$ is the probability with which the model predicts that choice j occurs on trial i, and x_{ji} is the choice that actually took place on trial i.

3.4.5 Leaky-integrator model

We also developed a leaky-integrator model to predict behavior in the dynamic foraging task inspired by earlier work (Sugrue, Corrado, and Newsome 2004b). This model determines choices on a given trial by comparing values assigned to each option the agent has to choose between based on choice and reward history.

The values (Q)were calculated for a given trial t using the following equations. If OCT is chosen by the model, values are updated according to

$$Q_{OCT}(t+1) = \tau \cdot Q_{OCT}(t) + (1-\tau) \cdot R(t) \ , \tag{5}$$

$$Q_{MCH}(t+1) = \tau \cdot Q_{MCH}(t) \ , \tag{6}$$

where τ is a constant related to the learning rate. Similarly, if MCH is chosen by the model, values are updated according to

$$Q_{OCT}(t+1) = \tau \cdot Q_{OCT}(t) \ , \tag{7}$$

$$Q_{MCH}(t+1) = \tau \cdot Q_{MCH}(t) + (1-\tau) \cdot R(t) \ . \tag{8}$$

These values are then compared and passed through a sigmoidal nonlinearity to determine the probability of each choice,

$$P(C(t) = OCT) = \frac{1}{1 + \mathrm{e}^{-\beta \cdot (Q_{OCT}(t) - Q_{MCH}(t))}} \ . \tag{9}$$

The probability of choosing MCH was one minus that of OCT. The probability generated by this operation is compared with a value drawn from a uniform distribution over the [0,1] interval to determine whether the resulting choice is OCT or MCH. These predicted choices could be compared to fly behavior to compute the model's fraction deviance explained. The parameters β and τ are fit for each fly so as to maximize the percentage deviance explained (values of these parameters can be seen in Fig. 3.3G and Fig. 3.4C,D).

3.4.6 Win - stay, Lose - switch model

A third model to predict behavior incorporated information only about the most recent choice made by the fly unlike the logistic regression and leaky-integrator alternatives. In this "win-stay, lose-switch" model the agent chooses randomly on the first trial. If the chosen option produces a reward the agent picks the option again on the next trial (stays). If it doesn't produce a reward the agent picks the other option on the next trial (switches). This procedure repeats to generate a sequence of choices. The accuracy of this model was calculated by observing correctly predicted switches and stays as well as incorrectly predicted switches as stays, shown in Fig. 3.3D as a probability matrix. To calculate this matrix the model was made to predict the behavior of flies on every trial

of the dynamic foraging task. For each individual fly four values were calculated, i) the probability that the model would predict that the fly would stay given the past outcome and the fly stayed, ii) the probability that model predicts a stay and the fly switched, iii) the probability the model predicts a switch and the fly stayed, and iv) the probability the model predicts a switch and the fly switched. The average values across 18 flies run in the dynamic foraging task is presented in the matrix in Fig. 3.3D

3.4.7 Neural circuit model of dynamic foraging

We designed a neural circuit model, inspired by work from Loewenstein and Seung (Loewenstein and Seung 2006), that was used to simulate behavior in a dynamic foraging task. Two versions of this model were used.

Replicating Loewenstein and Seung's Model

The first version aimed to directly replicate the model used by Loewenstein and Seung (Fig. 3.6A). It generated behavior on a trial-by-trial basis in the dynamic foraging task. The number of trials to simulate were input by the user prior to simulation (typically 60, 240 or 2000 trials). The model consisted of two sensory neurons (S_1 and S_2) whose activity was drawn at the beginning of each trial from a normal distribution with mean 1 and standard deviation 0.1. These neurons synapse with weights (W_1 and W_2) onto two motor neurons (M_1 and M_2) ,whose activity was given by the weighted sum of sensory neuron activity. The activity of M_1 and M_2 were compared and the choice was driven by whichever neuron had the larger activity.

Once the choice was made, rewards were provided as determined by the reward contingencies of the task. These were input by the user prior to running the simulation. The weights between S and M were updated after each choice and followed the following rules

$$\Delta W_i(t) = \eta \cdot \mathcal{R}(t) \cdot \mathcal{S}_i(t) \ , \tag{10}$$

$$W_i(t) = W_i(t-1) + \Delta W_i(t) \ , \tag{11}$$

where $\mathcal{R} = R$ or $\mathcal{R} = R - E(R)$ and $\mathcal{S}_i = S_i$ or $\mathcal{S}_i = S_i - E(S_i)$ based on the learning rule, and i iterates over odors. Note that $E(R)$ and $E(S_i)$ depended on time and were calculated in one of two ways: i) by calculating the mean over the last 10 trials, ii) by filtering the entire history with an exponential filter with exponential timescale (τ) of 3.4 trials. The various covariance and non-covariance rules were achieved by selecting the appropriate combinations of $\mathcal{R}$ and $\mathcal{S}_i$.

Task and mushroom body inspired version

The second version incorporated modifications to the model that made it more appropriate for the task we designed for fruitflies (Fig. 3.6B). This model consisted of two sensory inputs that represented activity of populations of Kenyon cells (KCs). However, this version of the model looped through odor experiences, rather than looping through trials determined by two-alternative forced choices. Therefore the activity of the sensory neurons was drawn differently. Rather than both values being drawn from normal distribution with mean 1 and standard deviation $\sigma = 0.1$, this was only true for the odor that was deemed to have been "experienced" by the model on a given odor experience. The activity of the other neuron was drawn from a normal distribution with mean $\alpha = 0.1$ (Fig. 3.7, 3.8, 3.9, 3.10) and standard deviation $\sigma = 0.1$. Here, α represents the similarity, or overlap, between the two inputs. This was included because the KC representations

of the two different odors used in our task are thought to have some amount of overlap (Robert A. A. Campbell et al. 2013). However, we found that modulating this term did not affect the resulting matching behavior (Fig. 3.10D-G) and so for Fig. 3.12, we chose $\alpha = 0$. We also explored incorporating noise covariance between the two sensory inputs (with correlation coefficient *c* = 0.1), but this correlation was empirically unimportant and we usually set *c* = 0.

Another difference is that an odor experience could lead to either an approach (choice) or a turn away. The behavior chosen by the model on any given odor experience depended on the response of the single output neuron incorporated into this model. The activity of this output neuron (M) was the weighted sum of the two inputs. This was then passed through a sigmoidal nonlinearity

$$P(A = Accept) = \frac{1}{1 + e^{\beta \cdot (M - b)}} , \tag{12}$$

where $\beta = 4$, $b = 1$(this value was chosen to encourage exploration at the beginning of learning) and A is the action produced by the model. When $A = 0$ the odor is accepted and when $A = 1$ the odor is rejected. A random number from the interval [0,1] was drawn and compared to Y to determine whether an approach/choice or turn was made. If a turn was made, no reward was provided, and the weights remained unchanged. The model then experienced a new odor and the process repeated. If a choice was made then a reward was provided based on the choice contingencies and weights were updated according to the rules in eqs. 10 and 11. The only difference here from eq. 10 and 11 is that the weights change following a depression rule instead of a potentiation rule and have a lower bound of 0.

3.4.8 Plasticity requirements of operant matching in the MB model

Relating operant matching to the covariance of neural activity and reward

We begin by reproducing the key theoretical argument provided by Loewenstein and Seung. Consider a sequence of trials where an animal chooses between two options and receives feedback via reward. Specifying an element in this trial sequence requires three random variables, the choice (C), the reward (R), and the underlying neural activity (N). Note that N is very general in this argument; it can be any quantification of neural activity whose mean depends on choice. We further assume that both options are sometimes chosen and that neural activity and reward are conditionally independent on choice. Under these assumptions, Loewenstein and Seung show that Herrnstein's operant matching law is satisfied if and only if the covariance between N and R vanishes over the trial sequence.

The proof begins by recalling the definition of covariance,

$$Cov(N, R) = E(\delta N R), \tag{13}$$

where E denotes the expectation over the trial sequence (i.e., $C, R,$ and N) and $\delta N = N - E(N)$. By the product rule of probability, $P(C, R, N) = P(C)P(R, N|C)$, so we can rewrite this expectation as

$$Cov(N, R) = E_C(E_{R,N|C}(\delta N R)) = E_C(E_{N|C}(\delta N)E_{R|C}(R)), \tag{14}$$

where the subscripts on E denote the probability distributions over which the expectations are computed, and we used the conditional independence assumption, $P(R, N|C) = P(R|C)P(N|C)$, in

the second step. Writing out the expectation over choice explicitly, the expression for the covariance becomes

$$Cov(N,R) = P(C=1)E_{N|C=1}(\delta N)E_{R|C=1}(R) + P(C=2)E_{N|C=2}(\delta N)E_{R|C=2}(R). \quad (15)$$

To simplify this expression, note that

$$E(\delta N) = E_C(E_{N|C}(\delta N)) = 0$$
$$\implies P(C=2)E_{N|C=2}(\delta N) = -P(C=1)E_{N|C=1}(\delta N), \quad (16)$$

where we again used the product rule of probability, and the $\implies$ notation means that the lefthand equation implies the righthand one. It follows that

$$Cov(N,R) = P(C=1)E_{N|C=1}(\delta N)(E_{R|C=1}(R) - R_{R|C=2}(R)). \quad (17)$$

From this expression, we can conclude that

$$E_{R|C=1}(R) = R_{R|C=2}(R) \implies Cov(N,R) = 0. \quad (18)$$

The lefthand equation is the matching law, as it says that the expected reward is independent of choice, so the matching law implies a vanishing covariance between neural activity and reward. Moreover, it follows that $E_{N|C=1}(\delta N) \neq 0 \neq E_{N|C=2}(\delta N)$ from the assumption that the average neural activity has a choice dependence (*i.e.*, $E_{N|C=1}(\delta N) \neq E_{N|C=2}(\delta N)$). Consequently, neither $P(C=1)$ nor $E_{N|C=1}(\delta N)$ is equal to zero, and

$$Cov(N,R) = 0 \implies E_{R|C=1}(R) = E_{R|C=2}(R). \quad (19)$$

This equation says that the matching law follows from the vanishing covariance of choice-related neural activity and reward. This completes the proof.

Casting the mushroom body model in the framework of Loewenstein and Seung

Our mushroom body model consists of a sequence of odor presentations and accept/reject decisions. Specifying an element in this decision sequence requires four random variables, the odor option experienced (O), the odor-induced neural activity in the KCs (S), the accept-reject decision provided by the MBON (A), and the reward received from this action (R). Reward delivery and plasticity only occur when an odor option is accepted, and we refer to the accepted odor option as the choice (C). Note that C is undefined when the odor is rejected. Therefore, specifying an element in the choice sequence only requires three random variables, S, C, and R. Setting $N = S$, this choice sequence satisfies the assumptions of Loewenstein and Seung's theory. We therefore expect flies to obey Herrnstein's operant matching law if and only if the covariance between KC activity and reward is equal to zero over the choice sequence.

It is important to recognize that the matching law is generally inconsistent with vanishing covariance between KC activity and reward over the decision sequence (rather than the choice sequence). Our assumption that plasticity only occurs following the decision to accept (*i.e.*, the choice) is thus critical for obtaining matching behavior from covariance-based plasticity rules.

Vanishing covariance does not imply matching between more than two alternatives

The preceding analyses assumed binary choices between two options. However, Herrnstein's operant matching law can also be satisfied with more than two options, and the general form of the matching law is

$$E_{R|C=1}(R) = E_{R|C=2}(R) = \cdots = E_{R|C=n}(R), \tag{20}$$

where $n \geq 2$ is the number of options. We can write this condition more succinctly as

$$E_{R|C}(R) = E(R). \tag{21}$$

Here we show that this more general form of the matching law implies that the covariance between neural activity and reward vanishes. However, the converse is not true, as it's possible for the covariance to vanish without behavior that produces the matching law. The biologically important consequence of this result is that covariance-based plasticity rules may not lead to matching when the animal is deciding between more than two options.

In this more general decision making task, we express the covariance between the neural activity and reward over the choice sequence as

$$Cov(N, R) = E_C(E_{N|C}(\delta N)E_{R|C}(R)) = \sum_{c=1}^{n} P(C = c)E_{N|C=c}(\delta N)E_{R|C=c}(R), \tag{22}$$

where we've made the same conditional independence assumption as in the binary analysis. If the matching law is satisfied, then we can take $E_{R|C=c}(R)$ out of the sum and we find

$$E_{R|C}(R) = E(R)$$

$$\implies Cov(N, R) = E(R) \sum_{c=1}^{n} P(C = c) E_{N|C=c}(\delta N) = E(R) E(\delta N) = 0. \tag{23}$$

Therefore, the matching law implies that the covariance between neural activity and reward is zero. To see that the converse need not be true, we construct a specific counter example. Consider the $n = 3$ case where $P(C = c) = 1/3$, $E_{N|C=1}(\delta N) = 2$, $E_{N|C=2}(\delta N) = E_{N|C=3}(\delta N) = -1$, $E_{R|C=1}(R) = 1/2$, $E_{R|C=2}(R) = 1$, and $E_{R|C=3}(R) = 0$. These numbers were constructed to ensure that the covariance between neural activity and reward vanishes,

$$Cov(N, R) = \frac{1}{3}((2\cdot) - (1 \cdot 1) - (1 \cdot 0)) = 0 \tag{24}$$

Nevertheless, the matching law is not satisfied because the expected reward depends on the choice.

3.4.9 Logistic regression model for estimating learning rules

To determine the learning rules that best predict fly behavior, we designed a logistic regression model that made use of the known relationship between MBON activity and behavior. This model predicted behavior between input and weights that give rise to MBON activity following the relationships

$$\hat{A}(t) = \sum_{i=1}^{2} W_i(t) \cdot S_i(t) \ , \tag{25}$$

$$W_i(t) = W_i(0) + \sum_{T|A=0} \Delta W_i(T) \ , \tag{26}$$

where $\hat{A}(t)$ is the predicted action on odor experience t, $T|A=0$ indicates all past odor experiences where the fly chose to accept the odor, and $W_i(t)$ represents the synaptic weights associated with neurons representing odor i at time t. Now the change in synaptic weights $\Delta W_i(t)$ depends on the learning rule that is used by the circuit. It was here that we wanted to have the regression model identify the rule that provided the best fit to actual data. To do this we allowed the model to use a learning rule with 4 different terms whose coefficients could be modified,

$$\Delta W_i(T) = a + b \cdot \mathcal{S}_i(T) + c \cdot \mathcal{R}(T) + d \cdot \mathcal{R}(T) \cdot \mathcal{S}_i(T) \ . \tag{27}$$

Here, a, b, c, and d are the coefficients assigned to each component of the learning rule. The regression model takes the sensory stimuli and synaptic weights at a given time as inputs to predict the output action. However, when fitting this model to behavior we have only sensory stimulus and reward information readily available. We therefore used eq. 26 and 27 to convert synaptic weights and sensory stimuli to inputs that consisted of sensory stimuli and rewards and a constant input that serves as a bias term. The resulting inputs could be represented as

$$I_0(T) = 1 \ , \tag{28}$$

$$I_1(T) = \sum_{i=1}^{2} (T-1) \cdot S_i(T) \ , \tag{29}$$

$$I_2(T) = \sum_{i=1}^{2} S_i(T) \cdot \sum_{t=1}^{T-1} \mathcal{S}_i(t) \ , \tag{30}$$

$$I_3(T) = \sum_{i=1}^{2} S_i(T) \cdot \sum_{t=1}^{T-1} \mathcal{R}(t) \ , \tag{31}$$

$$I_4(T) = \sum_{i=1}^{2} S_i(T) \cdot \sum_{t=1}^{T-1} \mathcal{S}_i(t) \cdot \mathcal{R}(t) \ . \tag{32}$$

The coefficients assigned to each of the five inputs (Bias, a,b,c,d) could then be used to identify the learning rule that the model predicted as the best estimate for producing the behavior that was tested. Of course the values of these coefficients varied from fly to fly. To examine if pairs of coefficients changed in a correlated manner across flies we estimated the correlations between the terms by using the Matlab function *corrcoef*, that produces a matrix of correlation coefficients.

Chapter 4. Input density and cell number of the Kenyon cell "expansion layer" modulates olfactory discrimination capabilities of *Drosophila*

Adapted from Ahmed, Rajagopalan et al. 2023, Current Biology 33, 1–19

Dr. Maria Ahmed and Dr. Josephine E. Clowney conceptualized the study. Dr. Maria Ahmed led a team of several co-authors in characterizing developmental manipulations and conducted calcium imaging to understand manipulations' effects on KC activity. This work is summarized in the Introduction (section 4.1) and constitutes a key part of the cited paper. Following up on these findings, Adithya Rajagopalan conceptualized the behavioral experiments used to test the effect of developmental manipulations along with Dr. Maria Ahmed, Dr. Josephine E. Clowney and Dr. Glenn C. Turner, and conducted behavioral experiments, additionally leading a team including Dr. Maria Ahmed, Dr. Kari C. Close and Dr. Christina P. Christoforou in analyzing behavioral data presented in the Results below. Dr. Maria Ahmed, Dr. Josephine E. Clowney Adithya Rajagopalan, Yijie Pan and Dr. Glenn C. Turner contributed to writing the paper listed above.

4.1 Introduction

As we introduced in section 1.5 of this thesis, the richness of the sensory world must be considered when attempting to understand the neural circuitry involved in decision-making. One important aspect of the sensory stimuli is that they very often share some similarity. The ability to discriminate sensory stimuli with overlapping features in several brain regions is thought to arise in structures called expansion layers, where neurons carrying information about sensory features make combinatorial connections onto a much larger set of cells (Marr 1969; Albus 1971; Babadi and Sompolinsky 2014; Litwin-Kumar et al. 2017). An increase in the number of expansion layer neurons is theorized to improve discrimination (Babadi and Sompolinsky 2014; Cayco-Gajic and

Silver 2019), while the number of sensory inputs is thought to have an optimal intermediate value (Jortner, Farivar, and Laurent 2007; Litwin-Kumar et al. 2017; Rajagopalan and Assisi 2020). However, the perceptual and behavioral effects of these theories have not been experimentally tested as tools for altering hard-wired circuit parameters have not been available.

To do so, our collaborators Dr. Maria Ahmed and Dr. E. Josephine Clowney's at the University of Michigan initiated a project of developmental circuit hacking. In the mushroom body calyx, individual KCs, from a population of roughly 2000, receive a median of 5–6 discrete inputs from among 52 types of olfactory PNs (Gruntman and Turner 2013; Caron et al. 2013). PN-KC connections consist of multisynaptic ‘‘microglomerular’’ structures formed by presynaptic ‘‘boutons’’ from PNs and dendritic ‘‘claws’’ from KCs (Caron et al. 2013; Zheng et al. 2018; Scheffer et al. 2020). As KCs fire only when multiple inputs are active, they act as coincidence detectors, allowing individual KCs to respond to diverse odors with the potential to expand the animal's perception from single channels to combinations. Leveraging their knowledge of MB structure and development, changes were made to the quantitative relationships between presynaptic PNs and postsynaptic KCs in vivo (Elkahlah et al. 2020; Puñal et al. 2021; Ahmed et al. 2023).

To reduce KC number, flies were treated with the mitotic poison hydroxyurea (HU) 8 hours after larval hatching. This time window was chosen because KC neuroblasts continue to divide in this period while most other neuroblasts have halted division (lateral PNs continue to divide and could be affected, so appropriate controls are considered, see section 4.2.1) allowing KCs to be specifically ablated (de Belle and Heisenberg 1994). As KCs derive from 4 neuroblasts, this approach led to flies with either 0, 500, 1000, 1500 or 2000 KC. As KC number reduced PNs reduced the number of boutons they produced leading to the PN-KC connectivity to be retained to similar levels as in the wild type. This manipulation therefore only led to a decrease in expansion

layer cell number without changing input connection density. When KC responses to odors were imaged with Ca^{2+} sensors in the ablated flies that had at least 1 surviving neuroblast, no statistically significant changes in population response to odors could be observed (Elkahlah et al. 2020), suggesting that the PN-KC connectivity adjusts when KC populations are reduced so as to maintain sparse and discriminable representations of odors.

To increase KC number the spindle orientation protein *mud* (mushroom body defect) was knocked down in mutant flies using RNA interference (RNAi) in KCs specifically. This leads some neuroblasts to divide symmetrically giving rise to two neuroblasts rather than the usual asymmetric division that produces one neuroblast and a differentiated neuron. This leads to a robust increase in KC number, sometimes more than doubling the number of KC in the mutant flies (Elkahlah et al. 2020). Again, PN-KC connections adjusted to ensure that the increased expansion layer cell number did not affect the input density. When responses to odors were observed the population showed a trend towards more sparse and discriminable representations (though this was not statistically significant) (Ahmed et al. 2023).

Next, to increase the density of inputs from PNs to KCs, *tao* a protein that regulates microtubule dynamics and dendritic branching was knocked out using RNAi in the KCs specifically. This led to a 50% increase in KC claws suggesting a large increase in input density. It must be noted that 25% decreases in KC number was also observed along with this manipulation, but given the fact that no odor representational changes were observed upon reducing KC numbers with the HU ablation this likely does not contribute to any affects observed with *tao* RNAi. Significant changes were observed when the activity of KCs were imaged. Twice as many KCs responded to odors compared to controls, with 40% of cells responding to all 4 odors that were tested. This would suggest that these flies likely cannot use the KC representation to discriminate odors and would perform poorly at a behavioral discrimination task.

Finally, to reduce the density of inputs to the KCs, DScam a protein previously shown to change calyx morphology was over-expressed in the KCs (J. Wang et al. 2004), and was found to reduce KC claw number by more than 75% with the median number of claws per KC reduced to 1. When these flies were imaged, KCs showed a significantly reduced response profile with 60% of cells responding to none of the odors tested.

These living animals, produced with hypothesis-driven variations on natural expansion layer wiring parameters show KC responses that are in line with the theoretical work that predict increased discrimination (equivalent to sparser coding at the KC level, with more cells responding to only single odors) with increased KC number, a reduced discrimination with changes in PN-KC connectivity. But do these KC level changes transfer to behavioral changes?

In this chapter, we described work that leverages the novel Y-arena described in chapter 2 as well as commonly used group learning circular olfactory arena to ask exactly this question. One of the advantages of using the Y-arena in this case is that it is a single fly assay. The KC cell number manipulation manipulations described above, particularly mud RNAi, cause variable changes to the KC circuit and so correlating behavior to the extent of change in KC number is important to quantify the behavioral effect of the change. The input density manipulations on the other hand produce more reliable changes and their behavioral effect is more easily tested in a high throughput group learning assay (4.4 Methods).

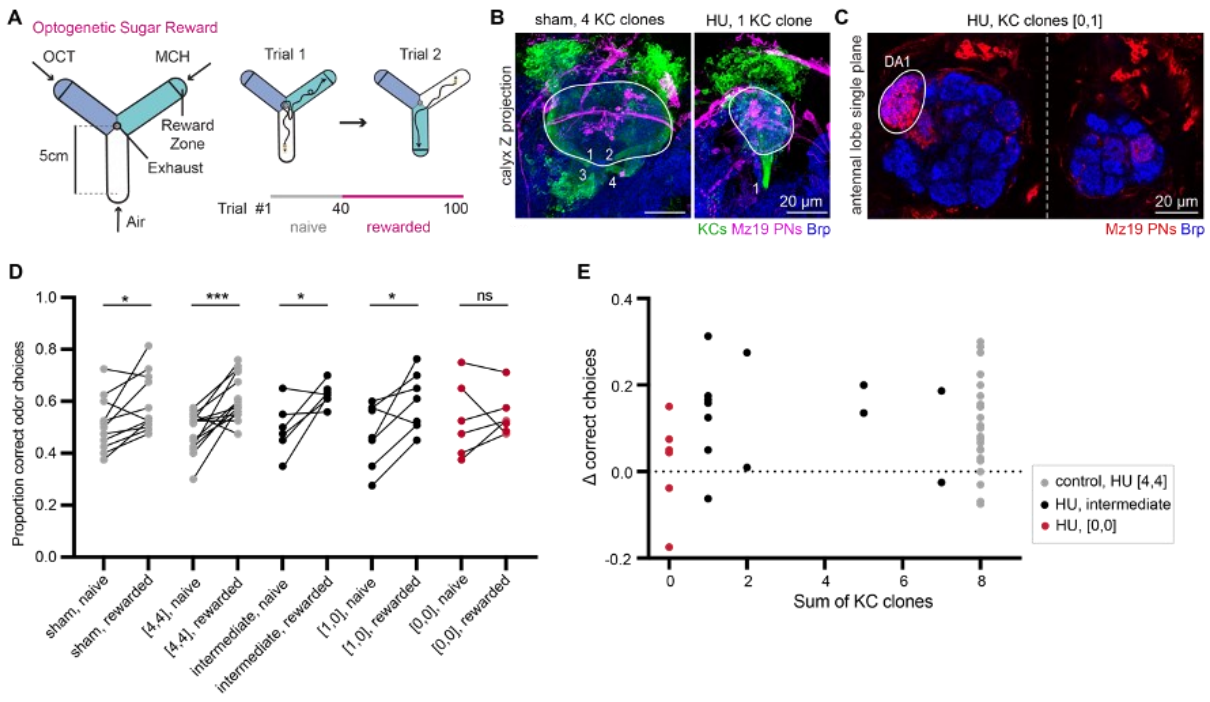

Figure 4.1 Associative learning and feed-forward functional responses in flies with reduced Kenyon cell numbers. (A) Schematic of Y-arena and 2AFC task. Entering the reward zone is considered a choice, then the two odors are randomly re-assigned to the arms for the next trial. During rewarded trials, entering reward zone for a specific odor triggers a 500 ms pulse of red light (optogenetic reward). (B) Maximum intensity Z-projections of calyces from sham- and HU-treated animals. 58F02 labels αβ core KCs (green). Mz19 labels a subset of ~16 PNs (magenta). Brp (blue) marks synapses. White outline: calyx. Numbers count 58F02^{+} neurite bundles (i.e. KC clones) innervating the pedunculus when traced through the stack. (C) Single slice of the antennal lobes from a HU-treated brain with 1 KC clone in the left hemisphere, and no clones in the right hemisphere. Dotted line: midline of the central brain. Mz19 labels the DA1 glomerulus (circled), allowing scoring of the lNB/BAlc PN neuroblast; it is ablated in the right hemisphere and DA1 glomerulus is lost. (D) Proportion of correct odor choices in naïve and rewarded trials. Each data point is one fly. Sham-treated and unaffected HU-treated animals are displayed in gray, intermediate HU-treated clone numbers in black, and fully ablated "[0,0]" HU-treated animals in red. Jitter added in this plot and (E) to display all the data points. Significance: paired t-test. See also Fig. 4.2B. (E) Relationship between Δ correct choices between rewarded vs naïve trails, and sum of KC clones from both hemispheres (0-8). No correlation was observed. See Fig. 4.2C-D for relation with lNB/BAlc ablation.

4.2 Results

4.2.1 Animals with reduced KC numbers learn equally well as controls

We tested individual HU-treated flies in the Y-maze two alternative forced-choice task (2AFC) described in chapter 2, assaying preference for 3-octanol versus 4-methylcyclohexanol in the 100:0 protocol, before and after one odor was randomly paired with optogenetic reward via stimulation of protocerebral anterior medial (PAM)-DANs (Fig. 4.1A). Animals performed 40/60 choices for the naive and rewarded trials respectively. Each fly was dissected post-hoc and immunostained (Fig. 4.1B). This allowed us to take advantage of the range of phenotypes produced by our manipulations, instead of merging animals with heterogeneous circuit anatomy in a group assay. We labeled the latest born ‘‘ab core’’ KCs, allowing us to score KC clone number by counting the groups of labeled soma or axon tracts. As HU ablation sometimes affects the neuroblast that gives rise to lateral PNs, we included the Mz19 marker to track the lateral DA1 glomerulus (Fig. 4.1C) (Stocker et al. 1997; Lai et al. 2008). Almost all animals retained the lateral PN neuroblast in at least one hemisphere (Fig. 4.2C, D), whereas the number of KC clonal units varied (Fig. 4.2E).

With PAM-DAN reinforcement, sham-treated control animals chose the rewarded odor in 59% of trials on average, compared with 51% prior to training (Fig. 4.1D, Fig. 4.2B; the reasoning for this low performance with PAM-DAN reinforcement is detailed in section 3.2.5). HU-treated but unaffected animals (‘‘[4,4]’’ set) learned similarly to controls, whereas animals with all 8 clones lost (‘‘[0,0]’’) failed to learn (Fig. 4.1D, Fig. 4.2B). Flies with at least one remaining KC clone (‘‘[1,0]’’) were still able to increase their preference for the odor paired with PAM-DAN reward. As a result, learning performance did not correlate with number of KC clones (Fig. 4.1E). This is

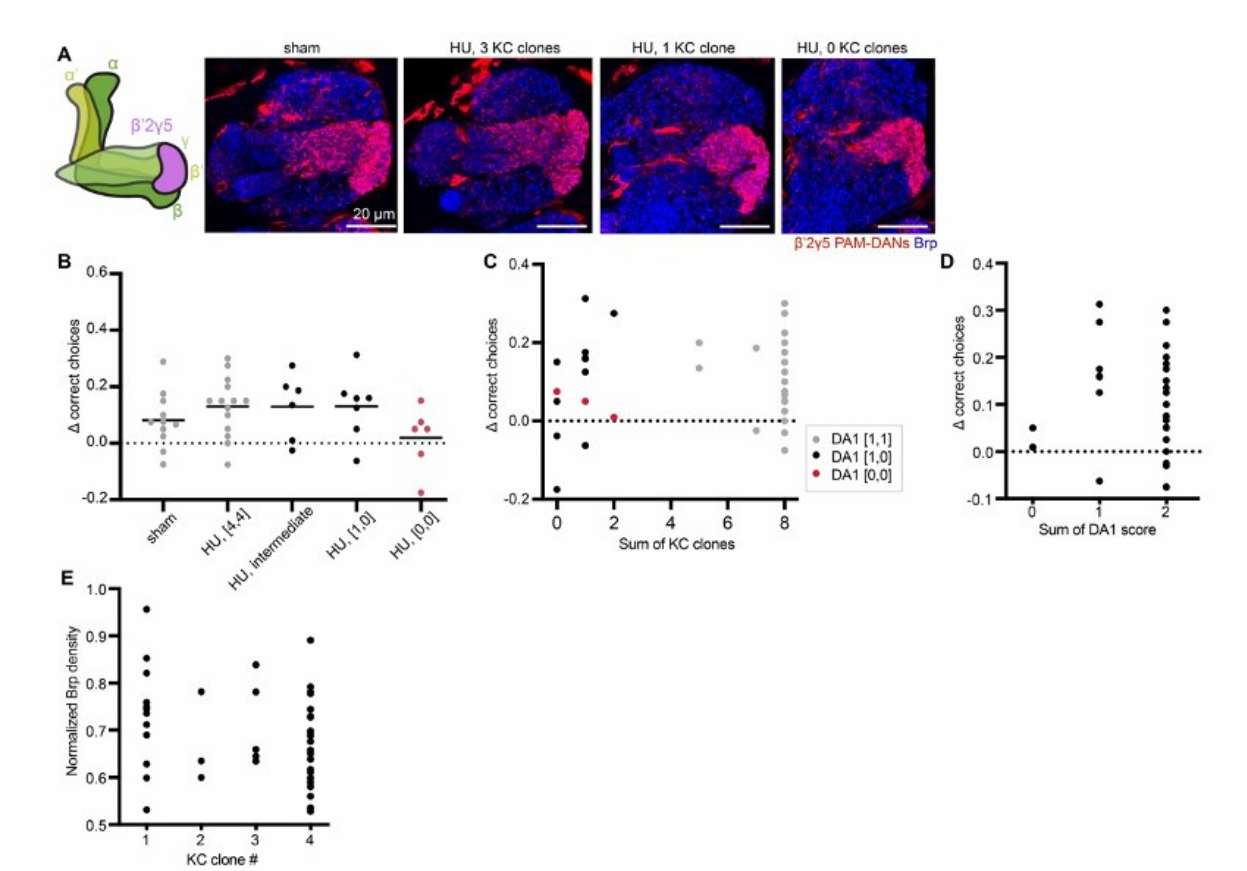

Figure 4.2 Learning as a function of PN and KC clone variability in flies with reduced KC number. (A) Left: Schematic of the mushroom body lobe anatomy with KCs in green and β'2γ5 PAM-DANs in magenta. Axons of β'2γ5 DANs in the lobe compartment are shown. Right: Single confocal slices of the MB lobe (identified by location and Brp staining shown in blue). Mz19 driver labels β'2γ5 DANs (red). Representative images shown of sham-treated and HU-treated hemispheres with 3, 1, or 0 KC clones. (B) Δ correct choices of sham-treated and HU-treated animals shown in Fig. 4.1D. Black bars indicate the medians. In B-D, each data point is an individual fly. (C) Relation of Δ correct choices, sum of KC clone number from both hemispheres and lNB/BAlc ablation status. DA1 present in both hemispheres is indicated as "DA1 [1,1]", presence in one hemisphere is indicated as "DA1 [1,0]", and absent in both hemispheres is indicated as "DA1 [0,0]". (D) Relation of Δ correct choices to sum of DA1 score (presence/absence). The data shown excludes fully KC ablated animals. Jitter added in (C-E) to display all the data points. (E) Relation of normalized Brp density in the mushroom body calyx to Kenyon cell clone number in sham-treated and HU-treated animals, excluding fully ablated animals as there is no calyx present.

consistent with previous results that indicate that as few as 25 KCs can be adequate for computationally distinguishing different odors (Robert A. A. Campbell et al. 2013) as well as

behavioral effects of partial mushroom body loss in honeybees (Malun et al. 2002). This supports our previous findings that population-level odor responses in these animals are similar to controls (Elkahlah et al. 2020). We note that the odor choices we used for this task were very distinct. Because PAM-DAN optogenetic reinforcement produced relatively weak learning scores in control animals, we did not attempt to subject KC-ablated animals to more difficult discrimination tasks.

4.2.2 Animals with increased KC numbers learn better than controls

We next tested the learning capabilities of individual OK107>mud RNAi flies, with expanded KC repertoire. As with HU-ablated flies, mushroom body anatomy was evaluated by post-hoc staining. In this set of experiments, we provided optogenetic reward using Gr64f+, sugar-sensing gustatory neurons, which improved learning in controls (Fig. 4.3A; discussed in section 3.2.5). Both controls and mudRNAi flies showed robust, odor-specific learning (Fig. 4.3B,C). In agreement with the trends observed when KC populations of mud RNAi flies were imaged, learning improved with increasing calyx size in the mud RNAi animals, suggesting that increasing the size of the KC repertoire gives animals more power to discriminate odors, form learned associations, or use learned associations to guide decisions (Fig. 4.3D).

4.2.3 Fine odor discrimination of highly similar odors is reduced in flies with increased input connection density

When we tested the ability of the tao RNAi mutant flies with increased input connection density, we observed robust and selective odor-reward associations (Fig. 4.3B,C). This is surprising, as the weakened KC odor selectivity in these animals should make odor discrimination more difficult. One possibility is that these two odors are still sufficiently different at the level of KCs to be decodable.

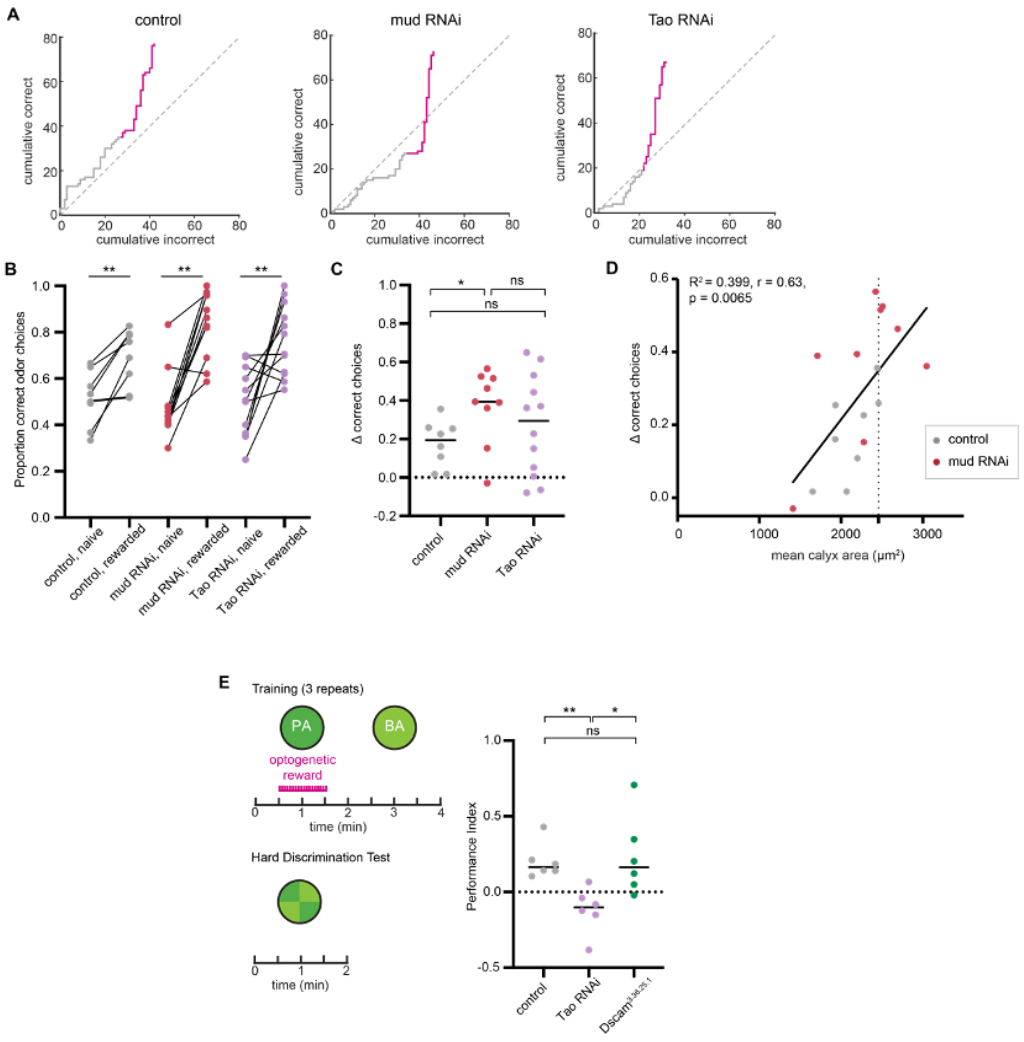

Figure 4.3 Effect of increased Kenyon cell number or claw number on associative learning and feedforward functional responses. (A) Example learning curves of individual flies from control, KC>*mud* RNAi, and KC>*Tao* RNAi are plotted to show cumulative correct and incorrect odor choices across all trials. Grey and pink lines indicate naïve and rewarded trials respectively (40 or 60 trials each). Dotted line (y=x) displays equal preference. (B) Proportion of correct odor choices made in naïve and rewarded trials, in control (gray), KC>*mud* RNAi (red), and KC>*Tao* RNAi (purple) animals. Each data point is an individual fly. Jitter added in this plot and (C) to display all the data points. Significance: paired t-test. (C) Difference (Δ) in correct choices made in rewarded vs naïve trials in control (gray), KC>*mud* RNAi (red), and KC>*Tao* RNAi (purple) animals. Significance: unpaired t-test. (D) Δ correct choices plotted against inter-hemisphere mean calyx area for each animal in control and KC>*mud* RNAi conditions. Black line: linear fit. Dotted line: largest control calyx. (E) Left: Schematic of circular arena and hard discrimination task. In training, one of the odors (PA: pentyl acetate or BA: butyl acetate) is paired with an optogenetic reward. Right: Change in performance index is calculated by comparing distribution of animals in rewarded-odor quadrants to their distribution before odor onset (4.4 Methods). Each dot is a group assay. Behavior traces over time provided in Fig. 4.4.

As Tao RNAi animals have relatively consistent calyx anatomy, we used a group learning assay, the circular arena (Aso and Rubin 2016). We chose two chemically similar odors that activate overlapping sets of KCs, pentyl acetate and butyl acetate. During the learning phase, one odor was paired with optogenetic reward via Gr64f neuron stimulation (Fig. 4.3E, top). In test trials, the distribution of animals in the quadrants of the arena was measured before and after the odors were presented (Fig. 4.3E, bottom). Upon odor onset, control animals redistributed into quadrants containing the rewarded odor, demonstrating that they can discriminate it from the similar odor in the other quadrants, as observed previously. Animals with excess claws per KC were distributed randomly among odor quadrants (Fig. 4.3E). The chance-level performance of excess-claw, *tao* RNAi animals is consistent with the expectation that increased representation overlap will degrade discrimination behavior and with the theoretical work that shows that increasing connectivity beyond an optimal value worsens discrimination (Jortner 2013; Litwin-Kumar et al. 2017).

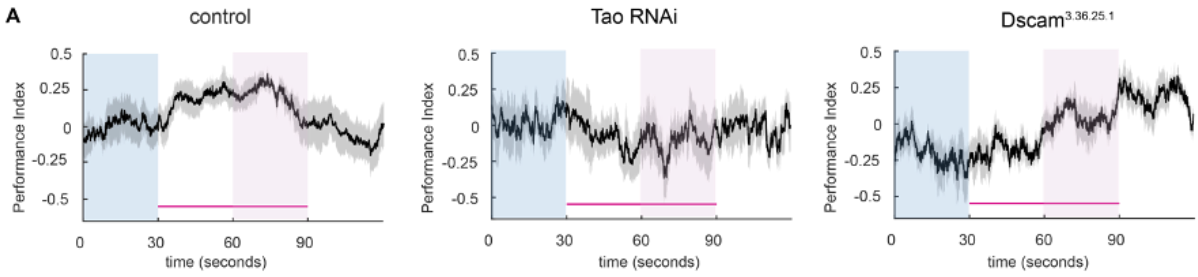

Figure 4.4 Time-course of behavior in animals with excess Kenyon cells or altered Kenyon cell dendrites. (A) Mean and standard error of group performance indices over time in the circular arena “hard discrimination” assay. Positions of all animals were monitored for the 30 seconds prior to odor onset (blue), at t=30s. Odor was presented for one minute (red bar, t=30-90s). Performance index at each time point is calculated as ((animals in quadrants assigned to paired odor)-(animals in quadrants assigned to unpaired odor))/total animals. In Fig. 4.3E we subtract the animals’ time-averaged PI before odor onset (blue shading) from the time-averaged PI during odor test (t=60-90s, pink shading). n=5-6 groups per genotype of 15-20 animals each.

4.2.4 Fine odor discrimination of highly similar odors is similar to controls but requires more time in flies with decreased input connection density

Animals with fewer claws per KC, via DScam over-expression were also tested in the group choice assay described above for tao RNAi flies. These animals moved into quadrants containing the rewarded odor in proportions similar to controls (Fig. 4.3E). However, they showed delayed choice kinetics (Fig. 4.4). The slow responses of these animals with diminished claws suggest these animals may have dampened odor detection. This could be explained by the fact that very few KCs respond in these animals. However, the cells that do respond typically respond to a limited number of odors and could be sufficient for discrimination. If these features are thought of in the context of a drift diffusion like model of decision making introduced in chapter 1. The amount of sensory information coming in is small as very few cells are responding but the information is still appropriately biased towards the rewarding odor. So, the DDM still leads to correct decisions but requires more time to reach the decision threshold.

4.3 Discussion

Here, we have initiated a new method, developmental hacking of circuit wiring, to test how specific circuit parameters of the arthropod expansion layer influence cognitive computations. Using chemical and genetic approaches, we increase and decrease KC number and KC input density. We find that changing KC number, and thus expansion ratio, improved two- choice odor learning of animals with larger numbers of KCs, and did not affect the learning of animals with diminished KC repertoires. These findings suggest that the developmental algorithms we identified allow nervous systems to be computationally robust. Second, we confirm the Marr-Albus theory in vivo. We find that flies with lowered input connectivity to the expansion layer remain adept at two-choice odor learning if the stimuli are very distinct. However, these animals fail at discriminating

chemically similar odors, whereas animals with fewer claws per KC were able to achieve this discrimination. This highlights the significance of having a low overlap of odor representations in KCs in order to allow odor discrimination at the behavioral level. Future behavioral analyses of animals with these diverse mushroom body architectures will allow us to assess in detail the effects on sensory sensitivity, discrimination, generalization, and memory storage.

4.4 Methods

4.4.1 Circular arena behavioral task structure and design

Groups of approximately 15-20 females, aged 4-10 day post-eclosion were anesthetized on a cold plate and collected two days prior to experiments. They were transferred to starvation vials containing nutrient-free agarose. Starved females were trained and tested at 25 C at 50% relative humidity in a dark circular arena described in (Aso and Rubin 2016). The arena consisted of a circular chamber surrounded by four odor delivery ports that divide the chamber into quadrants. The input flow rate through each port was 100 mL/min, which was actively vented out a central exhaust at 400 mL/min. Odors were pentyl acetate (PA), butyl acetate (BA) (Sigma-Aldrich product numbers 109584 and 287725, respectively). These odors were diluted by an odor delivery system which utilizes air dilution of saturated odorant vapor and delivered odors at a 1:16 dilution of saturated vapor. Flies were aspirated into the arena via a small port and allowed 60s to acclimatize before training commenced. Training consisted of exposing the flies to either PA or BA while providing optogenetic stimulation via a square array of red LEDs (617 nm peak emission, Red-Orange LUXEON Rebel LED, 122 lm at 700mA) which shone through an acrylic diffuser to illuminate flies from below. LED activation consisted of 30 pulses of 1s duration with a 1s inter-flash interval, commencing 5s after switching on the odor valves and terminating 5s after valve shut-off. The other odor was then provided for 60s without LED activation. Three training repeats, separated by 60 seconds, were used. Following training, testing was carried out with PA and BA.

In the test configuration, the two different odor choices are presented in opposing quadrants for 60s. Videos of fly behavior were captured at 30 fps using BIAS (http://archive.iorodeo.com/content/basic-image-acquisition-software-bias.html) and analyzed using custom-written code in MATLAB.

4.4.2 Behavior analysis

Y-arena analysis is described in chapter 2.

Circular arena videos recorded during the test phase were analyzed using custom-written MATLAB code. The centroid of each fly was identified and the number of centroids in each quadrant computed for every frame of the experiment. For discrimination experiments, a Performance Index (PI) was calculated as the number of flies in the quadrants containing the paired odor minus the number in the quadrants with the unpaired odor, divided by the total number of flies.83 This value was calculated for every frame of the movie, and the change in average PI values over the final 30 s of the test period.

4.4.3 Image analysis

We analyzed males for immunohistochemistry in our manipulations. For functional imaging experiments, we used mixed-sex populations, and did not observe any correlation with sex (not shown). The Y-arena and circular arena behavior used females due to the size of the arena not being optimal for males. Therefore, the flies dissected post-behavior and used for calyx area quantification in those animals and Brp staining were also females. Sex differences in the fly are well-documented, and anatomic and physiologic sex differences have not been observed in the mushroom body (Brovkina et al. 2021; Clowney et al. 2015). Any brains that appeared damaged

from dissections, or those with the mushroom body region obscured due to insufficient tracheal removal, were not included in the analysis. Researchers performing quantification could not generally be blinded to experimental condition due to the overt changes in neuron numbers and brain structures induced by our manipulations. However, analysis was performed blind to the goals of the experiment when possible, and quantitation of features on the anterior and posterior sides of the brain were recorded independent of one another and merged after all quantifications were completed. Moreover, many of our analyses make use of variation within an experimental condition or genotype, providing an additional bulwark against observational bias.

Calyx area

To measure the size of the mushroom body calyx, we used genetically encoded fluorescence driven in Kenyon cells by OK107-Gal4. In cases where the fluorescent marker was not added, we used markers such as ChAT to visualize the structure. We then identified its largest extent in Z (i.e. along the A-P axis), outlined it in FIJI and calculated the cross-sectional area using the 'Measure' command. We have previously shown the calyx area to positively correlate with KC number and hence, serve as a readout for KC number in the hydroxyurea-treated, KC-reduced and mud RNAi-driven, KC-increased conditions (Elkahlah et al. 2020).

Kenyon cell numbers

To count Kenyon cells, we again used genetically encoded fluorescence driven by OK107-Gal4. We counted labeled somata in every third slice in the stack (every third micron along the A-P axis), with reference to DAPI to distinguish individual cells from one another. We initially determined that somata in slice 0 could also be seen in slices 2, –1, +1, and +2 but not in slice 3 or +3. To avoid doublecounting, we therefore counted every third micron. For Fig. 4, Kenyon cells were counted using a cellular segmentation tool called Cellpose (Stringer et al. 2021). The confocal stack for each hemisphere was split into single planes every third micron, and those

slices were cropped to where the KC somata are present and given as input image into Cellpose. The count from each slice was then summed up to get the total count. The software used the Kenyon cell fluorescence, nuclear signal (DAPI) in the Kenyon cells, and an automatically calibratedcell diameter to identify individual cells. We verified that the counts obtained matched manual counts.

Kenyon cell and projection neuron neuroblast state

For quantifying Kenyon cell neuroblast state,as was done previously (Elkahlah et al. 2020), we used 58F02 to fluorescently label late-born Kenyon cells and counted clumps of labeled somata surrounding the calyx as well as groups of labeled neurites leaving the calyx and entering the pedunculus. These estimates usually matched; in the few cases where they did not, we used the number of axon clumps, as somata are closer to the surface of the brain and more susceptible to mechanical disruption during dissection. This resulted in numbers between 0-4 for each hemisphere. Any additional labeling of neurons by 58F02 was easy to discriminate from KCs as those neurons did not enter the calyx or pedunculus. For scoring presence of the PN neuroblast (PN lNB/BAlc) in sham-treated and hydroxyurea-treated animals, we included Mz19 driving a fluorescent reporter in PNs from this lineage.128,129 In the absence of lNB/BAlc, 12 glomeruli lose their typical PN partners; 40 glomeruli are innervated by the anterodorsal PN neuroblast, which is not affected by HU ablation.130,131 Mz19 labels PNs that innervate DA1 and VA1d glomeruli on the anterior side, and DC3 on the posterior side of the antennal lobe; DA1 is innervated by lNB/BAlc PNs.128 We quantified the presence or absence of the most distinctly labeled glomerulus – DA1, as a way to score the PN lNB/BAlc getting ablated. In some cases, we observed mislocalization of DA1. For simplification, we scored this in the ‘‘absence’’ Category.

Bruchpilot density

We measured Bruchpilot intensity in the calyx as a readout of synaptic density. First, we identified the Z plane with the largest extent of the calyx in the A-P axis, and then took three measurements of the average fluorescence signal in the Brp channel in a defined ROI measuring 40 mm2. The three measurements were taken randomly in different locations in the calyx to account for any variability in intensity. For normalization, the calyx Brp signal was divided by Brp signal measured in an unmanipulated brain region, the protocerebral bridge. We chose this brain region as it was the closest to the calyx and was in the field of view in all calyx images taken. These quantifications were done in the ‘‘63X hi-res-1" and ‘‘63X hi-res-2’’ sets of images (defined in the ‘‘Immunostainings’’ section above).

Chapter 5. Flexible discrimination or generalization between similar odors in *Drosophila* depends on the presented alternative

Adapted from Modi, Rajagopalan et al. 2023, eLife 12:e80923

The study described in the publication referenced above involved a cohesive program of behavioral and neural imaging experiments that fed back on each other, directing how the project was conceptualized. Dr. Mehrab Modi, Adithya Rajagopalan, Dr. Yoshi Aso and Dr. Glenn C. Turner conceptualized and Adithya Rajagopalan conducted the initial behavioral experiments (described in Results section 5.2.1) Dr. Modi and Dr. Turner conceptualized and Dr. Modi performed the imaging experiments that followed up on these behavioral results (described briefly in section 5.3). Adithya Rajagopalan then conducted follow up behavioral experiments to test hypotheses formed from imaging data (described in Results section 5.2.2). Adithya Rajagopalan and Dr. Mehrab Modi conducted the analysis of the behavioral data.

5.1 Introduction

In chapter 4 we discussed one aspect of the complexities that arise when animals must make decision in natural settings that involve sensory stimuli with overlapping representations, i.e. the neural principles that allow these stimuli to be represented distinctly. But as we mentioned in the Introduction section 1.5, animals do not always want to discriminate between similar sensory stimuli. It is often beneficial for the brain to maintain overlapping sensory representations and associated values, flexibly categorizing options into the same or different groups based on the available options and tailoring decision-making processes as a result of the categorization.

This flexibility can be studied by examining how animals use an associative memory in two different tasks: discrimination and generalization (Mackintosh 1974). In a discrimination task, the animal has to choose between a cue associated with reward and a second cue that could either be similar to (hard discrimination) or distinct (easy discrimination) from the original cue. In the generalization task, the flies have to choose between a cue that is perceptually similar to the trained cue and a cue that is very different. The correct choice in this task depends on the animal generalizing its learned response to the similar cue. So, the response to the perceptually similar cue differs between the two tasks - the animal chooses it when generalizing and chooses against it when discriminating. Despite the need to switch choices, performance can be extremely high on both these types of tasks (Xu and Südhof 2013), suggesting that comparisons between available alternatives have a strong impact on animals' behavioral responses.

We asked if Drosophila can incorporate such choice dependent flexibility into their decision-making process, employing these paradigms using aversive olfactory conditioning. If flies can perform these behaviors, the well-studied MB circuit could provide a strong framework to understand the neural basis of flexible categorization.

5.2 Results

5.2.1 Flies perform flexible discrimination and generalization depending on the available options in a MB compartment dependent manner

Previous work has shown that flies are capable of high levels of performance on both hard discrimination and generalization tasks (R. A. A. Campbell et al. 2013). This study identified a trio of odors to use for experiments on the specificity of memory, based on the degree of overlap of KC response patterns: pentyl acetate (PA) butyl acetate (BA) and ethyl lactate (EL) (Fig. 5.1A,

left). PA and BA are chemically similar and elicit highly overlapping response patterns in the KC population (R. A. A. Campbell et al. 2013). EL is distinct, both chemically and in terms of KC response patterns. Choices between different combinations of these cues can be used to test flies' ability to flexibly classify odors and measure memory specificity. Take, for example, an experiment where flies are trained to form an association with PA. We can present flies with a difficult discrimination task by giving them a choice between the similar odors (PA and BA), or an easy discrimination with a choice between the paired odor (PA) and the dissimilar odor (EL) (Fig. 5.1A, right). We can also test whether the association with PA generalizes to the similar odor BA, by giving flies a choice between BA and EL. Since we use these odors in many different combinations for different task structures, with and without reciprocal design, here we will use A to refer to the paired odor (PA or BA) and A' to refer to the other similar odor, which is unpaired, while B always refers to the dissimilar odor, EL. With this nomenclature, hard discrimination involves an A versus A' choice, easy discrimination is A versus B and generalization is A' versus B (Fig. 5.1A).

Although previous work showed flies can flexibly categorize odors and learn both generalization and discrimination tasks using these odors, electric shock was used as the reinforcement (R. A. A. Campbell et al. 2013). Consequently, the synaptic changes responsible were likely distributed across many areas of the mushroom body, and possibly elsewhere. To confine plasticity to a more restricted region of the brain, we used optogenetic reinforcement, pairing the activation of specific DANs with odor presentation (Fig. 5.1B)(Schroll et al. 2006; Claridge-Chang et al. 2009). We used drivers to express CSChrimson in specific DANs from the PPL1 cluster that target different compartments involved in aversion learning: α3 (MB630B) and γ2α'1 (MB296B)(Aso, Hattori, et al. 2014; Aso and Rubin 2016). Since compartments have different time courses for memory acquisition and recall (Aso and Rubin 2016), the number of repetitions of odor-

reinforcement pairing and the time between training and testing differed depending on the compartment tested (see 5.4 Methods).

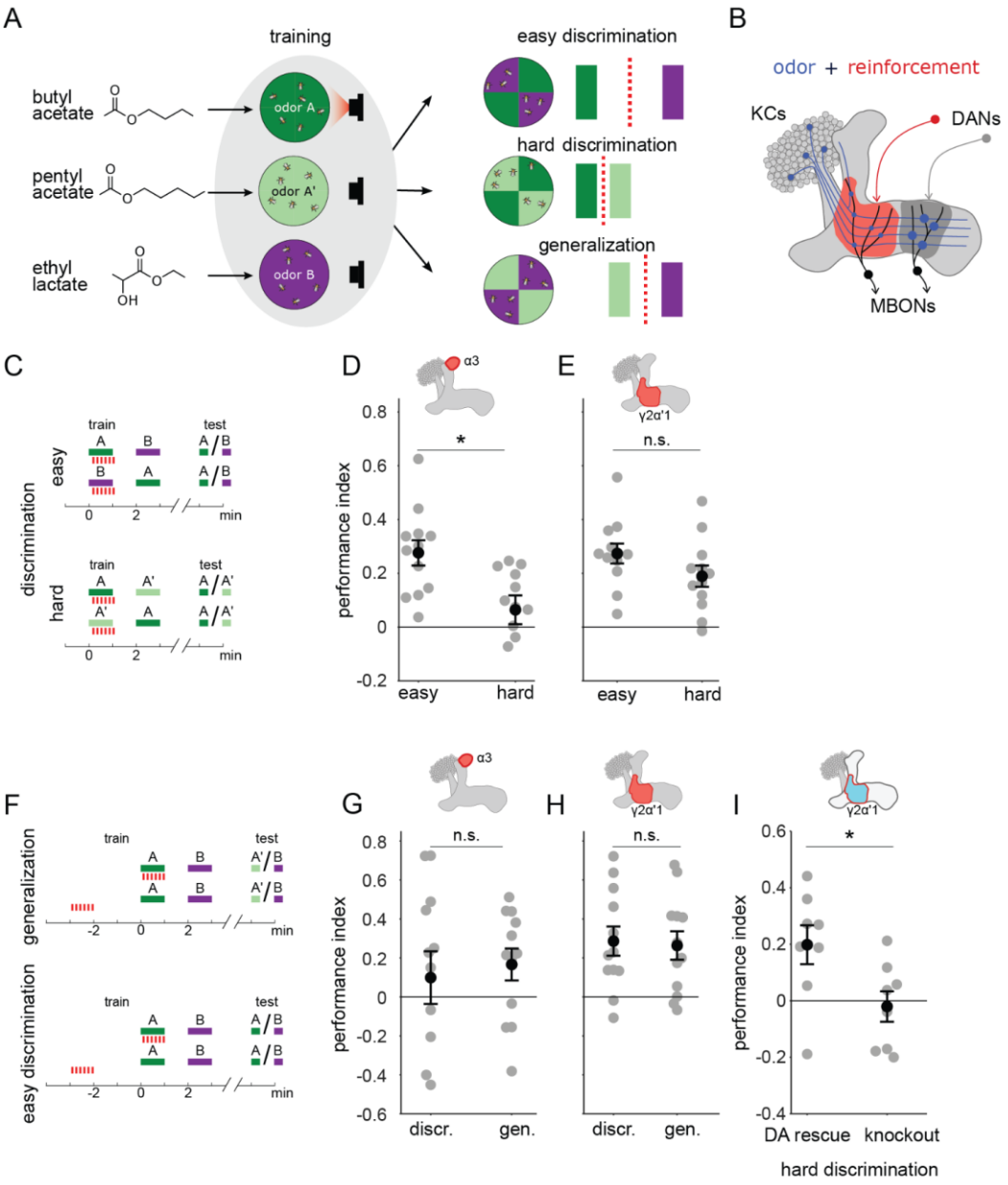

Figure 5.1 A single set of synapses can result in generalization or discrimination. (A) Left: Chemical structures of the three odors used in the study, the similar odors butyl acetate (BA) and pentyl acetate (PA) and the dissimilar odor ethyl lactate (EL). Middle: During training the similar odors are interchangeably used as the odors that are paired (A) or unpaired (A') with optogenetic reinforcement (LED). Right: Trained flies are then given one of three different choices between odors in opposing arena quadrants. These choices represent the three kinds of tasks used here

to study memory specificity. Performance index measures the bias in the distribution of flies across the different quadrants (see Methods). The circles depict fly population behavior in our arenas and the vertical bars depict stimulus choices. The dashed, red line depicts the discrimination boundary in each choice. This boundary shifts relative to the light-green stimulus, depending on the options. (B) Mushroom body learning schematic. KCs activated by an odor (blue) form synapses on MBONs in two compartments (red and gray shading). Reinforcement stimulates the DAN projecting to one compartment (red) leading to synaptic depression. (C) Behavior protocols for discrimination tasks at two levels of difficulty. Colored bars represent odor delivery periods, red dashes indicate LED stimulation for optogenetic reinforcement. A represents the paired odor, A' the similar odor and B the dissimilar odor. (D) Significantly lower performance on the hard discrimination task with reinforcement to α3 ($p = 0.007$, $n = 12$). Flies received 10 cycles of training and were tested for memory 24 hours later. CsChrimson-mVenus driven in DAN PPL1-α3 by MB630B-Gal4. (E) No significant difference in performance on easy versus hard discrimination with reinforcement to γ2α'1 ($p=0.08$, $n = 12$ reciprocal experiments). Flies received 3 cycles of training and were tested for memory immediately after. CsChrimson-mVenus driven in DAN PPL1 γ2α'1 by MB296B-Gal4. (F) Behavior protocol for generalization. Scores here are compared to a control protocol where light stimulation is not paired with odor presentation in time. (G) No significant difference in performance on generalization and easy discrimination with reinforcement to α3 ($p = 0.84$, $n = 12$). Flies received 10 cycles of training and were tested 24 hours later. (H) No significant difference in performance on generalization and easy discrimination with reinforcement to γ2α'1 ($p = 0.89$, $n = 12$ unpaired control performance scores). Flies received 3 cycles of training and were tested immediately after. (I) Rescue of the dopamine biosynthesis pathway in DAN PPL1-γ2α'1 is sufficient for performance on the hard discrimination task ($p = 0.04$, $n = 8$). Black circles and error bars are mean and SEM. Statistical comparisons made with an independent sample Wilcoxon rank sum test.

We found that these two compartments exhibited contrasting properties in the easy and hard discrimination tasks (Fig. 5.1C). Flies that received reinforcement from DAN PPL1-α3 were poor at the hard discrimination although they performed significantly better on the easy task (Fig. 5.1D, $p = 0.007$, $n = 12$). On the other hand, flies that received optogenetic reinforcement via DAN PPL1-γ2α'1 performed the hard discrimination as effectively as the easy discrimination (Fig. 5.1E, $p = 0.08$, $n = 12$). Empty driver controls performed no better than chance at either easy or hard discrimination (5.2A, $p = 0.052$, $p = 0.38$, $n= 12$). These results show that these two compartments have different capacities for discrimination, with α3 weakly discriminating and γ2α'1 stronger.

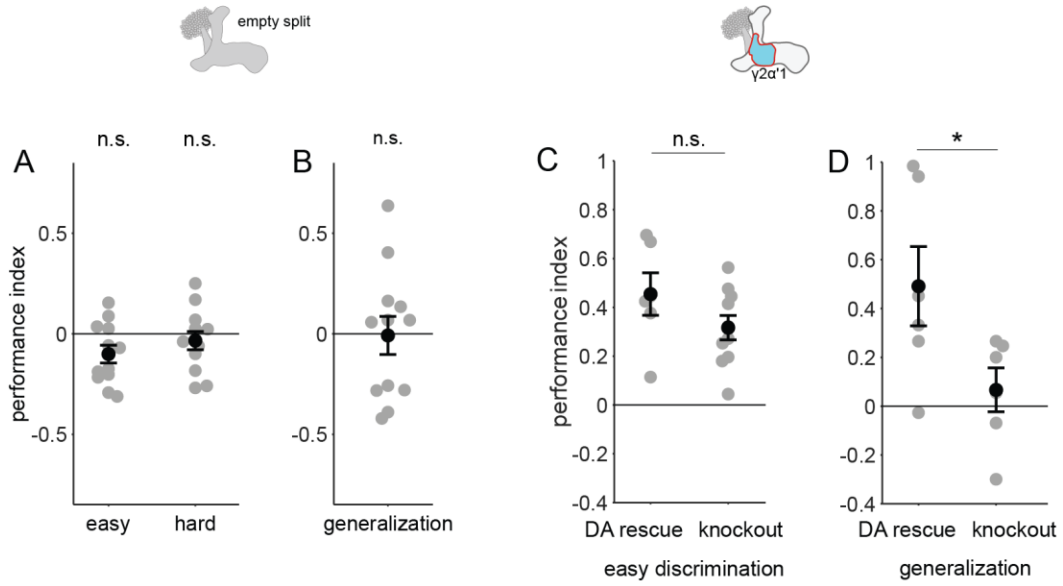

Figure 5.2 Control behavior experiments. (A) and (B) Flies with Chrimson driven by an empty split-Gal4 driver, ie. with no Chrimson expression did not learn any of the tasks. Performance indices not significantly different from 0 (n = 12, p = 0.052, 0.38 and 0.91 and for easy, hard discrimination and generalization; single-sample, Wilcoxon signed rank test). (C) Flies with no expression of dopamine in their nervous systems (knockout) are capable of performing the easy discrimination task as well as flies with dopamine rescued in DAN PPL1-γ2α'1 (rescue n = 6, knockout n = 12, p = 0.22, Wilcoxon rank sum test). (D) Flies with dopamine expression rescued in DAN PPL1-γ2α'1 alone are capable of generalization, while flies with dopamine knockout are not (n = 6 for rescue and knockout, p = 0.041, Wilcoxon rank sum test).

The difference in ability to support fine discrimination between these two compartments raises the question of whether and how they differ in a generalization task. In a simple model where performance reflects overlap between the test stimulus and the trained odor, the harder the discrimination, the easier the generalization. Does the weakly discriminating α3 compartment support strong performance on generalization, while the strongly discriminating γ2α'1 compartment does not? Or, does the γ2α'1 compartment somehow have the flexibility to support strong performance on both tasks?

We tested this by examining relative performance on generalization and easy discrimination tasks in these two compartments. We kept a parallel structure between the two types of tasks by

quantifying performance against control experiments where optogenetic stimulation was delivered unpaired to odor delivery (Fig. 5.1F; see 5.4 Methods; note that since these experiments did not have reciprocal controls the performance scores in Fig. 5.1G-H are computed differently than in Fig. 5.1D-E). As expected, training flies using DAN PPL1-α3 yielded similarly high performance on both generalization and easy discrimination tasks (Fig. 5.1G, $p = 0.84$, $n = 12$), while empty driver controls performed no better than chance (Fig. 5.2A,B, $p = 0.052$, $p = 0.91$, $n = 12$). However, performance on the generalization task was also high in the strongly discriminating compartment γ2α'1., with a performance level indistinguishable from that in the easy discrimination task (Fig. 5.1H $p = 0.89$, $n = 12$).

Although the experiments above target optogenetic punishment to specific sites within the MB, there is the possibility that there are secondary sites of plasticity that contribute to the behavioral performance we observe, via indirect connections between MB compartments. To more rigorously confine plasticity to γ2α'1, we performed an experiment where dopamine production is restricted solely to DAN PPL1-γ2α'1 within the fly. Dopamine is necessary for flies to show any measurable aversive learning (Kim, Lee, and Han 2007; Qin et al. 2012; Aso et al. 2019), and its production requires the *Drosophila* tyrosine hydroxylase enzyme, DTH (Neckameyer and White 1993; Riemensperger et al. 2011; Cichewicz et al. 2016). So we examined performance of flies lacking DTH throughout the nervous system (Cichewicz et al. 2016), but with production rescued specifically in PPL1-γ2α'1 by driving expression of UAS-DTH using the split hemidrivers TH-DBD and 73F07-AD (Aso et al. 2019). Performance was significantly higher for the DTH-rescue flies than for the mutants in hard discrimination (Fig. 5.1I, $p = 0.038$, $n = 8$) and generalization tasks (Fig. 5.2D, $p = 0.041$, $n = 6$), indicating that plasticity in this set of synapses is sufficient for both behaviors (For control experiments with easy discrimination, see Fig. 5.2C).

These results show that a single memory trace formed via plasticity confined to γ2α’1 supports strong performance on the hard discrimination and generalization tasks. We note that the choice outcomes of these paradigms are opposite: in the generalization experiments flies distribute away from odor A’, while in the hard discrimination task, flies accumulate in the A’ quadrant.

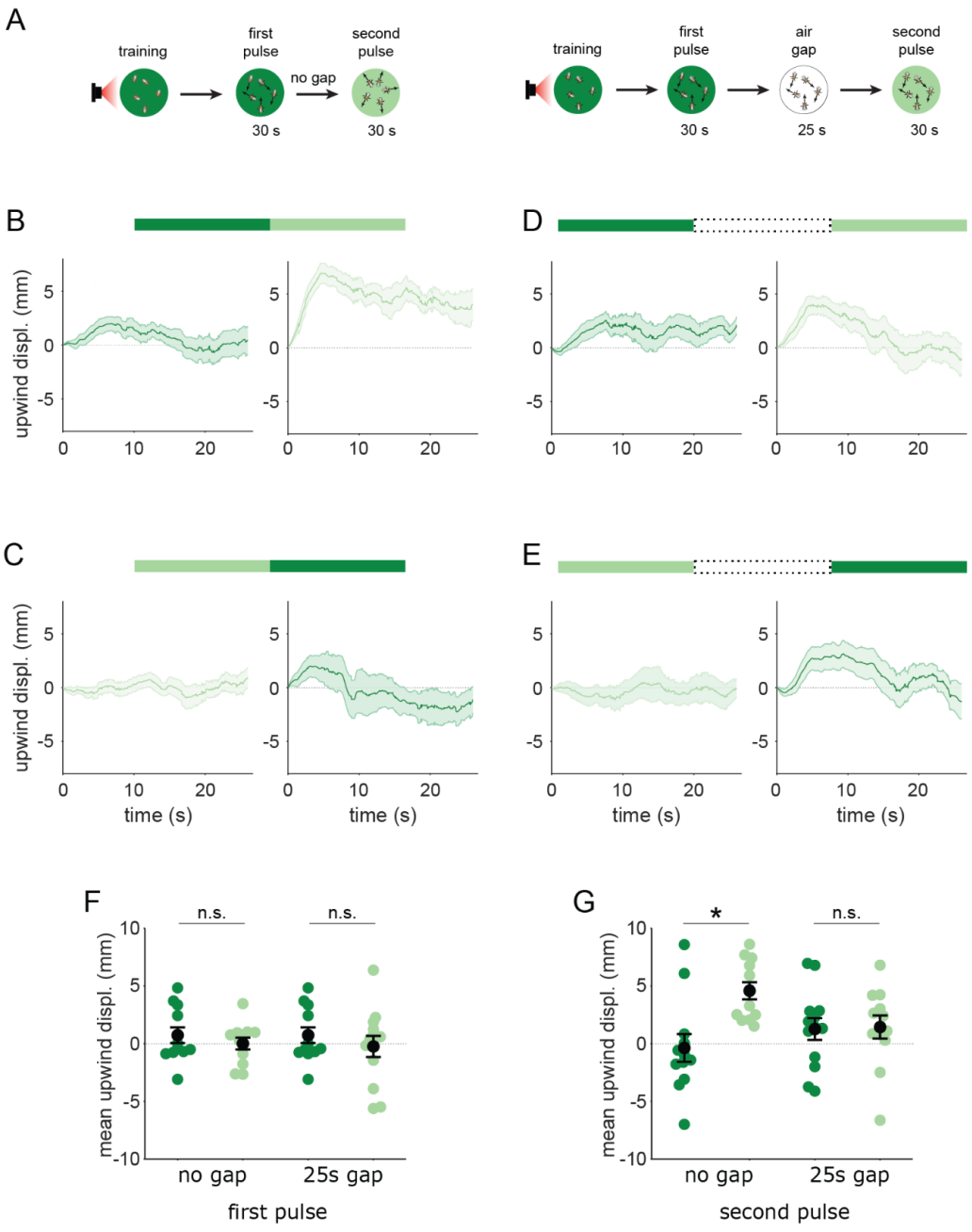

Figure 5.3 Flies are attracted to the unpaired odor only in transitions. (A) Experimental strategy for measuring behavioral responses to odor transitions. Flies were trained by pairing one of the similar odors with optogenetic activation of DAN PPL1-γ2α'1. They were then tested with 30s odor pulses presented either as direct transitions (left) or interrupted by a 25s air period (right). Schematics illustrate A-A' transitions but both sequences were tested, as indicated by the bars on top of panels (B-E). (B) Upwind displacement during the first and second pulses of an A-A' odor transition, as indicated by the green bars up top. This was computed as the increase in each fly's distance from the arena center over the odor delivery period, then averaged across all flies in an arena (approximately 15 flies per arena). Traces in dark and light green are responses to A and A' respectively. Plots are mean +/- SEM (n = 12 arena runs for all stimulus types). (C) Upwind displacement for the reverse odor transition i.e. A'-A. (D) Upwind displacement for A-gap-A' interrupted transition. (E) Upwind displacement for the reverse A'-gap-A interrupted transition. (F) Upwind displacement in response to the first odor pulse, averaged across flies in each arena experiment. Mean displacement was not significantly different between unpaired and paired odors for experiments with no gap (n = 12, 12 experiments for paired and unpaired odors, p = 0.70) and a 25 s gap (n = 12, 13, p = 0.40). (G) As in (F) except for responses to the second odor pulse. Displacement was significantly different for transitions with no gap between pulses (n = 12, 12 experiments for paired and unpaired odors, p = 0.004) but not different when transitions were interrupted by a 25s gap (n = 12, 13, p = 0.85). Statistical comparisons in (F) and (G) were made with the independent-sample Wilcoxon rank sum test with a Bonferroni-Holm correction for multiple comparisons.

5.2.2 Odor sequences show that a temporal comparison contributes to flexible discrimination.

Based on results obtained from experiments imaging KCs and MBONs led by Dr. Mehrab Modi (Detailed in section 5.3 Discussion and further experiments), we predicted that flies' behavioral response to similar odors should also be indistinguishable, unless they are encountered as transitions. Suggesting that the flexible behavior we observed is due to the presence of options that flies experience sequentially in time. Further, since MBON-γ2α'1 signals positive valence (Aso, Sitaraman, et al. 2014), our activity measurements predict that flies might be attracted to A' if they encounter an A to A' transition.

To test these predictions, we examined behavioral responses to temporal sequences of odor, converting the spatial odor border flies encountered in our earlier behavioral experiments, into an

odor transition in time. Flies were trained in the circular arena, and then tested by flooding the entire arena with a sequence of odor pulses. We then compared their behavioral response to direct odor transitions to their response when we interrupted the transition with 25s of clean air. We determined the timing of odor pulse transitions using photo-ionization detector measurements at arena exhaust (Fig. 5.4A) and analyzed behavior around these times.

Attraction to an odor was quantified by how much the flies move upwind; in the arena odors flow inwards from the periphery so we measured displacement away from the center of the arena. We examined the time course of upwind displacement for direct and interrupted transitions (Fig. 5.3A-E). We observed strong upwind displacement during the second pulse of an A-A' transition, which was significantly larger than during the reverse A'-A sequence (Fig. 5.3 B,C,F, $p = 0.004$, $n = 12$). This contrasted with results observed with a 25s gap in between the two odor pulses. In these interrupted transitions, responses to the second pulse were not significantly different depending on transition order (Fig. 5.3D,E,G $p = 0.85$, $n = 12$ for A-gap-A', $n = 13$ for A'-gap-A,), and showed a similar degree of upwind displacement to that evoked during the first pulse, as expected. Note that starting locations at the onset of the second odor pulse were not significantly different in any condition, ruling out the possibility that flies go more upwind with the A-A' transition because they start from further downwind in the arena (Fig. 5.4B, $n = 12$ for A-A', $n = 12$ for A'-A, $p = 0.08$ for direct; $n = 12$ for A-gap-A', $n = 13$ for A'-gap-A, $p = 0.39$ for interrupted). Additionally, we ruled out the possibility that the increased upwind displacement during such a transition comes from a linear combination of the response to the end of the first odor pulse and the beginning of the second (Fig. 5.4C-G). Similar results were not observed in control flies where no learning takes place suggesting that the behaviors, we observe following odor transitions were induced by the learning of association between odors and reward (Fig. 5.5). Together, these results show that behavioral responses to A' are distinct only when it immediately follows the paired odor A.

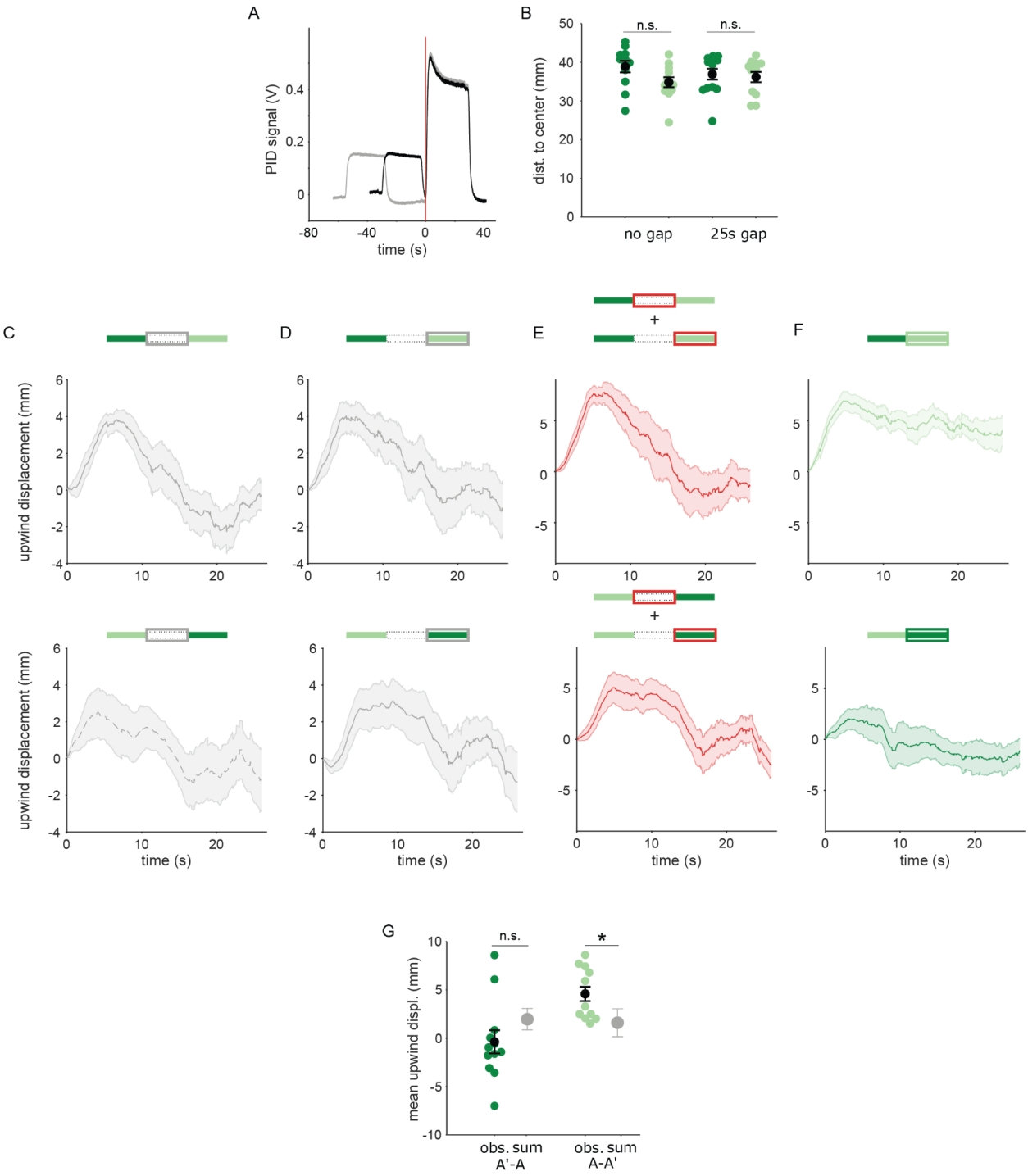

Figure 5.4 Transition dependent attraction is not a result of linearly summed, single-pulse responses. (A) PID measurements of odor time courses in the behavioral arena, measured at the central exhaust port. The black trace depicts odor pulses delivered with no gap and the gray trace depicts pulses interrupted by a 25s air period. The vertical red line indicates the time point identified as the onset of the second odor pulse. (B) Mean distances of flies from the arena center at the onset of the second odor pulse in each experiment. Distances from the center were not

significantly different between experiments with the unpaired and paired odor, with a 0s gap (n = 12, 12 resp. p = 0.08) or with a 25s gap (n = 12, 13, p = 0.39). Statistical comparisons made with the independent samples; Wilcoxon's rank sum test followed by a Bonferroni-Holm correction for multiple comparisons. (C) The off-response trace after the first odor pulse for A-A' transitions without an air gap (top) and A'-A transitions with an air gap (bottom). (D) The response at the onset of the second odor pulse for A-A' transitions without an air-gap (top) and A'-A transitions with an air gap (bottom). (E) The linear sum of upwind displacement traces in (C) and (D). This is the null model to compare with responses to an odor transition. (F) Actual displacement during the second odor pulse in unpaired to paired transition stimuli with no air-gap. Traces in C-F are mean +/- SEM. (G) Mean upwind displacement over the duration of the second odor pulse for the two transitions (dark green: A' to A, light green; A to A'). In gray is the mean and bootstrapped SEM of the linear sum of component responses in (C). The measured displacements are significantly greater than the sum when odor A' is second (n=12, p = 0.001) but not when odor A is second (n=12, p = 0.09). Displacements were compared with the Wilcoxon's signed rank test followed by a Bonferroni-Holm correction for multiple comparisons.

The upwind displacement during the A-A' transition is consistent with our observation that MBON-γ2α'1, a positive valence MBON that drives upwind behavior, is highly active during these transitions. In fact, the mean upwind displacement after an A-A' transition was similar to that caused by optogenetic activation of MBON-γ2α'1 in the arena (unpublished communication - Y. Aso). Overall, these results show that flies compare available alternatives "side-by-side" in time and that stimulus history is important for flexible categorization and behavior.

5.3 Discussions

The behavior experiments described in the section above were performed in combination with neural imaging experiments led by Dr. Mehrab Modi and in the context of a scientific line of thinking initiated by Mehrab in collaboration with Dr. Yoshi Aso and Dr. Herve Roualt. A complete picture of the insight gained from the experiments detailed above can only be ascertained within the context of these imaging and modeling results. I will therefore briefly summarize in this upcoming section these additional results that make up the remainder of the paper Modi,

Rajagopalan et al. published in eLife, drawing heavily from and adapting the text in that publication (Modi et al. 2023).

Upon observing that flies could achieve high levels of performance for both discrimination and generalization tasks and identifying a single MB compartment capable of supporting both (Fig. 5.1 and 5.2), we sought to understand how plasticity in this one compartment can result in the flexible categorization of an odor depending on its context. Surprisingly, MBON responses in this compartment showed no measurable stimulus-specificity to simple pulses of the two similar odors we used, despite being able to distinguish them behaviorally. However, when we presented odors in sequence, one transitioning immediately into the other, similar to what flies experience in the behavioral task, we found that MBON responses to these odors were clearly distinct. These findings show that MBON activity is modulated by a temporal comparison of the alternatives presented to the fly, allowing for switches in the categorization of odor stimuli. Importantly, KC representations did not show categorization switching to either simple stimuli or transitions suggesting the involvement of downstream mechanisms.

Both imaging and behavior provide complementary evidence that comparing available alternatives "side-by-side" in time is important for flexible categorization. These results show that the MB circuit implements a comparison, augmenting small differences between overlapping sensory representations to guide flexible stimulus categorization and choice behavior.

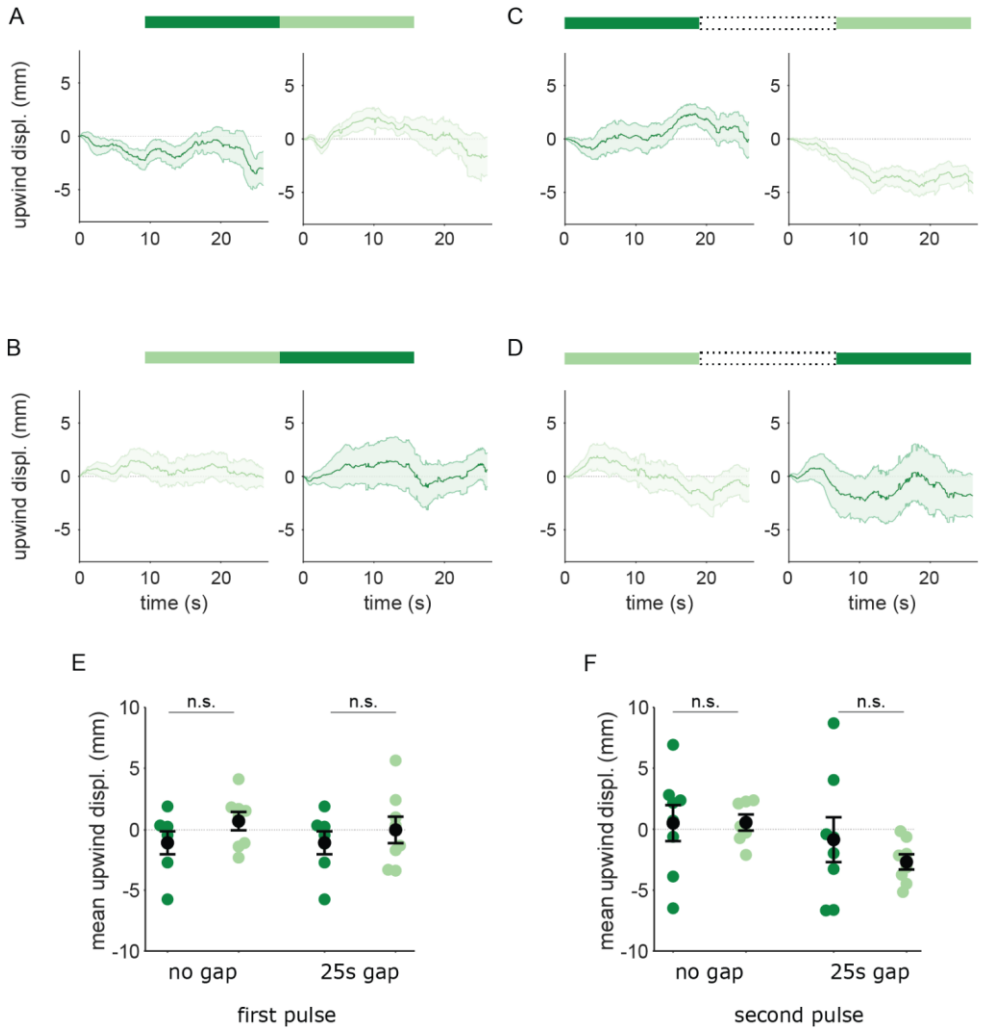

Figure 5.5 Flies reinforced via DAN PPL1-α3 do not respond to transitions between A and A'. Flies were trained by pairing one of the similar odors with optogenetic activation of DAN PPL1-α3. They were then tested with 30s odor pulses presented either as direct transitions (A,B) or interrupted by a 25s air period (C,D), as in Fig. 5.3. Transition order indicated by the bars on top of panels (A-D). (A) Upwind displacement in the first and second pulses of an A-A' odor transition. Traces in dark and light green are responses to A and A' respectively. Plots are mean +/- SEM (n = 8 arena runs for all stimulus types). (B) Upwind displacement for the reverse odor transition i.e. A'-A. (C) Upwind displacement for A-gap-A' interrupted transition. (D) Upwind displacement for the reversed, A'-gap-A interrupted transition. (E) Upwind displacement in response to the first odor pulse, averaged across flies in each arena experiment. Mean displacement was not significantly different between unpaired and paired odors for experiments with no gap (n = 8 experiments each for paired and unpaired odors, p = 0.28) and a 25 s gap (n = 8, p = 0.69). (F) As in (E) except for responses to the second odor pulse. Unlike with DAN PPL1-γ2α'1, displacement was not significantly different for transitions with no gap between pulses (n = 8, p = 0.78) or when transitions were interrupted by a 25s gap (n = 8, p = 0.57). Statistical comparisons in (E) and (F) were made with the independent-sample Wilcoxon rank sum test with a Bonferroni-Holm correction for multiple comparisons.

5.4 Methods

5.4.1 Fly strains and rearing

Drosophila melanogaster were raised on standard cornmeal food at 21℃ at 60% relative humidity on standard cornmeal food on a 12-12h light-dark cycle. For optogenetics behavior experiments, crosses were set on food supplemented with 0.2mM all-trans-retinal and moved to 0.4mM after eclosion and kept in the dark throughout.

Table 5.1 Fly genetic lines

Transgene	**Expression target/reporter description**	**Bloomington stock number, reference**
MB296B split Gal4	DAN PPL1-γ2α'1	BDSC:68253(Aso and Rubin 2016)
MB630B split Gal4	DAN PPL1-α3	BDSC:68290(Aso and Rubin 2016)
R82C10-LexA	DANs PPL1-γ2α'1, α2, α3	BDSC:54981(B. D. Pfeiffer et al. 2013)
20XUAS-CsChrimson-mVenus attp18	Optogenetic activation for behavior	BDSC:55134(Klapoetke et al. 2014)

Expression patterns of split-GAL4 lines produced by Janelia FlyLight(Jenett et al. 2012) can be viewed online (http://splitgal4.janelia.org/cgi-bin/splitgal4.cgi). DAN driver split Gal4 in the table above was crossed with 20XUAS-CsChrimson-mVenus attp18

TH-rescue experiments (genetic strategy as in (Aso et al. 2019))

Knockout

w, 20XUAS-CSChrimson-mVenus attP18; +; ple^{2}, DTHFS± BAC attP2, TH-ZpGAL4DBD VK00027/ TM6B

crossed with

w; R73F07-p65ADZp attP40/ CyO; ple^{2}, DTHFS± BAC attP2/ TM6B

Knockout and rescue in DAN PPL1-γ2α'1

w, 20XUAS-CSChrimson-mVenus attP18; UAS-DTH1m; ple^{2}, DTHFS± BAC attP2, TH-ZpGAL4DBD VK00027/ TM6B

crossed with

w; R73F07-p65ADZp attP40/ CyO; ple^{2}, DTHFS± BAC attP2/ TM6B

5.4.2 Behavior Experiments

Odor quadrant choice assay: Groups of approximately 20 females, aged 4-10 d post-eclosion were anaesthetized on a cold plate and collected at least two days prior to experiments. After a day of recovery on 0.4 mM all-trans-retinal food, they were transferred to starvation vials containing nutrient-free agarose. Starved females were trained and tested at 25℃ at 50% relative humidity in a dark circular arena described in(Aso and Rubin 2016). The arena consisted of a circular chamber surrounded by four odor delivery ports that divide the chamber into quadrants. The input flow rate through each port was 100 mL/min, which was actively vented out a central exhaust at 400 mL/min. Odors were pentyl acetate, butyl acetate and ethyl lactate (Sigma-Aldrich product numbers 109584, 287725, and W244015 respectively). Except for the TH-rescue experiments shown in Fig. 5.1I, these odors were diluted 1:10000 in paraffin oil (Sigma-Aldrich product number 18512). For the experiments in Fig. 5.1I, we used a different odor delivery system which utilizes air dilution of saturated odorant vapor and delivered odors at a 1:16 dilution of saturated vapor.

Flies were aspirated into the arena via a small port and allowed 60 s to acclimatize before training commenced. Training consisted of exposing the flies to one of the odors while providing optogenetic stimulation via a square array of red LEDs (617 nm peak emission, Red-Orange LUXEON Rebel LED, 122 lm at 700mA) which shone through an acrylic diffuser to illuminate flies from below. LED activation consisted of 30 pulses of 1s duration with a 1s inter-flash interval, commencing 5s after switching on the odor valves and terminating 5s after valve shut-off.

To optimize learning scores, we used different training regimes depending on the compartments receiving optogenetic reinforcement, according to(Aso and Rubin 2016). A single training session was used for MB296B, TH-mutant, TH-rescue, while 3 training sessions, separated by 60 seconds, were used for some MB296B experiments, as indicated in the text. For MB630B we used 10 training sessions separated by 15 minutes.

Following training, testing was carried out with the appropriate odors for each task. In the test configuration, the two different odor choices are presented in opposing quadrants for 60 s. Videos of fly behavior were captured at 30 frames per second using MATLAB (Mathworks, USA) and BIAS (http://archive.iorodeo.com/content/basic-image-acquisition-software-bias.html) and analyzed using custom-written code in MATLAB.

Odor attraction assay: For the odor attraction assay, the outputs of odor machines were re-configured to inject the output of a single odor machine into all four quadrants. We switched output from one machine to the other to deliver rapid odor transitions in time. About 15 flies were introduced into the arena for each experiment. The rest of the behavioral procedures were identical to those used in the quadrant choice assay.

Optogenetic MBON-activation assay: For this assay, a clean air stream was delivered into all four arena quadrants throughout the experiment. Flies expressed CSChrimson in MBON γ2α'1. Flies received six 10 s long LED flashes, separated by 60s of darkness. The rest of the behavioral procedures were identical to those used in the quadrant choice assay.

5.4.3 Behavior Analysis

Videos recorded during the test phase were analyzed using custom-written MATLAB code. The centroid of each fly was identified and the number of centroids in each quadrant computed for every frame of the experiment.

For discrimination experiments, a Performance Index (PI) was calculated as the number of flies in the quadrants containing the paired odor minus the number in the quadrants with the unpaired odor, divided by the total number of flies(Tully and Quinn 1985). This value was calculated for every frame of the movie, and the values over the final 30 s of the test period averaged to compute a single PI. Discrimination experiments employed a reciprocal design where the identity of the paired and unpaired odors was swapped, and a single data point represents the averaged PI from two reciprocally trained groups of flies.

Generalization experiments could not employ a reciprocal design, so instead we compared scores against control experiments where flies were exposed to LED stimulation that was not paired with odor delivery; instead, stimulation preceded odor by 2 min. In this case the PI score reported as a single data point is the PI observed from the generalization experiment minus the PI observed in the unpaired control, after both PIs were corrected for biases in initial quadrant occupancies by subtracting away the pre-odor baseline.

Statistical testing was done as described in figure legends. We used the non-parametric, independent sample, Wilcoxon rank sum test to compare performance indices across treatment groups. Statistical testing was performed with custom code written in Matlab (Mathworks, USA). The appropriate sample size was estimated based on the standard deviation of performance indices in previous studies using the same assay(Aso and Rubin 2016).

For the odor attraction and the MBON-activation assays, computing upwind displacement required us to track each fly's trajectory in time. We used the Caltech Fly Tracker (Eyjolfsdottir et al. 2014) to automatically extract fly trajectories from videos. Odor stimulus onset time in the arena was determined from PID measurements of odor concentration at the arena exhaust port. For the MBON-activation assay, stimulus onset was set as the moment the LED turned on. Upwind displacement was computed as the increase in the distance from the center for each fly, relative to its location at stimulus onset, for each time-point over the entire stimulus window. The displacement for all flies in an arena experiment were then averaged before plotting and statistical testing.

Chapter 6. Discussion and future outlook

6.1 Summary of results

The decision-making strategies used by animals play a key role in their survival. As mentioned in the introduction a large body of work has focussed on understanding the computations underlying decision making in a variety of animal brains and have identified key implementational elements such value and RPE signals represented in neural activity. However, despite this progress, a complete head-to-toe understanding of decision-making algorithms has been hard to arrive at. This has been due to the complexity of the brains commonly studied to understand decision-making and when simple brains are used the behavioral tasks have remained too simple to differentiate between different hypothetical algorithms. In the work described in this thesis we address this two-fold problem by detailing a novel dynamic-foraging task for fruit flies. This task combined with the well-mapped *Drosophila* brain allows us to test and prove general theories regarding the neural algorithms underlying decision making in i) dynamic and probabilistic environments and ii) in the face of sensorily-overlapping options.

Operant matching is one simple and ubiquitous behavioral strategy, utilized in dynamically changing and probabilistic environments. Despite the ubiquity of this strategy and strong theoretical background, little was known about the underlying biological mechanisms. We leveraged the growing body of knowledge regarding learning in the fruit fly and the plethora of available anatomical tools to tackle this knowledge gap. We developed a foraging task that allowed us to monitor choices of individual fruit flies and showed, for the first time, that flies follow Herrnstein's operant matching law. Combining experimental results with computational modeling, we found that this behavior requires synaptic plasticity and uses a rule that incorporates expectation of reward. Follow-up experiments manipulating neural circuitry found that reward

expectation signals were incorporated via the rewarding PAM DANs. Our results provide the first mapping of the learning rule underlying operant matching onto the plasticity of specific synapses – the KC-MBON synapses in the MB.

The MB utilizes an 'expansion-layer' architecture first put forward by Marr and Albus in the context of the cerebellum. Such a structure allows for accurate discrimination between sensorily-overlapping options for appropriate decision-making by tuning input connectivity and expansion cell number. Using developmental and chemical manipulations we modified these elements of the MB circuit to prove the theoretical predictions about 'expansion-layer' circuits, showing that increasing KC number improves discrimination, but changing PN-KC connectivity (thought to be optimal) reduces discrimination capabilities. In addition to this 'expansion-layer' architecture, we find that MBONs downstream of the KC 'expansion-layer' further aid flexible decision-making by modifying neuronal activity, as well as behavior, on the order and identity of the options perceived by the animal.

6.2 Future experiments

These findings provide answers to some important questions that have existed in the field but more importantly open the door to answering even more questions and more thoroughly understanding the neural algorithms underlying decision-making.

6.2.1 Future experiments related to Chapter 3

The most prominent unexplored direction in the work described in Chapter 3 is that the obvious predictions made by the behavior and modeling regarding neural activity have not been tested.

6.2.1.1 MBON imaging experiments

The first direction to explore would be to observe MBON activity in the context of this task. This can be done in one of two ways.

The most direct measurements would involve translating the behavioral paradigms described in Chapter 2 and 3, into a virtual reality setup where head-fixed flies walk on a revolving spherical treadmill. This would allow us to observe neural activity as the fly is performing the task. In such a setup, one could use a fluorescent calcium sensor such as GCaMP to record from specific MBONs of interest. Of particular interest is the MBON that innervates the gamma 5 compartment. Preliminary data collected in the lab shows that optogenetic reward via the Gr64f sugar neurons activates the PAM DANs that innervate this compartment. Our covariance-based learning rule model for operant matching makes clear predictions regarding how MBON activity will change following rewarded and unrewarded trials contingent on the reward expectation. If the expectation for reward following a given odor cue is high and a reward is provided, then on the next experience of that odor cue, the MBON response should be slightly lower or show no change compared to the previous experience. On the other hand, if the reward predicting odor is followed by no reward, then on the next experience of that cue, the MBON activity should be higher compared to the previous experience. This can be compared to the traditional R· S non-covariance product rule previously assumed to exist in the MB. Such a rule would never predict increases in MBON response to an odor, even after a missed reward.

Such a test of the covariance rule is complicated for multiple reasons. First, the fly on the ball version of the Y-arena task will need to be designed and tested. Second, the dynamic foraging paradigm leads to multiple different expectation values that change from trial to trial given the

probabilistic nature of the reward. Getting a sufficient number of responses where flies have experienced the same history of choices to account for noisiness in the data will be difficult.

A simpler option that offers a more direct way to test the prediction of the covariance-based learning rule. Here flies can be imaged in a traditional head-fixed setup previously described for recording from MB related neurons (Murthy and Turner 2013; Robert A. A. Campbell et al. 2013). In this non-behavioral set-up, flies would experience odors in a fixed order independent of its own behavior allowing us to replicate this experience across flies. A fixed sequence of odor presentation followed by reward or not would allow us to test if MBON activity can show both depression and potentiation as predicted by the covariance rule hypothesis. This option, however, suffers from the fact that the fly in this case is not performing the complete task. If any part of our observations in Chapter 3 depends on feedback from the fly's own movement and control over its world we may not obtain interpretable imaging data.

6.2.1.2 DAN imaging experiments

A similar and equally important experiment involves the imaging of DA release. If DANs do in fact represent a reward minus expected reward signal, then the amount of DA released following a reward that is expected (preceded by a reward predictive odor) should be less than the DA released following an unpredicted reward or a reward that is less expected (preceded by an odor with low expectation of reward).

This experiment could be performed in the same two ways as described for the MBON experiments in the previous section. Again, the ease of replicability and stimulus control offered by the non-behavioral head-fixed set-up offers advantages while still allowing the experimenter to test important theoretical predictions, while suffering from the same disadvantages. An important

difference between this experiment and the MBON that I would suggest would be to use a DA sensor, such as GRABDA (Sun et al. 2018) expressed in the axons of KCs rather than a calcium sensor expressed in the DANs as the DA released is a more proximal effector of changes at the KC-MBON synapse than DAN activity itself.

6.2.1.3 Calculation of reward expectation

There are many possible hypotheses by which reward expectation can be calculated in this circuit, see Chapter 3.3.5. One of these ideas, inspired by work in the mammalian midbrain (Keiflin and Janak 2015), suggests that reward expectation can arrive at the DANs through a GABAergic interneuron. The activity of MBONs upon experiencing an odor represent the value of that odor, this is highly correlated with the reward expectation. If a GABAergic interneuron can transfer this signal to the DANs with a sign opposite to that of the incoming reward signal from the sugar sensory neurons, that would provide the necessary mechanisms to carry out the reward minus expected reward computation required by the covariance rule theory. Such a GABAergic neuron has in fact been identified in the connectome that received input form the gamma 5 MBON and feeds back to the PAM DANs. Preliminary data using the same tissue specific CRISPR approach used to K.O. DA receptors in KCs in Chapter 3 to K.O. GABA receptors in the PAM DANS has suggested that this model may indeed be correct but follow up and control experiments are required to test this idea.

This mechanism differs in another important way from the models used in Chapter 3. Here reward expectation will be odor dependent as the MBON activity to two different odors is different. However, the model in Chapter 3 and Loewenstein and Seung's original work makes use of a reward expectation averaged over both options. If the GABA interneuron mechanism holds up, new modeling efforts will need to be undertaken to understand the effect of this difference on the

theory. It is possible that the undermatching we observe could be explained when models incorporate such option specific reward expectations.

6.2.1.4 Experiments to study risky-choice

The experiments described in Chapter 3 have focussed on theories about two key economic variables necessary for decision making, value and reward prediction error. However, factors such as confidence and risk-attitude play equally important roles in decision-making behaviors in probabilistic environments. Future work should explore if flies can be a useful organism for the study of such behaviors. For example, a simple change to the dynamic foraging paradigm described in Chapter 3 can convert the task into one that studies risk attitude. Let's say odor 1 is rewarded every time it is chosen, odor 2 on the other hand is rewarded on 50% of trials but flies are provided with a reward that is twice as long or twice as intense on those trials. Does this affect choice behavior like it does in mammals? Such early experiments are crucial to identify which theories about risk-attitude and decision-making can be tested in the fly, if any can at all.

6.2.1.5 Explaining undermatching

Future modelling and experimental efforts should also focus on explaining undermatching. The models we provide in this dissertation are 0th order models that explain a large trend in the behavioral strategies observed in flies and other animals, however they do not explain all the nuances. Undermatching is one such nuance that has been observed across animals. There are two dominant ideas for why undermatching could exist (see Chapter 3.3.3). They suggest that either reward expectation is calculated incorrectly or that learning over multiple time-scales leads to hysteresis that causes undermatching. Both of these options could be implemented in the fly. Different MB compartments could be estimating reward expectation differently or to different

extents in agreement with the first option. Alternatively, these compartments are known to learn at different timescales. Behavioral experiments should focus on inducing learning only in one compartment at a time. This can be done by pairing Gr64f reward with DA rescue only in specific compartment in DA deficient flies (like the ones used in Chapter 5). Model efforts should expand on the models in Chapter 3 by incorporating different timescales of learning similar to the work of Iigaya and colleagues (Iigaya et al. 2019).

6.2.2 Future experiments related to Chapter 4

The behavioral experiments described in Chapter 4 test hypotheses regarding the role of input connectivity and cell number of 'expansion layers' in stimulus discrimination. However, the circuitry of the MB that allows for learning and decision-making consists of MBONs and other neurons downstream of the MB as well as PNs and neurons upstream. Future experiments should apply the developmental manipulations described in Ahmed et al. (Ahmed et al. 2023) to PNs and relevant MBONs to understand the effect of cell number and connectivity related changes on these cells to decision-making processes. MBONs for example are likely locations of information integration and changing the number of inputs could allow for faster/slower decisions. Further, it would be interesting to assess how the changes in cell number at these regions affect the architecture of neural circuits downstream, does having more than one MBON for a given MB compartment affect the number of cells downstream? Finally, changing the number of PNs is a final test of the Marr-Albus expansion layer idea. Theoretical work suggests that increased connectivity between PNs and KCs could lead to worsened discrimination (as our results show). Does increasing PN number lead to increased connectivity and hence worse discrimination? These are all questions that are worth delving into and I believe that this approach of combining developmental and systems neuroscience could be of significant importance to our understanding of brain structure and function in the years to come.

6.2.3 Future experiments related to Chapter 5

The experiments detailed in Chapter 5 and the accompanying research paper (Modi et al. 2023), expand on our understanding of the mechanisms by which the MB circuit implements stimulus discrimination for flexible decision-making. While this work characterizes a clear behavioral reality and connects with a neural basis, the mechanisms that could give rise the flexible neural responses observed in the g2a'1 MBON remains unclear. Our work has identified that this neural flexibility involves a mechanism at or downstream of the KC- MBON synapses modified during learning.

One mechanism that could satisfy these criteria, would involve an explicit comparison of MBON activities in time, much like the delay-lines in the auditory pathways of owls and crickets (Sullivan and Konishi 1986; Schöneich, Kostarakos, and Hedwig 2015). This could be implemented via a downstream neuron that receives a real-time and a delayed copy of MBON-γ2α'1 activity, which then provides a positive feed- back signal to the MBON to amplify small increases in activity. Both of these motifs have been observed in the EM connectome (Li et al. 2020). As the extent of depression of KC >MBON synapses is inevitably slightly weaker for any odor that overlaps imperfectly with the learned odor, this mechanism would sensitize the circuit to small differences in MBON activity that arise around an odor transition. Another class of mechanisms centers on the observation that the KC population exhibits a distinct pattern of responses to odor offset (Tanaka, Tanimoto, and Ito 2008; Lüdke et al. 2018). Offset responses in other MBONs can be potentiated (Vrontou et al. 2021), presumably due to the timing of reinforcement (Cohn, Morantte, and Ruta 2015; Handler et al. 2019), suggesting a similar mechanism might operate in MBON-γ2α'1 to augment responses to the second odor in a transition. An additional candidate mechanism is plas- ticity of inhibitory input to KCs from the APL neuron. Activity of this inhibitory neuron is reduced by training (Liu and Davis 2009; Zhou et al. 2019), so when a non- overlapping

set of KCs is activated at an odor transition, that excitation may more effectively drive the downstream MBON. Inhibition at odor offset is a particularly prominent feature of the α'/β' KCs that are input to MBON-γ2α'1 (Inada, Tsuchimoto, and Kazama 2017), so this effect could act in combination with potentiation of KC offset responses to create pronounced changes in KC output at an odor transition. Future work will be needed to resolve these different possibilities.

6.3 Concluding remarks

David Marr, in his most famous work (Marr 2010), put forward a framework of understanding information processing systems that has been widely adopted to explain how the brain functions. In this work he proposed that a system like the brain must be understood at three layers. First, we must understand the computational problem that our brain region(s) of interest is solving. Second, we must identify the algorithm that the system uses to solve the problem and finally, we can investigate the specific instantiation of the algorithm in a given brain.

Now multiple algorithms could solve a potential problem and each algorithm can be instantiated in a variety of ways. However, this would only be true if each solution we observe in the animals we study were created in isolation. Animal brains have in fact evolved from common ancestors and share structure and functions to extent that most scientists agree that there are likely common algorithms underlying the solution of a given cognitive problem even if the specific instantiations may be different (see (Pereira-Obilinovic et al. 2022) for an example).

A large body of theoretical work has identified important algorithmic steps or computations that must be performed to give rise to a particular behavior. I provide examples of such theories in the earlier chapters. Loewenstein and Seung's covariance-based learning and operant matching (Loewenstein and Seung 2006), and the efforts of multiple scientists including David Marr,

himself, on the theory of expansion layers (Marr 1969; Albus 1971; Jortner, Farivar, and Laurent 2007; Babadi and Sompolinsky 2014; Litwin-Kumar et al. 2017; Rajagopalan and Assisi 2020). However, a roadblock in front of these important theories that provide strong and general bounds on our understanding of neural computations has been the difficulty in testing them. Mapping these theories onto neural circuits and manipulating these circuits to test whether they agree with the theory has not been possible (for reasons explained in Chapter 1). Recent work in the central complex of the fly has shown that *Drosophila melanogaster* can serve as a tool to jump this hurdle (Hulse et al. 2021; Turner-Evans et al. 2020; Noorman et al. 2022).

In the work described in this thesis my colleagues and I have used the fruit fly to get past this roadblock in the context of the cognition underlying decision-making. By demonstrating that flies can perform complex cognitive behaviors previously observed primarily in mammals, we have brought the powerful genetic tools available in this model organism into the frame as a way to test theories developed based on interesting behaviors and neural responses observed in mammals. In yet unpublished work developing on this project, others in Dr. Glenn Turner's lab have begun to explore even more complex behaviors such as risky decision-making with gambling (introduced in Chapter 1) as well as attempting to explain behavior in a reinforcement learning framework.

In addition to setting up a novel framework in which complex decision-making behaviors and their underlying neural computations can be understood, the work detailed in Chapter 3 has contributed a novel approach, to a growing body of work (Lim et al. 2015; Ashwood et al. 2020; Confavreux et al. 2020), aimed at estimating plasticity rules at particular synapses based on behavior. Given the growing impact of naturalistic behavior and tools to analyze behavior(Mathis et al. 2018; Pereira et al. 2019, 2022) in neuroscience, such theoretical tools will be important for understanding how our brains function.

Together the work in this thesis has combined multiple important avenues of neuroscience, from complex cognitive behavior and novel theoretical approaches to optogenetic and developmental manipulations to validate important theories regarding the neural principles underlying decision-making and map algorithmic elements of these theories onto circuit structures in the fruit fly. Providing an important demonstration of the linking of Marr's second and third levels.

Bibliography

Adel, Mohamed, and Leslie C. Griffith. 2021. “The Role of Dopamine in Associative Learning in Drosophila: An Updated Unified Model.” *Neuroscience Bulletin* 37 (6): 831–52.

Ahmed, Maria, Adithya E. Rajagopalan, Yijie Pan, Ye Li, Donnell L. Williams, Erik A. Pedersen, Manav Thakral, et al. 2023. “Input Density Tunes Kenyon Cell Sensory Responses in the Drosophila Mushroom Body.” *Current Biology: CB* 33 (13): 2742–60.e12.

Albus, James S. 1971. “A Theory of Cerebellar Function.” *Mathematical Biosciences* 10 (1): 25–61.

Aleskerov, Fuad, Denis Bouyssou, and Bernard Monjardet. 2007. *Utility Maximization, Choice and Preference*. Springer Science & Business Media.

Ashwood, Zoe, Nicholas A. Roy, Ji Hyun Bak, and Jonathan W. Pillow. 2020. “Inferring Learning Rules from Animal Decision-Making.” In *Advances in Neural Information Processing Systems*, edited by H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, 33:3442–53. Curran Associates, Inc.

Aso, Yoshinori, Daisuke Hattori, Yang Yu, Rebecca M. Johnston, A. Nirmala, Teri-T B. Ngo, Heather Dionne, et al. 2014. “The Neuronal Architecture of the Mushroom Body Provides a Logic for Associative Learning.” *eLife* 3:e04577: 1–47.

Aso, Yoshinori, Robert P. Ray, Xi Long, Daniel Bushey, Karol Cichewicz, Teri-Tb Ngo, Brandi Sharp, et al. 2019. “Nitric Oxide Acts as a Cotransmitter in a Subset of Dopaminergic Neurons to Diversify Memory Dynamics.” *eLife* 8 (November). https://doi.org/10.7554/eLife.49257.

Aso, Yoshinori, and Gerald M. Rubin. 2016. “Dopaminergic Neurons Write and Update Memories with Cell-Type-Specific Rules.” *eLife* 5 (JULY): 1–15.

Aso, Yoshinori, Divya Sitaraman, Toshiharu Ichinose, Karla R. Kaun, Katrin Vogt, Ghislain Belliart-guérin, Pierre-Yves Plaçais, et al. 2014. “Mushroom Body Output Neurons Encode Valence and Guide Memory-Based Action Selection in Drosophila.” *eLife*, no. 3: 1–42.

Babadi, Baktash, and Haim Sompolinsky. 2014. “Sparseness and Expansion in Sensory Representations.” *Neuron* 83 (5): 1213–26.

Bari, Bilal A., Cooper D. Grossman, Emily E. Lubin, Adithya E. Rajagopalan, Jianna I. Cressy, and Jeremiah Y. Cohen. 2019. “Stable Representations of Decision Variables for Flexible Behavior.” *Neuron*, June. https://doi.org/10.1016/j.neuron.2019.06.001.

Bayer, Hannah M., and Paul W. Glimcher. 2005. “Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal.” *Neuron* 47 (1): 129–41.

Belle, J. S. de, and M. Heisenberg. 1994. “Associative Odor Learning in Drosophila Abolished by Chemical Ablation of Mushroom Bodies.” *Science* 263 (5147): 692–95.

Bennett, James E. M., Andrew Philippides, and Thomas Nowotny. 2021. “Learning with Reinforcement Prediction Errors in a Model of the Drosophila Mushroom Body.” *Nature Communications* 12 (1): 2569.

Berns, G. S., S. M. McClure, G. Pagnoni, and P. R. Montague. 2001. “Predictability Modulates Human Brain Response to Reward.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 21 (8): 2793–98.

Beron, Celia C., Shay Q. Neufeld, Scott W. Linderman, and Bernardo L. Sabatini. 2022. “Mice Exhibit Stochastic and Efficient Action Switching during Probabilistic Decision Making.” *Proceedings of the National Academy of Sciences of the United States of America* 119 (15): e2113961119.

Berry, Jacob A., Anna Phan, and Ronald L. Davis. 2018. “Dopamine Neurons Mediate Learning and Forgetting through Bidirectional Modulation of a Memory Trace.” *Cell Reports* 25 (3): 651–62.e5.

Blaug, Mark. 1997. *Economic Theory in Retrospect*. Cambridge University Press.

Brembs, B., and M. Heisenberg. 2000. “The Operant and the Classical in Conditioned Orientation of Drosophila Melanogaster at the Flight Simulator.” *Learning & Memory* 7 (2): 104–15.
Brovkina, Margarita V., Rachel Duffié, Abbigayl E. C. Burtis, and E. Josephine Clowney. 2021. “Fruitless Decommissions Regulatory Elements to Implement Cell-Type-Specific Neuronal Masculinization.” *PLoS Genetics* 17 (2): e1009338.
Buchanan, Sean M., Jamey S. Kain, and Benjamin L. de Bivort. 2015. “Neuronal Control of Locomotor Handedness in Drosophila.” *Proceedings of the National Academy of Sciences of the United States of America* 112 (21): 6700–6705.
Bush, R. R., and F. Mosteller. 1951a. “A Mathematical Model for Simple Learning.” *Psychological Review* 58 (5): 313–23.
———. 1951b. “A Model for Stimulus Generalization and Discrimination.” *Psychological Review* 58 (6): 413–23.
Campbell, R. A. A., K. S. Honegger, H. Qin, W. Li, E. Demir, and G. C. Turner. 2013. “Imaging a Population Code for Odor Identity in the Drosophila Mushroom Body.” *Journal of Neuroscience* 33 (25): 10568–81.
Campbell, Robert A. A., Kyle S. Honegger, Hongtao Qin, Wanhe Li, Ebru Demir, and Glenn C. Turner. 2013. “Imaging a Population Code for Odor Identity in the Drosophila Mushroom Body.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 33 (25): 10568–81.
Carello, Christopher D., and Richard J. Krauzlis. 2004. “Manipulating Intent: Evidence for a Causal Role of the Superior Colliculus in Target Selection.” *Neuron* 43 (4): 575–83.
Caron, Sophie J. C., Vanessa Ruta, L. F. Abbott, and Richard Axel. 2013. “Random Convergence of Olfactory Inputs in the Drosophila Mushroom Body.” *Nature* 497 (7447): 113–17.
Cayco-Gajic, N. Alex, and R. Angus Silver. 2019. “Re-Evaluating Circuit Mechanisms Underlying Pattern Separation.” *Neuron* 101 (4): 584–602.
Cichewicz, K., E. J. Garren, C. Adiele, Y. Aso, Z. Wang, M. Wu, S. Birman, G. M. Rubin, and J. Hirsh. 2016. “A New Brain Dopamine-Deficient Drosophila and Its Pharmacological and Genetic Rescue.” *Genes, Brain, and Behavior* 16 (3): 394–403.
Claridge-Chang, Adam, Robert D. Roorda, Eleftheria Vrontou, Lucas Sjulson, Haiyan Li, Jay Hirsh, and Gero Miesenböck. 2009. “Writing Memories with Light-Addressable Reinforcement Circuitry.” *Cell* 139 (2): 405–15.
Clowney, E. Josephine, Shinya Iguchi, Jennifer J. Bussell, Elias Scheer, and Vanessa Ruta. 2015. “Multimodal Chemosensory Circuits Controlling Male Courtship in Drosophila.” *Neuron* 87 (5): 1036–49.
Cohen, Jeremiah Y., Sebastian Haesler, Linh Vong, Bradford B. Lowell, and Naoshige Uchida. 2012. “Neuron-Type-Specific Signals for Reward and Punishment in the Ventral Tegmental Area.” *Nature* 482 (7383): 85–88.
Cohn, Raphael, Ianessa Morantte, and Vanessa Ruta. 2015. “Coordinated and Compartmentalized Neuromodulation Shapes Sensory Processing in Drosophila.” *Cell* 163 (7): 1742–55.
Confavreux, Basile, Friedemann Zenke, Everton Agnes, Timothy Lillicrap, and Tim Vogels. 2020. “A Meta-Learning Approach to (re) Discover Plasticity Rules That Carve a Desired Function into a Neural Network.” *Advances in Neural Information Processing Systems* 33: 16398–408.
Dahanukar, Anupama, Ya-Ting Lei, Jae Young Kwon, and John R. Carlson. 2007. “Two Gr Genes Underlie Sugar Reception in Drosophila.” *Neuron* 56 (3): 503–16.
Davis, Ronald L. 2005. “Olfactory Memory Formation in Drosophila : From Molecular to Systems Neuroscience,” 275–304.
De Martino, Benedetto, Stephen M. Fleming, Neil Garrett, and Raymond J. Dolan. 2013.

"Confidence in Value-Based Choice." *Nature Neuroscience* 16 (1): 105–10.
Dennis, Emily Jane, Ahmed El Hady, Angie Michaiel, Ann Clemens, Dougal R. Gowan Tervo, Jakob Voigts, and Sandeep Robert Datta. 2021. "Systems Neuroscience of Natural Behaviors in Rodents." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 41 (5): 911–19.
Doya, Kenji, and Michael N. Shadlen. 2012. "Decision Making." *Current Opinion in Neurobiology* 22 (6): 11–913.
Dylla, Kristina V., Georg Raiser, C. Giovanni Galizia, and Paul Szyszka. 2017. "Trace Conditioning in Drosophila Induces Associative Plasticity in Mushroom Body Kenyon Cells and Dopaminergic Neurons." *Frontiers in Neural Circuits* 11 (June): 42.
El Hady, Ahmed, Jacob D. Davidson, and Deborah M. Gordon. 2019. "Editorial: An Ecological Perspective on Decision-Making: Empirical and Theoretical Studies in Natural and Natural-Like Environments." *Frontiers in Ecology and Evolution* 7. https://doi.org/10.3389/fevo.2019.00461.
Elkahlah, Najia A., Jackson A. Rogow, Maria Ahmed, and E. Josephine Clowney. 2020. "Presynaptic Developmental Plasticity Allows Robust Sparse Wiring of the Drosophila Mushroom Body." *eLife* 9 (January). https://doi.org/10.7554/eLife.52278.
Eschbach, Claire, Akira Fushiki, Michael Winding, Casey M. Schneider-Mizell, Mei Shao, Rebecca Arruda, Katharina Eichler, et al. 2020. "Recurrent Architecture for Adaptive Regulation of Learning in the Insect Brain." *Nature Neuroscience* 23 (4): 544–55.
Eyjolfsdottir, Eyrun, Steve Branson, Xavier P. Burgos-Artizzu, Eric D. Hoopfer, Jonathan Schor, David J. Anderson, and Pietro Perona. 2014. "Detecting Social Actions of Fruit Flies." In *Computer Vision – ECCV 2014*, 772–87. Springer International Publishing.
Felsenberg, Johannes, Oliver Barnstedt, Paola Cognigni, Suewei Lin, Scott Waddell, and Tinsley Building. 2017. "Re-Evaluation of Learned Information in Drosophila." *Nature Publishing Group*. https://doi.org/10.1038/nature21716.
Felsenberg, Johannes, Pedro F. Jacob, Thomas Walker, Oliver Barnstedt, Amelia J. Edmondson-Stait, Markus W. Pleijzier, Nils Otto, et al. 2018. "Integration of Parallel Opposing Memories Underlies Memory Extinction." *Cell*, September. https://doi.org/10.1016/j.cell.2018.08.021.
Gkanias, Evripidis, Li Yan McCurdy, Michael N. Nitabach, and Barbara Webb. 2022. "An Incentive Circuit for Memory Dynamics in the Mushroom Body of Drosophila Melanogaster." *eLife* 11 (April). https://doi.org/10.7554/eLife.75611.
Glimcher, Paul W. 2011a. *Foundations of Neuroeconomic Analysis*. Oxford University Press, USA.
———. 2011b. "Understanding Dopamine and Reinforcement Learning: The Dopamine Reward Prediction Error Hypothesis." *Proceedings of the National Academy of Sciences of the United States of America* 108 Suppl 3 (September): 15647–54.
Glimcher, Paul W., and Ernst Fehr. 2013. *Neuroeconomics: Decision Making and the Brain*. Academic Press.
Gold, Joshua I., and Michael N. Shadlen. 2007. "The Neural Basis of Decision Making." https://doi.org/10.1146/annurev.neuro.29.051605.113038.
Gomez-Marin, Alex, Joseph J. Paton, Adam R. Kampff, Rui M. Costa, and Zachary F. Mainen. 2014. "Big Behavioral Data: Psychology, Ethology and the Foundations of Neuroscience." *Nature Neuroscience* 17 (11): 1455–62.
Grabenhorst, Fabian, Edmund T. Rolls, Benjamin A. Parris, and Arun A. d'Souza. 2010. "How the Brain Represents the Reward Value of Fat in the Mouth." *Cerebral Cortex* 20 (5): 1082–91.
Gratz, Scott J., Fiona P. Ukken, C. Dustin Rubinstein, Gene Thiede, Laura K. Donohue, Alexander M. Cummings, and Kate M. O'Connor-Giles. 2014. "Highly Specific and Efficient CRISPR/Cas9-Catalyzed Homology-Directed Repair in Drosophila." *Genetics* 196 (4): 961–

71.
Greggers, Uwe, and Randolf Menzel. 1993. “Memory Dynamics and Foraging Strategies of Honeybees.” *Behavioral Ecology and Sociobiology* 32 (1): 17–29.
Groschner, Lukas N., Laura Chan Wah Hak, Rafal Bogacz, Shamik DasGupta, and Gero Miesenböck. 2018. “Dendritic Integration of Sensory Evidence in Perceptual Decision-Making.” *Cell* 173 (4): 894–905.e13.
Gruntman, Eyal, and Glenn C. Turner. 2013. “Integration of the Olfactory Code across Dendritic Claws of Single Mushroom Body Neurons.” *Nature Neuroscience* 16 (12): 1821–29.
Haberkern, Hannah, Melanie A. Basnak, Biafra Ahanonu, David Schauder, Jeremy D. Cohen, Mark Bolstad, Christopher Bruns, and Vivek Jayaraman. 2019. “Visually Guided Behavior and Optogenetically Induced Learning in Head-Fixed Flies Exploring a Virtual Landscape.” *Current Biology: CB* 29 (10): 1647–59.e8.
Handler, Annie, Thomas G. W. Graham, Raphael Cohn, Ianessa Morantte, Andrew F. Siliciano, Jianzhi Zeng, Yulong Li, and Vanessa Ruta. 2019. “Distinct Dopamine Receptor Pathways Underlie the Temporal Sensitivity of Associative Learning.” *Cell* 178 (1): 60–75.e19.
Harris, Justin A., and Joanne S. Carpenter. 2011. “Response Rate and Reinforcement Rate in Pavlovian Conditioning.” *Journal of Experimental Psychology. Animal Behavior Processes* 37 (4): 375–84.
Hart, Andrew S., Robb B. Rutledge, Paul W. Glimcher, and Paul E. M. Phillips. 2014. “Phasic Dopamine Release in the Rat Nucleus Accumbens Symmetrically Encodes a Reward Prediction Error Term.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 34 (3): 698–704.
Hattori, Ryoma, Bethanny Danskin, Zeljana Babic, Nicole Mlynaryk, and Takaki Komiyama. 2019. “Area-Specificity and Plasticity of History-Dependent Value Coding During Learning.” *Cell* 177 (7): 1858–72.e15.
Hayden, Benjamin Y. 2018. “Economic Choice: The Foraging Perspective.” *Current Opinion in Behavioral Sciences* 24 (December): 1–6.
Hayden, Benjamin Y., and Yael Niv. 2021. “The Case against Economic Values in the Orbitofrontal Cortex (or Anywhere Else in the Brain).” *Behavioral Neuroscience* 135 (2): 192–201.
Hayden, Benjamin Y., and Mark E. Walton. 2014. “Neuroscience of Foraging.” *Frontiers in Neuroscience* 8 (April): 81.
Heisenberg, Martin. 2003. “Mushroom Body Memoir: From Maps to Models.” *Nature* 4 (April): 266–75.
Heisenberg, M., A. Borst, S. Wagner, and D. Byers. 1985. “Drosophila Mushroom Body Mutants Are Deficient in Olfactory Learning.” *Journal of Neurogenetics* 2 (1): 1–30.
Herrnstein, Richard J. 1997. *The Matching Law: Papers in Psychology and Economics*. Harvard University Press.
Herrnstein, R. J. 1961. “Relative and Absolute Strength of Response as a Function of Frequency of Reinforcement.” *Journal of the Experimental Analysis of Behavior* 4 (July): 267–72.
Hige, Toshihide, Yoshinori Aso, Mehrab N. Modi, Gerald M. Rubin, and Glenn C. Turner. 2015. “Heterosynaptic Plasticity Underlies Aversive Olfactory Learning in Drosophila Article.” *Neuron* 88 (5): 985–98.
Honegger, Kyle S., Robert A. A. Campbell, and Glenn C. Turner. 2011. “Cellular-Resolution Population Imaging Reveals Robust Sparse Coding in the Drosophila Mushroom Body.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 31 (33): 11772–85.
Honegger, Kyle S., Matthew A-Y Smith, Matthew A. Churgin, Glenn C. Turner, and Benjamin L. de Bivort. 2020. “Idiosyncratic Neural Coding and Neuromodulation of Olfactory Individuality in *Drosophila*.” *Proceedings of the National Academy of Sciences of the United*

States of America 117 (38): 23292–97.
Hulse, Brad K., Hannah Haberkern, Romain Franconville, Daniel B. Turner-Evans, Shinya Takemura, Tanya Wolff, Marcella Noorman, et al. 2021. “A Connectome of the Drosophila Central Complex Reveals Network Motifs Suitable for Flexible Navigation and Context-Dependent Action Selection.” *eLife* 10 (October). https://doi.org/10.7554/eLife.66039.
Hunt, Laurence T., and Benjamin Y. Hayden. 2017. “A Distributed, Hierarchical and Recurrent Framework for Reward-Based Choice.” *Nature Reviews. Neuroscience* 18 (3): 172–82.
Ichinose, Toshiharu, Yoshinori Aso, Nobuhiro Yamagata, Ayako Abe, Gerald M. Rubin, and Hiromu Tanimoto. 2015. “Reward Signal in a Recurrent Circuit Drives Appetitive Long-Term Memory Formation.” *eLife* 4 (November): e10719.
Iigaya, Kiyohito, Yashar Ahmadian, Leo P. Sugrue, Greg S. Corrado, Yonatan Loewenstein, William T. Newsome, and Stefano Fusi. 2019. “Deviation from the Matching Law Reflects an Optimal Strategy Involving Learning over Multiple Timescales.” *Nature Communications* 10 (1): 1466.
Inada, Kengo, Yoshiko Tsuchimoto, and Hokto Kazama. 2017. “Origins of Cell-Type-Specific Olfactory Processing in the Drosophila Mushroom Body Circuit.” *Neuron* 95 (2): 357–67.e4.
Itti, L., and C. Koch. 2001. “Computational Modelling of Visual Attention.” *Nature Reviews. Neuroscience* 2 (3): 194–203.
Jenett, Arnim, Gerald M. Rubin, Teri-T B. Ngo, David Shepherd, Christine Murphy, Heather Dionne, Barret D. Pfeiffer, et al. 2012. “A GAL4-Driver Line Resource for Drosophila Neurobiology.” *Cell Reports* 2 (4): 991–1001.
Jevons, William Stanley. 1879. *The Theory of Political Economy*. Macmillan.
Jiang, Linnie, and Ashok Litwin-Kumar. 2021. “Models of Heterogeneous Dopamine Signaling in an Insect Learning and Memory Center.” *PLoS Computational Biology* 17 (8): e1009205.
Jiao, Yuchen, Seok Jun Moon, Xiaoyue Wang, Qiuting Ren, and Craig Montell. 2008. “Gr64f Is Required in Combination with Other Gustatory Receptors for Sugar Detection in Drosophila.” *Current Biology: CB* 18 (22): 1797–1801.
Jortner, Ron A. 2013. “Network Architecture Underlying Maximal Separation of Neuronal Representations.” *Frontiers in Neuroengineering* 5 (January): 19.
Jortner, Ron A., S. Sarah Farivar, and Gilles Laurent. 2007. “A Simple Connectivity Scheme for Sparse Coding in an Olfactory System.” *Journal of Neuroscience* 27 (7): 1659–69.
Juavinett, Ashley L., Jeffrey C. Erlich, and Anne K. Churchland. 2018. “Decision-Making Behaviors: Weighing Ethology, Complexity, and Sensorimotor Compatibility.” *Current Opinion in Neurobiology* 49: 42–50.
Kahneman, Daniel, and Amos Tversky. 1979. “Prospect Theory: An Analysis of Decision under Risk.” *Econometrica: Journal of the Econometric Society* 47 (2): 263–91.
Keiflin, Ronald, and Patricia H. Janak. 2015. “Dopamine Prediction Errors in Reward Learning and Addiction: From Theory to Neural Circuitry.” *Neuron* 88 (October): 247–63.
Kepecs, Adam, Naoshige Uchida, Hatim A. Zariwala, and Zachary F. Mainen. 2008. “Neural Correlates, Computation and Behavioural Impact of Decision Confidence.” *Nature* 455 (7210): 227–31.
Kim, Young-Cho, Hyun-Gwan Lee, and Kyung-An Han. 2007. “D 1 Dopamine Receptor dDA1 Is Required in the Mushroom Body Neurons for Aversive and Appetitive Learning in Drosophila.” *Journal of Neuroscience* 27 (29): 7640–47.
Klapoetke, Nathan C., Yasunobu Murata, Sung Soo Kim, Stefan R. Pulver, Amanda Birdsey-Benson, Yong Ku Cho, Tania K. Morimoto, et al. 2014. “Independent Optical Excitation of Distinct Neural Populations.” *Nature Methods* 11 (3): 338–46.
Komura, Yutaka, Akihiko Nikkuni, Noriko Hirashima, Teppei Uetake, and Aki Miyamoto. 2013. “Responses of Pulvinar Neurons Reflect a Subject’s Confidence in Visual Categorization.” *Nature Neuroscience* 16 (6): 749–55.
König, Christian, Afshin Khalili, Thomas Niewalda, Shiqiang Gao, and Bertram Gerber. 2019.

"An Optogenetic Analogue of Second-Order Reinforcement in Drosophila." *Biology Letters* 15 (7): 20190084.
Krajbich, Ian, Carrie Armel, and Antonio Rangel. 2010. "Visual Fixations and the Computation and Comparison of Value in Simple Choice." *Nature Neuroscience* 13 (10): 1292–98.
Krakauer, John W., Asif A. Ghazanfar, Alex Gomez-Marin, Malcolm A. MacIver, and David Poeppel. 2017. "Neuroscience Needs Behavior: Correcting a Reductionist Bias." *Neuron* 93 (3): 480–90.
Kudryavitskaya, Elena, Eran Marom, Haran Shani-Narkiss, David Pash, and Adi Mizrahi. 2021. "Flexible Categorization in the Mouse Olfactory Bulb." *Current Biology: CB* 31 (8): 1616–31.e4.
Lai, Sen-Lin, Takeshi Awasaki, Kei Ito, and Tzumin Lee. 2008. "Clonal Analysis of Drosophila Antennal Lobe Neurons: Diverse Neuronal Architectures in the Lateral Neuroblast Lineage." *Development* 135 (17): 2883–93.
Lau, Brian, and Paul W. Glimcher. 2005. "Dynamic Response-by-Response Models of Matching Behavior in Rhesus Monkeys." *Journal of the Experimental Analysis of Behavior* 84 (3): 555–79.
———. 2008. "Value Representations in the Primate Striatum during Matching Behavior." *Neuron* 58 (3): 451–63.
Lesar, Amanda, Javan Tahir, Jason Wolk, and Marc Gershow. 2021. "Switch-like and Persistent Memory Formation in Individual Drosophila Larvae." *eLife* 10 (October). https://doi.org/10.7554/eLife.70317.
Lewis, Sara A., David C. Negelspach, Sevag Kaladchibachi, Stephen L. Cowen, and Fabian Fernandez. 2017. "Spontaneous Alternation: A Potential Gateway to Spatial Working Memory in Drosophila." *Neurobiology of Learning and Memory* 142 (Pt B): 230–35.
Li, Feng, Jack W. Lindsey, Elizabeth C. Marin, Nils Otto, Marisa Dreher, Georgia Dempsey, Ildiko Stark, et al. 2020. "The Connectome of the Adult Drosophila Mushroom Body Provides Insights into Function." *eLife* 9 (December). https://doi.org/10.7554/eLife.62576.
Lim, Sukbin, Jillian L. McKee, Luke Woloszyn, Yali Amit, David J. Freedman, David L. Sheinberg, and Nicolas Brunel. 2015. "Inferring Learning Rules from Distributions of Firing Rates in Cortical Neurons." *Nature Neuroscience* 18 (12): 1804–10.
Litwin-Kumar, Ashok, Kameron Decker Harris, Richard Axel, Haim Sompolinsky, and L. F. Abbott. 2017. "Optimal Degrees of Synaptic Connectivity." *Neuron*. https://doi.org/10.1016/j.neuron.2017.01.030.
Liu, Xu, and Ronald L. Davis. 2009. "The GABAergic Anterior Paired Lateral Neuron Suppresses and Is Suppressed by Olfactory Learning." *Nature Neuroscience* 12 (1): 53–59.
Loewenstein, Yonatan, and H. Sebastian Seung. 2006. "Operant Matching Is a Generic Outcome of Synaptic Plasticity Based on the Covariance between Reward and Neural Activity." *Proceedings of the National Academy of Sciences of the United States of America* 103 (41): 15224–29.
Lüdke, Alja, Georg Raiser, Johannes Nehrkorn, Andreas V. M. Herz, C. Giovanni Galizia, and Paul Szyszka. 2018. "Calcium in Kenyon Cell Somata as a Substrate for an Olfactory Sensory Memory in Drosophila." *Frontiers in Cellular Neuroscience* 12 (May): 128.
Mackintosh, N. J. 1974. "The Psychology of Animal Learning" 730. https://psycnet.apa.org/fulltext/1975-11296-000.pdf.
Malun, Dagmar, Martin Giurfa, C. Giovanni Galizia, Niels Plath, Robert Brandt, Bertram Gerber, and Beate Eisermann. 2002. "Hydroxyurea-Induced Partial Mushroom Body Ablation Does Not Affect Acquisition and Retention of Olfactory Differential Conditioning in Honeybees." *Journal of Neurobiology* 53 (3): 343–60.
Marr, David. 1969. "A Theory of Cerebellar Cortex." *The Journal of Physiology* 202 (2): 437–70.
———. 2010. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. MIT Press.

Mathis, Alexander, Pranav Mamidanna, Kevin M. Cury, Taiga Abe, Venkatesh N. Murthy, Mackenzie Weygandt Mathis, and Matthias Bethge. 2018. “DeepLabCut: Markerless Pose Estimation of User-Defined Body Parts with Deep Learning.” *Nature Neuroscience* 21 (9): 1281–89.
Milosavljevic, Milica, Jonathan Malmaud, Alexander Huth, Christof Koch, and Antonio Rangel. 2010. “The Drift Diffusion Model Can Account for the Accuracy and Reaction Time of Value-Based Choices under High and Low Time Pressure.” *No.* 6: 437–49.
Modi, Mehrab N., Adithya E. Rajagopalan, Hervé Rouault, Yoshinori Aso, and Glenn C. Turner. 2023. “Flexible Specificity of Memory in Drosophila Depends on a Comparison between Choices.” *eLife* 12 (June). https://doi.org/10.7554/eLife.80923.
Modi, Mehrab N., Yichun Shuai, and Glenn C. Turner. 2020. “The Drosophila Mushroom Body: From Architecture to Algorithm in a Learning Circuit.” *Annual Review of Neuroscience* 43 (July): 465–84.
Mohandasan, Radhika, Fathima Mukthar Iqbal, Manikrao Thakare, Madhav Sridharan, and Gaurav Das. 2021. “Enhanced Olfactory Memory Performance in Trap-Design Y-Mazes Allows the Study of Novel Memory Phenotypes in Drosophila.” *bioRxiv*. https://doi.org/10.1101/2020.11.18.386128.
Murthy, Mala, and Glenn Turner. 2013. “Dissection of the Head Cuticle and Sheath of Living Flies for Whole-Cell Patch-Clamp Recordings in the Brain.” *Cold Spring Harbor Protocols* 8 (2): 134–39.
Mysore, Shreesh P., Ali Asadollahi, and Eric I. Knudsen. 2011. “Signaling of the Strongest Stimulus in the Owl Optic Tectum.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 31 (14): 5186–96.
Mysore, Shreesh P., and Eric I. Knudsen. 2011. “Flexible Categorization of Relative Stimulus Strength by the Optic Tectum.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 31 (21): 7745–52.
Neckameyer, W. S., and K. White. 1993. “Drosophila Tyrosine Hydroxylase Is Encoded by the Pale Locus.” *Journal of Neurogenetics* 8 (4): 189–99.
Neumann, John von, and Oskar Morgenstern. 2007. *Theory of Games and Economic Behavior (Commemorative Edition)*. Princeton University Press.
Noorman, Marcella, Brad K. Hulse, Vivek Jayaraman, Sandro Romani, and Ann M. Hermundstad. 2022. “Accurate Angular Integration with Only a Handful of Neurons.” *bioRxiv*. https://doi.org/10.1101/2022.05.23.493052.
Ober, J. 2022. *The Greeks and the Rational: The Discovery of Practical Reason*. Sather Classical Lectures. University of California Press.
Owald, David, Johannes Felsenberg, Clifford B. Talbot, Gaurav Das, Emmanuel Perisse, Wolf Huetteroth, and Scott Waddell. 2015. “Activity of Defined Mushroom Body Output Neurons Underlies Learned Olfactory Behavior in Drosophila.” *Neuron* 86 (2): 417–27.
Padoa-Schioppa, Camillo, and John A. Assad. 2006. “Neurons in the Orbitofrontal Cortex Encode Economic Value.” *Nature* 441 (7090): 223–26.
Padoa-Schioppa, Camillo, and Katherine E. Conen. 2017. “Orbitofrontal Cortex: A Neural Circuit for Economic Decisions.” *Neuron* 96 (4): 736–54.
Pareto, Vilfredo. 2014. *Manual of Political Economy: A Variorum Translation and Critical Edition*. OUP Oxford.
Pearce, John M. 2008. *Animal Learning and Cognition: An Introduction*. 3rd ed. Psychology Press.
Pereira-Obilinovic, Ulises, Han Hou, Karel Svoboda, and Xiao-Jing Wang. 2022. “Brain Mechanism of Foraging: Reward-Dependent Synaptic Plasticity or Neural Integration of Values?” *bioRxiv*. https://doi.org/10.1101/2022.09.25.509030.
Pereira, Talmo D., Diego E. Aldarondo, Lindsay Willmore, Mikhail Kislin, Samuel S-H Wang, Mala Murthy, and Joshua W. Shaevitz. 2019. “Fast Animal Pose Estimation Using Deep

Neural Networks.” *Nature Methods* 16 (1): 117–25.
Pereira, Talmo D., Nathaniel Tabris, Arie Matsliah, David M. Turner, Junyu Li, Shruthi Ravindranath, Eleni S. Papadoyannis, et al. 2022. “SLEAP: A Deep Learning System for Multi-Animal Pose Tracking.” *Nature Methods* 19 (4): 486–95.
Pfeiffer, Barret D., Teri-T B. Ngo, Karen L. Hibbard, Christine Murphy, Arnim Jenett, James W. Truman, and Gerald M. Rubin. 2010. “Refinement of Tools for Targeted Gene Expression in Drosophila.” *Genetics* 186 (2): 735–55.
Pfeiffer, B. D., T. B. Ngo, K. L. Hibbard, C. Murphy, A. Jenett, and J. W. Truman. 2013. “LexA Driver Collection of Rubin Laboratory at Janelia Farm.” *Personal Communication to FlyBase*.
Pierce, W. D., and W. F. Epling. 1983. “Choice, Matching, and Human Behavior: A Review of the Literature.” *The Behavior Analyst / MABA* 6 (1): 57–76.
Plassmann, Hilke, John O’Doherty, and Antonio Rangel. 2007. “Orbitofrontal Cortex Encodes Willingness to Pay in Everyday Economic Transactions.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 27 (37): 9984–88.
Platt, M. L., and P. W. Glimcher. 1999. “Neural Correlates of Decision Variables in Parietal Cortex.” *Nature* 400 (6741): 233–38.
Port, Fillip, and Simon L. Bullock. 2016. “Augmenting CRISPR Applications in Drosophila with tRNA-Flanked sgRNAs.” *Nature Methods* 13 (10): 852–54.
Pouget, Alexandre, Jan Drugowitsch, and Adam Kepecs. 2016. “Confidence and Certainty: Distinct Probabilistic Quantities for Different Goals.” *Nature Neuroscience* 19 (3): 366–74.
Puñal, Vanessa M., Maria Ahmed, Emma M. Thornton-Kolbe, and E. Josephine Clowney. 2021. “Untangling the Wires: Development of Sparse, Distributed Connectivity in the Mushroom Body Calyx.” *Cell and Tissue Research* 383 (1): 91–112.
Qin, Hongtao, Michael Cressy, Wanhe Li, Jonathan S. Coravos, Stephanie A. Izzi, and Joshua Dubnau. 2012. “Gamma Neurons Mediate Dopaminergic Input during Aversive Olfactory Memory Formation in Drosophila.” *Current Biology: CB* 22 (7): 608–14.
Quinn, W. G., W. A. Harris, and S. Benzer. 1974. “Conditioned Behavior in Drosophila Melanogaster.” *Proceedings of the National Academy of Sciences of the United States of America* 71 (3): 708–12.
Raghuraman, Anantha P., and Camillo Padoa-Schioppa. 2014. “Integration of Multiple Determinants in the Neuronal Computation of Economic Values.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 34 (35): 11583–603.
Rajagopalan, Adithya, and Collins Assisi. 2020. “Effect of Circuit Structure on Odor Representation in the Insect Olfactory System.” *eNeuro* 7 (3). https://doi.org/10.1523/ENEURO.0130-19.2020.
Ratcliff, Roger, and Gail McKoon. 2008. “The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks.” *Neural Computation* 20 (4): 873–922.
Rescorla, R. A., and Allan R. Wagner. 1972. “A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement.” In *Classical Conditioning: Current Research and Theory*. Vol. 2. unknown.
Riabinina, Olena, David Luginbuhl, Elizabeth Marr, Sha Liu, Mark N. Wu, Liqun Luo, and Christopher J. Potter. 2015. “Improved and Expanded Q-System Reagents for Genetic Manipulations.” *Nature Methods* 12 (3): 219–22, 5 p following 222.
Riemensperger, Thomas, Guillaume Isabel, Hélène Coulom, Kirsa Neuser, Laurent Seugnet, Kazuhiko Kume, Magali Iché-Torres, et al. 2011. “Behavioral Consequences of Dopamine Deficiency in the Drosophila Central Nervous System.” *Proceedings of the National Academy of Sciences of the United States of America* 108 (2): 834–39.
Riemensperger, Thomas, Thomas Völler, Patrick Stock, Erich Buchner, and André Fiala. 2005. “Punishment Prediction by Dopaminergic Neurons in Drosophila.” *Current Biology: CB* 15 (21): 1953–60.

Rohrsen, Christian, Aida Kumpf, Kader Semiz, Ferruh Aydin, Benjamin deBivort, and Björn Brembs. 2021. “Pain Is so close to Pleasure: The Same Dopamine Neurons Can Mediate Approach and Avoidance in Drosophila.” *bioRxiv*. https://doi.org/10.1101/2021.10.04.463010.
Rushworth, Matthew F. S., Nils Kolling, Jérôme Sallet, and Rogier B. Mars. 2012. “Valuation and Decision-Making in Frontal Cortex: One or Many Serial or Parallel Systems?” *Current Opinion in Neurobiology* 22 (6): 946–55.
Rustichini, Aldo, and Camillo Padoa-Schioppa. 2015. “A Neuro-Computational Model of Economic Decisions.” *Journal of Neurophysiology* 114 (3): 1382–98.
Sabandal, John Martin, Jacob A. Berry, and Ronald L. Davis. 2021. “Dopamine-Based Mechanism for Transient Forgetting.” *Nature* 591 (7850): 426–30.
Samuelson, P. A. 1938. “A Note on the Pure Theory of Consumer’s Behaviour.” *Economica* 5 (17): 61–71.
———. 1948. “Consumption Theory in Terms of Revealed Preference.” *Economica* 15 (60): 243–53.
Sayin, Sercan, Jean-Francois De Backer, K. P. Siju, Marina E. Wosniack, Laurence P. Lewis, Lisa-Marie Frisch, Benedikt Gansen, et al. 2019. “A Neural Circuit Arbitrates between Persistence and Withdrawal in Hungry Drosophila.” *Neuron*, August. https://doi.org/10.1016/j.neuron.2019.07.028.
Scheffer, Louis K., C. Shan Xu, Michal Januszewski, Zhiyuan Lu, Shin-Ya Takemura, Kenneth J. Hayworth, Gary B. Huang, et al. 2020. “A Connectome and Analysis of the Adult Drosophila Central Brain.” *eLife* 9 (September). https://doi.org/10.7554/eLife.57443.
Schoenbaum, Geoffrey, Barry Setlow, Michael P. Saddoris, and Michela Gallagher. 2003. “Encoding Predicted Outcome and Acquired Value in Orbitofrontal Cortex during Cue Sampling Depends upon Input from Basolateral Amygdala.” *Neuron* 39 (5): 855–67.
Schöneich, Stefan, Konstantinos Kostarakos, and Berthold Hedwig. 2015. “An Auditory Feature Detection Circuit for Sound Pattern Recognition.” *Science Advances* 1 (8): e1500325.
Schroll, Christian, Thomas Riemensperger, Daniel Bucher, Julia Ehmer, Thomas Völler, Karen Erbguth, Bertram Gerber, et al. 2006. “Light-Induced Activation of Distinct Modulatory Neurons Triggers Appetitive or Aversive Learning in Drosophila Larvae.” *Current Biology: CB* 16 (17): 1741–47.
Schultz, Wolfram. 1997. “Dopamine Neurons and Their Role in Reward Mechanisms.” *Current Opinion in Neurobiology* 7 (2): 191–97.
———. 2004. “Neural Coding of Basic Reward Terms of Animal Learning Theory, Game Theory, Microeconomics and Behavioural Ecology.” *Current Opinion in Neurobiology* 14 (2): 139–47.
Schultz, Wolfram, Peter Dayan, and P. Read Montague. 1997. “A Neural Substrate of Prediction and Reward.” *Science* 275 (5306): 1593–99.
Seger, Carol A. 2008. “How Do the Basal Ganglia Contribute to Categorization? Their Roles in Generalization, Response Selection, and Learning via Feedback.” *Neuroscience and Biobehavioral Reviews* 32 (2): 265–78.
Seger, Carol A., and Earl K. Miller. 2010. “Category Learning in the Brain.” *Annual Review of Neuroscience* 33: 203–19.
Seidenbecher, Sophie E., Joshua I. Sanders, Anne C. von Philipsborn, and Duda Kvitsiani. 2020. “Reward Foraging Task and Model-Based Analysis Reveal How Fruit Flies Learn Value of Available Options.” *PloS One* 15 (10): e0239616.
Séjourné, Julien, Pierre-Yves Plaçais, Yoshinori Aso, Igor Siwanowicz, Séverine Trannoy, Vladimiros Thoma, Stevanus R. Tedjakumala, et al. 2011. “Mushroom Body Efferent Neurons Responsible for Aversive Olfactory Memory Retrieval in Drosophila.” *Nature Neuroscience* 14 (7): 903–10.
Shuai, Yichun, Ying Hu, Hongtao Qin, Robert A. A. Campbell, and Yi Zhong. 2011. “Distinct

Molecular Underpinnings of Drosophila Olfactory Trace Conditioning." *Proceedings of the National Academy of Sciences of the United States of America* 108 (50): 20201–6.
Simonnet, Mégane M., Martine Berthelot-Grosjean, and Yael Grosjean. 2014. "Testing Drosophila Olfaction with a Y-Maze Assay." *Journal of Visualized Experiments: JoVE*, no. 88 (June). https://doi.org/10.3791/51241.
Smith, Matthew A-Y, Kyle S. Honegger, Glenn Turner, and Benjamin de Bivort. 2022. "Idiosyncratic Learning Performance in Flies." *Biology Letters* 18 (2): 20210424.
Soltani, Alireza, and Xiao-Jing Wang. 2006. "A Biophysically Based Neural Model of Matching Law Behavior: Melioration by Stochastic Synapses." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 26 (14): 3731–44.
So, Na-Young, and Veit Stuphorn. 2010. "Supplementary Eye Field Encodes Option and Action Value for Saccades with Variable Reward." *Journal of Neurophysiology* 104 (5): 2634–53.
Song, H. Francis, Guangyu R. Yang, and Xiao-Jing Wang. 2017. "Reward-Based Training of Recurrent Neural Networks for Cognitive and Value-Based Tasks." *eLife* 6 (January). https://doi.org/10.7554/eLife.21492.
Springer, Magdalena, and Martin Paul Nawrot. 2021. "A Mechanistic Model for Reward Prediction and Extinction Learning in the Fruit Fly." *eNeuro* 8 (3). https://doi.org/10.1523/ENEURO.0549-20.2021.
Stauffer, William R., Armin Lak, Peter Bossaerts, and Wolfram Schultz. 2015. "Economic Choices Reveal Probability Distortion in Macaque Monkeys." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 35 (7): 3146–54.
Stauffer, William R., Armin Lak, and Wolfram Schultz. 2014. "Dopamine Reward Prediction Error Responses Reflect Marginal Utility." *Current Biology: CB* 24 (21): 2491–2500.
Stephens, David W., and John R. Krebs. 1986. *Foraging Theory*. Princeton University Press.
Stocker, R. F., G. Heimbeck, N. Gendre, and J. S. de Belle. 1997. "Neuroblast Ablation in Drosophila P[GAL4] Lines Reveals Origins of Olfactory Interneurons." *Journal of Neurobiology* 32 (5): 443–56.
Stringer, Carsen, Tim Wang, Michalis Michaelos, and Marius Pachitariu. 2021. "Cellpose: A Generalist Algorithm for Cellular Segmentation." *Nature Methods* 18 (1): 100–106.
Sugrue, Leo P., Greg S. Corrado, and William T. Newsome. 2004a. "Matching Behavior and the Representation of Value in the Parietal Cortex." *Science* 304 (5678): 1782–87.
———. 2004b. "Matching Behavior and the Representation of Value in the Parietal Cortex." *Science* 304 (5678): 1782–87.
Sullivan, W. E., and M. Konishi. 1986. "Neural Map of Interaural Phase Difference in the Owl's Brainstem." *Proceedings of the National Academy of Sciences of the United States of America* 83 (21): 8400–8404.
Sun, Fangmiao, Jianzhi Zeng, Miao Jing, Jingheng Zhou, Jiesi Feng, Scott F. Owen, Yichen Luo, et al. 2018. "A Genetically Encoded Fluorescent Sensor Enables Rapid and Specific Detection of Dopamine in Flies, Fish, and Mice." *Cell* 174 (2): 481–96.e19.
Sutton, Richard S., and Andrew G. Barto. 2018. *Reinforcement Learning: An Introduction*. MIT Press.
Tanaka, Nobuaki K., Hiromu Tanimoto, and Kei Ito. 2008. "Neuronal Assemblies of the Drosophila Mushroom Body." *The Journal of Comparative Neurology* 508 (5): 711–55.
Tsutsui, Ken-Ichiro, Fabian Grabenhorst, Shunsuke Kobayashi, and Wolfram Schultz. 2016. "A Dynamic Code for Economic Object Valuation in Prefrontal Cortex Neurons." *Nature Communications* 7 (September): 12554.
Tully, T., and W. G. Quinn. 1985. "Classical Conditioning and Retention in Normal and Mutant Drosophila Melanogaster." *Journal of Comparative Physiology. A, Sensory, Neural, and Behavioral Physiology* 157 (2): 263–77.
Turner-Evans, Daniel B., Kristopher T. Jensen, Saba Ali, Tyler Paterson, Arlo Sheridan, Robert P. Ray, Tanya Wolff, et al. 2020. "The Neuroanatomical Ultrastructure and Function of a

Biological Ring Attractor.” *Neuron*, September. https://doi.org/10.1016/j.neuron.2020.08.006.
Turner, Glenn C., Maxim Bazhenov, and Gilles Laurent. 2008. “Olfactory Representations by Drosophila Mushroom Body Neurons.” *Journal of Neurophysiology* 99: 734–46.
Villar, Maria E., Miguel Pavão-Delgado, Marie Amigo, Pedro F. Jacob, Nesrine Merabet, Anthony Pinot, Sophie A. Perry, Scott Waddell, and Emmanuel Perisse. 2022. “Differential Coding of Absolute and Relative Aversive Value in the Drosophila Brain.” *Current Biology: CB*, September. https://doi.org/10.1016/j.cub.2022.08.058.
Vrontou, Eleftheria, Lukas N. Groschner, Susanne Szydlowski, Ruth Brain, Alina Krebbers, and Gero Miesenböck. 2021. “Response Competition between Neurons and Antineurons in the Mushroom Body.” *Current Biology: CB* 31 (22): 4911–22.e4.
Wang, Jane X., Zeb Kurth-Nelson, Dharshan Kumaran, Dhruva Tirumala, Hubert Soyer, Joel Z. Leibo, Demis Hassabis, and Matthew Botvinick. 2018. “Prefrontal Cortex as a Meta-Reinforcement Learning System.” *Nature Neuroscience* 21 (6): 860–68.
Wang, Jian, Xiaojun Ma, Jacob S. Yang, Xiaoyan Zheng, Christopher T. Zugates, Ching-Hsien J. Lee, and Tzumin Lee. 2004. “Transmembrane/juxtamembrane Domain-Dependent Dscam Distribution and Function during Mushroom Body Neuronal Morphogenesis.” *Neuron* 43 (5): 663–72.
Wickens, Jeffery R., John N. J. Reynolds, and Brian I. Hyland. 2003. “Neural Mechanisms of Reward-Related Motor Learning.” *Current Opinion in Neurobiology* 13 (6): 685–90.
Wiggin, Timothy D., Yungyi Hsiao, Jeffrey B. Liu, Robert Huber, and Leslie C. Griffith. 2021. “Rest Is Required to Learn an Appetitively-Reinforced Operant Task in Drosophila.” *Frontiers in Behavioral Neuroscience* 15 (June): 681593.
Wilson, Rachel I. 2014. “Early Olfactory Processing in Drosophila: Mechanisms and Principles.” *Annual Review of Neuroscience*, 217–41.
Xu, Wei, and Thomas C. Südhof. 2013. “A Neural Circuit for Memory Specificity and Generalization.” *Science* 339 (6125): 1290–95.
Yamada, Daichi, Daniel Bushey, Feng Li, Karen L. Hibbard, Megan Sammons, Jan Funke, Ashok Litwin-Kumar, Toshihide Hige, and Yoshinori Aso. 2023. “Hierarchical Architecture of Dopaminergic Circuits Enables Second-Order Conditioning in Drosophila.” *eLife* 12 (January). https://doi.org/10.7554/eLife.79042.
Yamada, Hiroshi, Agnieszka Tymula, Kenway Louie, and Paul W. Glimcher. 2013. “Thirst-Dependent Risk Preferences in Monkeys Identify a Primitive Form of Wealth.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (39): 15788–93.
Zheng, Zhihao, J. Scott Lauritzen, Eric Perlman, Camenzind G. Robinson, Matthew Nichols, Daniel Milkie, Omar Torrens, et al. 2018. “A Complete Electron Microscopy Volume of the Brain of Adult Drosophila Melanogaster.” *Cell* 174 (3): 730–43.e22.
Zhou, Mingmin, Nannan Chen, Jingsong Tian, Jianzhi Zeng, Yunpeng Zhang, Xiaofan Zhang, Jing Guo, et al. 2019. “Suppression of GABAergic Neurons through D2-like Receptor Secures Efficient Conditioning in Drosophila Aversive Olfactory Learning.” *Proceedings of the National Academy of Sciences of the United States of America* 116 (11): 5118–25.
Zolin, Aryeh, Raphael Cohn, Rich Pang, Andrew F. Siliciano, Adrienne L. Fairhall, and Vanessa Ruta. 2021. “Context-Dependent Representations of Movement in Drosophila Dopaminergic Reinforcement Pathways.” *Nature Neuroscience* 24 (11): 1555–66.

ADITHYA ECHAMBADI RAJAGOPALAN

Present Position: PhD Student, Solomon H. Snyder Department of Neuroscience, Johns Hopkins University & Howard Hughes Medical Institute, Janelia Research Campus
Present Lab: Lab of Dr. Glenn Turner. **Email:** rajagopalana@janelia.hhmi.org
ORCID ID: https://orcid.org/0000-0002-3184-3647

PAST EDUCATION:

BS-MS Integrated Degree, Majoring in Biology
Indian Institute of Science Education and Research (IISER), Pune, India

PREVIOUS RESEARCH EXPERIENCE:

Topic	Advisor	Location and Date
Prefrontal projections to striatum encode decision variables	Dr. Jeremiah Cohen	Johns Hopkins University, January 2018 – April 2018
Stimulus competition among many stimuli in the barn owl midbrain	Dr. Shreesh Mysore	Johns Hopkins University, August 2017 – August 2018
The role of AL-MB connectivity in insect odor representation	Dr. Collins Assisi	IISER Pune, India January 2016 – August 2017

ACADEMIC PUBLICATIONS:

1)**Adithya E. Rajagopalan**, Ran Darshan, Karen L. Hibbard, James E. Fitzgerald, and Glenn C. Turner. 2023. "Reward expectation direct learning and drive operant matching in *Drosophila.*" Proceedings of the National Academy of Sciences (in press)

2)Ahmed, Maria, **Adithya E. Rajagopalan**, Yijie Pan, Ye Li, Donnell L. Williams, Erik A. Pedersen, Manav Thakral, et al. 2023. "Input Density Tunes Kenyon Cell Sensory Responses in the Drosophila Mushroom Body." *Current Biology: CB* 33 (13): 2742–60.e12.

3)Modi, Mehrab N., **Adithya E. Rajagopalan**, Hervé Rouault, Yoshinori Aso, and Glenn C. Turner. 2023. "Flexible Specificity of Memory in Drosophila Depends on a Comparison between Choices." *eLife* 12 (June). https://doi.org/10.7554/eLife.80923.

4)**Adithya Rajagopalan** & Collins Assisi. 2020. "Effect of circuit structure on odor representation in the insect olfactory system." eNeuro 28 April 2020.

5)Bari, Bilal A., Cooper D. Grossman, Emily E. Lubin**, Adithya E. Rajagopalan**, et al. 2019. "Stable Representations of Decision Variables for Flexible Behavior." Neuron 103 (5): 922–33.e7.

INVITED TALKS:

Topic	**Conference**	**Location and Date**
Learning through surprise: Expectation guides foraging decisions in the fruit fly	Annual Meeting of the Central Virginia Chapter, Society for Neuroscience	Harrisonburg, Virginia April 2023
Reward expectations direct learning and drive operant matching in *Drosophila*	NeuroMEETS Seminar Series, organized by the Max Planck Florida Institute	Jupiter, Florida April 2023
Learning through surprise: Expectation guides foraging decisions in the fruit fly	Baltimore Brain Series, organized by University of Maryland, Baltimore City	Baltimore, Maryland November 2022
Learning rules underlying dynamic foraging in *Drosophila melanogaster*	Junior Scientist Workshop on Mechanistic Cognitive Neuroscience at Janelia Research Campus	Ashburn, Virginia November 2022

ACCEPTED POSTERS:

Topic	**Conference**	**Location and Date**
Reward expectations direct learning and drive operant matching in *Drosophila*	Gordon Research Conference - Neuroethology	Mount Snow, Vermont August, 2023
Reward expectations direct learning and drive operant matching in *Drosophila*	Gordon Research Seminar - Neuroethology	Mount Snow, Vermont August, 2023
Learning rules underlying dynamic foraging in *Drosophila melanogaster*	Neuroscience (organized by Society for Neuroscience)	San Diego, California November, 2022

ACCEPTED POSTERS (Cont)

Topic	Conference	Location and Date
Learning rules underlying operant matching in *Drosophila melanogaster*	Computational and Systems Neuroscience (Cosyne)	Lisbon, Portugal March 2022
Stimulus competition among more than two stimuli in the barn owl midbrain	Neuroscience (organized by Society for Neuroscience)	San Diego, California November, 2022
Functional implications of connectivity in two olfactory circuits – the fruit fly and the locust	Neuroscience (organized by Society for Neuroscience)	San Diego, California November, 2022

SCIENCE COMMUNICATION:

1)**Adithya E. Rajagopalan**, 'Person, Woman, Man , Camera, TV', Nautilus, December 2020.

2)**Adithya E. Rajagopalan**, 'The Brain Cells That Guide Animals', Facts So Romantic – Nautilus, January 2020.

3)**Adithya E. Rajagopalan**, 'New Evidence for the Strange Geometry of Thought', Facts So Romantic - Nautilus, February 2019.

4)**Adithya E. Rajagopalan**, 'Does a Bigger Brain Mean a Higher IQ? Nope, and This Is Why', The Wire (India), April 2018.

5)**Adithya E. Rajagopalan**, 'The Surprising Relativism of the Brain's GPS', Nautilus, March 2018.

AWARDS & HONOURS:

1) Selected to participate in the Santa Fe science writing workshop organized by authors and New York Times science writers Sandra Blakeslee and George Johnson. **(May 2019)**

2) Selected for the Janelia Undergraduate Scholars Program at the HHMI Janelia Research Campus. **(May-August 2016, May-August 2015)**

3) Awarded the 2nd rank at the Foundations of Ecology and Evolution winter school, held at Indian Institute of Science Education and Research (IISER), Pune. **(December 2014)**

4) Summer Research Fellowship by the Indian Academy of Sciences. **(May-July, 2014)**

5) Represented India at the Asian Science Camp in Tsukuba Japan, organized by the Japan Science and Technology Agency. **(August, 2013)**

6) Awarded the C.N.R Rao prize for academic excellence by IISER, Pune. **(March 2013)**

7) Awarded the INSPIRE fellowship by the Department of Science and Technology, Govt. of India, given to the top 1% of Indian undergraduates. **(August 2012 – May 2017)**

LEADERSHIP EXPERIENCE:

Lead Organizer / Mimamsa - Science Quiz Jan 2015 —Apr 2015
Indian Institute of Science Education and Research - Pune Pune, MH, India

• Served as lead organizer of Mimamsa, India's largest student organized annual science quiz which focuses on questions designed to promote deep thought and investigation
• Contacted sponsors and interacted with outreach arms of multiple scientific companies; secured funding for event from two sponsors
• Organized and scheduled 17 prominent research scientists to serve as judges for two-day final event; four finalists competed
• Coordinated team of 200+ students to conduct preliminary written quiz with 400+ participating teams at 10 centers across India; led smaller team of 20 students to produce innovative biology related questions for both rounds of the event

Co-Organizer / Johns Hopkins Dept. of Neuroscience Retreat July 2019 —Sep 2019
HHMI Janelia Research Campus Ashburn, VA

• Member of 7-person organizing committee for annual Johns Hopkins Department of Neuroscience retreat, with 250+ participants
• Invited 18 students and faculty in Department of Neuroscience to speak at two-day event held at HHMI Janelia Research Campus
• Scheduled talks during two-day event and organized local leisure activities for attendees between sessions.
• Moderated one of six event sessions, introducing speakers and coordinating Q&A sessions

TEACHING/MENTORING EXPERIENCE:

Teaching Assistant / Biol 341 Animal Physiology Sept 2022—Dec 2022
Howard University, Washington DC

• Served as a laboratory instructor for students performing behavioral experiments with fruit flies
• Instructed students on the use of MATLAB for analysis of behavioral videos

Laboratory Mentor / Master's Student Jan 2022—Jan 2023
HHMI Janelia Research Campus Ashburn, VA

• Involved in multiple meetings focused on planning structure of thesis project and research direction
• Instructed student on use of existing behavioral equipment in lab and advised student on design of novel experimental setup

Teaching Assistant / Math Methods for Neuro & ML Jun 2019—Aug 2019
HHMI Janelia Research Campus Ashburn, VA

• Served on 11-person organizing group, planned course structure, topics, tutorials and office hours
• Prepared lecture notes and problem sets for week 1 of course, introducing calculus
• Led tutorials and office hours to discuss problem sets and field student questions

Laboratory Mentor / Summer Undergraduate Intern May 2018 —Aug 2018
HHMI Janelia Research Campus Ashburn, VA

• Served as mentor for summer undergraduate intern in lab, enrolled through the Janelia Undergradaute Scholars program
• Instructed student on use of olfactory-learning apparatus in lab to study the role of extended experience on memory
• Evaluated code written by student to analyze data from behavioral experiments
• Advised on structure and design of poster presented at internship conclusion