On adaptive behaviour and its neural basis: Insights from computational models and the mouse auditory cortex



Yves Weissenberger

Madgalen College, University of Oxford

Thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2017/2018

For my mother

Abstract

Yves Weissenberger M.Sc.

Magdalen College, University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

Michaelmas 2017/2018

The processes of learning and decision making are at the core of intelligence. Although these phenomena have been studied for centuries, their nature and biological implementation are still unclear. In this thesis, I approach these issues from multiple perspectives.

First, I developed a novel infrastructure for performing large-scale behavioural and optogenetics experiments in rodents. Additionally, I performed validating experiments showing that auditory cortex is necessary for sound detection in mice.

To gain further insights into mouse behaviour, I developed a novel model of head-fixed mouse behaviour. I applied this model to data from mice performing a sound detection task and showed that it accurately describes behaviour. Additionally, I used this method to obtain a deeper understanding of behavioural strategies. Next, I extended this model to account for changing strategies over learning and demonstrated that it is able to infer changing strategies on a trial-by-trial basis. By applying it to behavioural learning data, I showed that mice exploit simplicity in task-structure to rapidly acquire an effective behavioural strategy.

Finally, I investigated the neural basis of sound-driven behavioural strategies by performing two-photon calcium imaging in layer 2/3 of the mouse auditory cortex during behaviour. Neuronal activity, which covaried with environmental events, importantly covaried with behavioural choice, potentially reflecting its role in driving behaviour. Modelling of population activity showed that choice-related activity is best accounted for by both changes in population gain and local functional connectivity. These data suggest that local computations in layer 2/3 of mouse auditory cortex contribute to sound-driven behaviour.

Together, this work has generated novel tools and represents an almost end-to-end study of the processes of learning and decision making, providing a deeper understanding of behaviour and its neural basis.

Acknowledgements

First and foremost I would like to express my deepest gratitude to Andy King and Johannes Dahmen.

Andy, I cannot express how grateful I am for having been a part of your lab for the past three years, I have had amazing time and learned an incredible amount. None of it would have been possible without your kindness, help and guidance; it has truly been inspiring to work with you.

Johannes, it has been amazingly fun working with you and I am very grateful to have had you as a supervisor and as a friend. I have learned an incredible amount under your supervision and this work would not have been possible without you. I would also like to thank Peter Keating who, despite being busy travelling between Oxford and London, was always happy to offer advice, support and guidance.

I'm also deeply indebted to the whole auditory neuroscience group, for advice, guidance and fun christmas lunches. In particular, I would like to thank Fernando Nodal, Victoria Bajo Lorenza, Kerry Walker and Nicol Harper for numerous valuable discussions, particularly related to mice, modelling and

many other things.

Martin Kahn and Michael Lohse. The last three years would not have been the same without our many hallway discussions, climbing sessions, drinks at the pub, but mainly all the late nights we spent together at the lab. You guys have made an incredible contribution to this work and the last few years of my life. Thank you!

I would also like to thank my office-mates Sebastian Vasquez-Lopez and Samuel Picard for great company and friendship over the last few years and listen to me moan about this thesis for the past few months. I'd also like to thank Ed Mann, Pantelis Antonoudiou, Alexei Bigrave, Jacqueline Stacey, Will Maddern, Julian Bartram and all the rest of the Oxford gang, you have made these last few years amazing!

To James Phillips. Your friendship and our work together over the last 7 years has been an incredible. I can honestly say that my knowledge, understanding and love of neuroscience would not be the same if we hadn't met. Our discussions and collaborations always inspire me and I hope they will continue for the rest of our careers.

Also to my family: Dinah, Adrian, Oma and Mama. Thank you for all your support, love and advice over the years. I would not have been able to even start producing this work without you.

Finally, to Lydia. The last three years we have spent together have been some of the best of my life. Your support and love have made the good times better and the hard times easier. You know, as well as I do, the many reasons

that the writing of this thesis would not have been possible without you. I am always amazed and inspired by you and astonished at how patient and supportive you have been these last few months. I can only hope I will be somehow able to repay you. I love you.

Scientific Acknowledgements

The electrophysiological data in chapter 2 were collected together with Michael Lohse. I thank Peter Keating for providing code for presenting pure tone stimuli for tonotopic mapping performed in chapter 5. I am also greatly indepted to the former and current members of the DPAG mechanical workshop, Mark Howard, Daniel Heaton and Marcas O'Bardain. Finally, I would like to express an additional thanks to Johannes Dahmen for guidance on surgical procedures, imaging and technical design of the training apparatus. I would also like to thank Stacey Jackson for advice on typesetting.

Contents

1 Introduction **35**

1.1 Some problems are harder than others 35

1.1.1 Interactions between neuroscience and artificicial intelligence 36

1.2 Behaviour 41

1.2.1 The paradigm for quantitative behaviour 43

1.2.2 The difficulty of understanding behaviour 45

1.3 Learning 50

1.3.1 The reinforcement learning framework 50

1.3.2 Representation is key 53

1.4 Understanding behaviour by understanding the nervous system 56

1.4.1 Understanding how the brain implements behaviour . . 58

1.5 Current approach and contributions 64

2 A platform for large scale behavioural experiments **66**

2.1 Introduction 66

2.2 Methods 70
2.2.1 Behavioural Experiments 71
2.2.1.1 Water Regulation 71
2.2.1.2 Stimuli 71
2.2.1.3 Optogenetic Behavioural Experiments 72
2.2.1.4 Freely Moving Frequency-Discrimination . . . 73
2.2.1.5 Head-fixed Frequency-Discrimination 74
2.2.2 Surgeries for optogenetics experiments 74
2.2.3 Electrophsyiological Recordings 76
2.2.4 Analysis 77
2.3 System Architectre 78
2.3.1 Overview 78
2.3.2 Server side 79
2.3.3 Client side 82
2.3.4 Apparatus 85
2.3.5 Optogenetics 85
2.4 System Validation 87
2.4.1 Two-alternative forced choice task 87
2.4.2 Optogenetics 94
2.5 Discussion 97
2.5.1 Frequency Discrimination Behaviour 97
2.5.2 Comparison to other Behaviour Infrastructures 98
2.5.3 Optogenetics 99

2.5.4 Conclusion 100

3 Encoding and Decoding head-fixed Behaviour 101

3.1 Introduction 101

3.2 Methods 107

3.2.1 Behavioural Paradigm 107

3.2.2 Basic Model 108

3.2.3 Fitting procedure -Estimating the animal's policy . . . 110

3.2.4 Empirical Bayes-ish estimation 112

3.2.5 Decoding the causes of behaviour 115

3.2.6 Statistical analysis 118

3.3 Results 120

3.3.1 Posterior Predictive Checks 124

3.3.2 Influence of non-stimulus factors on reward rate 127

3.3.3 Decoding the causes of behaviour 133

3.3.4 Quality of Fits 136

3.4 Discussion 138

3.4.1 Detailed descriptions of behaviour 138

3.4.2 Understanding the value of behavioural strategies . . . 142

3.4.3 Future work 144

3.4.4 Conclusion 145

4 Dynamical Models of Behavioural Change 146

4.1 Introduction 146

4.2 Methods . . . 150
4.2.1 Animals and Behavioural Paradigm . . . 150
4.2.2 Generative Model . . . 151
4.2.3 Inference and Learning . . . 153
4.2.3.1 E-step . . . 154
4.2.3.2 M-step . . . 155
4.2.3.3 Fitting . . . 156
4.2.4 Policy Gradient Methods . . . 157
4.2.5 Statistical analysis . . . 160
4.3 Results . . . 160
4.3.1 Validation on simulated data . . . 161
4.3.2 Tracking behavioural change . . . 166
4.3.3 Mice use structure inherent to task to quickly reach high performance levels . . . 170
4.4 Discussion . . . 173
4.4.1 Conclusion . . . 178

5 Neural Correlates of Behaviour 179

5.1 Introduction . . . 179
5.2 Methods . . . 183
5.2.1 Animals and Behaviour . . . 183
5.2.2 Window surgery . . . 184
5.2.3 Two-photon imaging . . . 184

5.2.4 Analysis . . . 185
5.2.4.1 Behaviour . . . 185
5.2.4.2 Two-Photon Data Preprocessing . . . 185
5.2.4.3 Triggered averages . . . 186
5.2.4.4 Choice Probabilities . . . 186
5.2.4.5 Generalized Linear Model fitting . . . 186
5.2.4.6 Analysis of Generalized Linear Model Parameters . . . 188
5.2.4.7 Choice Decoding . . . 190
5.2.5 Statistical Analysis . . . 190
5.3 Results . . . 191
5.3.1 Model free analysis of neural activity . . . 195
5.3.2 Model based analysis of neural activity . . . 201
5.3.3 Functional connectivity in pA1 . . . 204
5.3.4 Model based analysis of decision related activity . . . 208
5.4 Discussion . . . 216

6 Discussion **221**

6.1 Overview . . . 221
6.2 Large scale behaviour . . . 222
6.3 Policy-based descriptions of learning . . . 224
6.4 Cortical contributions to sound driven behaviour . . . 227
6.5 Conclusion . . . 238

A Behaviour Apparatus **295**

List of Figures

2.1 Overview of the system architecture. See main text for details. Abbreviations: U - user; Rpi - Raspberry Pi 80

2.2 Example psychometric functions from single sessions of head-fixed mouse behaviour. Three or four behavioural sessions from each mouse are shown. Top two rows are sessions from one mouse. Next row and first panel of third row are from a second mouse. Remaining are from the third mouse. Error bars show 95% binomial confidence intervals. 89

2.3 Example psychometric functions from single sessions of home-cage mouse behaviour. Error bars show 95% binomial confidence intervals. 92

2.4 Summary of behaviour. Mean ± 2 s.e.m lapse rates A, biases B and thresholds (where 75% correct) C averaged across all sessions from all animals. **D** distribution of self-paced inter-trial intervals from the training of mice shows most inter-trial intervals are short. **E** Mice behaviour in the home-cage is largely restricted to the dark phase of their day-night cycle. . 93

2.5 Optogenetic experiments showing dependence of sound detection on auditory cortex. **A** Histological validation of expression. **B** Reduction in fraction correct (stimulus vs catch-trials) as a result of optogenetic stimulation. Each pair of points denotes a single session. Each colour denotes behaviour of a different mouse. **C** Occluding auditory cortex from the LED rescues behavioural performance. **D** Electrophysiological validation of the effects of LED stimulation. Red dashed lines indicate onset of a 100ms 80dB SPL noise burst. Blue lines denote boundary between onset and offset of optogenetic stimulation. 96

3.1 Figure illustrating the structure of the model. *(Top)* Output of the model is lick probability (blue line) which accurately tracks licks (black line). Yellow vertical lines indicates clicks, red vertical lines indicate rewards (rews). *(Bottom left)* Parameters of the model which are fit to the data are β weights of the regression model. *(Bottom Middle)* Structure of model predicting licks. Events (vertical lines) are convolved with inferred parameters of the model (grey horizonal lines). The resulting time series are added together, and with the inferred offset parameter, and passed through a sigmoid whose output is the probability of licking at each timepoint. 114

3.2 Example behaviour on a single session shown as a lick raster. Each row represents a trial, with black horizontal lines indicating licks. Blue line indicates stimulus onset. 118

3.3 Consistent and variable elements of mice' behavioural strategies. Effect of click **A**, licks **B** and rewards **C** on future licking. **D** Range of bout initiation rates observed. **E** Range of peak positive values of lick kernels. **F** range of peak positive values of stimulus kernels. **G** range of positive durations of licks on future licks. **H** Correlation between inferred bout initiation rates and false alarm rates. **I** Correlation between stimulus sensitivity and hit rates. **J** Example fit to a training session shows prediction (blue) of licks (black) based on licks, rewards (red) and clicks (yellow) 121

3.4 Summary statistics of mouse licking structure. *Left top*, example of a complex reaction-time distribution for measured data and data simulated from the model fit to the corresponding dataset showing good correspondence between the two. Curves are kernel density estimates while vertical lines on axes show individual cases. *Left bottom* summary of pearson-r^2 between kernel density estimates and simulated and measured reaction time distributions. *Centre top* Example empirical inter-lick interval distributions from measured data and simulated from corresponding fit. *Centre bottom* correlation coefficients between measured and simulated empirical distributions. *Right top* Peri-stimulus time histogram (in 20ms bins) of licks from measured example measured dataset and simulated from fit to the same dataset. *Right bottom*, mean-squared error between measured and corresponding simulated dataset. 125

3.5 Comparison of traditional performance metrics derived from simulation and measured in real datasets. For each dataset and single animal, our model was fitted to the data and an equal number of trials to those in the measured dataset were simulated from our model. Comparison between measured and simulated session hit-rates **A** false-alarm rates **B** and overall lick-rates **C**. 127

3.6 Characterising the value of policies. **A** exponential fit to decaying phase of lick-kernel. **B** Summary of quality of fits to all lick-kernels. **C** Distribution of decay time constants of lick-kernels. **D** Effect of varying policy parameters on fraction of rewarded trials (i), rewards per lick (ii), as well as lick-rates (iii). Lick rates above 10Hz are blacked out in the top two rows as they were never observed in the real data. 129

3.7 Linear fits to functions of reward. **A** Reward rate. **B** Reward-efficacy. 131

3.8 Compensation for poor stimulus sensitivity by adaptation of non-stimulus related parameters **A** across all sessions Bout initiation rates and stimulus sensitivity are anticorrelated. **B** Low stimulus sensitivity is associated with longer autocorrelations of licking behaviour in mice **C** No statistically significant relationship between the peak-height of the lick-policy filter and stimulus sensitivity 132

3.9 Decoding the causes of behaviour **A** Reaction time distribution across all animals and sessions. Red shading indicates region heuristically defined as a hit-trial in the following analysis. **B** Decoding performance separated into hit-rates (normalised to number of responses to stimuli), correct-reject rates (normalised to total number of lick bouts and fraction correct (mean of the previous two metrics). Results for behaviour with one-and two-lick spouts shown separately. **C** Example of decoding performance **D** Example decoding of stimuli from lick trains. Black and green vertical lines indicate lick times. Red line indicate stimulus times and purple lines indicate decoded stimulus times. 134

3.10 Comparison of encoding quality between GP-GLM and L2-GLM137

4.1 see below 162

4.1 Example of all parameter fits to simulated data. Each panel shows the changing value of 1 behavioural parameter. In blue, the ground truth parameter value changing over the course of 200 trials of simulated learning. In black, the estimate from the ssglm. grey shaded regions indicate the 99% confidence intervals based on a Gaussian approximation to the posterior. The indices appended to events at the top of individual panels indicate the time in blocks of 100ms after the event indicated above (e.g. clicks2 is the response to parameter affecting lick probability 200ms after onset of a click) 163

4.2 **A** Distribution of mean correlation coefficients between estimated and ground truth parameters, weighted according to the fraction of the variance each parameter contributed **B** distribution of correlation coefficients. each dot represents a single parameter from a single animal, colour-coded according to the parameter type. **C** distribution of mean-squared errors between estimated and ground truth parameter values. **D** same as **C** except sorted by parameter, marginalising across animals 164

4.3 Inferred step sizes plotted against actual step sizes in the simulated dataset. Overall there is a weak correlation between the two 165

4.4 Parameter fits identified as poor on the basis of correlation coefficients are accurate in terms of mean-squared error. **A** Randomly generated selection of example fits where actual and inferred parameters are poorly correlated. **B** Distribution of parameter standard deviations for all parameters (white) and poorly inferred parameters (red) 166

4.5 example of a set of policy filters fitted to left licks and to right licks over the course of learning for an example animal. Dashed line is at 0. lickL: lick at left spout; lickR: lick at right spout; rewL: reward to left spout; rewR: reward to right spout 167

4.6 **A** reconstruction quality of fold change in lick probability pre vs post stimulus from the simulated data **B** Fold probability of licking pre vs-post stimulus over the course of learning for two example animal. **C** same as **B** for different example animal. **D** Mean fold-lick-probability pre vs post-learning for all animal's whose data were fitted. **E** Across learning, bout-initiation rates (see chapter 3) decrease **F** Across learning, there is no change in the peak of the lick autocorrelation kernel **G** Distribution of fractional change of stimulus sensitivity between vs within sessions. **H** Standard-deviation of the distribution of parameter values over the learning process for each session. Error bars are ± 2 s.e.m 168

4.7 **A** Cumulative variance explained as a function of number of principal components included for each animal in our dataset. Vertical lines indicate number of principal components required to explain 80% of the variance. **B** cumulative variance explained as a function of the number of principal components included when performing PCA on the dataset of all animals. **C** accuracy in the reconstruction, based on simulated data, of the loadings of the principal components for the first principal component **D** same as **C** except for the second principal component. **E** Loadings of the first principal component separated by feature **F** same as **E** except for the second principal component. 172

5.1 Summary of behavioural experiment **A** Schematic showing the state structure of our behavioural task. **B** Data from chapter 2 showing that optogenetic inhibition of auditory cortex impairs performance on our behavioural task. **C** Four example psychometric functions demonstating level-dependence of sound-driven behaviour. NS denotes hit rates where no stimulus was presented (catch trials) **D** Schematic of mouse performing our behavioural task **E** Distribution of reaction times **F** Linear correlation between stimulus level and mean reaction time; error-bars are 2 s.e.m. 192

5.2 Functional identification of auditory cortex. **A** Example imaging region. Scale bar denotes 200µm **B** Tuning curves of 6 example neurons as a function of stimulus frequency. **C-E** Example tonotopic maps acquired by mapping best frequencies onto neural location. Only neurons with significant frequency tuning (one-way ANOVA $p < 0.001$ are included in the figure. M: medial; R: rostral 194

5.3 see below 197

5.3 Diverse responses to external events in auditory cortex. **A** Example fluorescence traces ($\Delta F/F$) in black and overlayed inferred spike rates (red). **B** Triggered average traces from all **(i)** or example **(ii)** neurons tuned to stimuli, rewards or licks. Responses are triggered to onset of events named in text above figure panels. **C** Representative example of evoked population responses as a function of stimulus level. **D** Evoked responses averaged over the grey window in C show sigmoidal level dependence of population responses. **E** Distribution of response selectivities, as assessed using the triggered average method. **F** Joint feature tuning assessed using the triggered average method. Data are normalised along rows, such that each row is the fraction of neurons active following the occurrence of a given event which also respond to other events. 198

5.4 Choice related activity in primary auditory cortex. **A,B** Example inferred spike rates of two putative highly choice selective neurons. **C** Distribution of choice probabilities calculated using the first ∼250ms following stimulus onset. **D** Same as **C** except calculated comparing hit and early-response trials. **E** Correlation CP calculated based on hit-vs-miss trials and CP calculated based on hit-vs-early trials. 200

5.5 Poisson Generalized Linear Models capture neuronal firing rates. **A** Schematic of Generalized Linear Models fitted to imaged neural data. **B** Cross-validated Pearson-r of individual neurons fit with a coupled GLM or uncoupled GLM shows coupled GLM provides better fits. **C** Example GLM fits to single trial neural activity. Two randomly selected examples from the top tenth percentile of fits are shown together with three examples close to median (as determined by cross-validated pearson-r). **D** Horizontal smoothed histograms showing the marginal distributions of parameters across all imaged populations. The height of the histogram indicates the number of neurons with a level of feature preference β-values were calculated by summing thresholded β parameters of the model for each neuron. **E** Example coupling matrix showing sparse coupling of populations in pA1. 202

5.6 Tuning Covariance Structure mouse pA1. In off diagonal plots, each dot represents the tuning of a single neuron to pairs of features. In diagonal plots, the marginal distribution of tuning to single features. β.s were thresholded and include only β. where $p < 0.05$. 205

5.7 Coupling structure between neurons in pA1. **A** Schematic showing random coupling between neurons with diverse feature preferences. Colors match histograms. **B** A priroi predicted structure of coupling in pA1 populations in layer 2/3. Neurons with similar feature preference are more strongly coupling. **C** Pairwise coupling probability as a function of source and target selectivity demonstrates that neurons preferentially couple with neurons sharing the same dominant feature. **D** Same as **C** except based on coupling strength rather than probability. 207

5.8 Population decoding from pA1. **A** Population choice probability as a function of time post stimulus **B** Cumulative explained variance of decoding weights accounted for by each principal component. **C** Decoding weights of individual neurons (grey) and the first two principal components 209

5.9 see below 212

5.9 Origins of choice related activity in pA1. **A** Schematic illustrating stimulus drive contributes to pA1 activity. **B** Cumulative histogram showing differences in stimulus drive between hit and miss trials does not differentiate decision+ and decision– neurons. **C** Stimulus drive to decision related neurons is larger on hit trials than on miss trials. **D** The difference in stimulus drive between hit and miss trials is larger for decision related neurons than neurons without significant decoding weights. **E** Schematic illustrating changes in functional coupling. **F** Coupling is stronger on hit trials than on miss trials across the population of decision related neurons. **G** Cumulative histgram of differences between hit trial coupling and miss trial coupling show that changes in coupling differentiate decision+ from decision– neurons. **H** Decision+ neurons increase activity on hit trials. **I** Coupling to decision+ neurons selectively increases on hit trials. **J** Selective increase of coupling from decision+ neurons to other decision+ neurons. **K-M** Same as *H-J* except for decision– neurons. Error bars represent 2 × s.e.m. 213

6.1 Schematic of a canonical decision making circuit . Processing stages which undergo representational plasticity are colored in blue; states where changes cause readout plasticity are coloured green. 230

6.2 Differential predictions of the structure of plasticity under the assumptions of representational and read-out plasticity. *(Top)* Post-learning, we can accurately read-out population activity (chapter 5) to predict animal's behaviour on single trials. We can do so by estimating the decoding weights and using them to project population activity on a 1-dimensional axis. By convention, values greater than zero along this axis predict that a given trial is a hit trial; values below 0 indicate that a given trial is a miss trial. The distance from zero corresponds with the confidence of the prediction. *(Middle)* An idealised form of representational plasticity, in which readout weights are intact from early learning. Under this assumption, neural activity on trials during initial learning mostly look like miss trials when onto the decision axis (defined post-learning). However, rare trials whose projected value exceeds 0 would result in hit trials. *(Bottom)* An idealised model of readout plasticity predicts that activity along the decision axis is distributed just as it is post-learning (i.e. no representational learning), but that activity along the decision axis, defined post learning, is unrelated to behaviour. 231

6.3 Ability to predict mice' behaviour on single trials on the first day of learning and the first day of testing in an example animal. Just as above, neural activity on trials projected to positions greater than zero along the decision axis are predicted to be hits, trials projected to below zero are predicted to be misses. A single decoder trained on post-learning data is able to perfectly predict hit and miss trials. Although there were substantially more miss trials in early learning, equal numbers of hit and miss trials on each day were selected for this analysis 233

6.4 Cross validated prediction of neural activity including a single trial estimate stimulus sensitivity inferred using the state space GLMs described in chapter 4. Sessions correspond to the greatest change in fraction correct between the first and final 20 trials of training. In both cases this was the second day of training. 234

6.5 Pooled data from two animals shows that the decision readout weights are highly correlated between the first day of training and during testing 235

6.6 Differential behaviour pre-hit and pre-miss trials. **A** Prior to miss trials, animal's lick rates are higher. **B** Prior to miss-trials mean inter-lick-intervals are longer 237

A.1 Behaviour apparatus. **A** Example of three-port behavioural apparatus. Central speaker is protected by a grill. **B** Behavioural apparatus inset into an animals home-cage. **C** Example of a head-fixed apparatus for auditory experiments, in sound shielded enclosure. **D** Example of a head-fixed apparatus used for visual discrimination experiment in which mice performed an orientation discrimination task. **E** Rack of individually ventilated cages with behaviour apparatus inset and electronics on top. Image illustrates the small footprint of our apparatus. **F** Room for behavioural experiments with 5 enclosures for head-fixed experiments visible. 296

A.2 Circuit schematics and apparatus for optogenetic experiments. **A** Circuit used for lick detection. **B** Circuit for reward delivery. **C** Printed circuit board of circuit shown in B together with connectors, shows small footprint of apparatus. **D** Constant current LED driver circuit used for optogenetics. **E** Head-cap used for head-fixed optogenetic experiments. **F** Head-fixed optogenetic apparatus showing masking lights. . . 297

List of Tables

2.1 Python libraries for implementing behavioural scripts 84

2.2 Experimental hardware 86

Chapter 1

Introduction

The topic of this thesis is the study of behavioural strategies, how they change with experience and their neural correlates. By studying this, we hope to gain some understanding of the origin of intelligent, adaptive behaviour. In the introduction, I will discuss my subjective view of the field. The goal is to place the work that is to be done into context and provide and argument for why I chose to pursue the problems addressed in this thesis.

1.1 Some problems are harder than others

The study of intelligence is at the intersection of artificial intelligence and neuroscience (Hassabis et al., 2017). There is a long tradition of trying to build intelligent systems that can solve problems and theories of learning have benefited substantially from this work. A major contribution of this

work, at least historically, has been to differentiate those problems that are difficult to solve from those which are not. This labels problems as interesting from an information processing point of view. In describing the history of vision research Marr (Marr, 1982) recounts:

> The first great revelation was that the problems are difficult [...] In the 1960s almost no one realised that machine vision was difficult [...] The notion of a feature detector was well established by Barlow and by Hubel and Wiesel and the idea that extracting edges and lines from images might be at all difficult simply did not occur to those who had not tried to do it.

If we consider the problems neuroscience is addressing now, many are at the level of the implementation of elementary computations (Yuste, 2008). It is not clear, a priori, whether this is a useful perspective from which to proceed to understand cognition (Krakauer et al., 2017). In light of the suggestion above, at least we can infer that the problem of the implementation of computations is interesting, regardless of whether it ultimately leads us to understanding the inner workings of the mind or not.

1.1.1 Interactions between neuroscience and artificial intelligence

Many of the major contemporary principles that are core to the success of artificial-intelligent systems are derived from neuroscience (Hassabis et al.,

2017). Examples of this abound and often referenced by researchers in the field (listening to any talk by Geoffrey Hinton would provide a well-rounded picture). Perhaps the main reason why neuroscience is central for developing artificial-intelligent systems and why so many insights have been derived from the study of psychology and neuroscience, is that we know that humans and animals are highly effective information processing systems. They can solve an incredible array of problems and how they achieve these feats has been the subject of extensive research.

For instance, the hierarchical organisation of information processing, which has been a core organising principle for the study of biological sensory systems Hubel and Wiesel, 1965; Hubel and Wiesel, 1962; Van Essen and Maunsell, 1983 for many decades, has recently found its way into artificial-neural networks in the form of deep-learning (LeCun et al., 2015; Hinton, 2006). Temporal-difference learning, now a core algorithm in artificial intelligence (Sutton and Barto, 1998; Li, 2017), was originally derived as a theory to explain learning of experimental animals (Sutton and Barto, 1981). Dropout (Srivastava et al., 2014), a major regularisation technique applied to artificial neural networks, is inspired by stochastic spiking of cortical neurons (Hinton, 2014). Additionally, the modular organisation of cognition, attention, episodic memory, working-memory and hippocampal-replay, now applied in machine-learning, are ideas originating from psychology and neuroscience (Hassabis et al., 2017).

Large numbers of artificial-systems are thus based around ideas and prin-

ciples derived from neuroscience (I will henceforth use neuroscience as an umbrella for neuroscience, psychology and related disciplines studying biological intelligence). In turn, we can view their applications as experiments of what happens when artificial systems, based on neuroscientists notion of cognition, attempt to tackle real world problems. Doing so has both emphasised the importance of our ideas and highlighted their limitations.

From this work, it is clear that we have made substantial progress towards understanding the basis of intelligence (LeCun et al., 2015). In the past decades, there has been rapid progress, largely driven by a growing appreciation of the power of biologically inspired artificial neural networks. Problems whose solution seemed decades away have been solved, at rapid pace, including accurate visual classification (He et al., 2015; LeCun et al., 2015), semantically constrained image synthesis (Zhang, Xu, Li, et al., 2016), superhuman performance on computer games (Mnih et al., 2015a), and the ancient Chinese board-game GO (Silver, Huang, et al., 2016; Silver, Schrittwieser, et al., 2017). All of this work has highlighted that we have uncovered at least some important operating principles of intelligent systems.

Contrary to this comes an observation that seems to be the flip side of the above. Although we have been able to solve these problems, the solution methods do not really work as well as they could (Hassabis et al., 2017; Li, 2017; LeCun et al., 2015; Tsividis et al., 2017). For example, to date it has not been possible to build robots with motor control systems which compare favourably to those of humans or animals (Deisenroth, Neumann, and Peters,

2011; Li, 2017). This is true despite the fact that many sophisticated motor behaviours are learnable in artificial environments (). This difference may arise because in physical environments, only certain types of failure lead to continued learning. As an example, falling off a cliff will prevent one from learning to run in the future. On the contrary, artificial agents in artificial environments are freed from this constraint of safe exploration and are thus solving simpler problems.

Another well-known problem is the vastness of data required to train artificial systems (Vinyals, Blundell, and Lillicrap, 2016), particularly the amount of labelled data necessary. This is in stark contrast to biological agents, which can make rapid use of information (; Lake, Salakhutdinov, and Tenenbaum, 2015; Tsividis et al., 2017). It could be argued that this is not a real problem, but just a case of lesser experience, though two powerful arguments repudiate this view. Firstly, humans learn simple computer games faster than artificial neural networks, even at the peak of their learning rate (). Secondly, learning in artificial neural networks does not generalise across tasks but instead largely leads to catastrophic forgetting (Kirkpatrick et al., 2017).

Beyond putting neuroscience ideas to test, artificial intelligence also provides new ideas for neuroscience. Although many examples exist (see Hassabis et al., 2017 for a review) attractor networks provide a nice example because of their broad application in neuroscience. Borrowing from work on dynamical systems and thermodynamics, it was shown that networks with

specific connectivity patterns (now known as attractor networks) displayed dynamics dynamics tended to converge on a certain pattern of activity from various starting points (Hopfield, 1982). These network dynamics might be useful in many scenarios, such as completing partial input patterns or the maintenance of persistent activity (Dayan and Abbott, 2013). Today, a plethora of activity patterns in biological networks, across species, are understood as operating under this umbrella (Bathellier, Ushakova, and Rumpel, 2012; Inagaki et al., 2017; Seung, 1996; Kim, Rouault, et al., 2017).

Finally, computational and mathematical work can provide a normative framework for neuroscientific theories. The field of reinforcement learning has probably provided the most substantial work on normative accounts of neuroscientific ideas. Derived from the observation of animals in natural and experimental conditions (Sutton and Barto, 1998), reinforcement learning provides useful mathematical descriptions of human and animal behaviour. However, experimental observations alone are not sufficient to answer why this kind of learning might be efficient or evaluate how good it is. However, mathematical analysis can provide guarantees for situations in which learning algorithms will converge to the right solution (Watkins and Dayan, 1992) or how long learning might take (Liu, Liu, and Zhao, 2013; Pineau et al., 2015). Such analyses may provide reasonable answers to the questions of why organisms process information the way they do.

As a final point, an interesting question arises from this discussion: if we have domain specific superhuman artificial-intelligences, why do we need

to study biological systems in order to understand intelligence? Would it not suffice to understand the artificial systems we have and improve them in places where they are still limited by engineering?

Two points should be raised with regards to these questions. Firstly, as noted above, biological systems still outperform artificial ones in many respects and are likely to provide continued insight. Secondly, this is not the first time in history when humans were outperformed by computers. It is interesting to note that the 1997 defeat of Garry Kasparov by Deep Blue came as a huge surprise to most technologists and chess enthusiasts alike and was considered to have proven the superhuman intelligence of the computer. "It has the impact of a Greek tragedy" reported Monty Newborn, chairman of the committee officiating the match (Weber, 1997). The Guardian (Harding and Barden, 1997) considered it a "depressing day for humankind in general". However, for at least a decade following the victory, little progress was made and comparable successes were absent. It does not seem unlikely to me that the same thing will happen again. Perhaps then again, new discoveries from neuroscience will push us forward.

1.2 Behaviour

A priori the most obvious way to understand how and why we behave and learn is to study behaviour directly. We can treat the brain as a black-box input-output function; observe and manipulate inputs (e.g. external stimuli,

previous behaviour) and observe the behavioural outputs (i.e. movement). In this terminology, studying some steady-state behaviour would be investigating the properties of the function: what behavioural output is generated in response to what input? Studying learning would be investigating how the function changes over time and experience. Two distinct approaches to understanding behaviour have emerged over recent years.

The first approach is purely observational; it does not manipulate an animal's environment and thus tries to understand the structure of an animal's spontaneous behaviour. In this case, the inputs to the function are either absent or often the animal's previous behaviours. These types of methods have been fruitfully applied to understand Drosophila behaviour (Egnor and Branson, 2016), segment mouse behaviour (Wiltschko et al., 2015) and find usable descriptions of the behavioural repertoire of *C. Elegans* (Stephens, Johnson-Kerner, et al., 2008).

Contemporary instantiations of these methods draw heavily on complex statistical analyses and machine learning (Gomez-Marin, Paton, et al., 2014). They have been used to develop discrete clusters of behaviour, which provide detailed and structured descriptions. The key strength of this approach is that it allows experimenters to look for structure even in cases when the form of such structure might be unknown, or in contexts where there is no obvious way to impose the desired structure experimentally. Examples of the latter include structures defined by the internal state of an animal or enactment of some spontaneous behaviour. These computational methods have come to

the fore more recently and their application has thus been less exhaustive.

Contrary to observational methods, the second approach is based on manipulating subjects' environment and its key advantages are three-fold. Firstly, as an experimenter one can control the aspect of information processing that is required and study a wide range of phenomena by using the right type of stimuli. For instance, If one wants to understand visual processing, it is likely that visual stimuli will be presented. Secondly, repeated manipulation of a subject's environment in the same way allows us to obtain reliable estimates of potentially noisy behaviour, which can be achieved by averaging over multiple repetitions of the same context. Finally, by manipulating environments along different specified axes, it is possible to test specific hypotheses regarding the basis of observed behavioural phenomena (Green and Swets, 1966; Gomez-Marin, Paton, et al., 2014; Tervo, Tenenbaum, and Gershman, 2016).

1.2.1 The paradigm for quantitative behaviour

Psychophysics exemplifies the application of quantitative behavioural approaches to the understanding of information processing (Green and Swets, 1966). In this method, subjects (humans or experimental animals) are repeatedly presented with stimuli and report their judgements about what stimulus they think was presented. By presenting many identical stimuli, one can obtain accurate estimates of a subject's sensitivity to different stimuli. These methods were originally used to quantify the acuity of human perceptual sys-

tems () and have by now been applied to dozens of independent dimensions of perception, across many different modalities.

In such a paradigm, as applied to experimental animals, a set of stimuli along a perceptual axis, such as sound frequency, are repeatedly presented in a randomised order and subjects are asked to give different behavioural responses for different sets of stimuli. The stimulus space, in this experiment, is then divided into two sets, one above 1kHz in frequency and one below. Subjects are then required to give different behavioural responses based on which set the presented stimulus belongs to. Stimuli are then presented varying from 200Hz to 10kHz and subjects respond. The experimenter then asks how close to 1kHz stimuli can be whilst subjects are still able to assign them to the correct set. Although these methods ask conceptually simple questions about our perceptual system, they can provide comprehensive, quantitative answers.

Even though the example above is trivial, these methods can be applied to complex problems. Doing so allows experimenters to disentangle the features relevant to subjects' behaviour (Gomez-Marin, Paton, et al., 2014). In artificial environments, for example, humans and experimental animals exhibit biases in their behaviour (Bak et al., 2016; Busse et al., 2011; Fründ, Wichmann, and Macke, 2014a).

For instance, in the earlier example, they might preferentially give one of the two responses, or might prefer responding to the side where they received the previous reward. Depending on the structure of the task or

subjects' models thereof, these might be adaptive or mal-adaptive (Iigaya et al., 2017). If one stimulus is presented more frequently than another and the task is difficult, then giving the response that is more frequently rewarded for being correct might be an adaptive strategy.

Indeed, previous work has pitted behaviour based on different strategies against one-another, asking whether subjects are capable of finding the optimal strategy to solve a given problem (Herrnstein, 1961; Faisal and Wolpert, 2009). These kinds of approaches ask interesting questions about the computational mechanisms underlying the solution of the task. Therefore, they allow experimenters to measure and quantify behaviour of humans and experimental animals in an efficient way.

As a more intuitively interesting example, real-world stimuli vary in complex ways along multiple axes. However, it is unclear which axes are relevant to making the perceptual decision about a stimulus *a priori*. For instance, it is not obvious at all what visual features we use to identify faces. Psychophysical methods can be applied to answer these questions, yielding results that are consistent across individual perceptual tasks (Houlsby et al., 2013; Sanborn, Griffiths, and Shiffrin, 2010).

1.2.2 The difficulty of understanding behaviour

The approaches outlined above have transformed our understanding of how behaviour is enacted and methods based on these principles are likely to continue to do so. What the above discussion belies is the extent to which

this programme of behavioural description has been systematic.

Even the simplest behavioural paradigms suffer from much of the complexity discussed earlier. As an illustrative example, let us consider pure tone frequency discrimination as illustrated above. A pubmed search (9 Nov 2017) for "pure tone discrimination" filtered between 1996-2016 yields 1180 hits, or 50 per year. Many of these studies investigate the involvement of specific brain regions (e.g. Ohl et al., 1999; Dykstra et al., 2012; Gimenez, Lorenc, and Jaramillo, 2015) or the mechanisms of how these stimuli are encoded by the brain in relation to behaviour (e.g. Kuchibhotla et al., 2016a; Carcea et al., 2017; Jaramillo, Borges, and Zador, 2014). Concomitantly, it is unclear and seldom investigated what physical features of sounds are used to detect and discriminate pure tones (Mao and Carney, 2015; Mao, Vosoughi, and Carney, 2013).

The detection of pure tones can be achieved by several different mechanisms (Mao and Carney, 2015). One could try to detect a sudden rise in the power at a specific frequency or detect prolonged elevated power at a specific frequency band or compare power across frequency bands. Although resolving this question might be irrelevant for many studies, it is highly relevant for many others. For instance if one is trying to understand how the brain implements the detection of pure tones in background noise (McDermott and Simoncelli, 2011) knowing the algorithm being implemented is key. Similarly, if we are using some sensory task to infer the functional role of some sensory brain region, knowing what computations are implemented is necessary

if we are to understand the contribution of that brain region. However, In many cases, without really knowing what the brain is doing, we are trying to understand details of how it does it.

Trying to understand the contribution of activity in some part of the nervous system to the solution of a task, without knowing how the task is solved, will lead to nonsensical conclusions (Krakauer et al., 2017). (Tervo, Tenenbaum, and Gershman, 2016) recount a pertinent cautionary-tale where pigeons were trained to perform complex same-different judgements, but in reality were simply comparing the luminance of images. In sensory tasks. This example highlights the vital importance of properly understanding and controlling the stimulus space. An inability to understand stimulus space is still a pertinent problem for neuroscience, particularly for cognitive paradigms (Garner et al., 2006).

However, the difficulties of understanding behaviour do not end with understanding how subjects perceive the stimulus space. In cases where decision making is studied, problems may end up arising from a core assumption that most quantitative descriptions of choice behaviour make: the assumption that decisions are derived from a single decision process.

When subjects attempt to discriminate very similar sensory stimuli, it is unclear whether they provide their best estimate based on the available information (Brunton, Botvinick, and Brody, 2013) or whether they guess randomly (Tervo, Proskurin, et al., 2014)? Do subjects always follow either of these strategies or do they stochastically switch between guessing and

estimating? While this problem may be insignificant for the estimation acuity thresholds, it is incredibly important when ones tries to understand the neural implementation of these processes.

Studying a mixture of two neural processes as if they are one is not the same as observing a set of noisy observations of a single process. If decisions are assumed to arise from a single process (noting that even something as simple random guessing is a defined neural process ()), but are not, then the function of observed patterns of neural activity will be distorted. This notion is bolstered by the observation that even identical behavioural outputs are driven by qualitatively distinct patterns of neural activity depending on the underlying decision process (Siniscalchi et al., 2016; Durstewitz et al., 2010).

Similar methods to the above have also been used to study more complex cognitive phenomena such as working memory (Luck and Vogel, 2013; Li, Daie, et al., 2016), uncertainty quantification (Faisal and Wolpert, 2009; Ernst and Banks, 2002), multi-decision planning (Akam, Costa, and Dayan, 2015; Daw et al., 2011a), and evidence accumulation (Brunton, Botvinick, and Brody, 2013).

In these cases, there are also plenty of difficulties curtailing certainty in the conclusions drawn by these studies. Such tasks are often highly complex and require substantial training both in human subjects and in experimental animals. Understanding how the subject might be performing these tasks or acquiring relevant information is unclear without careful testing (Park, Lueckmann, et al., 2016; Brunton, Botvinick, and Brody, 2013). This prob-

lem is particularly severe in experimental animals, which cannot be given direct instructions and in many cases plateau at moderate performance levels in difficult tasks (e.g. Schmitt et al., 2017; Howe et al., 2013.

Worse still, these complex tasks may allow for multiple solution methods. (Akam, Costa, and Dayan, 2015) elegantly emphasise this problem in the context of multi-step decision making. They use a systematic computational approach to assess performance in a multi-step decision making task. The task they studied was previously used to show model-based planning in humans (Daw et al., 2011a) and adapted and changed by multiple groups for use in experimental animals. Akam, Costa, and Dayan, 2015 show showed that in the adapted versions of the task simple, model-free behavioural strategies yield very similar behaviour to model-based behaviour and could thus easily be mistaken for such, undermining naîve use of the paradigm.

In this case, a substantial body of work was required in order to understand different solution methods used for this particular task. This is problematic. Due to the lack of consistent methods to actually narrow down the strategies animals can use to solve complex tasks, simpler tasks should be favoured when trying to tease apart an animal's behavioural strategies.

In summary, I have illustrated that quantifying behaviour is a very difficult task. The underlying theme has been that although the interpretation of behaviour may seem straightforward, there are often many different ways to explain why behaviour looks the way it does. Disentangling which of these explanations is appropriate relies on careful design of experimental paradigms

and a detailed quantitative analysis of the behaviour.

1.3 Learning

Let us return again to animals, performing simple tasks in artificial environments. Consider an environment in which an animal receives a reward for pressing a lever, if some specific visual symbol is presented and no reward, if some other stimulus is presented. We can delineate multiple aspects about the structure of the environment that need to be learned to behave optimally: the relevance of the two stimuli for outcomes; the role of the lever press; the absolute value of the outcome;, the differential value of the two stimuli; the irrelevance of everything else (such as where the animal was when the stimulus was presented or the previously presented stimulus).

This is not the only way to partition the set of 'things to be learned'. The correct way to partition them will only become apparent when the problem of learning is solved, but the above example hopefully provides an intuition that there are multiple qualitatively distinct processes involved in procedural learning.

1.3.1 The reinforcement learning framework

Although many approaches to describing learning exist, reinforcement learning presents the most developed framework to think about learning based on seen outcomes. Reinforcement learning is an umbrella term for the study

of how agents learn to optimize behaviour by interacting with their environment. It has seen extensive applications in neuroscience (Sutton and Barto, 1998; Doll, Simon, and Daw, 2012; Kolling and Akam, 2017; Gershman and Daw, 2017; Gershman and Niv, 2010) and proven powerful when used in artificial systems (**Silver2017a**; Silver, Huang, et al., 2016; Mnih et al., 2015b).

Other frameworks for understanding learning agents exist. Among cognitive science the major two others are Bayesian inference (Lake, Salakhutdinov, and Tenenbaum, 2015; Friston et al., 2010) and foraging theory (Stephens and Krebs, 1986). However, recent work suggests that these frameworks could be subsumed under the reinforcement learning umbrella (Weber et al., 2015; Kolling and Akam, 2017). For instance, it has been suggested, that the Bayesian inference framework shares deep mathematical similarities to reinforcement learning (Weber et al., 2015). Additionally, recent research is beginning to explore how reinforcement learning may explain the origins of certain types of foraging behaviour (see Kolling and Akam, 2017 for a review).

In computational reinforcement learning, two broad classes of methods may be delineated (Sutton and Barto, 1998; Deisenroth, 2011): value-based methods and policy-based methods. The former class works by explicitly learning the value of all available actions in experienced contexts and choosing actions by picking the best one in any given context. In contrast the policy-based approach directly tries to learn how to behave effectively by

simply following different strategies and seeing how effective they are (policy is roughly synonymous with the intuitive notion of strategy). Although there are exceptions (e.g. (**lia**; Desrochers et al., 2010)), the vast majority of work in neuroscience is based on value-function methods (Sutton and Barto, 1998; Kolling and Akam, 2017; Gershman and Daw, 2017; Doll, Simon, and Daw, 2012).

The most thoroughly researched aspect of learning in experimental animals is learning about the (different) value of stimuli and actions, typically in the form of conditioning (Maia, 2009)and forced choice paradigms (Sutton and Barto, 1998; Kolling and Akam, 2017; Gershman and Daw, 2017; Doll, Simon, and Daw, 2012). Methods from reinforcement learning have been applied extensively to such paradigms (Gershman and Daw, 2017; Elber-Dorozko and Loewenstein, 2017; Huh et al., 2009; Daw, 2011; Samejima, 2005) and have led us to a substantial understanding of how animals learn to improve their behaviour. Seminal work in the 1980s and '90s showed that temporal difference learning provided a parsimonious explanation for many observed phenomena (Sutton and Barto, 1981). The key observation was that once the association between a neutral stimulus and a reward was created, this stimulus could act in itself as a reinforcer.

A didactic example for this kind of work is that of Ito and Doya, 2009. In essence the task consisted of rats placing their snout in a central nose-poke, whereupon one of two tone was played. One tone was associated with an absence of response, while the other required animals to engage in further

behaviour. Following the offset of the second tone, the rats were required to nose poke in either a left or right port to obtain reward. Rewards were delivered at each port at different frequencies (e.g. 50% left; 10% right; 90% left, 50%right etc) in blocks which lasted several dozen trials. Thus, in order to obtain as much reward as possible, rats had to continuously update their beliefs about the value of poking left or right. The model they developed, a Q-learning based model with some bells and whistles, was based on maintaining estimates of the value of each port after the onset of tone A. They found that their model was able to accurately describe the choice behaviour of rats compared with sophisticated competitors in a forced choice environment where the appropriate action for a stimulus changed.

1.3.2 Representation is key

While this is a rather brief review of reinforcement learning largely focussing on historical approaches, I hope it provides sufficient background for the following discussion. The main problem with the general applicability of reinforcement learning methods to understanding learning in biological agents is that assumptions must be made. Principal among these are assumptions as to how agents abstract the world. In practice, the structure of the world is defined *a priori* according to the experimenter. In many cases the structure is highly reductive. In the example outlined earlier (), for example, the world was essentially reduced to the value of the left and right spouts. However it is unclear whether this is an accurate depiction of how the rat's brain conceptu-

alises the world. This problem applies very generally to the (reinforcement) learning work done in neuroscience as has been noted by practitioners (Gershman and Daw, 2017; Tervo, Tenenbaum, and Gershman, 2016).

This severity of this problem is compounded by recent computational work on learning, which suggests that the development of efficient task representations is vital to successful learning (Li, 2017). If this finding holds, it implies that the most important problem of learning is not studied.

In my opinion, two facts speak to whether the above statement is indeed true. Firstly, once backed by the sophisticated representations provided by artificial neural networks, a wide variety of different learning algorithms, including some of the most naïve algorithms imaginable (Salimans et al., 2017), can solve complex real-world tasks (Li, 2017). Secondly, variants of the reinforcement learning algorithms used in artificial neural networks have been around for decades, but have failed to achieve similar results. Together, these two facts demonstrate that the interesting problem of learnings is not the learning rule itself, but how an effective model of the learning environment is built, a problem largely unaddressed by neuroscience thus far (Tervo, Tenenbaum, and Gershman, 2016). Moreover, it is self-evident that learning effective abstractions of some environment is highly task-dependent and thus an integral part of the reinforcement learning problem itself.

Transfer learning, or learning to learn paradigms, has been the most forceful demonstration that biological agents do learn the structure of their environment (Harlow, 1949). For example, a skilled pianist will more rapidly

be able to learn to play the trombone than someone who has never played an instrument. This is also true in experimental animals. It has been shown that mice trained to perform one sensory discrimination task will learn subsequent sensory tasks with identical structure almost twice as fast (McKenzie et al., 2014). Interestingly, artificial agents are notoriously bad at this, indicating that there is much to learn about how to build flexible yet useful representations (Kirkpatrick et al., 2017).

These observations naturally raise the question as to why this has not been studied more extensively. Indeed, it is possible to envision computational approaches that are capable of modelling the learning of structure, particularly when combined with the wealth of previous knowledge acquired about learning. There are three reasons for why the process of building flexible representations has not been studied more. Firstly, learning the structure of the environment is not a repeatable process and is therefore not amenable to many of the classical methods used. Secondly, if there is only one chance to observe the learning process, this provides very little data with which to constrain models of this process. Indeed, in many of the studies above, the sole outcome variable is the left/right choice of the animal, which provides only 1 bit of information about the animal's behaviour per trial. This limited information cannot be sufficient to inform a model that is expressive enough to capture the diverse range of the animal's possible internal models of the world. Finally, few paradigms have been developed which divulge the structure of already learned abstractions (Tervo, Tenenbaum, and Gershman,

2016). Without knowing how subjects end up abstracting their environment, the question of how they learn this abstraction in the first place is virtually impossible to address.

In sum, reinforcement learning provides a powerful framework for understanding the process of learning in experimental animals. Substantial progress has been made in understanding how animals learn the contextual value of actions. However, how animal's structure the world to implement learning is a key unsolved problem. The core difficulty here arises from the fact that the processes of structure learning is not repeatable. This means that very little data are available to constrain models of this learning process.

1.4 Understanding behaviour by understanding the nervous system

The nervous system is almost universally accepted to be the seat of cognition. This idea originated historically from the observation that information from all sensory organs travels to the brain the brain in turn supplies dense fibre tracts through the spinal cord to innervate the muscles. Situated between input and output then, the brain is viewed as the seat of cognition.

Seminal work over the course of the last century is foundational to current thinking about activity in the nervous system. Early pioneers performed single-unit recordings in the sensory periphery using microelectrodes and made two major discoveries that have been retained to this day. Firstly,

this work showed that there are systematic relationships between the firing rate of individual neurons and presented sensory stimulation (Adrian and Zotterman, 1926; Hartline, 1938), giving rise to the idea of rate-coding. Secondly, it showed that the activity of single neurons can be interpreted as representing meaningful features of the environment, giving rise to the notion of receptive fields ().

The reason these ideas have been so influential, is that they have found applicability throughout the nervous system. Venturing deeper into the brain Hubel and Wiesel, 1965; Hubel and Wiesel, 1962 showed that rate code and receptive fields were useful for understanding cortical function and later O'Keefe and Dostrovsky, 1971 used these ideas to reveal the encoding of space by the hippocampus.

These principles can be applied directly to behaviour and learning. For example, although different learning algorithms can make very similar predictions about behaviour (Akam, Costa, and Dayan, 2015), their implementation involves the computation of different quantities in response to identical inputs (Sutton and Barto, 1998). If we assume the representational framework, we would therefore predict that we should find neurons encoding these quantities in the brain. This provides an alternative way of distinguishing between hypotheses.

The discovery that reward prediction errors are encoded in midbrain dopaminergic neurons (Schultz, Apicella, and Ljungberg, 1993; Schultz, 1998; 2013; Montague, Dayan, and Sejnowski, 1996) illustrates this idea

beautifully. Following these discoveries, the idea that temporal-difference learning based algorithm is implemented by the brain, has become almost universally accepted (Schultz, 2013; Sutton and Barto, 1998; Daw, 2011; Gershman, Norman, and Niv, 2015).

Related reward prediction error work also illustrates another way in which the understanding the brain is useful in understanding behaviour. An open question, as mentioned above, is how animal's construct a model of their environment. A recent theoretical model (Daw, Courville, and Touretzky, 2006) proposed a mechanism for this, which made strong predictions about how reward prediction errors should look in certain environments. By knowing that midbrain dopaminergic neurons encode reward prediction errors, Starkweather et al., 2017 then tested this hypothesis by monitoring the activity of midbrain dopaminergic neurons. This example illustrates, knowledge the variables encoded by a neuronal population can further our understand computational processes relating to those variables.

1.4.1 Understanding how the brain implements behaviour

Another way in which studying the nervous system can elucidate behaviour is in understanding its neural implementation. Perhaps the most studied paradigm in this context is the sensory decision making task (**Poort2014**; Parker and Newsome, 1998; Guo, Li, et al., 2014; Bizley, Walker, Nodal, et al., 2013; Znamenskiy and Zador, 2013a; Park, Meister, et al., 2014; Burgess, Lak, Steinmetz, Zatka-Haas, Bai Reddy, et al., 2017; Guo, Li,

et al., 2014; Ding and Gold, 2012; Law and Gold, 2008). Mechanistically one of the core question studied, has been how the decision-related areas in the brain reads out information from sensory areas in order to make such decisions (Parker and Newsome, 1998; Shadlen, Britten, et al., 1996; Guo, Li, et al., 2014; Haefner et al., 2013). In particular, a commonly asked question is what the readout weight of particular neurons is (; Parker and Newsome, 1998; Shadlen, Britten, et al., 1996). In other words, what is the influence of the spiking of particular neurons on a decision.

Historically, studying the role of sensory areas in decision making was partly guided by views originating in Barlow's *Neuron Doctrine for Perceptual Psychology* (Barlow, 1972), one of whose 'dogmas' states:

> Perception corresponds to the activity of a small selection from the very numerous high-level neurons, each of which corresponds to a pattern of external events [...]

Taken at face value, this enables investigation of the sensory decision making process by recording and manipulating the activity of the groups of relevant neurons. While identifying these neurons may seem like a challenging task *a priori*, the representational principles used above guide the search. For example, in trying to understand the neural basis for motion discrimination, those neurons that encode motion in the firing patterns were (Salzman, Britten, and Newsome, 1990; Britten et al., n.d.).

Application of these approaches led to tremendous initial progress. It was

shown that in monkeys performing visual motion discrimination task, electrical stimulation of neurons, in area MT, tuned to a particular motion-direction could bias an animal's decisions towards the tuned direction (Salzman, Britten, and Newsome, 1990). This provided evidence for the involvement of these groups of neurons in the perceptual decision making process.

In a similar vein, it was shown that when animal's were presented with ambiguous stimulus fluctuations, the recorded activity of single neurons in these areas could predict perceptual report in primates (Britten et al., n.d.; Parker and Newsome, 1998). Strikingly, neurons' choice-related activity was related to their tuning (i.e. increased firing in neurons tuned to the leftward motion was associated with a bias towards leftward decisions) and thresholds (neurons more sensitive to motion showed more choice related activity). Quantifying this covariation between neural activity and perceptual report (termed choice probability), was thought to provide a way to determine the fine structure of how decisions were read-out from neural populations(Shadlen, Britten, et al., 1996).

The inference from this was this body of work was that fluctuations in activity of these populations of neurons drove differences in perception, when stimuli were ambiguous. Together, these observations were combined into a model where small pools of neurons, localised in (task-) specific sensory cortices provide information to decision related areas (; Yang and Shadlen, 2007) which then integrate the information to drive behaviour.

Recent evidence has challenged and this view and expanded our under-

standing of how perceptual decisions are made. In part, this has been driven by immense technological progress over the course of the last 20 years. In this time, our experimental access to the nervous system has increased substantially. Where previously the electrical activity of neurons had been blindly recorded with single electrodes, today the activity of hundreds or thousands of neurons can be recorded simultaneously using large arrays (Jun et al., 2017) of electrodes or sophisticated microscopes (Denk, Strickler, and Webb, 1990; Svoboda and Yasuda, 2006; Sofroniew et al., 2016) combined with fluorescent dyes or indicators (Chen, Wardill, et al., 2013; Dana et al., 2016) sensitive to the occurrence of action potentials.

Firstly, it has become clear that most of the correlation between the activity of single neurons and behavioural choice is not the result of fluctuations in activity driving different perception (the bottom-up view) (Pitkow, Liu, et al., 2015; Katz et al., 2016; Allen et al., 2017; Wimmer et al., 2015). Probably the strongest demonstration of this is the occurrence of very strong choice related activity in the lateral intraparietal area of the macaque monkey (Roitman and Shadlen, 2002). This is evident despite the fact that transient inactivation of this structure leaves decisions unaffected (Katz et al., 2016). Another piece of evidence arguing against this view is choice-related activity is observed in cortical regions of the wrong sensory modality for performing the task (Allen et al., 2017). This somewhat undermines the ability to use the correlations between activity and choice to infer the structure of population readout (though see Haefner et al., 2013 and/or Chicharro).

More recent views have emphasised the heterogeneity of the mechanistic origins of this type of activity. For example, by carefully aligning trial by trial behaviour and neural activity Nienborg and Cumming, 2009 provide evidence that choice related activity is largely the result of top down mechanisms. Consistent with this, Allen et al., 2017 show that removal of frontal areas leads to significant diminishing of choice related activity. Similarly, Kwon et al., 2016 recorded activity in multiple cortical somatosensory areas of the mouse and argued that choice-related activity arises, at least in part, from recurrent activity in loops between these areas.

Finally, even neurons whose tuning is orthogonal to the motion direction being discriminated (and hence should have no bearing on the decision) can show choice related activity Cohen and Newsome, 2009; 2008. These authors argued that that, in these cases, choice-related activity arises because the firing of these neurons is correlated with the activity of neurons contributing to the decision.

Thus choice related activity is thought to arise from three sources: The correlation structure of populations, top down input and bottom up effects. Consequently, interpreting the contribution of neurons to decisions using this recording approach is not straightforward (Haefner et al., 2013; Pitkow, Liu, et al., 2015).

A second important factor involvement of redundant parallel pathways in driving sensory behaviour. Work in simple model systems has extensively shown that identical patterns of population activity can be stably generated

by different mechanisms (Marder, Goeritz, and Otopalik, 2015), suggesting a priori that parallel pathways could heterogeneously implement function.

In the auditory cortex, choice related activity has also robustly been observed (Bizley, Walker, Nodal, et al., 2013; Runyan et al., 2017) and inactivation of the auditory cortex has been shown to impair function, even on tasks as simple as tone detection (Talwar, Musial, and Gerstein, 2001). At the same time, chronic lesions of the auditory cortex do not impair function, even on adaptive sound processing tasks (Gimenez, Lorenc, and Jaramillo, 2015), suggesting that other circuits are sufficient to drive behaviour. TA potential anatomical substrate of the parallel function are parallel projections to subcortical decision related areas (LeDoux, Farb, and Romanski, 1991), which are increasingly viewed as crucial to sensory based decision making (Znamenskiy and Zador, 2013b; Xiong, Znamenskiy, and Zador, 2015; Ding and Gold, 2012; 2010).

A final major conceptual challenge to the approach relates to the interpretation of neural activity. There is an increasing recognition that neural activity is the sum of its parts, with increasing recognition that neural activity in the sum of its parts (Pillow, Shlens, et al., 2008; Abeles and Gerstein, 1988). This contrasts with Barlow's doctrine, where the hope was that the brain could be dissected and understood by examining the contribution of single neurons, one at a time (**Marr1981**; Barlow, 1972).

An illustrative example is provided by work in motor cortex (for a more general review see Stevenson and Kording, 2011) which has served as a

testbed for developing conceptual and analytical approaches to neural population activity (**Elsayed**; Yu et al., 2009; Churchland et al., 2012; Kao et al., 2015). Unlike activity in sensory areas, on which most of Barlow's doctrine was based, activity in motor cortex is not well described on a neuron-by neuron, feature by feature basis (Fetz, 1992; Scott, 2008). This poses a major challenge for Barlow's doctrine in it suggests that the important aspects of neural activity are not necessarily the activity of individual cells, but the global covariation and dynamical structure of the population activity. Indeed, when analysed from this perspective activity in motor cortex is simple and easily relatable to motor output (Churchland et al., 2012; Kao et al., 2015). From this view, the entire structure of population activity needs to be considered when trying to understand neural activity.

1.5 Current approach and contributions

As stated in the introduction, the goal of this thesis is to develop a better understanding of the basis of behaviour. We tackled this goal using several different approaches:

In chapter 2 we describe a platform for performing behavioural experiments and optogenetics at scale, which will greatly facilitate the process of experimental design and data collection.

Using these methods we proceed, in chapter 3 to develop models to describe the behaviour of head-fixed mice. Specifically build Bernoulli-Generalized

linear models which model the detailed licking pattern of head-fixed mice and show they that mice use multiple adaptive strategies to obtain reward.

In chapter 4 we extend the models from chapter 3 to be able to track changing behaviour as mice learn. We build state-space generalized linear models capable of tracking mice strategies with single trial resolution. Using the developed methods, we show that mice exploit the simplicity of the structure of their environment while they learn to behave optimally.

Finally, in chapter 5, we use two-photon imaging to characterise the activity in mouse auditory cortex during the performance of this task. We show that the activity in auditory cortex is dominated by the presented auditory stimuli, but at the same time, shows robust encoding of other behavioural variables, including the choice of the animal.

Chapter 2

A platform for large scale behavioural experiments

2.1 Introduction

Understanding the neural basis of behaviour is a key goal of modern neuroscience (Krakauer et al., 2017). Indeed, the principal function of the brain is to produce behaviour (Gomez-Marin, Paton, et al., 2014). In accordance with this notion, the study of behaviour continues to exert a powerful influence on neuroscience. This influence stems in part from the fact that behavioural experiments can identify computations performed by organisms (Sutton et al., 1999a; Shadlen and Kiani, 2013; Harlow, 1949; Tervo, Tenenbaum, and Gershman, 2016; Montague, Dayan, and Sejnowski, 1996; Herrnstein, 1961; Li and Dudman, 2013), and hence the brain. In doing so, they guide the

search for principles which can be used to explain neural activity (Marr, 1982; Krakauer et al., 2017; Gershman, Norman, and Niv, 2015). In addition, once a computation has been identified, behavioural paradigms provide an experimental context in which to understand the neural-circuit mechanisms by which these computations are implemented (Shadlen and Kiani, 2013; Schultz, 2013; Krakauer et al., 2017).

The basic biology of mice makes them an attractive model organism for behavioural and neuroscience experiments. The reasons for this are numerous. Firstly, the gross organisation of the mouse nervous system is similar to that of humans and other primates (Brown and Bowman, 2002; Kaas, 2007; Sereno and Allman, 1991), making them an appropriate model for neural function in humans. Secondly, mice are small in size and thus minimise housing costs and maximise space-efficiency. Thirdly, their short generation time (Silver, 1995) and well understood genome (Waterston et al., 2002) mean that they are highly genetically tractable. Combined with modern genetic tools this has been invaluable in both the study of disease (Rosenthal and Brown, 2007) and basic neural function (Yizhar et al., 2011).

In addition to this, there is a growing appreciation of the richness of mouse behaviour and cognition. As well as displaying sophisticated motor and sensory abilities (Carandini and Churchland, 2013; Mathis, Mathis, and Uchida, 2017; Busse et al., 2011; Brown, Martin, and Dudman, 2016; Krakauer et al., 2017), recent evidence suggests that mice use many cognitive functions crucial to human intelligence (Gomez-Marin and Mainen, 2016). These

include long-term memory (Silva et al., 2015; Huang et al., 2013), working memory (Li, 2017), motor learning (Mathis, Mathis, and Uchida, 2017), conceptual abstraction (McKenzie et al., 2014), probabilistic inference (Li and Dudman, 2013) and evidence integration (Koay et al., 2015). Together these findings demonstrate that mice are a viable model system for studying the basis of cognition.

However, exploring these cognitive functions and their underlying computations in mice can be challenging. In many cases, mice need to be trained to perform a behavioural task. This is often necessary to be able to systematically probe some computation or process of interest. The development of paradigms in which sophisticated cognitive functions can be studied in mice is challenging. It is often associated with extensive training or highly specific training regiments (Brunton, Botvinick, and Brody, 2013; Guo, Hires, et al., 2014; Koay et al., 2015). For example, while animals must constantly integrate information in order to survive (Gold and Shadlen, 2007), studying this in natural behaviour is too complex to be amenable to scientific investigation. Instead, mice must be trained to integrate evidence in a highly controlled experimental paradigm (Britten et al., n.d.) which requires months of training (Koay et al., 2015). Importantly, the development and application of appropriate training protocols are usually highly labour intensive and time consuming, curtailing scientific progress (Poddar, Kawai, and Ölveczky, 2013; Schaefer and Claridge-Chang, 2012).

These same problems apply to behaviours which require no or only little

training. In these cases, it is often the number, rather then the duration or difficulty of experiments that limit the speed of progress. Such large-scale behavioural experiments are clasically performed during drug screens (Werren et al., 2010) or the development of transgenic disease models (Kile and Hilton, 2005; Silver, 1995; Jhuang et al., 2010). More recently, however, similar large scale behavioural experiments have found their way into basic science. Combining high-throughput automated behavioural experiments with genetic targeting of effector proteins (Lima and Miesenböck, 2005; Boyden et al., 2005; Zemelman et al., 2003; Yizhar et al., 2011) in *Drosophila* has led to the description of whole brain neuro-behavioural maps (Vogelstein et al., 2014; Robie et al., 2017).

These examples highlight the importance of acquiring large amounts of behavioural data for continued progress in neuroscience, and emphasise that speed and labour-intensiveness remain limiting factors for the acquisition of such data in rodents.

Two major factors contribute to these limitations (Poddar, Kawai, and Ölveczky, 2013; Gomez-Marin, Paton, et al., 2014; Schaefer and Claridge-Chang, 2012). The first is the amount of manual human labour involved in the execution of individual experiments. The second is the extent to which such experiments can be performed in parallel. An important issue contributing to both these problems is the cost of experimental equipment, which can prohibit automation and parallelisation of behavioural experiments.

Here we present a novel infrastructure for performing both behavioural

and optogenetic experiments on a large scale. Our infrastructure is scalable, cheap, easy to implement and use, as well as highly adaptable. We consider the systems' modularity and the problems each module addresses. We have thus far trained over 200 mice using this infrastructure on over 10 different tasks. In this chapter we present the architecture of the system and validation behavioural data of mice performing basic operant tasks, as well as preliminary optogenetic experiments.

2.2 Methods

All experiments were approved by the local ethics review committee at the University of Oxford and licensed by the UK Home Office. To obtain the results reported in this chapter used eight male C57BL/6J (Harlan Laboratories) mice, aged 4-10 weeks for behavioural experiments. Young C57BL/J mice were selected because C57BL/6J mice have previously been reported to suffer from sensorineural hearing loss later in life (Ison, Allen, and O'Neill, 2007). We additionally used two b6;129p2-Pvalb$^{\mathrm{tm1(cre)Arbr}}$/J$^{+/+}$ (Jackson Laboratory; stock number 008069), aged eight weeks at experiment onset, for piloting optogenetic experiments (data not shown) and four Gad2$^{\mathrm{tm2(cre)Zjh}}$/J$^{+//+}$ (Jackson Laboratory; stock number 010802) mice, aged seven weeks at onset of experiment, for optogenetics experiments presented here. Data from one of the Gad2 mice were excluded as after seven days of training, performance did not exceed chance-levels.

2.2.1 Behavioural Experiments

2.2.1.1 Water Regulation

During behavioural experiments, mice had ad libitum access to food, but their water intake was regulated. Body weight and other health signs were monitored daily (Guo, Hires, et al., 2014) and body weights were kept above 85% of starting weight. If mice received insufficient water through behavioural experiments (<0.5ml), water was supplemented to a total volume of 0.7ml.

2.2.1.2 Stimuli

Stimuli were presented using equipment described in subsection 2.3.4. Stimuli were callibrated using an M500 microphone (Pettersson), which was itself referenced to a sound-level calibrator (Iso-Tech SLC-1356). We used two types of stimuli: clicks and pure-tones. Clicks were 5ms in duration and biphasic, presented at 80dB SPL. Click volumes were calibrated by integrating the recorded RMS of clicks over the mouse hearing range (1-100kHz) and comparing it to the RMS of stimuli from the reference sound-level calibrator. Pure tones were 200ms in duration, presented at a mean level of 70dB SPL and phased in and out using 25ms cosine ramps. Levels of pure tones were roved between 60dB SPL and 80dB SPL.

2.2.1.3 Optogenetic Behavioural Experiments

Our behavioural protocol for optogenetics consisted of mice performing a simple click-detection task. Mice were initially acclimatised to handling and head-fixation following (). Thereafter, mice were trained to lick in response to a 5ms biphasic click stimulus presented at 80dB SPL. There were two types of trials: stimulus trials (80dB SPL click; reward for licking) and catch trials (no stimulus; no reward for licking). These were randomly interleaved at an inter-trial interval drawn from a uniform distribution between 6s and 12s. If mice licked during a 1s window following onset of the stimulus, a water drop (3μl) was delivered immediately.

After multiple days of training, mice began to lick exclusively in response to clicks. Once the fraction of correct trials (lick for stimulus trial; no-lick for catch trial) exceeded 80% (<5 days in all cases) we began performing bilateral optogenetic stimulation (20Hz; 10ms pulses) of the auditory cortex on 50% of trials. Optogenetic stimulation lasted for approximately 700ms starting 50ms prior stimulus onset. Stimulation intensities were estimated using photodiode power sensors (Thorlabs S120C) in combination with a digital power meter (Thorlabs). Stimulation intensity was $\sim$ 0.2mWmm^{-2}. To prevent visual detection of optogenetic stimulation we used masking flashes, intended to flash blind mice, in accordance with existing protocols (Guo, Li, et al., 2014; Huber, Petreanu, et al., 2008). Specifically, we placed two ultra-bright LEDs (60mW light intensity) approximately 7cm in front of the the animal's snout, which were pulsed using identical parameters to the stimulation LEDs but

were activated during trials both with and without optogenetic stimulation. We included all behavioural sessions in which the animal's performance in masking condition was above chance.

2.2.1.4 Freely Moving Frequency-Discrimination

Freely-moving mice had ad-libitum access to water by engaging in the behavioural task. Tasks were performed in the mice's home-cages, which were fitted with a custom-built three-port behavioural chamber. The ports were designed for nose-poking and were thus elevated from the ground by 15mm, 20mm in diameter and equipped with an infra-red beam-break to detect poking. Water was delivered through blunted 19 gauge needles embedded in the back of the nose-pokes.

Prior to performing the full frequency discrimination task, the behaviour was shaped in two phases. First, mice were trained to nose-poke centrally and received reward (3 µl) at both outer ports for doing so. Once poking became reliable (i.e. mice centrally poked more than 200 times in a single day), they were trained to perform a simple pure done discrimination task. At this stage, central poking led to the presentation of one of two pure tone stimuli (4kHz or 16kHz). Upon poking centrally mice had to choose to poke one of the two outer ports, depending on the frequency of the sound that was presented. If a high frequency stimulus was presented mice were required to poke right. If a low frequency stimulus was presented, mice poked left. If mice responded by moving to the correct lateral port, a reward (3µl) was delivered.

Stimuli were roved between 60 dB SPL and 80 dB SPL to discourage mice from using level as a cue. Once they reached criterion (>80% correct), mice were moved to the final task version. Here, they were presented with twelve stimuli logarithmically spaced between 4 and 16kHz and had to perform the task described above.

2.2.1.5 Head-fixed Frequency-Discrimination

We trained mice to perform a left-spout/right-spout frequency discrimination task. In the head-fixed version of the task, mice also underwent initial behavioural shaping. Specifically, they initially underwent a simple operant conditioning protocol in which they were required to lick to a click stimulus, similar to subsubsection 2.2.1.3. The sole difference from the protocol outlined in subsubsection 2.2.1.3 is that mice could lick either of the two available spouts during the response window to receive bilateral reward. Thereafter, mice were trained to perform the same two-tone discrimination task described in subsubsection 2.2.1.4. In the testing phase, 6 stimuli, logarithmically spaced between 4 and 16kHz, were presented.

2.2.2 Surgeries for optogenetics experiments

Prior to surgery, mice were premedicated with intraperitoneal injections of eexamethasone (Dexadreson, 3 µmg), atropine (Atrocare, 1 µmg) and caprofen (Rimadyl, 0.15 µmg). General Anaesthesia was induced with an intraperitoneal injection of fentanyl (Sublimaze, 0.05 mg/kg), midazlom (Hypnovel

5mg/kg) and medetomidine (Dormitor, 0.5mg/kg. Thereafter, mice were placed in a stereotaxic frame (Model 900LS, David Kopf Instruments) and held in place using a mouth-piece and ear bars, under sterile conditions. Body temperature was kept constant at 37 °C through a closed loop temperature control system with a rectal thermometer (FHC). Depth of anaesthesia was continuously monitored using the paw-pinch reflex.

Following induction of anaesthesia both eyes were covered with ointment (Maxitrol, Alcon) to prevent corneal desiccation during anaesthesia. We then shaved the mouse's scalp and applied a local anaesthetic (Marcain) to the scalp and removed a 8mm flap of skin exposing the parietal and temporal bones, bilaterally. We removed the muscle overlying the right temporal bone with a scalpel and pushed it ventrally. We next demarcated a 3mm diameter region whose center was approximately 2.5mm posterior to bregma and 4.5mm lateral to the midline, for subsequent window implantation. Cyanoacrylate glue (Pattex Classic, Henkel, Germany) was applied to the surrounding tissue to seal wound margins and prevent further bleeding. We then used a drill equipped with a 0.4mm tip to remove a 3mm disc of bone and expose the underlying cortex. Saline was applied to wash blood from the dura and prevent drying of the cortex.

We then injected approximately 500 nl of AAV-EF1a-DIO-hChR2(H134R)-EYFP (Guo, Li, et al., 2014; Lin et al., 2009) into the exposed cortex. Injections were distributed across 8 injection sites (4 locations, two depths [600µm, 200m]). Injections were performed using a custom-built pressure

injection system.

Following injections, we placed a glass cover slip 3mm in diameter in contact with the surface of the cortex. The cover slip was attached to the edges of the skull with cyanoacrylate glue (Pattex Ultra Gel, Henkel, Germany).

The same procedure was then performed on the left hemisphere. Following drying of the glue, we attached a custom head-post to the skull for head-fixation using Super-Bond C&B dental acrylic (Sun Medical, Japan).

Following drying of the dental acrylic, anaesthesia was reversed via an intraperitoneal injection of naloxone (1.2 mg/kg), flumazenil (Anexate, 0.5mg/kg) and atipamezol (Adntisedan, 2.5mg/kg. Burprenorphine (Vetergesic, 1ml/kg) was injected postoperatively.

2.2.3 Electrophsyiological Recordings

Following behavioural optogenetics sessions we performed terminal extracellular electrophysiological experiments to validate our optogenetic experimental protocol. Electrophysiological experiments were performed as described previously (Vasquez-Lopez et al., 2017).

General anaesthesia was induced with ketamine (100mg/kg, Vetalar) and medetomidine (140µg). Thereafter, mice were maintained under anaesthesia using ketamine (50 mg/kg/h) and medetomidine (0.07 mg/kg/h), topped up at 30 minute intervals. Following induction of anaesthesia the glass coverslip was removed, revealing the auditory cortex. 32 channel recording probes (1 $\times$ 32) (Neuronexus) were inserted to record from the granular

layers of the auditory cortex. Data were acquired using an RZ2 BioAmp processor (Tucker-Davis Technologies) and collected and saved using Benware (www.github.com/beniamino38/benware). Stimuli were generated using an RX6 Multifunction Processor (Tucker-Davis Technologies), amplified by a TDT SA1 Stereo Amplifier (Tucker-Davis Technologies) and delivered through an ultrasonic loudspeaker (Vifa, Avisoft Bioacoustics) coupled to a tube positioned near the entrance of the mouse's left ear canal. Stimuli were **100ms** broadband noise-bursts presented at 60-90 dB SPL. Optogenetic stimuli were presented using the system described in subsection 2.3.5, synchronised to the RZ2. Stimuli consisted of 500ms of continuous illumination at $\sim$0.2mW/mm^2.

2.2.4 Analysis

All analysis was performed using custom scripts in the programming language Python. All reported p-values are two tailed.

2.3 System Architectre

2.3.1 Overview

Our system is modular and hierarchically structured, enabling flexible interchanging of the modules comprising the system (Figure 2.1).

Briefly, at the top-level of the infrastructure is a central, locally-hosted webserver. The web-server provides the user interface, stores the behavioural-data and harbours scripts for basic analysis.

The next level are Raspberry-Pis (RPis), which act as the computational controllers of the behaviour boxes. The RPis act as client computers in our architecture and are thus bi-directionally interfaced with the central web-server. RPis receive commands to run behavioural scripts, acquire data and send data to the central web-server for storage.

Finally, the lowest level, which we refer to under the umbrella-term apparatus, comprises a set of data-acquisition components (lick-detectors,video-acquisition etc.) and actuators (solenoids, loudspeakers).

Though the infrastructure we describe is end-to end, we appreciated the fact that many potential users are likely to have parts of a behavioural system established in their laboratories. The ability to integrate with existing hardware is an essential part of our system. The modularity of our infrastructure means that parts of our system may be used in combination with existing equipment, enabling flexible use of our infrastructure.

Finally, for parsimony and ease of customisation, all software is written

in the programming language Python. We will now proceed to discuss the different modular levels of the system.

2.3.2 Server side

The top-level of the system is a central web-server, through which users interact with the rest of the system. The user interface is in the form of a web-site hosted on the server. Code for our custom web-server is publically available at www.github.com/yves-weissenberger/behaviour_webserver. The core benefits of having a high-level server are two-fold. Firstly, by enabling central control and access to data from many behaviour setups, it greatly simplifies the parallelisation of experiments. Secondly, the central storage and unified data-structure facilitates aggregation of experimental data.

Our webserver is written using the popular Python web-development library Django (Django Software Foundation, 2017). We chose this platform for its high performance, flexibility and also because it is free and open-source. Together, these features enable simple customisation of the web-server and cost-effective implementation. For security reasons, the web-server is hosted on a local area network to which lower-level behavior modules are connected. Users can access the web-server from the host computer directly, or access it using personal devices when on the local area network. This greatly facilitates multi-user access and use.

A key factor enabling the flexible exchange of modules in our infrastructure is the provision of multiple communication protocols for server-client

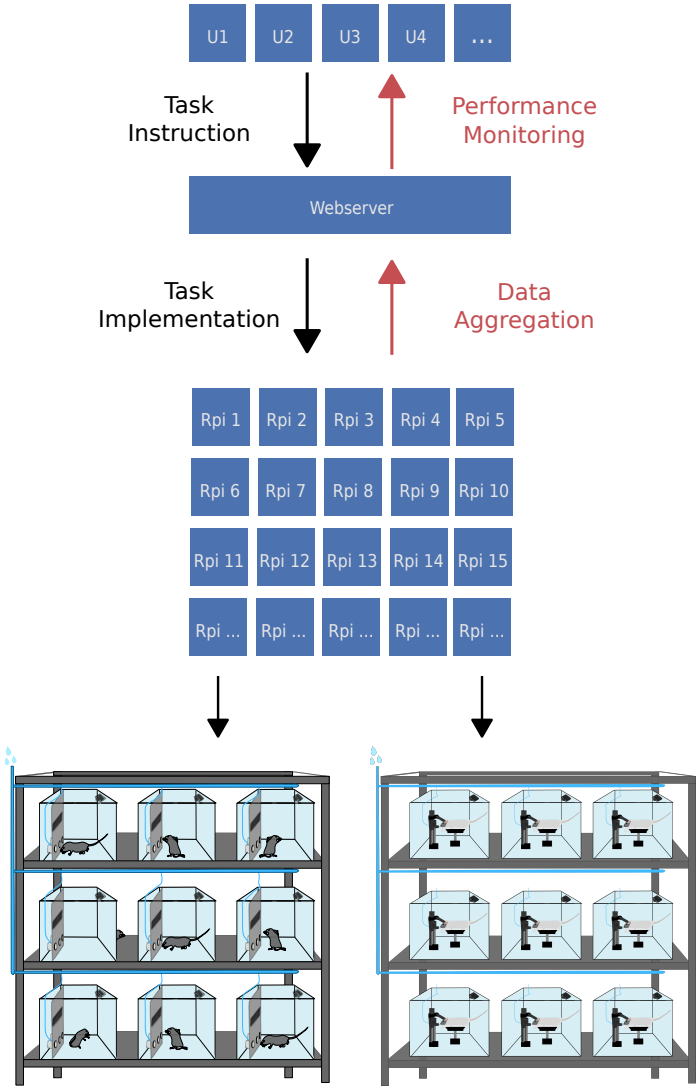


Figure 2.1: Overview of the system architecture. See main text for details. Abbreviations: U - user; Rpi - Raspberry Pi

communication. All communication with the central-web-server runs through the local area network and is mediated through two communication protocols: secure shell (ssh) and hyper-text transfer protocol (http). These protocols can be used to interface with virtually any electronic device with wifi capabilites (i.e. most computers and microcontrollers as well as many other devices). Importantly, most major programming languages (e.g. MATLAB, LabView, C++ etc.) provide implementations of these communication protocols, meaning that a host of devices other than RPis, can act as clients controlling individual behavioural apparatus.

In practice, we use transient ssh connections to initiate execution of behavioural protocols on client computers and http for sending data from the clients to the central web-server, with multiple built-in handshakes for reliability. We chose ssh for protocol initiation as it provides a simpler means of flexible, modifiable communication with clients. For client-to-server communication, we chose http as it allowed for fast sending of flexible data-types without the need for constantly active ssh connections. In our experiments, data are typically sent every 2-5s from client RPis running behavioural protocols. Training animals in practice we, similar to others (Guo, Hires, et al., 2014), found it vital to be able to analyse behaviour-data in real-time to adjust training protocols, particularly in head-fixed experiments.

Though we experimented with direct data-analysis running on the website we found that that this was not an effective solution. The major problem we encountered was that different experiments required different analyses

to be performed making static integration into the website difficult. As an alternative solution we host a separate Jupyter notebook server (Kluyver et al., 2016) on the server-computer, which can be accessed by multiple system users in parallel to analyse the data stored on the server.

The hardware requirements on the computer hosting the web-server are very low. In practice, we have found a budget laptop ($200) to be sufficient for typical use (data acquisition from >10 behavioural boxes by 5 users in parallel). Our unit tests indicate that the maximal capacity is approximately 200 client-computers with 20 parallel users, demonstrating that our system is highly scalable.

2.3.3 Client side

The client hardware, controlling the contingency structure of behavioural setups, are Raspberry Pis (RPis). RPis are low cost (<$40) computers (current instantiation: 4-core ARM-Cortex 1.2GHz CPU, 1GB DDR2 RAM, Broadcom Videocore IV GPU) running bare-bone linux distributions (we use Raspian). As controllers of behavioural apparatuses, RPis receive data from sensors (e.g. lick-detectors, manipulanda) and drive actuators (e.g. stimulus presentation, reward delivery) depending on contingencies in user specified protocols.

We chose RPis as they combined the computational power and broad functionality of standard computers with easy interfacing to low-level electronics. This enables implementation of a complete behavioural setup using

just an RPi and peripheral sensors and actuators. Thus, our RPi based training clients can be used in isolation by researchers who do not require a server based infrastructure for large-scale experiments. Separate advantages of the computational power are increased robustness of our system, as failure of any one unit will not impair function of the rest of the system, and the the distribution of computational requirements.

In our system, once a task is selected on the central web-server, RPis are instructed, though an ssh command, to run a Python script that implements behavioural tasks. An example of a script for a running two-alternative forced choice task, implemented on an RPi, may be found at www.github.com/yves-weissenberger/2AFC_Frequency_Discrimination. Python provides an extensive range of libraries for a variety of different requirements. Table 2.1 provides a list of libraries we routinely use for behavioural experiments and an overview of their functionality.

The RPi can be interfaced with peripheral apparatus via 40 general purpose input output pins (GPIOs), 4 USB-ports, 1 HDMI port and 1 DSI connector. The latter three interfaces allow for connections with a variety of different commercially produced hardware (e.g. webcams, digital-to-analog-converters etc). Complementing this connectivity, the GPIOs enable interfacing with low-level electronics equiptment, either through simple digital-read write operations or through a variety of communication protocols implemented on the RPis (e.g. I2C, UART). Previous training platforms using computers have typically required use of a second digital IO board (Poddar,

Table 2.1: Python libraries for implementing behavioural scripts

Python Library	**Function**
requests	Library implementing HTTP used for sending data to web-server
Numpy	Standard library for numerical computing. Useful for stimulus generation and randomisation
PsychoPy	Psychophysics library used for accurate and high quality visual stimulus presentation
RPi.GPIO	Library providing access to the RPis GPIO pins used to interface with low level electronics
PyGame	Standard Library for auditory stimulus presentation
billiard	Backport of Python 3 multiprocessing library (if using Python 2) used for distributing parallel processes across multiple cores on a single RPi
time	Library used for measuring timing of events (do not use for introducing delays as will prevent input processing!)
os	Library for interfacing with with operating system and running commands therein directly from Python. Used for auditory stimulus playback (via Linux aplay) in multi-speaker environments
csv	Library for writing data to csv file locally on RPi, useful for systems without server based data storage
PiCamera	Library for video acquisition using the RPi Camera module. Also useful for basic online image processing

Kawai, and Ölveczky, 2013; Francis and Kanold, 2017) complicating the system and increasing costs.

Compared to traditional computers the use of RPis has two major advantages. The first major advantages is that they are an order of magnitude

cheaper. This enables cost-effective large scale implementation of training apparatus. The second is that they enable direct interfacing with both low level electronics via digital input output and high level-electronics through standardised connectors.

Thus, RPis are an ideal choice of client computer for running behavioural protocols.

2.3.4 Apparatus

Table 2.2 provides an overview of the hardware used, in addition to the Raspberry Pis and Server Computer, for performing experiments. Further details can be found in appendix 1.

2.3.5 Optogenetics

Optogenetics actuators have revolutionised modern neuroscience (Lima and Miesenböck, 2005; Boyden et al., 2005; Zemelman et al., 2003; Yizhar et al., 2011). However, while large scale optogenetic experiments have been performed in other organisms, work in rodents is still performed at small scales. In part this is because experiments is labour intensive and parallelisation is curtailed by high costs. To alleviate these problems, we developed an ultra low-cost system for permorming optogenetic experiments. Similar to previous work, we characterised light transmission through adult mouse skulls. In our hands, light transmission was $\sim 50\%$, enabling performance

Table 2.2: Experimental hardware

Hardware	**Implementation**
Setup enclosure, head-fixed	Custom boxes constructed from 30mm MDF and Maxiboard boxes providing $\sim$ 30dB sound attenuation
Setup enclosure, freely-moving	Animal-facility cages and racks
Head-fixed behavioural setup	see (Guo, Hires, et al., 2014) for details
Head-bar and head-post for auditory experiments	Custom Design (see appendix 1)
Auditory Stimulus Presentation (standard)	StarTech USB digital to analog converter (96kHz; 16bit); SSME-SA-50 amplifier; Foundtek NEO-CD1 ribbon tweeter
Auditory Stimulus Presentation (high quality)	same as standard, replacing StarTech digital-to-analog converted with
Visual Stimulus Presentation (basic)	Custom Circuit (see optogenetics)
Visual Stimulus Presentation (high quality)	Raspberry Pi 7inch Capacitative touch screen
Somatosensory Stimulation	RS Pro Magnetic Lock with iron dust; 2000N
Solenoids for Reward Delivery	Burkert type 6650
Solenoid control circuit	Custom circuit (see appendix 1)
Reward spout	blunted 19 guage stainless steel needle
lick detector	Custom circuit (see appendix 1)
nose-poke detection	Custom circuit (see appendix 1)
Optogenetic Stimulation	Custom (see subsection 2.3.5 AND Y)
Video Recording	PiNoir camera

of transcranial optogenetics experiments, as has been previously reported (Guo, Li, et al., 2014; Burgess, Lak, Steinmetz, Zatka-Haas, Bai Reddy, et al., 2017).

Our light sources for head-fixed optical stimulation consisted of 2 ultra-bright blue leds (CREE-XP-E2 M3; $\lambda = 470$) mounted on heat-absorbent PCBs (CREE-XP Star PCB; LED-Tech; part # LT-1566). These LEDs are relatively large (= 3mm) and emit a maximal intensity of 350mW. For mounting LEDs on the circuit board, we recommend applying solder to the board, placing the LED on top and heating using a hot-plate. For head-mounted LEDs for freely moving experiments, we recommend Luxeon Z LEDs mounted on 5mm aluminium PCBs (Luxeon Star; part # SM-01-H3).

The LEDs were driven by a constant current LED driver (Buck V2 700mA, 42V; LED-tech; part # LT-2785) and controlled by an Arduino MEGA 2650 using a custom circuit design. The total price of an LED driver and electronics required for a complete apparatus is approximately $50, meaning that large numbers of these can be implemented in parallel and again, experiments may be performed at scale.

2.4 System Validation

2.4.1 Two-alternative forced choice task

As our validation test, we used a perceptual decision-making task. We chose this paradigm because of its wide use (see Carandini and Churchland, 2013

for a review) in mice, and because signal detection theory provides a standardised, theoretically grounded and quantitative framework for its description (Green and Swets, 1966; Wichmann and Hill, 2001). The specific task we chose to train mice to perform was a pure-tone frequency discrimination task. Frequency discrimination has been thoroughly investigated and is widely used to study hearing in mice (De Hoz and Nelken, 2014; Ehret, 1975; Clause, Nguyen, and Kandler, 2011; Heffner and Masterton, 1980; Jaramillo and Zador, 2014; Kuchibhotla et al., 2016b). Thus, there are extensive datasets to validate against.

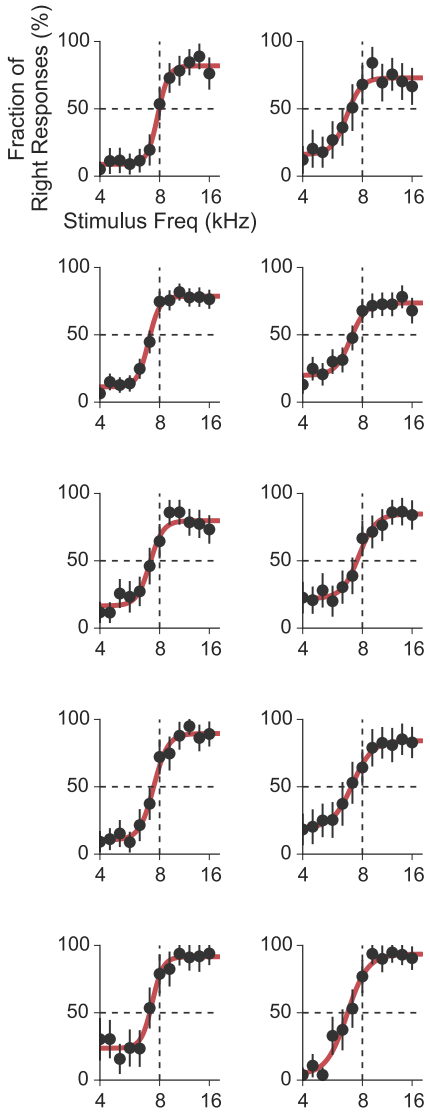


Figure 2.2: Example psychometric functions from single sessions of head-fixed mouse behaviour. Three or four behavioural sessions from each mouse are shown. Top two rows are sessions from one mouse. Next row and first panel of third row are from a second mouse. Remaining are from the third mouse. Error bars show 95% binomial confidence intervals.

We initially trained freely moving (n=5) and head-fixed (n=3) mice to discriminate 4kHz from 16kHz stimuli. We found that both groups of mice successfully acquired our behavioural task, reaching high levels of performance in terms of fraction correct (μ=91%; s.e.m 5%). However, we found that time to criterion (fraction correct >0.8) were substantially different between home-cage trained (μ=4.2 days; s.e.m. 0.6 days) and head-fixed mice (μ=16.5 days; s.e.m. 3.1 days).

Following acquisition of the simple discrimination task, we moved mice to the testing phase. Here, we used the method of constant stimuli to assess animal's frequency sensitivity. We presented mice with pure tones whose frequencies logarithmically tiled the space between 4kHz and 16kHz and required them to respond right for frequencies higher than 8kHz and left for frequen-

cies below 8 kHz. We used 6 tones for head-fixed mice and 12 for freely-moving based on the preliminary observation that freely moving mice performed substantially more trials per session.

To understand the discrimination behaviour, we used a standard psychometric function fitting approach (Wichmann and Hill, 2001) which has been widely applied to mouse behaviour (**Giminez2015**; Busse et al., 2011; Carandini and Churchland, 2013; Yang, Kwon, et al., 2015; Jaramillo and Zador, 2014; Katz et al., 2016). As expected, we found that psychometric functions accurately described the relationship between stimulus intensity and behaviour such that stimuli further from the 8kHz decision boundary were more easily discriminable. We further examined the structure of behaviour by assessing the three parameters of the psychometric function: bias, lapse-rates and stimulus sensitivity.

First, we quantified biases in mouse behaviour (Figure 2.4A). We defined bias as the absolute difference in the fraction of right choices from 50%, at the boundary stimulus (i.e. 8kHz). We found that there were large and varied biases in the behaviour of individual mice across sessions (range=0.1% - 30%). However, biases varied substantially from animal to animal and from session to session such that overall, the bias was was not significantly different from 0 (One Sample T-test; μ=0.02; p=0.69). Further the size of biases were not significantly different between freely moving and head-fixed mice (T-test;

p=0.3).

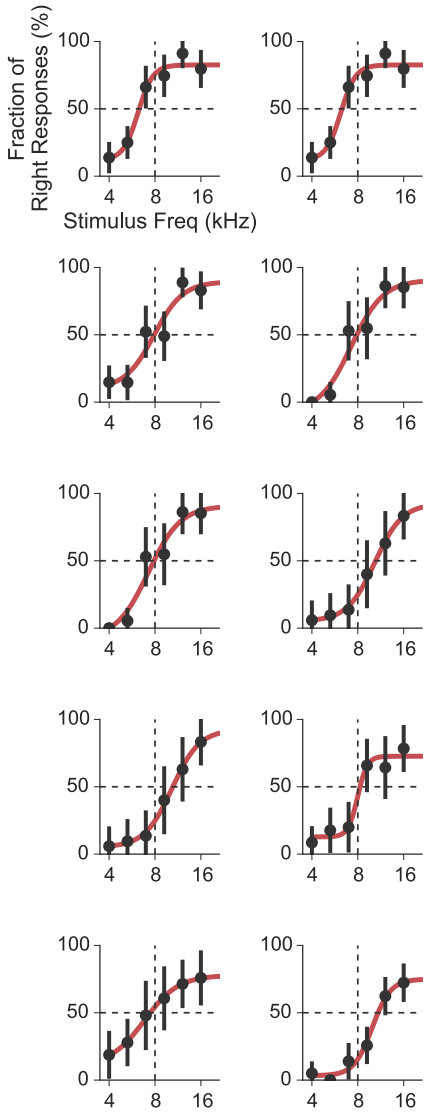


fig:2.3

Figure 2.3: Example psychometric functions from single sessions of home-cage mouse behaviour. Error bars show 95% binomial confidence intervals.

Similar to previous reports of mouse perceptual decision making (**Oconner**; Busse et al., 2011; Carandini and Churchland, 2013; Yang, Kwon, et al., 2015; Guo, Hires, et al., 2014), we found that sensory discrimination behaviour was characterised by high lapse-rates (Figure 2.4B). A subject's lapse rate is equal to the proportion of easily discriminable stimuli that animals miss. In both head-fixed and freely moving behaviour lapse rates were significantly larger than 0 (One-sample t-test p 0.0005). Although, lapse rates were higher in home-cage trained than head-fixed mice (p=0.023), the effect was small (4% difference in mean lapse rates). Overall, our observed lapse rates are in broad agreement with previous litera-

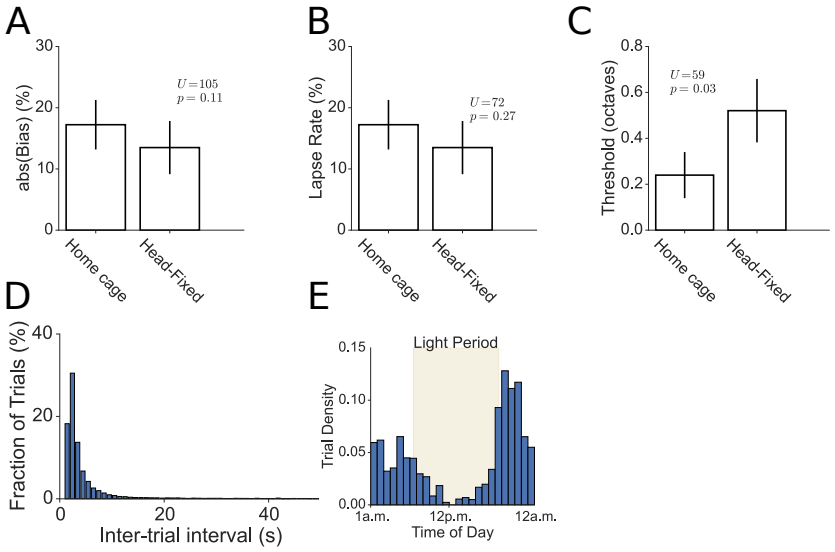


Figure 2.4: Summary of behaviour. Mean ± 2 s.e.m lapse rates A, biases B and thresholds (where 75% correct) C averaged across all sessions from all animals. **D** distribution of self-paced inter-trial intervals from the training of mice shows most inter-trial intervals are short. **E** Mice behaviour in the home-cage is largely restricted to the dark phase of their day-night cycle.

---

ture.

The final parameter of the psychometric function, and indeed the parameter of primary interest, is sensitivity to frequency. We quantified this sensitivity as the fraction of an octave required to change behaviour from 50% correct to 75% correct (approximately equivalent to a d-prime of 1) (Figure 2.4C). Here again, we observed a large range of sensitivities (range=0.1 - 1 octave). Further, we observed significant differences between head-fixed and freely moving mice (Mann-Whitney-U; p=0.029), whereby home-cage trained mice had lower average thresholds (μ = 0.23 octaves) than their

head-fixed counterparts (μ=0.52 octaves).

In addition to these psychometric parameters, we were interested in investigating the basic structure of behaviour in home-cage trained mice. Two striking features of behaviour appear. Firstly, mice perform behaviour such that the majority of trials are initiated at short inter-trial intervals (Figure 2.4D). This behavioural structure is similar to what we have observed when training mice in the same apparatus but restricted to 1-2h sessions during their light phase (data not shown). Secondly, we observed that when home cage training mice, they perform the majority (>80%) of trials during the dark phase (Figure 2.4E).

Thus, the data acquired from by both our home-cage training and head-fixed training is well described in terms of classical psychophysics, validating our methodological approaches.

2.4.2 Optogenetics

The final aspect of our system we sought to validate was the optogenetic apparatus. Previous reports have shown that optogenetic (Li, Liao, et al., 2017) and pharmacological inactivation (; Talwar, Musial, and Gerstein, 2001) of rodent primary auditory cortex impairs the detection of even basic sounds sounds. Thus, we used to a simpler behavioural task than described above in which animals simply had to lick in a 1s window after the onset of a 5ms biphasic click-stimulus. We then validated our apparatus for optogenetics experiments by showing that inhibiting auditory cortex during this operant

task impairs performance.

We transfected the auditory cortex of 4 GAD-IRES-CRE mice were cre-dependent channelrhodopsin (ChR2) and implanted windows bilaterally (Figure 2.5A). Since GABAergic interneurons form dense local arborisations (Tremblay, Lee, and Rudy, 2016) their activation leads to robust cortical inhibition (Guo, Li, et al., 2014). We found inhibition of auditory cortex led to a modest but highly consistent reduction in behavioural performance when stimulating GABAergic interneurons (Figure 2.5B). This effect was mediated by both a decrease in hit rates (μ=6%), as well as an increase in false alarm rates (μ=13%).

To ensure that this effect was not driven by mice's response to visual aspects of the optogenetic stimulation, we took two precautions. Firstly, we introduced a masking light which was switched on concurrently with the optogenetic stimulation. Secondly, in a subset of sessions, we occluded the windows bilaterally, preventing light from activating GABAergic interneurons. We found that this rescued the observed behavioural deficits (Figure 2.5C). After experiments had terminated, we also performed electrophysiological controls, confirming the efficacy of the optogenetic stimulation of GABAergic interneurons (Figure 2.5D). Overall, this work shows that our system mediates effective inhibition.

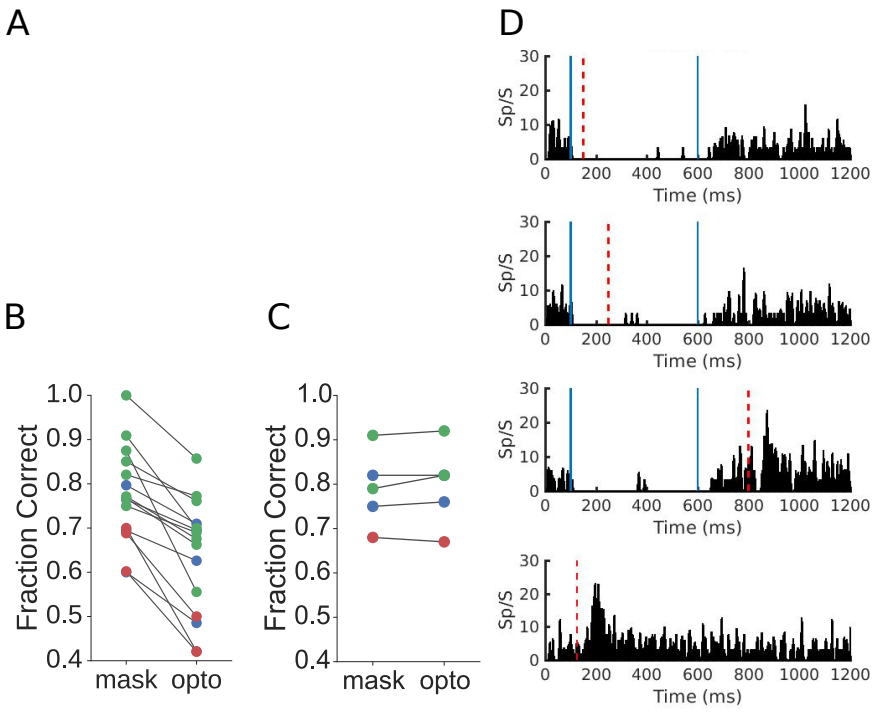


Figure 2.5: Optogenetic experiments showing dependence of sound detection on auditory cortex. **A** Histological validation of expression. **B** Reduction in fraction correct (stimulus vs catch-trials) as a result of optogenetic stimulation. Each pair of points denotes a single session. Each colour denotes behaviour of a different mouse. **C** Occluding auditory cortex from the LED rescues behavioural performance. **D** Electrophysiological validation of the effects of LED stimulation. Red dashed lines indicate onset of a 100ms 80dB SPL noise burst. Blue lines denote boundary between onset and offset of optogenetic stimulation.

2.5 Discussion

Here we have presented a high-throughput infrastructure for performing behavioural experiments and optogenetics in freely-moving and head-fixed mice. We have provided a description of the infrastructure and presented validation data from mice performing a two-alternative forced choice task in both home-cage freely moving and head-fixed conditions. To date, we have used this infrastructure for training over 200 mice one a variety of tasks. These include further work in the auditory modality (sound-level discrimination and sound localisation), work in other sensory modalities (visual and somatosensory), as well as non-sensory behaviours (matching behaviour and sleep-tracking). Together, these experiments illustrate the flexibility of our system for performing behavioural experiments.

2.5.1 Frequency Discrimination Behaviour

In comparison with previous work, the thresholds we have observed during frequency discrimination are larger (De Hoz and Nelken, 2014; Ehret, 1975; Clause, Nguyen, and Kandler, 2011; Heffner and Masterton, 1980; Kuchibhotla et al., 2016b). Our home-cage thresholds are approximately twice as large as previously reported, while our head-fixed thresholds are as much approximately five times higher than the thresholds previously reported De Hoz and Nelken, 2014. Multiple factors likely contribute to these differences. Firstly, our mice were performing a two-alternative forced choice, rather than

a go/nogo task (Ehret, 1975; De Hoz and Nelken, 2014) which has been shown previously to yield higher thresholds (Walker et al., 2011). Secondly, in the above studies, mice were specifically trained over long durations to respond to just noticeable differences. Finally, in contrast to other work (De Hoz and Nelken, 2014; Heffner and Masterton, 1980), we did not use aversive cues to punish incorrect behaviour. Thus, although our estimated thresholds are higher, differences in task design can readily account for the observed differences.

2.5.2 Comparison to other Behaviour Infrastructures

Recently, several infrastructures for large scale behavioural experiments have emerged (Schaefer and Claridge-Chang, 2012; Poddar, Kawai, and Ölveczky, 2013; Aoki et al., 2017; Gess et al., 2011; Francis and Kanold, 2017). Compared to previous behavioural infrastructures, ours has several notable advantages.

Firstly, while several systems offer automated high-throughput behavioural experiments, many are highly specialised, precluding widespread adoption (Mechan et al., 2009; Schaefer and Claridge-Chang, 2012; Gess et al., 2011; Francis and Kanold, 2017). In contrast, our system is highly flexible and extensible. This is mediated by a combination of structural modularity, open access software and flexible communication systems. Together, these features lead to flexible customisation and simple integration with existing hardware.

A second major advantage of our system is the ease of use. This is

largely due to the use of Python as a dynamically typed general-purpose programming language. As such, it is similar to MATLAB (Mathworks), probably the most commonly used programming language in neuroscience.

Finally, one of the most significant advantages of our system lies in its low pricepoint. Commercial systems are often very expensive, often costing thousands of dollars (Schaefer and Claridge-Chang, 2012; Gess et al., 2011; Mechan et al., 2009) for a single behavioural apparatus. In comparison, implementation of our system is highly affordable. A behaviour infrastructure consisting of 10 freely-moving and head-fixed cages, together with a central webserver costs approximately $6000 and a single behavioural apparatus can be assembled for less than $100. At this price point, it also remains almost an order or magnitude cheaper than comparable automated systems (Francis and Kanold, 2017; Poddar, Kawai, and Ölveczky, 2013). This low price also holds for use of optogenetic systems. While commercial systems typically cost thousands of dollars, implementation of our system costs less than $100.

2.5.3 Optogenetics

We have validated the use of our low cost optogenetics system in two ways. Firstly, through a set of electrophysiological recordings. Secondly, by using this system to demonstrate that auditory cortex plays an important role in sound detection, in line with previous work (Talwar, Musial, and Gerstein, 2001; Ji et al., n.d.) . The magnitude of our behavioural effects are similar or slightly smaller than those which have been reported previ-

ously (**Yang2015a**; Poort et al., 2015; Guo, Hires, et al., 2014; Ji et al., n.d.). Two factors likely contribute to this. Firstly, the surface irradiance used in our experiments ($\sim$ 0.2mW) is low compared to other use-cases (**Yang2015a**; Guo, Li, et al., 2014; Burgess, Lak, Steinmetz, Zatka-Haas, Harris, et al., 2017; Mathis, Mathis, and Uchida, 2017). We chose such a low intensity *a priori* as behavioural effects have been reported at these intensities (Poort et al., 2015) and to minimise tissue heating and/or damage (**Susjetke2014**). Nevertheless, we ere able to obtain robust electrophysiological and behavioural effects using our optogenetic system.

2.5.4 Conclusion

Overall, we have developed and deployed a novel infrastructure for large scale behavioural experiments and optogenetics. Our system is effective, flexible and cheap. Systems for automated and high throughput behavioural experiments, as described here, have the potential to substantially accelerate research in many areas of neuroscience.

Chapter 3

Encoding and Decoding head-fixed Behaviour

3.1 Introduction

Understanding behaviour and its computational and neural underpinnings is a core goal of modern neuroscience (Krakauer et al., 2017; Gomez-Marin and Mainen, 2016). A widely used set of protocols whose use has contributed substantially to our understanding of the neural and computational underpinnings of behaviour are operant conditioning tasks (Blackman, 1974; Skinner, 1948a; 1938). Operant conditioning tasks have been used successfully in primates (Oikonomidis et al., 2017; Rushworth et al., 2011), rodents (Guo, Hires, et al., 2014; Carandini and Churchland, 2013) and other organisms to tease apart the computational and circuit mechanisms involved in learning

and decision making.

Normative and descriptive statistical models have been essential to developing an understanding of behaviour in these tasks (Gomez-Marin and Mainen, 2016; Tervo, Tenenbaum, and Gershman, 2016; Gomez-Marin, Paton, et al., 2014; Krakauer et al., 2017). In the sensory domain (though also in other fields (Nevin, 1969)), learning and decision-making are typically analysed in terms of signal detection theory (Green and Swets, 1966; Stangor, 2011; Rippere, 2015; Shadlen and Kiani, 2013; Carandini and Churchland, 2013; Parker and Newsome, 1998). Signal detection theory provides a normative account of stimulus-driven behaviour whose remit has proven to be extraordinarily wide-ranging, offering powerful tools for interpreting both psychophysical and behavioural data (Green and Swets, 1966; Carandini and Churchland, 2013; Shadlen and Kiani, 2013).

When subjects perform perceptual tasks they often make suboptimal choices with respect to the prescribed task structure (Wearden, 1983; Wichmann and Hill, 2001; Fründ, Wichmann, and Macke, 2014b; Busse et al., 2011; Carandini and Churchland, 2013). This is particularly true in experimental animals, which are often reported to have high lapse rates when performing even simple sensory tasks (Busse et al., 2011; Guo, Hires, et al., 2014; O'Connor et al., 2010; Carandini and Churchland, 2013; Iigaya et al., 2017; Wearden, 1983; Yang, Kwon, et al., 2015). Signal detection theory explains this observation by asserting that subjects are ideal observers, corrupted by noise in their perceptual apparatus (Green and Swets, 1966;

Parker and Newsome, 1998; Shadlen and Kiani, 2013; Shadlen, Britten, et al., 1996). Through this assumption, signal detection theory is able to provide a parsimonious and predictive account of perceptual decisions on a trial-by-trial basis (Shadlen and Kiani, 2013).

However, there is a growing appreciation that what is typically considered noise in the signal-detection paradigm, is often highly structured behaviour (Fründ, Wichmann, and Macke, 2014b; Busse et al., 2011; Bak et al., 2016; Gold and Shadlen, 2007; Carandini and Churchland, 2013; Fründ, Wichmann, and Macke, 2014a; Smith et al., n.d.; Tervo, Proskurin, et al., 2014; Iigaya et al., 2017). Further, a variety of work has shown that ostensibly noisy behaviour may be adaptive when we have to account for factors such as incomplete knowledge of the environment (Tervo, Tenenbaum, and Gershman, 2016; Iigaya et al., 2017; Behrens et al., 2007; Kepecs et al., 2008), multiple task solution methods (Akam, Costa, and Dayan, 2015; Desrochers et al., 2010), the cost of actions (Walton, Bannerman, et al., 2003; Walton, Kennerley, et al., 2006; Klein-Flugge et al., 2016; Rudebeck et al., 2006) and other factors (Dayan, n.d.; Bathellier, Poh, et al., 2013; Kawai et al., 2015a). Thus, while signal-detection theory can be used to characterise sensory performance, it does not address a large part of the computational underpinnings of behaviour.

An alternative normative framework with which behaviour is often characterised is reinforcement learning (Sutton and Barto, 1998; Sutton et al., 1999a). Reinforcement learning provides an extensive set of tools for as-

sessing the value of actions. As such these methods can be used to derive testable predictions of subjects behaviour under different assumptions about the causes and structure of their behaviour.

The application of reinforcement learning as a tool for understanding behaviour has been broad. It has been applied to compare the explanatory power of different normative accounts of behaviour (Ito and Doya, 2009; Niv et al., 2015), and models of their environment (Daw et al., 2011b; Doll, Simon, and Daw, 2012; Gershman, Norman, and Niv, 2015; Tervo, Tenenbaum, and Gershman, 2016), and to explain stochasticity in behaviour (Charlesworth et al., 2011; Tervo, Proskurin, et al., 2014). Thus, reinforcement learning provides a way of asking whether certain underlying computations can account for subject's behaviour (Doll, Simon, and Daw, 2012; Gershman, Norman, and Niv, 2015; Ito and Doya, 2009; Daw et al., 2011b; Dayan, Kakade, and Montague, 2000; Daw, 2011; Niv et al., 2015).

At the core of reinforcement learning models used in neuroscience, much like the signal detection theory models, is the concept of action selection (Sutton and Barto, 1998). Accordingly, much of the application of reinforcement learning to neuroscience has centred on learning (; Ito and Doya, 2009; Samejima, 2005; Behrens et al., 2007; Lau and Glimcher, 2005) or inferring (Akam, Costa, and Dayan, 2015; Doll, Simon, and Daw, 2012; Daw et al., 2011b) the contextual values of performing a specific, discrete action. As such, most these models are largely restricted to the analysis of behavioural paradigms involving discrete decision points (Sejnowski, 2010), where the

subjects choose one of several *a priori* defined actions.

However, even in tasks with discrete decision points animals typically interact continuously with their environments. To make these data amenable to action-centred reinforcement learning approaches, behaviour outside the decision point is typically ignored. However, behaviour beyond repeated discrete decisions can form an integral parts of subject's strategies and their task solution methods (Kawai et al., 2015b; Desrochers et al., 2010; Skinner, 1948b). Additionally, in many cases, reward-guided and spontaneous behaviour may consistent of long sequences of actions (Wiltschko et al., 2015; Seeds et al., 2014; Desrochers et al., 2010). In these cases, existing action-centred analysis approaches are at best onerous and in the worst case intractable (Sejnowski, 2010).

An alternative framework, also derived from reinforcement learning, conceptualises behaviour in terms of enacted strategies (Sutton and Barto, 1998; Williams, 1992). Formally known as policy-based reinforcement learning (; Sutton et al., 1999a), this framework models agent's continuous behavioural output as arising from a parametric function, mapping sequences of raw inputs and previous behaviour to future behaviour.

The key opportunity provided by this of this approach stems from the fact that it can be used to model any behaviour that is describable by a parametric model (Williams, 1992; Sutton et al., 1999a). This enables modelling of tasks without discrete decision points. Additionally, it, at least in principle, affords the ability to apply reinforcement learning methods to the complete

observed behaviour, which may be vital to understanding subjects (reward-guided) behaviour (Kawai et al., 2015a; Gomez-Marin, Paton, et al., 2014; Krakauer et al., 2017). This knowledge is crucial for understanding how the brain implements behaviour (see for a review).

Here, we explore policy based models of behaviour in the context of our auditory detection task. We port methods initially developed to model neural spike trains (Paninski, 2004; Pillow, Ahmadian, and Paninski, 2010) and build models of head-fixed behaviour in mice. Rather than focus on discrete action selection, we consider behaviour in terms of policies (Sutton and Barto, 1998; Williams, 1992) - strategies that may be enacted over extended time-periods. We build regression models which map the animals' previous behaviour, stimuli and rewards to the detailed future licking patterns on a moment-by-moment basis. Using our methods, we show that we are able to capture a diverse array of behavioural strategies which depend either on auto-correlation in mice behaviour or conditioned stimuli. We demonstrate that by systematically altering subsets of parameters in our models we can examine the value of different aspects of the animals' strategies. In doing so, we find that in many cases, non-stimulus driven strategies affect reward rates and that mice make use of this to maximise their reward rate. Finally, in light of heterogenous goal-directed strategies, we show that the precise patterns of licks made by an animal contain enough information to reliably decode the causes of the animals behaviour, with high temporal resolution.

3.2 Methods

3.2.1 Behavioural Paradigm

For a detailed description of the behavioural-apparatus, surgical procedures and animal welfare see chapter 2. Head-fixed mice were trained to perform an auditory operant conditioning task.

8 wild-type mice C57BL6/J (6 male; 2 female), aged 6 weeks at onset of the experiments, were initially implanted with head-posts for head-fixation. After 1 week of recovery, behavioural training was started. Mice were initially acclimatised to handling and head-fixation as described by (Guo, Hires, et al., 2014). Thereafter, mice were trained to lick in response to a 5ms biphasic click stimulus presented at 80dB SPL. There were two types of trials: stimulus trials (click; reward for licking) and catch trials (no stimulus; no reward for licking). These were randomly interleaved at an inter-trial interval drawn from a uniform distribution between 6s and 12s. If mice licked during a 1s window following onset of the stimulus, a water drop (3μl) was delivered. Over multiple days, mice began to lick exclusively in response to stimuli. Once the fraction of correct trials (lick for stimulus trial; no-lick for catch trial) exceeded 80% (<5days in all cases), mice were moved to the testing phase, in which stimulus levels were reduced to 40dB SPL, which we found to be a challenging level for mice to detect, for 5 days. On each of these 5 days, mice performed approximately 200 trials per day.

We also trained 2 additional C57BL6/J mice (n=2) on a variation of the

task above. All task parameters were identical to those given above, except that two were spouts present. For licking either spout in the response period, mice received bilateral reward.

3.2.2 Basic Model

We assume that licks are generated probabilistically according to the behavioural strategy (henceforth referred to as the policy) of the animal. We define the policy as an input-output function, dependent on a set of parameters Θ mapping the history of a given mouse's experience to the probability of it licking at each given timepoint. We assume that the animal's policy is static, that is to say the parameters Θ are constant for the entire session. We discretise the total duration of the behavioural experiments into T timebins (in practice we set the length of a time-bin to 20ms) and denote the set of all time bins $\mathcal{T} = \{0 : T\}$. Let t be a particular timebin in $\mathcal{T}$, and let $y_t \in [0, 1]$ denote whether a lick occurred in a particular timebin $t \in \mathcal{T}$. Further, let $\mathcal{I} = \{i_k; k = 1 : N\}$ denote the set of all event types that the mouse's behaviour is sensitive to and $x^i_\cdot$ the time series of a single event type. Then, a mouse's experience $\mathbf{E} = \{x^i_t;\ \forall i \in \mathcal{I},\ \forall t \in \mathcal{T}\}$. Finally, let $\mathbf{y} = [y_0, y_1, \cdots, y_T]$ denote the array of all licks observed in the session. Then:

$$p(\mathbf{y}) = f(\mathbf{E}, \Theta) \tag{3.1}$$

We assume that each event in the mouse's environment has a time-limited influence on its licking behaviour going into the future and that the influence of each event is independent of others. At each timepoint after a particular event, the influence of the event may be different. Similarly, the influence of each event may be different from that of all others. For notational convenience, we thus define θ^i to be all parameters that relate a particular event i to the licking behaviour of the mouse and term θ^i, in analogy to the stimulus filters of neurons, the policy filter for the particular event type i.

This allows us, for each $x^i \in \mathcal{I}$, to define a matrix K_i implementing the projection of the event onto the mouse's corresponding policy filter. Since we assume independence between events, K_i is a Toeplitz matrix. Thus, $K_i x^i$ is an M×T matrix, where M is the number of bins into the future a particular event has influence. Finally let $\mathbf{X}$ denote the NM × T stacked array $[K_0 x^0, K_1 x^1, \ldots, K_x^n]^T$.

To complete our model of mouse's licking patterns, we need to specify the relationship between the parameters of the model and its inputs $\mathbf{Kx}$ and the licking behaviour. The natural form for this kind of model is that of a Bernoulli Generalized Linear Model (GLM). In the Bernoulli GLM, the probability of the mouse licking at a given timepoint is given by the dot product between the projection of the events onto policy filters and the policy filters themselves passed through a sigmoid:

$$p(y_t) = \frac{1}{1 + \exp\left(-\theta \mathbf{X}_t\right)} \tag{3.2}$$

3.2.3 Fitting procedure -Estimating the animal's policy

Estimating the policy of the animal, is then synonymous with knowing the parameters of this model. Here, we describe a maximum likelihood based approach to fitting the parameters of our model.

Fitting parameters to a GLM is a standard and solved statistical inference problem (McCullagh and Nelder, 1989; Bishop, 2006). It is known that although no analytical solution exists for the class of GLMs, we can make use of the iterative Newton-Raphson algorithm to efficiently fit parameters to the data (McCullagh and Nelder, 1989). At each iteration of the fitting procedure, the parameter vector θ is updated as follows:

$$\theta^{i+1} = \theta^i - \mathbf{H}^{-1}\frac{\partial \mathcal{L}}{\partial \theta^i} \tag{3.3}$$

where H denotes the matrix of second partial derivatives (the Hessian) of the log-likelihood function. Since this optimisation problem is convex with respect to θ, we are guaranteed to converge to a global optimum. However, the use of this method requires the gradient and Hessian. Which we now derive.

Following **??** the log-likelihood of the data is given by

$$\mathcal{L}(\mathbf{y}|\theta) = p(\mathbf{y}|\mathbf{X}, \theta) = \sum_{t=0}^{T} y_t \log \mu_t + (1 - y_t) \log (1 - \mu_t) \tag{3.4}$$

where

$$\mu_t = \frac{1}{1 + exp(-\theta \mathbf{X}_t)} \tag{3.5}$$

In order to perform gradient-descent on $\mathcal{L}$ we require the gradient. We can differentiate using the chain rule, **??** with respect to θ and μ separately.

$$\frac{\partial \mathcal{L}}{\partial \theta} = \frac{\partial \mathcal{L}}{\partial \mu} \frac{\partial \mu}{\partial \theta} \tag{3.6}$$

the derivative of μ with respect to θ, by the chain rule, is:

$$\frac{\partial \mu}{\partial \theta} = \mu(1 - \mu) \tag{3.7}$$

After simplification, the derivative of $\mathcal{L}$ with respect to μ is given by

$$\frac{\partial \mathcal{L}}{\partial \mu} = \sum_{t=0}^{T} \frac{y_t}{\mu_t} - \frac{1 - y_t}{1 - \mu} \tag{3.8}$$

Combining these expressions, the gradient is given by

$$\mathbf{J} = \frac{\partial \mathcal{L}}{\partial \theta} = \mathbf{X}^T(\mathbf{y} - \mu) \tag{3.9}$$

The Hessian, derived analogously from the $\mathbf{J}$ is given by

$$\mathbf{H} = -\mathbf{X}(diag((\mu)(1 - \mu)))\mathbf{X} \tag{3.10}$$

3.2.4 Empirical Bayes-ish estimation

The maximum likelihood (ML) estimator described in the previous section has several drawbacks. Notably, it requires large amounts of data to converge (Park and Pillow, 2011) and if event are not independent and identically distributed (i.i.d.) the estimates can be biased (Paninski, 2004). To counteract these problems, parameter-estimation (i.e. inference) can be biased biased towards the desired part of the parameter space. This is typically achieved by regularising parameter estimates in the fitting process (Bishop, 2006). In the following we will detail our approach towards this in the context of regularising parameter estimates, following (Park and Pillow, 2011) in our exposition.

In the Bayesian framework, regularisation is formalised in terms of priors placed on the parameters. Essentially, the purpose is to bias our estimate of the parameters towards parts of the parameter space that are more likely *a priori*. Formally, we assume that the parameters of a particular fit are drawn from a probability distribution $p(\theta)$. Estimating the model parameters is then equivalent to inferring the posterior distribution over parameters, given the data and the prior:

$$p(\theta|\mathbf{X}, \mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{X}, \theta)p(\theta)}{p(\mathbf{y}|\mathbf{X})} \tag{3.11}$$

This approach naturally gives rise to the question of how to choose an appropriate prior. Here we use two approaches. The first, which we do not

discuss further is simple L2 regularisation by cross-validation on the training set (Bishop, 2006). An alternative approach is Emprical Bayes.

Empirical Bayes, also known as evidence optimisation, essentially involves computing the probability given the set of hyperparameters by marginalising over model parameters. Let ϕ denote the set of hyperparameters that are determining the distribution of $p(\theta)$.

Finding the optimal hyperparameters of the model is equivalent to maximising $p(\mathbf{y}|\mathbf{X}, \phi)$ and explicit expression for which is given below.

$$p(\mathbf{y}|\mathbf{X}, \phi) = \int_{\Omega} p(\mathbf{y}|\mathbf{X}, \theta)\ p(\theta|\phi)\ d\theta \tag{3.12}$$

Following earlier work (Park and Pillow, 2011; Sahani and Linden, 2003a), we will assume that the parameters are drawn from a Gaussian distribution with mean 0 and covariance function $\mathcal{K}$ parameterised by the parameters ϕ:

$$p(\theta|\phi) = \mathcal{N}(0, K_\phi) \tag{3.13}$$

where the covariance between any 2 points x_1 and x_2 is given by

$$K_\phi(x_1, x_2) = k_1 \exp\left(-\frac{1}{2}\ \tau(x_1 - x_2)^2\right) \tag{3.14}$$

and $\phi = (k_1, \tau)$. Since the Gaussian prior we have chosen for our likelihood function is non-conjugate (Bishop, 2006), no closed form solution for this problem exists. Thus, instead of optimising **??** we will simply optimise

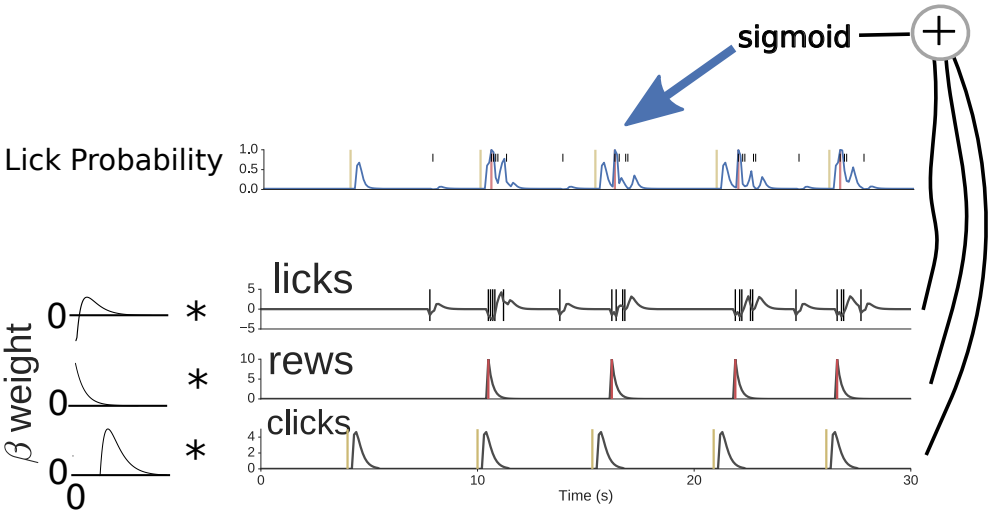


Figure 3.1: Figure illustrating the structure of the model. *(Top)* Output of the model is lick probability (blue line) which accurately tracks licks (black line). Yellow vertical lines indicates clicks, red vertical lines indicate rewards (rews). *(Bottom left)* Parameters of the model which are fit to the data are β weights of the regression model. *(Bottom Middle)* Structure of model predicting licks. Events (vertical lines) are convolved with inferred parameters of the model (grey horizonal lines). The resulting time series are added together, and with the inferred offset parameter, and passed through a sigmoid whose output is the probability of licking at each timepoint.

?? using our posterior estimated from the maximum likelihood case as a fixed approximation to $p(\mathbf{y}|\mathbf{X}, \theta)$. Thus we simply optimise the parameters of K_ϕ by maximising the marginal likelihood of each of the policy filters, using gradients given by (Rasmussen and Williams, 2006).

3.2.5 Decoding the causes of behaviour

After fitting the parameters describing the relationship between external events and licking, we may use Bayes Theorem to invert the model and attempt to decode the events in the environment that led a particular pattern of licks. We interpret decoding from the model as estimating the likely cause of the observed licking. Mathematically this means estimating $p(\mathbf{X}|\mathbf{y}, \theta)$. Our primary interest here is in decoding the posterior probability of a stimulus having occurred $p(x^{stim}|\mathbf{y}, \theta, \mathbf{X}^{\neq stim})$ where $\mathbf{X}^{\neq stim} = \{x^i \ \forall i \in \mathcal{I} \ s.t. \ i \neq stim\}$).

To compute this quantity efficiently, we make use of the fact that our likelihood function **??** is a convex function both with respect to the parameters θ and the inputs $\mathbf{X}$. This, in principle, allows us to invert the model and optimise the likelihood function with respect to the inputs. The probability of a certain value of an input at each timepoint is given by

$$p(\hat{x}^{stim}|\mathbf{X}^{\neq stim}, \mathbf{y}, \theta) = \frac{1}{Z} \ p(\mathbf{y}|\mathbf{X}^{\neq stim}, \hat{x}^{stim}, \theta)p(\hat{x}^{stim}) \tag{3.15}$$

where Z is the normalising constant that does not depend on $\hat{x}^{stim}$. In practice, the quantity we are maximising, the negative log-likelihood, is given by:

$$\log p(\hat{x}^{stim}|\mathbf{X}^{\neq stim}, \mathbf{y}, \theta) = \log p(\mathbf{y}|\mathbf{X}^{\neq stim}, \hat{x}^{stim}, \theta) + \log p(\hat{x}^{stim}) + k \tag{3.16}$$

Implementing this for the specific case of the Bernoulli log-likelihood, we arrive at:

$$\mathcal{L} = -\sum_{t=1}^{T} \theta\hat{\mathbf{X}}_t + \log\left(1 + \exp\left(\theta\hat{\mathbf{X}}_t\right)\right) - \log p(\hat{x}^{stim}) \tag{3.17}$$

where $\mathbf{X}$ is the stacked vector of projections onto the policy filters. Differentiating this expression gives us the gradient:

$$\nabla_{\hat{x}^{stim}} = \hat{\mathbf{K}}^{stim}(\mathbf{y} - (1 + \exp\left(\theta\hat{\mathbf{X}}\right)^{-1})) + \nabla_{\hat{x}^{stim}} p(x) \tag{3.18}$$

We can differentiate this again to obtain the Hessian

$$\mathcal{H} = ((1 + \exp\left(\theta\hat{\mathbf{X}}\right)^{-1})\hat{\mathbf{K}}^{stim})((1 + \exp\left(\theta\hat{\mathbf{X}}\right)^{-1})\hat{\mathbf{K}}^{stim})^T + \nabla_{\hat{x}^{stim}}\nabla_{\hat{x}^{stim}} p(x)$$

(3.19)

although possible, we have not exploited the tri-diagonal structure of the Hessian (Paninski et al., 2010), to more rapidly perform the matrix-inversion of the Newton-Raphson Step.

With an appropriate choice of prior, we can, similar to the above, perform inference on the stimulus input space. In order for the optimisation problem to remain convex, we require a convex prior to be placed on the stimulus

space. However, the stimulus space is bounded (i.e. the stimuli we present cannot take on negative values). Furthermore, for a large subset of our data, the distribution of stimuli is binary (a stimulus is either presented or not). This distribution is not convex.

In order to still exploit the fast optimization methods discussed above, we follow (Pillow, Ahmadian, and Paninski, 2010) and place a uniform prior, bound to the interval [0,1], as a surrogate distribution for decoding. To perform inference using this method, we use an interior point method (Boyd and Vandenberghe, 2004) (also known as the log-barrier method). Instead of hard constraints on the prior, we add a logarithmic barrier, which tends to $-inf$ as the bounds are approached. Thus, our prior and the gradients with respect to it become:

$$\log p(x) = -\sum_{t=1}^{T}(\log(x_t^{stim}) + \log(1 - x_t^{stim})) \tag{3.20}$$

$$\nabla \log p(x) = \frac{1}{x} - \frac{1}{1-x} \tag{3.21}$$

$$\nabla\nabla \log p(x) = diag\left(\frac{-1}{x^2} - \frac{1}{(1-x)^2} \right) \tag{3.22}$$

We additionally included L1 and L2 priors on x with standard derivatives (Bishop, 2006).

Now, decoding can be performed using the efficient methods discussed above.

3.2.6 Statistical analysis

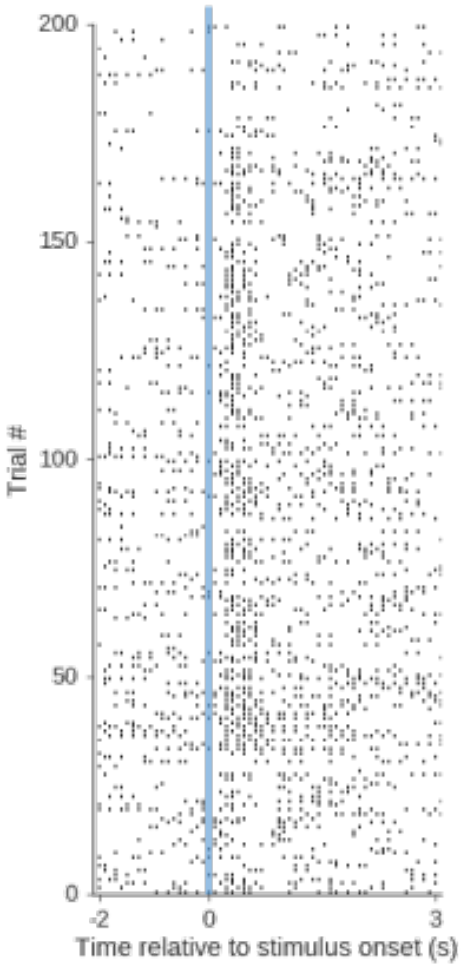


Figure 3.2: Example behaviour on a single session shown as a lick raster. Each row represents a trial, with black horizontal lines indicating licks. Blue line indicates stimulus onset.

All regression models were custom implemented in Python 2.7 using the numpy extension package (**VanDerwalt2011**). All analyses were performed using custom written scripts in Python 2.7 with statistical tests implemented in the Statsmodels (Seabold and Perktold, 2010) or Scipy (Oliphant, 2007) packages. P-values for correlation coefficients were determined using permutation tests with 10^6 permutations performed. All statistical tests were two-tailed. Errorbars in histograms, unless otherwise indicated, show deviations extending to two standard errors of the mean.

For the decoding analysis, the behavioural data were separated into trials and 20% of trials were held out for decoding. The model was then fitted on the remaining 80% of data and L1 and L2 regularisation coefficients were found

by maximising decoding performance. Optimal coefficients were found by maximising the fraction correct for decoding using L1 and L2 coefficients selected from a pre-defined grid, for each session individually. Results in figures sho.

To approximate stimulus-unrelated policies we used a two-step approach. First, we fit a single exponential decay function to the decay phase of lick-filters using a least-squares approach. We then used this decay, together with the initial starting value to constrain a Gaussian Process regression model (Rasmussen and Williams, 2006) that reconstructed the complete lick-filters. Gaussian Processes were fit using the Python Library GPy (ML, 2012). Using this approach we were able to reduce the dimensionality of the lick-filter to two.

To estimate the value of policies, we simulated 2^6 trials of behaviour and calculated the desired summary statistics based thereon (see subsection 3.3.2 for details).

For comparison of the L2-regularised GLM and the GP-GLM, we fit 80% of data and report performance on the held out 20%.

3.3 Results

We observed and modelled the behaviour of head-fixed mice as they performed a challenging auditory stimulus detection task. Our goal was to describe, understand and provide a normative account of detailed licking behaviour in head-fixed mice (for behaviour in an example session see Figure 3.2).

Figure 3.1 illustrates the structure of the model used to predict the licking activity of head-fixed animals. Briefly, we fitted a generalized linear model (GLM), which related the previous second of the animals' experience to the probability of future licks. Overall, we fitted models to the behaviour of 40 behavioural sessions from 8 mice, while they performed the challenging stimulus detection task.

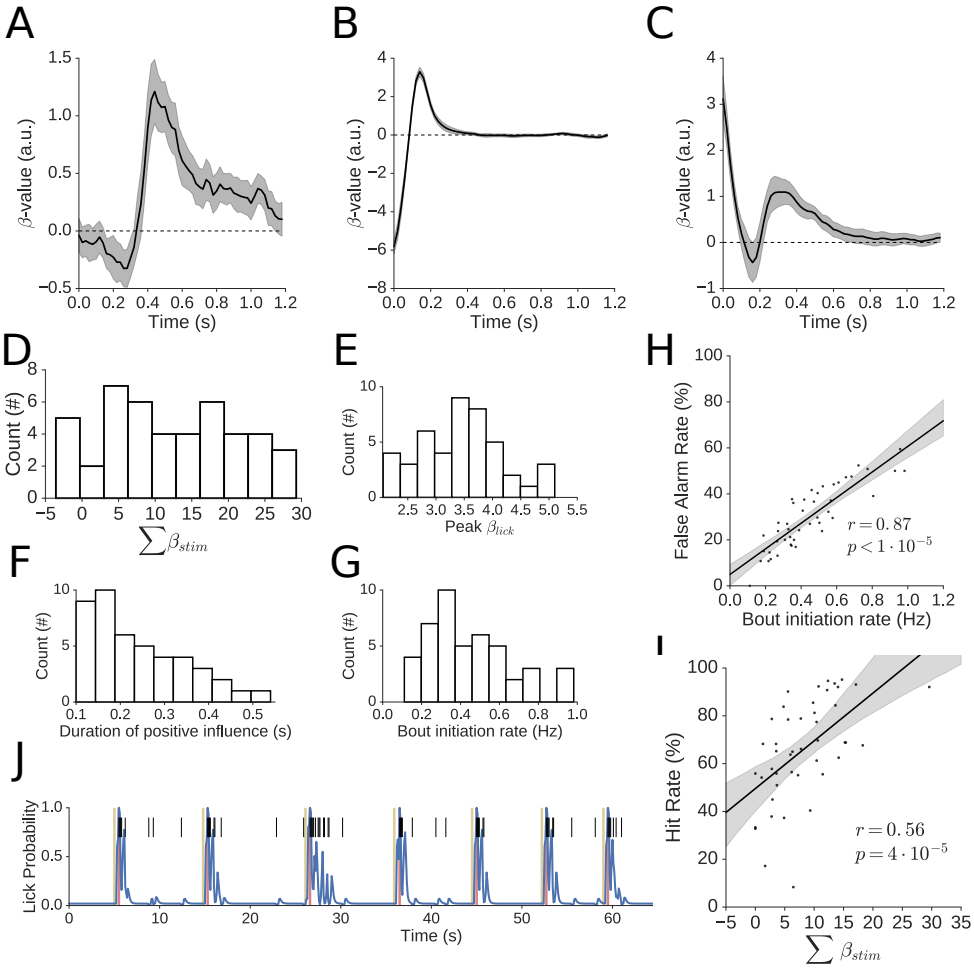


Figure 3.3: Consistent and variable elements of mice' behavioural strategies. Effect of click **A**, licks **B** and rewards **C** on future licking. **D** Range of bout initiation rates observed. **E** Range of peak positive values of lick kernels. **F** range of peak positive values of stimulus kernels. **G** range of positive durations of licks on future licks. **H** Correlation between inferred bout initiation rates and false alarm rates. **I** Correlation between stimulus sensitivity and hit rates. **J** Example fit to a training session shows prediction (blue) of licks (black) based on licks, rewards (red) and clicks (yellow)

To better visualise the influence of different events in the environment on mice behaviour, we grouped together parameters relating to the effect of a single type of event and plotted them sequentially in time. We refer to these groups of parameters as policy kernels (or policy filters). We modelled the influence of three factors on mice future licking behaviour: stimuli, previous licks and rewards and refer to the policy kernels in terms of their respective events.

Although the model was fit to individual behavioural sessions and individual mice, we observed several consistent features of the fits across sessions. Figure 3.3A-C show the mean policy kernels averaged across all mice and all sessions. Across mice we observed robust sensitivity to sound which evolved dynamically over one second following stimulus onset (Figure 3.3A). In the first $\tilde{2}$00ms following stimulus, the stimulus had little to no effect on licking, consistent with previously reported reaction times in mice (O'Connor et al., 2010; Guo, Hires, et al., 2014). Thereafter, typically 300-600ms after stimulus onset, we observed a large upwards deflection of the stimulus, indicating that lick-probability increased. After this peak, the stimulus-kernel decayed back to 0 over the course of 400-600ms.

The lick-kernels capture the effect of previous licks on future licks (Figure 3.3B). The mean lick-kernel has a sinusoidal profile. The initial negative deflection of the sinusoidal policy kernel for licks reflects the physical constraints on the animals' maximal lick rates ($\sim$10Hz) while the following positive deflection captures the bout-like structure of lick-trains that we ob-

served. By bout-like structure we refer to the observation that licks, rather than being uniformly distributed in time, tend to occur in in groups such that five to ten licks are likely to occur within a second. We reasoned that licks occurring with inter-lick intervals of less than 400ms to belong to a bout.

Finally the reward-kernel (Figure 3.3C) is initially strongly positive followed by a small negative deflection, in which mice consume the majority of the reward, followed by a prolonged increase in lick probability.

Across animals and sessions, we observed large variation across the values of these parameters and will describe our quantification of each in turn.

We characterised the stimulus-kernel in a given session by its sum. Across sessions, this varied substantially, indicating that stimulus sensitivity was substantially different across behavioural training sessions (Figure 3.3D). Importantly, in almost all cases the sum was above 0, indicating stimulus-induced licking occurring in most sessions.

We quantified the lick-kernels in two ways. We quantified the peak and the duration of the positive influence, both as a correlate of lick bout duration (i.e. the number of licks likely to occur at short inter-lick intervals). Both sets of parameters varied substantially across sessions and mice (Figure 3.3E,F).

Finally, we quantified the offset of our regression model. As we interpreted the lick kernel as a correlate of lick bout duration, we interpreted the offset parameter as the (lick-) bout initiation rate (i.e. the rate at which animals will spontaneously start licking, given that no events occurred in the preceding second). We calculated this by simply passing the fit offset

parameter through a sigmoid. This parameter varied over a 10-fold range from between 0.1 and 1Hz (Figure 3.3D).

We next validated our interpretation of the parameters describing the licking behaviour. We reasoned that if the offset parameter of our model mainly reflected bout initiation rates, its value would correlate strongly with false alarm rates. Consistent with this, we observed a strong correlation between between false-alarm rates and the value of the offset parameter, across animals and sessions (r=0.87; $p < 10^{-5}$) (Figure 3.3H).

Next, we wanted to confirm that the stimulus kernel (β_{stim}) reflected the probability of responding to the stimulus, rather than the vigour of the ensueing reward consumption. In order to examine this we correlated the sum of β_{stim} with the hit rate of the animal. Consistent with our interpretation of the stimulus kernel, we observed a robust correlation between the sum of β_{stim} and the hit-rates across sessions and animals (r=0.56; $p < 10^{-5}$) (Figure 3.3I). (see *see next section for further details*). Thus, the parameters of our model allowed us to capture significant variation in the strategies animals use to obtain reward and are relatable to separate, identifiable aspects of mice behaviour.

3.3.1 Posterior Predictive Checks

Posterior predictive checks, that is, simulating from fitted models to check correspondence with data, are an integral part of model validation (Gelman, Carlin, et al., 2004; Linderman and Gershman, 2017; Gelman and Shalizi,

2013). Fortuitously, simulating from the conditional distribution of our model is simple and has correspondingly been widely used for posterior predictive checks on related regression models of neural spike trains (Latimer et al., 2015; Pillow, Shlens, et al., 2008; Park, Bohner, and Macke, 2015).

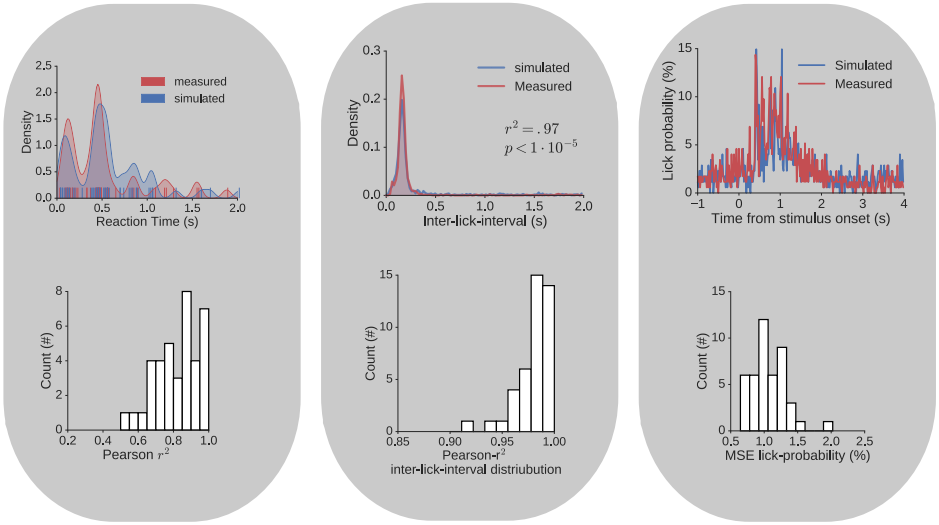


Figure 3.4: Summary statistics of mouse licking structure. *Left top*, example of a complex reaction-time distribution for measured data and data simulated from the model fit to the corresponding dataset showing good correspondence between the two. Curves are kernel density estimates while vertical lines on axes show individual cases. *Left bottom* summary of pearson-r^2 between kernel density estimates and simulated and measured reaction time distributions. *Centre top* Example empirical inter-lick interval distributions from measured data and simulated from corresponding fit. *Centre bottom* correlation coefficients between measured and simulated empirical distributions. *Right top* Peri-stimulus time histogram (in 20ms bins) of licks from measured example measured dataset and simulated from fit to the same dataset. *Right bottom*, mean-squared error between measured and corresponding simulated dataset.

Performing this kind of analysis relies on the comparison of summary

statistics computed from the real and simulated datasets (Gelman and Shalizi, 2013). For models of neural spike trains, validation is difficult, because the appropriate summary statistics depend on how population activity in the modelled area affects downstream areas, which is often unclear. Thus, it is unclear what features of the data are important to recapitulate.

In contrast to neural spike trains, the features of behaviour that our simulated data should recapitulate are clearer. A rich literature links various behavioural summary statistics to computations underlying behaviour as well as to neural function (Green and Swets, 1966; Carandini and Churchland, 2013; Heitz and Schall, 2012; Poort et al., 2015; Sutton and Barto, 1998; Shadlen, Britten, et al., 1996; Siniscalchi et al., 2016; Huber, Gutnisky, et al., 2012; Park, Meister, et al., 2014; Berditchevskaia, Cazé, and Schultz, 2016). We used features of behaviour identified as important by this work to guide further validation of our model.

Concretely, we therefore decided to validate our model by comparing the values measured in the real behaviour and the metrics derived from simulated data, for each behavioural session model fit pair. We chose 6 summary statistics to validate against (Figure 3.4, Figure 3.5), on the basis of either their relation to the classical sensory behaviour (i.e. hit-rate, false-alarm rate, reaction time) or as basic features of the behaviour (i.e. lick-rate, inter-lick intervals, lick peri-stimulus time histogram) and compared measured and simulated values.

Across our measures of interest, we observed a strong correspondence

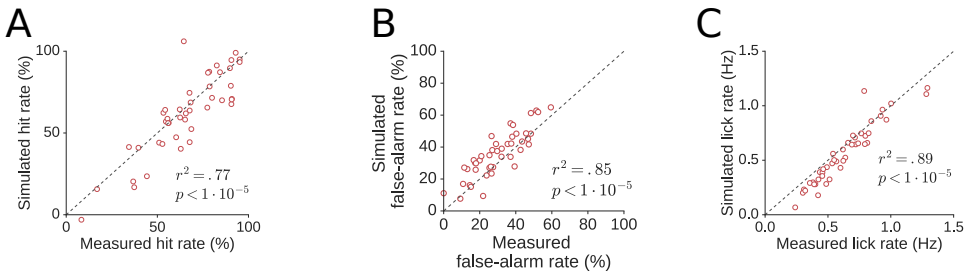


Figure 3.5: Comparison of traditional performance metrics derived from simulation and measured in real datasets. For each dataset and single animal, our model was fitted to the data and an equal number of trials to those in the measured dataset were simulated from our model. Comparison between measured and simulated session hit-rates **A** false-alarm rates **B** and overall lick-rates **C**.

---

between the summary statistics derived from the measured behaviour and from the simulation (Figure 3.4, Figure 3.5), further validating our models.

3.3.2 Influence of non-stimulus factors on reward rate

Although there were consistencies in policy parameters, we also observed substantial differences between parameter values across animals and sessions (Figure 3.3, Figure 3.6). Since in our paradigm, the response window was long (1s) and there was no punishment for excess licking, we reasoned that the structure of mouse spontaneous licking would affect reward rates and could be optimised. For example, if mice deterministically licked once per second, every second, they would receive every reward and would not need to perform many licks.

To quantify the effect of non-stimulus related parameters of our model

on the efficacy of mouse behaviour, we constructed two functions of reward and explored how variation in behavioural parameters affected each. The first function was the fraction of rewarded trials, which is proportional to the total reward, and hence we refer to this as *total reward*. With the second function, we aimed to measure the efficacy of the behaviour. We reasoned that excess licking was undesirable, hence defining the second function as the average number of licks required per reward. We refer to this function as *reward efficacy*.

We explored the space of possible policies in terms of these functions by generating a variety of stimulus unrelated policies (where all β_{stim} were set to 0), simulating behaviour using them and thereby observing the effects of varying policy-parameters on the two functions of reward.

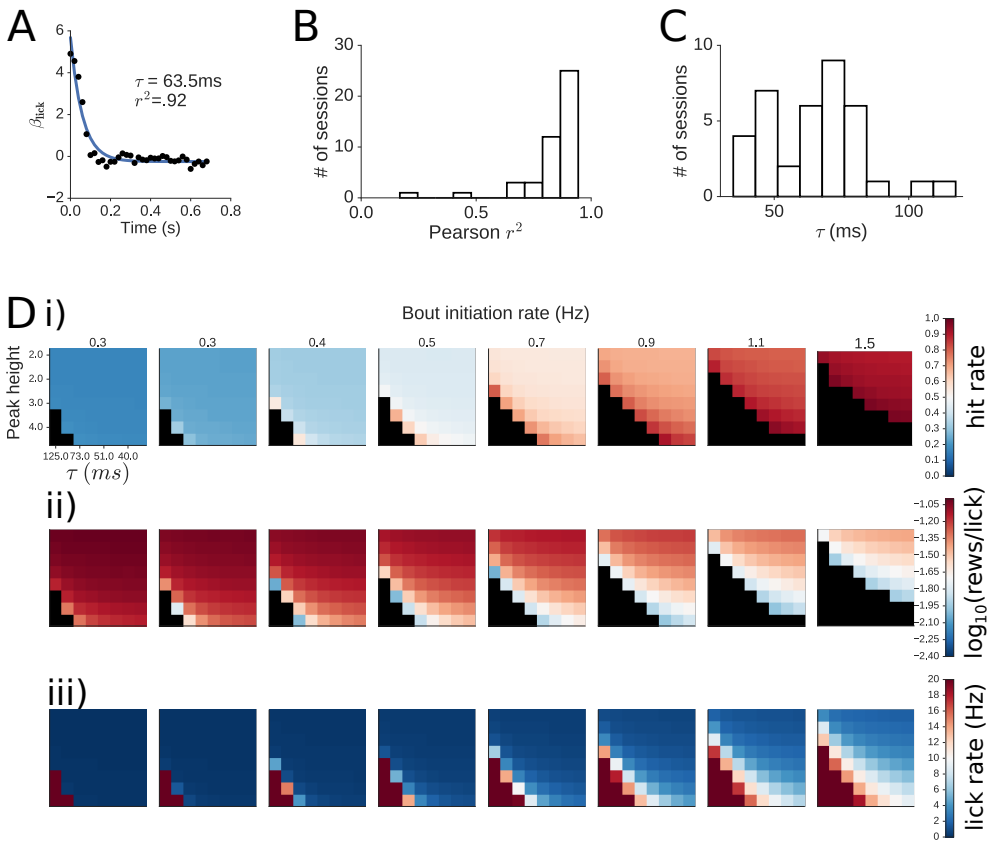


Figure 3.6: Characterising the value of policies. **A** exponential fit to decaying phase of lick-kernel. **B** Summary of quality of fits to all lick-kernels. **C** Distribution of decay time constants of lick-kernels. **D** Effect of varying policy parameters on fraction of rewarded trials (i), rewards per lick (ii), as well as lick-rates (iii). Lick rates above 10Hz are blacked out in the top two rows as they were never observed in the real data.

The space of stimulus-unrelated strategies is too high dimensional to study exhaustively and obtain meaningful insights. Thus we needed to reduce the dimensionality of the data in a consistent way that would allow us to easily generate artificial policies for simulation. We simplified stimulus-

unrelated parameter space to three dimensions:

1. The bout initiation rate
2. The height of the peak in the lick autocorrelation kernel
3. The decay of the lick autocorrelation kernel

While extraction of the first two of these parameters is trivial, the final parameter depends explicitly on many parameters of the model. To summarise these parameters, we fitted a single exponential to the decay of the β_{lick}. We found that single exponentials provided accurate summaries of decay time-courses of the fit parameters (Figure 3.6 A). Further, we found that the decay, together with the initial starting value and peak were sufficient to constrain a Gaussian process regression model of the complete shape of the lick kernels (Figure 3.6 B,C).

Equipped with our reward functions and reduced policy space, we proceeded to measure the influence of these three sets of behavioural parameters on reward rate. We generated artificial policies, constructed using the above three parameters which varied over the observed range in the data, used them to simulate behaviour and measured the effects of varying these parameters on our functions of reward. Consistent with our expectations, we found that all three parameters appeared to influence both functions in a systematic manner (Figure 3.6 D).

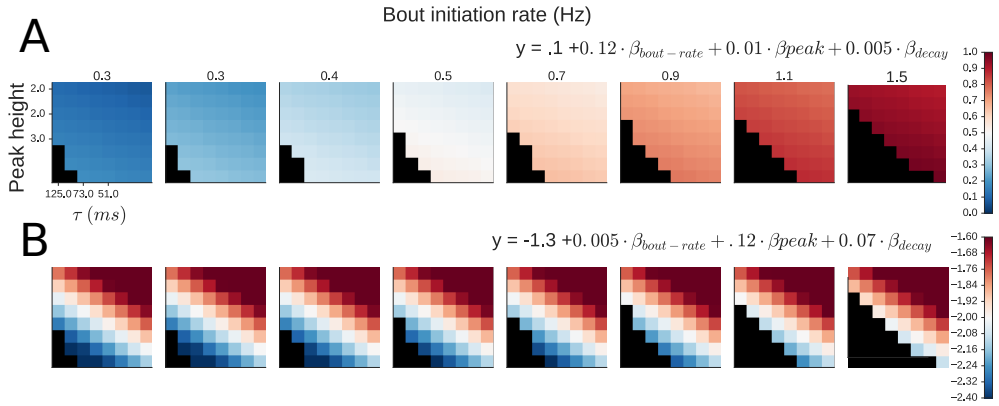


Figure 3.7: Linear fits to functions of reward. **A** Reward rate. **B** Reward-efficacy.

To quantify these results, we attempted to fit a linear model to the functions of reward. We found that a multilinear regression accurately captured the effect of varying these parameters on both total reward ($r^2 = 0.98$) and reward efficacy ($r^2 = 0.71$). This allowed us to quantify the effect of parameter variation on the two functions of reward by looking at the regression coefficients. Overall, the most effective way of increasing total reward while maximising reward efficacy was to increase bout initiation rates. Indeed, compared to increasing the lick autocorrelation, varying this parameter was 10-fold more effective at increasing total reward and reduced reward efficacy 10-fold less. Relative to extending the decay of the lick-kernel, it was 20-fold more effective and reduced efficacy 5-fold less. Overall then, increasing bout initiation rate is the most effective way to increase returns, though other parameters also contribute.

In light of the effects in parameter variations on reward rate, we hy-

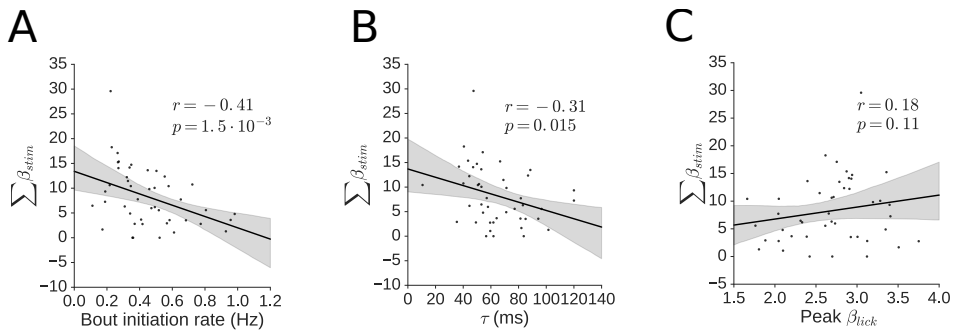


Figure 3.8: Compensation for poor stimulus sensitivity by adaptation of non-stimulus related parameters **A** across all sessions Bout initiation rates and stimulus sensitivity are anticorrelated. **B** Low stimulus sensitivity is associated with longer autocorrelations of licking behaviour in mice **C** No statistically significant relationship between the peak-height of the lick-policy filter and stimulus sensitivity

---

pothesised that the observed variations are not idiosynchratic features of the behaviour of different mice, but rather adaptive parts of their behavioural strategies. The strong test of this hypothesis would be to switch between different task difficulties (by decreasing stimulus levels) and see whether the behavioural parameters change in line with predictions from simulations. However, without having collected this data, we made do with what was available and present a preliminary analysis of data consistent with these predictions.

We reasoned that in sessions where mouse behaviour was less stimulus driven (i.e. the task was perceived as more difficult), they would compensate by engaging in effective sampling behaviour to maintain high reward rates. Based on the result of our stimulations, we reasoned that in sessions with

lower stimulus sensitivity, mice would show higher their bout initiation rates, increased peaks in the licking autocorrelation and longer positive influences on future licks.

Our predictions were partially born out. We observed significant negative correlations between stimulus sensitivity, as measure by $\sum \beta_{stim}$, and both bout initiation rates (r=-0.41, p=$1.5 \cdot 10^{-3}$) and decay constants (r=-0.31, p=0.015); but no effect on the peak of the autocorrelation (r=0.18, p=0.11). Thus, mice compensate for reduced stimulus sensitivity by altering non-stimulus related parameters, thereby using parallel approaches for optimising their behaviour.

3.3.3 Decoding the causes of behaviour

In summary, we found that mice use heterogenous strategies to optimise their behaviour in head-fixed environments. In light of this, behaviour on different trials is likely driven by different factors. Since the stimulus-related licking occurs intermingled with stimulus-unrelated licking, we wondered whether we could use our model to identify licking that was driven by the presence of the stimulus. In other words, we wondered whether we could decode the occurrence of stimuli from behaviour alone.

The stimuli we were interested in decoding were those that the animal responded to. To achieve this, we thus heuristically defined hit trials based on the reaction time distribution and inter-lick interval distributions. We exploited the bimodal shape of the reaction time distribution to constrain

stimulus driven hit-trials to those that occur between 300 and 800ms post stimulus onset (Figure 3.9A). We also aimed to exclude licks that were part of a lick bout. To do this, we observed the inter-lick interval distributions, noting that in all cases, the observed peak decayed to baseline by 400ms (Figure 3.4 *centre*). Thus, as a second criterion, we required that no lick had occurred 400ms prior to response. All other trials were considered false alarms.

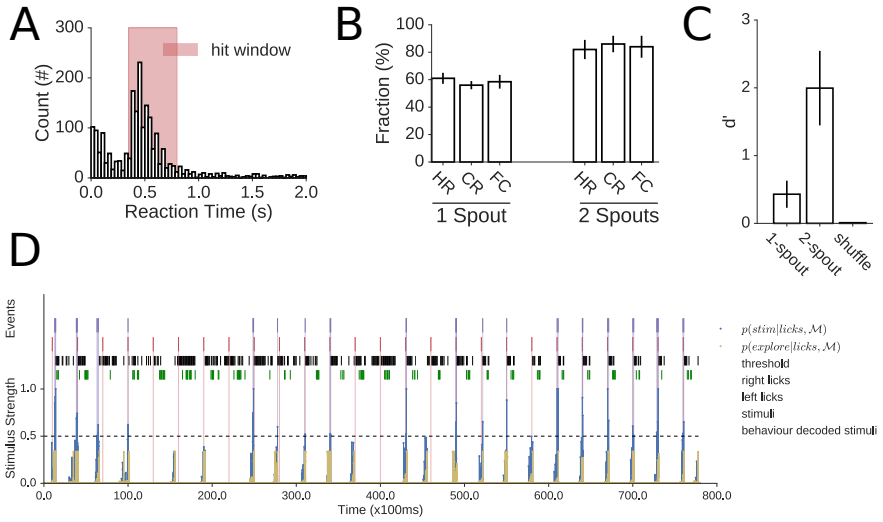


Figure 3.9: Decoding the causes of behaviour **A** Reaction time distribution across all animals and sessions. Red shading indicates region heuristically defined as a hit-trial in the following analysis. **B** Decoding performance separated into hit-rates (normalised to number of responses to stimuli), correct-reject rates (normalised to total number of lick bouts and fraction correct (mean of the previous two metrics). Results for behaviour with one-and two-lick spouts shown separately. **C** Example of decoding performance **D** Example decoding of stimuli from lick trains. Black and green vertical lines indicate lick times. Red line indicate stimulus times and purple lines indicate decoded stimulus times.

Using these criteria then, we attempted to decode those presented stimuli from behaviour. We separately decoded two potential causes of behaviour, stimuli and exploratory or spontaneous drive. The kernel for stimuli was that inferred from the training data for the corresponding session, while the exploratory drive was simple an impulse function lasting 50ms post stimulus onset. The introduction of the second kernel was necessitated to reduce the false alarm rate for spontaneous lick-initiation. As has been done previously (Pillow, Ahmadian, and Paninski, 2010), we decoded stimuli as present when the inferred stimulus strength (blue-line in Figure 3.9D exceeded a value of 0.5).

We first performed this analysis on cross-validated subsets on the data with animals licking a single spout. We calculated hit rates relative to the number of presented stimuli, and as a conservative estimate of the correct-reject rates relative to the number of non-stimulus driven lick bouts (we defined lick bout as successive licks with an inter-lick intervals of less than 400ms). To calculate an overall 'fraction correct' we averaged these two statistics. We also used these metrics to calculate d'. Using these criteria, we found that we were able to decode the occurrence of stimuli (hit-rate=61%) at with some false positives (correct-reject rate = 59%) (Figure 3.9B,C).

Although we could decode the occurrence of stimuli, we surmised that accuracy levels were not sufficiently high to use our estimates as a basis for further analysis of e.g. neuronal data. We therefore wondered if we could alter our experimental paradigm so as to obtain better decoding performance.

Since our goal here was, in essence, to make stimulus driven behaviour more discriminable from spontaneous licking, we reasoned that introducing a second lick spout, whose behaviour completely mirrored the first (in the sense that it obeyed the same task-contingencies as the first see *Methods* for details) might lead to greater differences in spontaneous and stimulus driven licking. provide us with additional information about the animal's behaviour, and might allow us to better decode behaviour.

We thus trained two additional mice on the click-detection task, but now with two lick spouts available. We again attempted to decode the lick patterns. Consistent with our predictions, we now achieved substantially better performance (Figure 3.9B,C) (hit rate=82% correct-reject rate = 86%). Two lick spouts are commonly used to train head-fixed mice to perform two-alternative forced choice discrimination tasks (Carandini and Churchland, 2013; Guo, Hires, et al., 2014) which our results suggest is a sensible strategy to use.

3.3.4 Quality of Fits

Finally, we attempted to improve the quality of our model fits by assuming additional structure in the policy kernels (Park and Pillow, 2011; Sahani and Linden, 2003b) and performing inference within a Bayesian framework. Specifically, we assumed that policy kernels would be smooth in time, with a degree of smoothness to be inferred from the data (see Empirical Bayes Section in Methods).

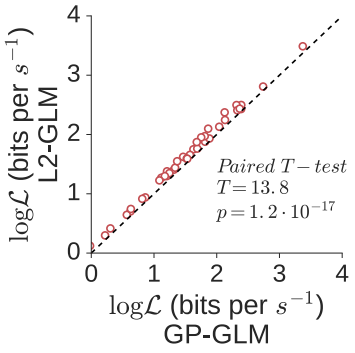


Figure 3.10: Comparison of encoding quality between GP-GLM and L2-GLM

To assess the quality of fits, we fitted our data to behavioural sessions (n=36 sessions from 6 animals). We assessed model fits using information based measures of model performance and found that our standard L2-regularised GLM outperformed the the smoothness regularised GLM (GP-GLM) (paired t-test p=$1.2 \cdot 10^{-17}$) (Figure **??**).

3.4 Discussion

Overall, we have developed and deployed Generalized-Linear Models for the description of behaviour in head-fixed mice. We have shown that we are able to describe the detailed licking patterns of mice performing simple operant tasks in head-fixed environments using these models. We then proceeded to show that in this environment, stimulus-unrelated aspects of mouse behaviour can have large influences on reward rates and efficiency. We proceeded to present preliminary evidence that head-fixed mice use this extended parameter space to optimise their behaviour by engaging in efficient sampling strategies. Finally, we show that in the face of the use of heterogenous strategies by head-fixed mice, we can accurately decode the causes of their behaviour.

3.4.1 Detailed descriptions of behaviour

The main opportunity provided by the policy based framework in the context of head-fixed behaviour is the description of a large part of their behavioural output (i.e. their complete licking patterns). In freely-moving animals, idiosynchratic motor behaviour forms an integral part of animal's behavioural strategies (Gomez-Marin, Paton, et al., 2014; Skinner, 1948b; Kawai et al., 2015b) and the natural analytical framework for its description is unclear (though see **Wiltshko2015**). In contrast, the large amount of homogenous (in the statistical sense of being drawn from one distribution) behavioural

data available per decision epoch (in our case, stimulus presentation) provides a rich dataset with which to describe mouse behavioural policies. The homogeneity suggests a natural analytical framework in which to analyse the behavioural data, making it amenable to a rich set of analysis methods (McCullagh and Nelder, 1989; Pillow, Ahmadian, and Paninski, 2010; Ahmadian, Pillow, and Paninski, 2011; Lewi, Butera, and Paninski, 2009; Lewi, Schneider, et al., 2011; Truccolo, 2004; Roweis and Ghahramani, 1999) which afford the ability to accurately describe behavioural strategies.

The work most similar to ours characterises behavioural strategies in tasks with discrete decision points (Busse et al., 2011; Carandini and Churchland, 2013; Fründ, Wichmann, and Macke, 2014b; Smith et al., 2004a; Ito and Doya, 2009). Consistent with this literature, we observe that substantial regularity in mouse stimulus-unrelated behaviour. Additionally, we observe substantial variation across sessions and individual mice in the strategies they use to obtain rewards similar to what has been reported previously (Busse et al., 2011; Ito and Doya, 2009).

We build on this work, showing that during performance of a perceptual task, structured, reward-driven behaviour takes place outside of the decision epoch. These insights are derived from our use of a policy-based framework for understanding behaviour, which enables use of data outside these epochs to estimate behavioural strategies.

Analysing behaviour in this framework is also likely to provide new insights into cognitive processes underlying behaviour. Consider a trained

mouse performing in a go/nogo paradigm. In an action-centred paradigm (**Berdit2017**;), the mouse's interpretation of the nogo stimulus is unclear: is the stimulus meaningless to the animal behaviourally or is it suppressing a go behavioural response? These are two distinct cognitive processes, subserved by distinct neural mechanisms (Bannerman et al., 2012; Schmidt et al., 2013; Obeso et al., 2014). By considering behaviour outside of the decision point, we can thus ask whether behaviour following presentation of the nogo stimulus affects behaviour, imbueing this important distinction with an operational definition. Thus, modelling behaviour in terms of policies is likely to provide novel insights into the cognitive mechanisms underlying behaviour.

This wealth of data also affords higher temporal resolution for the estimation of behavioural strategies. In action-based models of mice behavioural strategies are by necessity defined stochastically across trials. The basis of decisions on single trials remains unobserved (Shadlen and Kiani, 2013). In contrast, we have shown that we are able to identify the causes of behaviour with single lick-bout resolution, enabling inference on a trial-by-trial basis of the causes of the animal's behaviour. While we considered, as an example, blind decoding of stimulus times, other approaches are possible. For example, given two qualitatively different models of behaviour, one can easily use these methods to identify which better describes the behaviour, on a single trial basis, by a simple likelihood ratio test (Pillow, Ahmadian, and Paninski, 2010; Pillow, 2005). This could provide more novel insights into the struc-

ture of behaviour in contexts where it has been difficult to analyse previously e.g. during initial-task learning (Tervo, Tenenbaum, and Gershman, 2016).

The ability to interpret behaviour on a trial-by-trial basis is also essential to the analysis of neural data. Indeed, the inability of probabilistic models of behaviour to provide single trial descriptions has been a source of criticism (Shadlen and Kiani, 2013). We envision that the high temporal resolution analysis of behaviour could provide better understanding of neural activity.

Analysing neural data on a trial by trial basis was pioneered in sensory research in primates (Parker and Newsome, 1998; Britten et al., 1996). Today it is increasingly being used to study sensory behaviour in rodents (Inagaki et al., 2017; Yang, Kwon, et al., 2015; Runyan et al., 2017), whose behaviour is notoriously noisier (Busse et al., 2011; Carandini and Churchland, 2013) (see also chapter 1). Further, with increasingly sophisticated methods for recoding large populations of neurons (Jun et al., 2017; Denk, Strickler, and Webb, 1990; Sofroniew et al., 2016) and describing the resulting datasets (Stevenson and Kording, 2011; Cunningham and Yu, 2014; Rossant et al., 2016; Yu et al., 2009), neural data is increasingly being analysed in a trial-by-trial manner using more diverse paradigms (Churchland et al., 2012; Stevenson and Kording, 2011; Inagaki et al., 2017). We envision that analytical tools for behaviour, such as the methods presented here, will prove essential in furthering our understanding of the generation of the neural and computational bases of behaviour, particularly in the context of complex tasks (Akam, Costa, and Dayan, 2015; Brunton, Botvinick, and

Brody, 2013; Schmitt et al., 2017) or over learning, where the causes of behaviour may be less clear.

3.4.2 Understanding the value of behavioural strategies

In our paradigm, as in many others (Guo, Li, et al., 2014; Yang, Kwon, et al., 2015; Huber, Gutnisky, et al., 2012; Carandini and Churchland, 2013; Runyan et al., 2017), the response and reward delivery mechanisms are combined into a single physical apparatus. Concurrently, in head-fixed environments, mice continuously engage with the response apparatus (**Berdit2017**; Guo, Li, et al., 2014). This gives rise to a situation in which mice receive reward by sampling the decision apparatus. In turn, this, at least in principle, provides opportunity for adaptive sampling behaviour to emerge.

We constructed two functions of reward and showed that we were able to quantify the value of behavioural strategies using these functions. Doing so, we confirmed that stimulus-unrelated behaviour had a major effect on reward rates and the efficacy with which reward was obtained. Importantly, we were able to characterise this dependence using a low dimensional description of mouse behaviour. In doing so, we identified behavioural parameters which had a large influence on reward rates, finding that increasing bout initiation rates is the most effective means of increasing returns.

Building on this characterisation, we provide preliminary evidence that

mice use this stimulus-unrelated space to optimise their behaviour. We showed that in sessions where mice are less sensitive to stimuli they engage in more effective sampling behaviour. This suggests that mice may compensate for reduced sound sensitivity with improved sampling behaviour. These conclusions align with previous work showing that animal's can learn optimal sampling policies by trial and error (Desrochers et al., 2010; Sejnowski, 2010). However, our findings add to this work by showing that animals use efficient sampling in parallel with other adaptive behavioural strategies.

The fact that the parameters or behaviour are linearly related to the value of policies provides an interesting finding for future research. A major problem for reinforcement learning algorithms is that the space of possible inputs (e.g. all the individual pixels of an image) and outputs (e.g. the musculature of arms and fingers) are too large to learn a lookup table representation (Sutton and Barto, 1998). The way these algorithms typically work is by using functions that approximate the value of actions or strategies (Geramifard et al., 2013; Li, 2017). However, function approximation has, to the best of the author's knowledge, not previously been studied in biological agents. The fact that a linear function approximates the value of strategies in the studied space may, due to its analytical tractability, provide a useful paradigm in which to understand whether and how experimental animals approximate value-functions.

3.4.3 Future work

The most important piece of remaining work is to conclusively demonstrate that mice adapt multiple features of their behaviour in order to maximise their reward rate. To validate this beyond the preliminary analysis presented here, we have planned the following experiment. We plan to train mice to perform the task as outlined above. Then we plan to characterise animals' ability to detect these stimuli at different sound levels. Then, across blocks of variable numbers of stimuli, to present stimuli of greater and lower intensities in order to test whether animals compensate for reduced stimulus-levels by altering the stimulus unrelated parameters predicted by our analyses.

A second further piece of work will be to robustly benchmark our approach against existing methods, to see whether our methods provide better predictions of behaviour from an action-centred perspective.

The computational methods described here can also be easily and extended and improved. Firstly, identifying useful priors for model parameters (Park and Pillow, 2011; Sahani and Linden, 2003a) would allow more rapid and accurate estimation of parameter values. Secondly, extending these models to include nonlinearities might be useful. However, in preliminary analysis of fitting these datasets using artificial neural networks, we observed no improvement in performance (data not shown). Thirdly, the validity of the assumptions of the Gaussian posterior should be validated using Markov-Chain Monte-Carlo methods (**Adhmadian2010a**; **Pillow2011**; Park and Pillow, 2011). Finally, for freely-moving behaviour, there are dynamic mod-

els of behaviour which change over time (Smith et al., 2004a; n.d.; Chen, 2015). These are based on reinforcement learning or on statistical models of changing behaviour. We address this in future chapters.

3.4.4 Conclusion

In summary, our work makes two contributions. It provides a set of computational tools for characterising the structure and causes of behaviour in head-fixed mice, resolving each with high temporal resolution. Secondly, it provides a quantitative framework for understanding the value of different behavioural strategies in environments in which animals continuously engage with the response apparatus. Importantly, due to the multiple factors affecting reward rates, these environments are much more naturalistic and will may provide substantial insights into learning and decision-making. While our work has focussed entirely on the behaviour of head-fixed mice performing operant tasks, these methods are more broadly applicable to any task where behaviour is continuously monitored.

Chapter 4

Dynamical Models of Behavioural Change

4.1 Introduction

Humans and other organisms can often effortlessly learn to solve tasks which pose incredible challenges to artificial intelligences (Tsividis et al., 2017; Li, 2017). The difference is particularly pronounced for 'natural' real-world problems (Sutton and Barto, 1998; Li, 2017). One of the reasons that many problems in the natural world are difficult to solve is that they are often high dimensional; that is there are many *a priori* orthogonal dimensions of input which must independently be mapped to an appropriate, possibly itself high dimensional, output. The earliest pioneers of learning theory realised this problem, often referred to as the curse of dimensionality (Bellman, 1961).

How humans and other animals are able to solve these high-dimensional problems is an open question (Niv et al., 2015; Leong et al., 2017; Tervo, Tenenbaum, and Gershman, 2016; Gershman, Norman, and Niv, 2015).

The dominant framework for modelling learning and dynamic behaviour in experimental animals has been reinforcement learning (reinforcement learning) (Sutton and Barto, 1998; Gershman and Daw, 2017). The umbrella term reinforcement learning encompasses a large set of methods which enable learning to act effectively based on previous experience.

A subset of these algorithms, value-based methods, have been widely and successfully applied to modelling dynamic behaviour and learning in neuroscience (Doll, Simon, and Daw, 2012; Ito and Doya, 2009; Sutton and Barto, 1998; Daw, 2011; Daw et al., 2011b). In broad terms, these algorithms model the value of performing specific actions in specific environmental states (Sutton and Barto, 1998). These states are typically defined *a priori* as high level abstractions of animals' environment (e.g. the response-window in an operant conditioning paradigm).

Without these abstract states, states must be naïvely defined as unique sets of inputs to an agent (). From this naïve perspective, even the simplest environment gives rise to a dizzying number of states, each defined as a unique combination of sensory inputs and previous behaviour. For example, all the possible combinations of visual input a mouse might experience in a Skinner box already constitute an innumerable set of states, as visual inputs change dependent on the exact position and orientation of animals in their

environment, ambient lighting, etc. Under these circumstances, value-based reinforcement learning algorithms quickly succumb to the curse of dimensionality as they become incredibly inefficient when the number of states and/or available actions becomes large (Gershman and Niv, 2010; Niv et al., 2015; Sutton and Barto, 1998).

The experimental study of animals' representation of the state structure of their environment is in its infancy (Tervo, Tenenbaum, and Gershman, 2016). The majority of this work builds on the basic experimental observation that that when animals are successively trained to solve two tasks with a similar structure (e.g. two sensory discrimination tasks), the second task will be acquired more quickly than the first (Herrnstein, 1961; Braun, Mehring, and Wolpert, 2010; Scheggia and Papaleo, 2016; Garner et al., 2006; Baxter and Gaffan, 2007; Starkweather et al., 2017). The inference from this is that in addition to learning the specific contingencies in their environment, animals also learn about their environment's (state) structure.

However, how animals acquire the state structure of their environment, or whether they use state-based representations for initial learning, is unclear (Gershman and Niv, 2010; Tervo, Tenenbaum, and Gershman, 2016). Therefore, value-based reinforcement learning approaches, which rely on a defined state structure, may have limited applicability to initial task learning.

An alternative approach, known as policy based reinforcement learning, circumvents direct solution of the problem of abstract state inference. Rather than learning the values of actions in abstract state structures of the environ-

ment, this class of algorithms learns the parameters of a functional mapping from inputs to the outputs directly (Sutton et al., 1999b; Deisenroth, Neumann, and Peters, 2011; Williams, 1992). One of the reasons these methods can overcome the curse of dimensionality is that although many problems are a priori high dimensional, their intrinsic dimensionality is low (Salimans et al., 2017; Kemp and Tenenbaum, 2009). In light of this, algorithms directly optimising policies become useful, even in high dimensional problem spaces. In other words, rather than explicity learn an efficient state structure, the update rules these algorithms use implicitly exploit any state structure inherent in the world.

Guided by the policy-based reinforcement learning approach, we analyse data from head-fixed mice over the course of learning a simple operant task. We trained head-fixed mice to perform a simple operant task and observed how their goal-directed behaviour changed over the course of several days of training. To analyse the data in terms of changing policies, we developed Bayesian state-space models which track the animal's policies with single-trial resolution. Importantly, our algorithms are agnostic to learning rules and environmental models and are thus broadly applicable to initial task learning. We show, using simulation and on real data, that our models are able to accurately infer the high-dimensional function that maps the history of the animal's licking, as well as external stimuli to, future actions. Analysing the resulting fits further, we show that although the space is high-dimensional, animals efficiently optimise their behaviour on a low-dimensional manifold of

the high dimensional behaviour space. We conclude that mice use the low intrinsic dimensionality of the problem space to efficiently optimise behaviour.

4.2 Methods

4.2.1 Animals and Behavioural Paradigm

For detailed animal welfare considerations and surgical procedures, see chapter 2 and chapter 5.

7 mice (C57B6.129S-Gt(ROSA)26Sor$^{tm95.1(CAG-GCaMP6f)Hze}$ [Jackson Laboratories; stock number 024105] x C57B6.Cg-Tg(Camk2a-cre)T29-1Stl/J [Jackson Laboratories; stock number 005359]), aged 5-7 weeks, were surgically implanted with a 3mm cranial window over the right auditory cortex (Vasquez-Lopez et al., 2017) and a headpost for head-fixation. After a two week recovery period, during which animals were habituated to handling, animals were trained over the course of approximately 1 week to perform a simple click detection tasks while undergoing chronic two-photon imaging (see chapter 5 for details).

Following recovery, mice were head-fixed in the experimental setup and trained to perform a simple click detection task. Clicks were 5ms in duration, biphasic and presented at 80dB SPL. Mice were placed in a head-fixed behavioural setup (chapter 2 for details) with two available lick spouts centered on the animals snout. Lick spouts were individually positioned across days. Clicks were presented at random inter-stimulus intervals drawn from

a uniform distribution from 6-12s. If mice licked either spout within a 1s window following the onset of the stimulus bilateral water rewards ($3\mu l$ per spout) were delivered immediately. Animal's were trained on the task until satiation, which lasted variable durations.

The behavioural apparatus was controlled from a computer running Windows 7 using MATLAB interfaced with a National Instruments board (NI-DAQ USB-6008) for data acquisition. Stimuli were presented using MATLAB 2016a (Mathworks) running psychtoolbox. Stimuli were digital-to analog converted using a commercial soundcard (ASUS Xonar-U7), amplified (Portable Ultrasonic Power Amplified; Avisoft Bioacoustics) and played through a free-field electrostatic speaker (Vifa; Avisoft Bioacoustics), positioned approximately 15cm in front of the mouse's snout.

4.2.2 Generative Model

Our goal was to model the changing licking patterns of head-fixed mice over learning a simple auditory detection task. We observe behaviour in a time interval $(0, T^{end}]$ and binned this interval into T discrete bins $t \in \{0, 1, \cdots T\}$ such that each time bin contains at most one lick (in practice, we used time bins of 100ms). Let $x_{1:N,t} \in \mathbb{R}^N$ denote the set of N observed covariates at time t (i.e. stimuli presented, rewards received and history of licking) and $y_t \in \{0, 1\}$ denote whether the mouse licked in time bin t. Further, let $x_{1:N,1:T} \in \mathbb{R}^{N \times T}$ and $y_{1:T} \in \{0, 1\}^T$ denote the entire observation history of covariates and licks. We assume that the behaviour of mice is stochastically

governed by our set of covariates and defined a set of parameters $\theta_t \in \mathbb{R}^N$ which relate our measured covariates to $x_{1:N,t}$ to the mouse's behaviour y_t at each time point. Since mice are learning a task, we assume that the parameters $\theta_{1:N,t}$ are changing across time. Simplifying notation by denoting $\langle\cdot\rangle_{1:q}$ by $\langle\cdot\rangle_{:}$, we model the behaviour using a Bernoulli GLM, giving rise to the following likelihood function:

$$\mathcal{L}(\theta_{:,1:t}, x_{:,1:t}, y_{1:t}) = \sum_{t=1}^{T} p(y_t = 1|\theta_{:,t}; x_{:,t}) = \sum_{t=1}^{T} (1 + exp(-\theta_{:,t} x_{:,t}))^{-1} \quad (4.1)$$

Estimating the parameters of this model is a challenging problem given the limited amount of data available at each time-point. Given the natural trial structure of our data, we bin our data into P discrete trials of length M, such that $x \in \mathbb{R}^{P,N,M}$, and assume for tractability that $\theta_{:,t}$ are constant across timebins within a trial. Thus, $\theta \in \mathbb{R}^{P,N}$. Further, we assume that across trials, $\theta_{p,:}$ changes slowly. We incorporate this into the model by assuming that $\theta_{:,t}$ evolves with linear-Gaussian dynamics given by:

$$\theta_{:,1} \sim \mathcal{N}(\theta_{:,0}, Q_0) \quad (4.2)$$

and

$$\theta_{:,t} | \, \theta_{:,t-1} \sim \mathcal{N}(A\,\theta_{:,t} + Bu_{:,t},\, Q) \quad (4.3)$$

where $\theta_{:,0} = \mathbb{E}[\theta_{:,1}]$, $Q_0 = \mathbb{E}[\theta_{:,1}\theta_{:,1}^T]$, $Q = \mathbb{E}[\theta_{:,t}\theta_{:,t}^T]$, A is a $P \times P$ dynamics matrix describing the determinstic component of the state evolution, B denotes the deterministic input driven part of the dynamics and $u{:}, t$ denotes the inputs at timepoint t.

4.2.3 Inference and Learning

Inference of the latent states and generative parameters in this model is intractable. However, we can reformulate the problem as an inference problem over latent variables to obtain a way of inferring the parameters of the model. Specifically, if we treat the changing strategies, $\theta_{:,:}$, as latent variables, and the parameters of the change process, i.e. $\Phi = \{A, B, Q, Q_0, x_0\}$, as the parameters of the model, we can perform inference using the Expectation-Maximisation algorithm (Dempster, Laird, and Rubin, 1977).

Formally, the objective we seek to maximise then is

$$\mathcal{L}(\Phi) = \sum_{trial=1}^{P} \log p(y|\Phi) = \sum_{trial=1}^{P} \int \log p(y, \theta|\Phi) d\theta \tag{4.4}$$

This approach is effective in this situation because conditional on knowing the parameters of the model i.e. Φ, we can infer the strategies $\theta_{:,:}$, and knowing the strategies, we can infer the parameters. Parameter estimation in the EM algorithm iterates between two steps. First, inferring the latent variables, here $\theta_{:,:}$, given the current estimate of the parameters of the model (E-step), here Φ. Second, inferring the parameters of the model given the

best estimate of the latent variables (M-step). Practically, this involved iteratively inferring $p(\theta|y, \Phi)$ based on the current estimate of Φ and then maximising an approximation to **??** with respect to Φ, using $p(\theta|y, \Phi)$. In essence, this works because the constructed approximation provides a lower bound on the true likelihood function, which is tight at the current estimate of Φ. Maximising this function will thus monotonically increase the likelihood provided by **??**. Accordingly, repeated iterations are guaranteed to ascend to a local-optimum of the likelihood function.

Below we detail the mathematics of the E and M step in our specific case.

4.2.3.1 E-step

In the E-step we wish to infer $p(\theta|y, \Phi)$. By Bayes Rule:

$$\log p(\theta|y, \Phi) = \log p(y|\theta) + \log p(\theta|\Phi) = L_{data} + \mathcal{L}_{prior} \tag{4.5}$$

Thus, the likelihood decomposes into two parts. The first part is the Bernoulli likelihood function whose derivative and Hessian were derived previously in chapter 3. The only difference here is that the likelihood of the data for each trial is computed separately with corresponding parameters of each trial. Similarly, the second term is simply the likelihood θ under a Gaussian prior, whose likelihood and first and second derivatives are here given by:

$$\mathcal{L}_{prior} = \frac{1}{2}(\theta - \mu)_{\pi}^{T}\Sigma_{\pi}^{-1}(\theta - \mu_{\pi}) \tag{4.6}$$

$$\nabla\mathcal{L}_{prior} = \Sigma_{\pi}^{-1}(\theta - \mu_{\pi}) \tag{4.7}$$

$$\nabla\nabla\mathcal{L}_{prior} = \Sigma_{\pi}^{-1} \tag{4.8}$$

where μ_{π} and Σ_{π} are the prior mean and covariance, arising from Φ in closed form (Paninski et al., 2010).

As discussed previously, there is no analytical solution for this problem and no closed form posterior. However, since the posterior of this model is log-concave and unimodel (McCullagh and Nelder, 1989), we can perform inference using the Laplace Approximation, which approximates the posterior using a Gaussian (Bishop, 2006). The Laplace Approximation of the posterior can trivially be derived by approximating the true posterior with the first two terms of its Taylor Expansion (). From this, it is clear that the negative inverse of the Hessian corresponds to the covariance of the posterior.

Thus, to perform the E-step, we use the Newton-Raphson method and arrive at an approximation to $\log p(\theta|y, \Phi)$.

4.2.3.2 M-step

Armed with an approximation to $\log p(\theta|y, \Phi)$, we can now proceed to maximise our **??** with respect to Φ. Fortunately, as we have assumed linear-

Gaussian dynamics, we can update all parameters of Φ in closed form (Roweis and Ghahramani, 1999; Chen, 2015). The closed form updated are given by:

$$x_0 = \mu_1 \tag{4.9}$$

$$Q_0 = \Sigma_{11} + (x_0 = \mu_1)(x_0 - mu_1)^T \tag{4.10}$$

$$A = \frac{1}{P-1}\left(\sum_{t=2}^{P} M_{t,t-1}\right)\left(\sum_{t=2}^{P} M_{t-1,t-1}\right) \tag{4.11}$$

$$Q = \frac{1}{P-1}\sum_{t=2}^{P}\left(M_{t,t} + AM_{t-1,t-1}A^T - AM_{t-1,t} - Mt, t-1A^T\right) \tag{4.12}$$

where $M_{t,s} = \Sigma_{t,s} + \mu_t \mu s^T$ and $\Sigma_{t,s}$ is the posterior covariance of $\theta_{:,t}$ and $\theta_{:,s}$ (i.e. the posterior covariance of policy parameters between trials t and s) and μ_t is the posterior mean of $\theta_{:,t}$.

4.2.3.3 Fitting

Our ssGLM models were fitted separately to behavioural sessions on different days. Parameter initialisation was held constant for all experiments. Preliminary analyses indicated that within a reasonable range, different initialisations converged to the same final values. We performed a maximum of 50-EM iterations. Since our E-step was based on the Laplace Approximation,

the EM guarantee of a monotonically increasing likelihood was broken. We therefore used a lookahead procedure to monitor convergence (Macke et al., 2011). In practice, most fits were terminated by early stopping, rather than running through all 50 iterations.

4.2.4 Policy Gradient Methods

Here, we will derive the policy gradient methods following (Peters and Schaal, 2008; Peters, 2010). For a full derivation of the policy gradient theorem, see (Sutton et al., 1999b). Let t denote the current time step, $y_t \in 0, 1$ the chosen action (lick or no lick), and x_t denote the current state of the system. Further, we assume that the current state evolves according to some probability distribution:

$$x_{t+1} \sim p(x_{t+1}|x_t, y_t) \tag{4.13}$$

which may or may not depend on the current action. Further, let us define a policy, which is a conditional probability distribution over actions as:

$$y_t \sim \pi_\theta(y_t|x_t) \tag{4.14}$$

where $\theta \in \mathbb{R}^d$ denote parameters linking the current state of the world to a chosen action. Finally, at each time point t, the learner also receives a reward $r_t(x_t, y_t)$. The goal of the learner is to change these parameters in order to maximise its expected return:

$$J(\theta) = \mathbb{E}\left\{ C^{-1} \sum_{t=0}^{T} \gamma^t r_t \right\} \tag{4.15}$$

where γ is a discount factor which means that rewards further in the future are less valued, and C is simply a normalisation constant for the γs.

This problem can be conveniently reformulated, offering a simple solution method. Assume that trajectories τ are samples from a process defined over the space $\mathcal{T}$ such that

$$\tau \sim p_\theta(\tau) = p(\tau|\theta) \tag{4.16}$$

and let $r(\tau)$ denote the rewards generated by a one draw from that process. This enables us to redefine the expected returns $J(\theta)$ as the average over all possible trajectories that could be sampled, of the expected reward, weighted by the probability of occurring.

$$J(\theta) = \int_{\mathcal{T}} p_\theta(\tau) r(\tau) d\tau \tag{4.17}$$

Therefore, we can write the gradient:

$$\nabla_\theta J(\theta) = \int_{\mathcal{T}} \nabla_\theta p_\theta(\tau)\, r(\tau) d\tau \tag{4.18}$$

Using the following identity (the log-derivative trick):

$$\nabla_\theta p_\theta(\tau) p_\theta(\tau) = \nabla_\theta p_\theta(\tau) \tag{4.19}$$

we can rewrite $\nabla_\theta J_\theta$ as:

$$\nabla_\theta J_\theta = \int_{\mathcal{T}} p_\theta(\tau) \nabla \log p_\theta(\tau) r(\tau) d\tau = \mathbb{E}\{p_\theta(\tau) \nabla \log p_\theta(\tau) r(\tau)\} \tag{4.20}$$

Importantly, we can define the probability of any trajectory τ under θ, by the assumption that we dealing with a Markov decision process

$$p_\theta(\tau) = p(x_0) \prod_{t=0}^{T} p(x_{t+1}|x_t, y_t)\, \pi_\theta(y_t|x_t) \tag{4.21}$$

The only parameter in this equation that depends on θ is $\pi_\theta(y_t|x_t)$. This implies that

$$\nabla_\theta \log p_\theta(\tau) = \sum_{t=0}^{H} \nabla \log \pi_\theta(y_t|x_t) \tag{4.22}$$

where H is the horizon we are computing returns over. Hence we can rewrite

$$\nabla_\theta J(\theta) = \mathbb{E}\{\nabla_\theta \log p_\theta(\tau)\ r(\tau)\} \tag{4.23}$$

Although the expectation is intractable we can use a Monte-Carlo approximation to it, thereby deriving a functional form for the derivative, which can be computed together with the gradient provided in chapter 3.

4.2.5 Statistical analysis

All analyses were performed using custom-written scripts in Python 2.7 (Python Software foundation; www.python.org). Above algorithms were implemented in Python and NumPy (Van Der Walt, Colbert, and Varoquaux, 2011). Statistical tests were performed using standard libraries (Hunter, 2007; Oliphant, 2007; Van Der Walt, Colbert, and Varoquaux, 2011; Seabold and Perktold, 2010). All statistical tests were two-tailed. Unless otherwise indicated, errorbars correspond to 2 standard errors of the mean. P-values for correlations were determined using permutation tests.

4.3 Results

Our results will be presented in three parts. First we show, using simulation, that our models of head-fixed behaviour allow us to track the behavioural strategy of head-fixed mice, with single-trial resolution. Second, we apply the developed method to data from head-fixed mice learning to perform a simple operant task and provide a broad overview of the resulting fits. Finally, we further analyse the fits to show that over the course of learning, behavioural strategies are restricted to low-dimensional manifolds in the high dimensional policy space.

4.3.1 Validation on simulated data

Capturing the change in behavioural strategy that occurs in subjects across learning is a difficult problem because of the combination of a priori unknown rates of strategy change and small amounts of data to estimate them with. To overcome these challenges, building on previous work (Bak et al., 2016; Smith et al., 2004b; n.d.) from freely moving animals, we developed state-space generalized linear models (ssGLM) to describe changing head-fixed strategies. ssGLMs are able to fit strategies to data from single trials, despite the small amount of available data, by constraining the structure of the change process. Specifically, our models assume that the parameters defining the behavioural strategy change according to a Gaussian random walk, whose covariance structure is inferred together with the strategies themselves.

We tested the ability of our ssGLM to capture changing behavioural strategies. Since strategies are latent, there is no ground truth data that we could use to directly measure our model against. We therefore attempted to evaluate our quality of our models using simulated data. Two hundred trials of simulated data were generated per learner (n=10) by simulating behaviour on single trials using static policy parameters and updating parameters between trials using Monte-Carlo policy-gradient methods (Williams, 1992) (see subsection 4.2.4 for details). Learners were simulated as responding only to one lick spout. The Monte-Carlo updates together with the noise ensured that the learning trajectories were sufficiently different to robustly test the ability of our algorithm to infer the learning process.

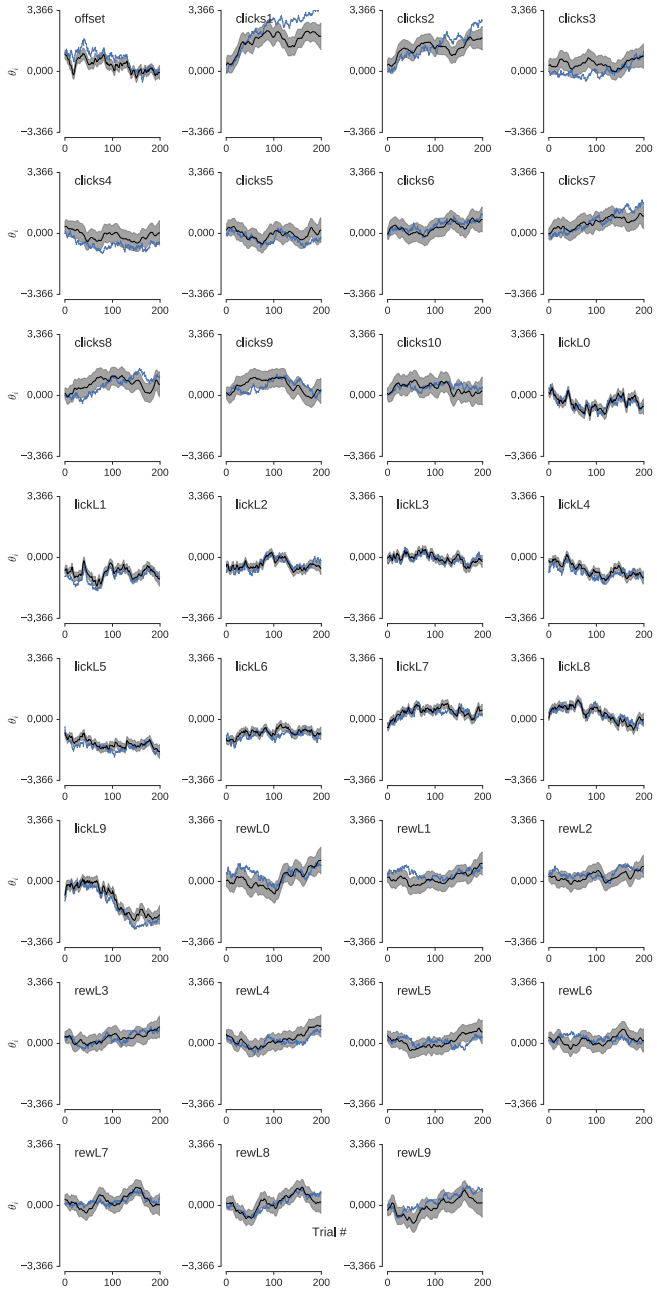


Figure 4.1: see below

Figure 6.3 shows an example of a fit, broken down into individual parameters. Each panel shows the changing value of a single parameter across 200 trials of simulated learning for a single simulated dataset.

Despite the fact that we fitted 32 parameters to the behaviour on each trial, we found the constraints placed on the change process were sufficient to accurately estimate the changing behavioural strategies.

To summarise the quality of fits across all 10 simulated behaviour sessions, we initially exam-

Figure 4.1: Example of all parameter fits to simulated data. Each panel shows the changing value of 1 behavioural parameter. In blue, the ground truth parameter value changing over the course of 200 trials of simulated learning. In black, the estimate from the ssglm. grey shaded regions indicate the 99% confidence intervals based on a Gaussian approximation to the posterior. The indices appended to events at the top of individual panels indicate the time in blocks of 100ms after the event indicated above (e.g. clicks2 is the response to parameter affecting lick probability 200ms after onset of a click)

---

ined the correlation coefficients between the trajectories of ground-truth and estimated parameters. We found that our models were able to explain most of the variance in behavioural strategy that occurred across learning ($\mu(r^2)$=0.68; SD=0.05) (Figure 6.4A). Marginalising across parameters, we observed that fits to all simulated datasets were accurate corroborating the reliability our fits (Figure 6.4 B,C).

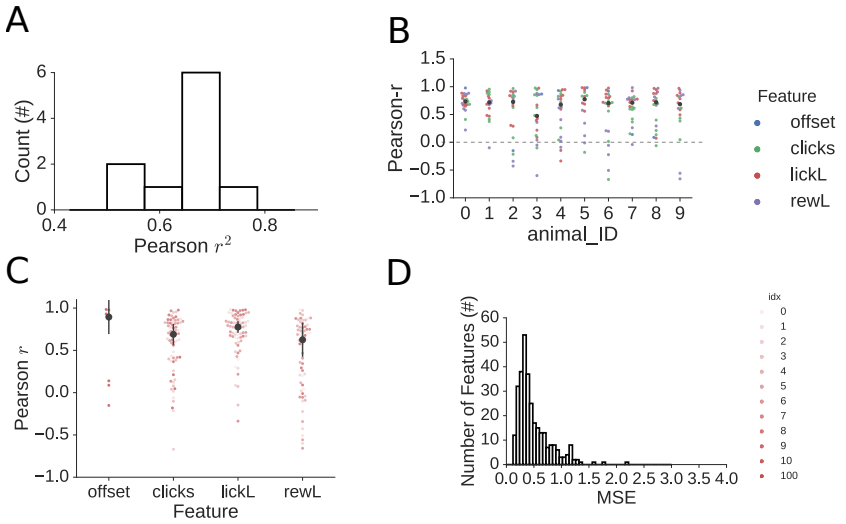


Figure 4.2: **A** Distribution of mean correlation coefficients between estimated and ground truth parameters, weighted according to the fraction of the variance each parameter contributed **B** distribution of correlation coefficients. each dot represents a single parameter from a single animal, colour-coded according to the parameter type. **C** distribution of mean-squared errors between estimated and ground truth parameter values. **D** same as **C** except sorted by parameter, marginalising across animals

---

Although the average quality of fits was high, in some cases we observed weak or even negative correlations between inferred and actual parameter values. However, when we calculated fit quality in terms of mean-squared error (MSE), these outliers were not observed (Figure 6.4C) and MSE were low (median=0.28; SD=0.29). To gain a better understanding of what was causing these poor fits, we visualised parameters and fits that were of poor quality (r¡0.2) (Figure 4.4A). We found that in many cases, fits were accurate in terms of MSE, the reason for the negative correlations were rather that the fluctuations in the simulated parameters, and hence the fit parameters were

small in amplitude. Indeed, many of these parameters fluctuated on scales that were not behaviourally relevant. Consistent with this, we found that the median variance of anitcorrelated parameters was approximately one third of those above 0 (0.08 vs 0.2) (Figure 4.4B). We therefore conclude that our models are able to capture the majority of variance in behavioural strategies.

Finally, we asked whether we could reliably infer changes in the values of parameters from one trial to the next. We therefore, for all parameters across all simulated runs of behaviour, looked at the parameter changes from one trial to the next and compared it against the corresponding changes in the fits. Overall, we found a highly significant, if weak, correlation between the two (r=0.22; p=10^{-10}) (Figure **??**).

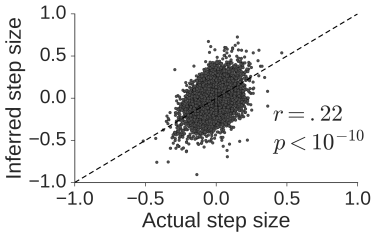


Figure 4.3: Inferred step sizes plotted against actual step sizes in the simulated dataset. Overall there is a weak correlation between the two

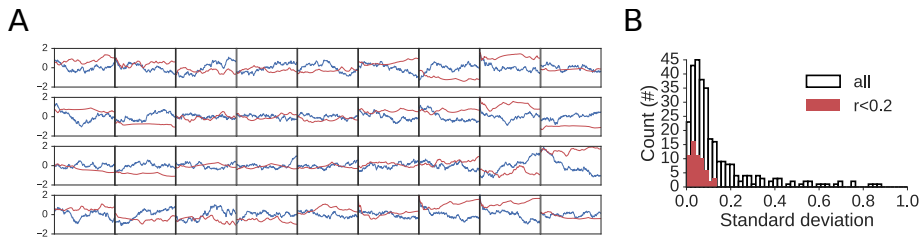


Figure 4.4: Parameter fits identified as poor on the basis of correlation coefficients are accurate in terms of mean-squared error. **A** Randomly generated selection of example fits where actual and inferred parameters are poorly correlated. **B** Distribution of parameter standard deviations for all parameters (white) and poorly inferred parameters (red)

4.3.2 Tracking behavioural change

We proceeded to fit our ssGLMs to the behaviour from head-fixed mice (n=7) as they learned to perform a head-fixed operant task. The task was an auditory detection task in which mice received rewards if they licked during a 1s window after the onset of a 5ms click stimulus presented. We then analysed the results of the fits to better understand changing behavioural strategies.

Figure 6.6 shows a representative example of the policy filters changing over the course of three sessions for an example animal. Each curve shows the effect of the respective event on licking behaviour at various timepoints after it has occurred, for a single trial. Values above 0 indicate that the event increases the probability of licking. Values below 0 indicate that the event decreases the probability of licking at that delay. Here, as we commonly observed, subjects licking following stimulus onset tended to focus on one

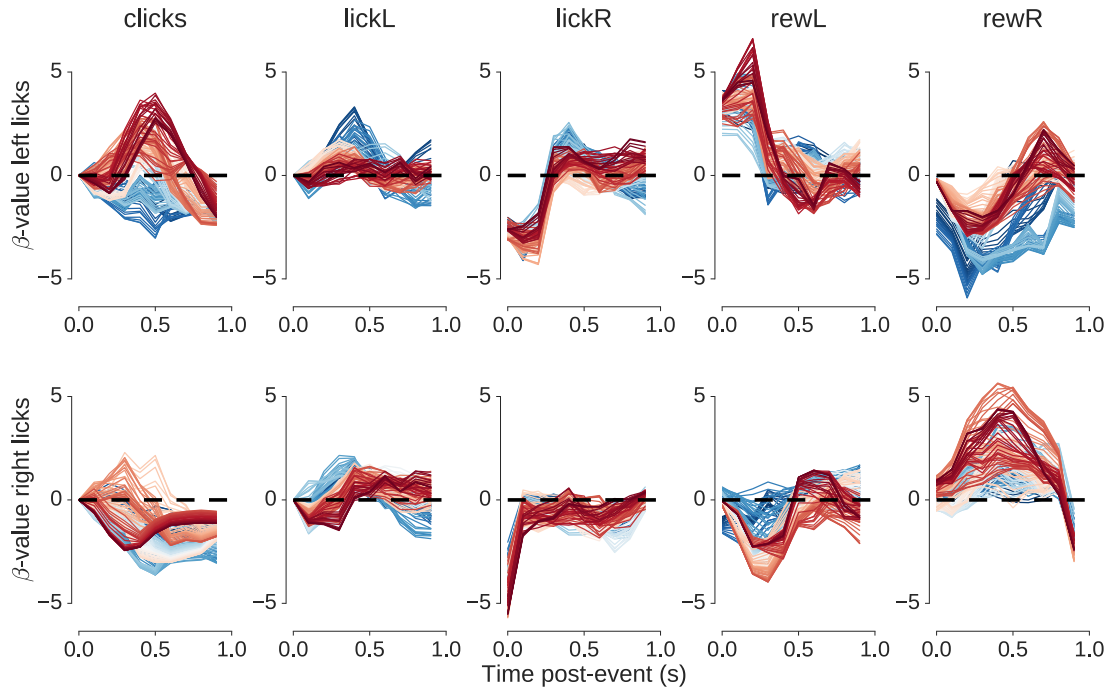
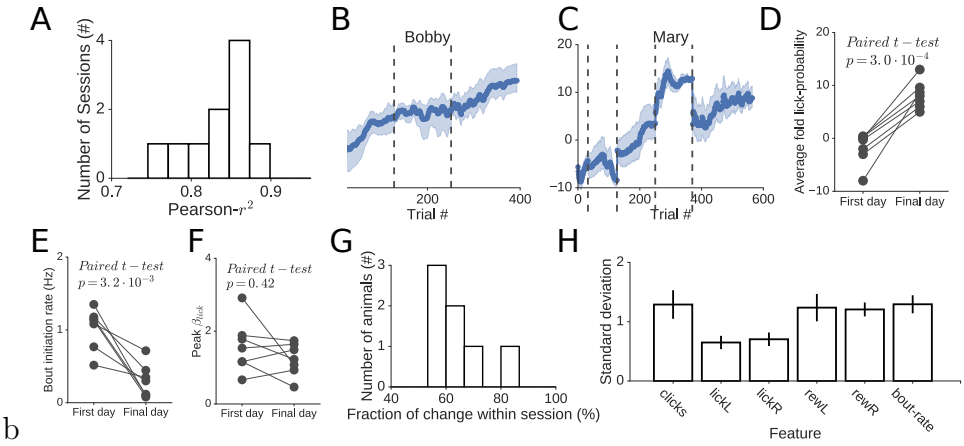


Figure 4.5: example of a set of policy filters fitted to left licks and to right licks over the course of learning for an example animal. Dashed line is at 0. lickL: lick at left spout; lickR: lick at right spout; rewL: reward to left spout; rewR: reward to right spout

---

of the two spouts, in this case the left one. As expected, over learning, stimulus driven behaviour changes: initially, licking is inhibited by stimuli (which is commonly observed for salient stimuli), but gradually increases over the course of learning. Other parameters of the behaviour also change. For instance, the autocorrelation in licking behaviour decreased.

The main behavioural feature typically examined in operant paradigms is sensitivity of the subject's behaviour to the presentation of stimuli (Green



b

Figure 4.6: **A** reconstruction quality of fold change in lick probability pre vs post stimulus from the simulated data **B** Fold probability of licking pre vs-post stimulus over the course of learning for two example animal. **C** same as **B** for different example animal. **D** Mean fold-lick-probability pre vs post-learning for all animal's whose data were fitted. **E** Across learning, bout-initiation rates (see chapter 3) decrease **F** Across learning, there is no change in the peak of the lick autocorrelation kernel **G** Distribution of fractional change of stimulus sensitivity between vs within sessions. **H** Standard-deviation of the distribution of parameter values over the learning process for each session. Error bars are $\pm$ 2 s.e.m

---

and Swets, 1966; Carandini and Churchland, 2013; Busse et al., 2011; Guo, Hires, et al., 2014). We therefore asked whether our model could recapitulate this sensitivity on a trial-by-trial basis. We quantified the sensitivity of to the stimulus as expected value (in the statistical sense) of the ratio of initiating a lick bout pre and post stimulus licking. For clarity, we linearised this metric around 0.

For instance, if the mouse's licking was strongly inhibited following stimulus onset the ratio might be -10, which would imply that the animal is 10

times less likely to lick pre-stimulus, compared with post stimulus. This ratio was calculated using the offset of stimulus parameters of our model. We first validated on our simulated data-sets that we are accurately able to infer this metric, and found we were able to do so (mean $r^2 = 0.86$; SD=0.06) (Figure 4.6A).

Figure 4.6 B,C show two the learning trajectories of two example animals. Two features are striking about these curves. As expected, we see an increase that was consistent across the population (Figure 4.6 D). Strikingly, we often observed continuity in subject's stimulus sensitivity across days although models were fitted to data from individual days, the stimulus sensitivity is not constrained to vary smoothly across days. Nonetheless, in both examples shown, we observed near-continuity in subject's stimulus sensitivity.

We then proceeded to quantify this phenomenon more robustly. Although the learning process was not strictly monotonic, we quantified how much change took place within as opposed to between days. By considering between day changes to be the change from the 20th to last trial of a given day to the 20th trial of the next day (20 trials either side were included to prevent contamination by factors such as loss of motivation or daily acclimatisation). Within session changes were those that occurred between the 20th and the 20th to last trial within a session. Behavioural sessions typically consisted of more than two hundred trials. Quantifying behaviour this way, we find that the majority of changes occur within sessions (Figure 4.6G).

In addition to changes in stimulus sensitivity, we also observed dynamic

changes in the behavioural strategy related to spontaneous licking behaviour (Figure 4.6D,E,F,G). Most strikingly, while stimulus-evoked licking increased, spontaneous bout-rate decreased (Figure 4.6E) indicating that mice switched from a sampling-based to a stimulus-guided strategy for obtaining reward.

To quantify changes in the parameters of behaviour more broadly, we calculated the standard deviation of each parameter in a group and took the mean of the standard deviations across parameters belonging to a group (e.g. all parameters relating to click onset). We found that the variation was approximately equal for stimulus, reward and bout-initiation parameters but reduced for the structure of the autocorrelation in licking. Thus, over the course of learning we observed changes in all major groups of behavioural parameters (Figure 4.6H).

4.3.3 Mice use structure inherent to task to quickly reach high performance levels

Without a priori knowledge of the structure of the task, the optimisation problem could be quite difficult to solve, due to its high dimensionality. However, as mentioned above, the structure in the task makes it effectively a one-or two-dimensional optimisation problem: how much should the mouse lick when no stimulus is present and how much should the mouse lick post-stimulus. We reasoned that if animals make use of this structure, either through structured representations of the world, or through efficient learning

algorithms, those parameters that define that higher-order structure would fluctuate together over the course of learning. To test this idea, we performed principal components analysis (PCA) (all analyses were repeated using factor analysis with qualitatively similar conclusions; data not shown) on the fluctuating strategies inferred by our models, to check whether mouse strategies are restricted to a low dimensional manifold of the high dimensional strategy space. If this structure is used, we would expect the dimensionality of the manifold to be low, hence we would expect that most of the variance of the PCA should be in the first few components.

We *a priori* defined 80% of the variance explained as capturing the majority of the variance in behavioural strategy of individual animals. We found that in all cases, the majority of the variance was explained by the few principal components ($\mu = 5$; $\sigma = 2$) (Figure 4.7A), which is 20-fold small than the 100-dimensional space our model was fit in, consistent with the use of structure in the task. To then test whether the fluctuations across all mice were in a similar manifold, we performed PCA on the concatenated behavioural strategies of all mice simultaneously (Figure 4.7B). Although the space now was larger (20 components explain ¿80% of the variance) it is still five times smaller than the original space. However, the fact that the space of strategies across animals is larger, suggests that individual mice change their behaviour along different axes, as no small number of axes can capture the combined variance across mice.

Finally, we were interested to view the axes along which behaviour changed

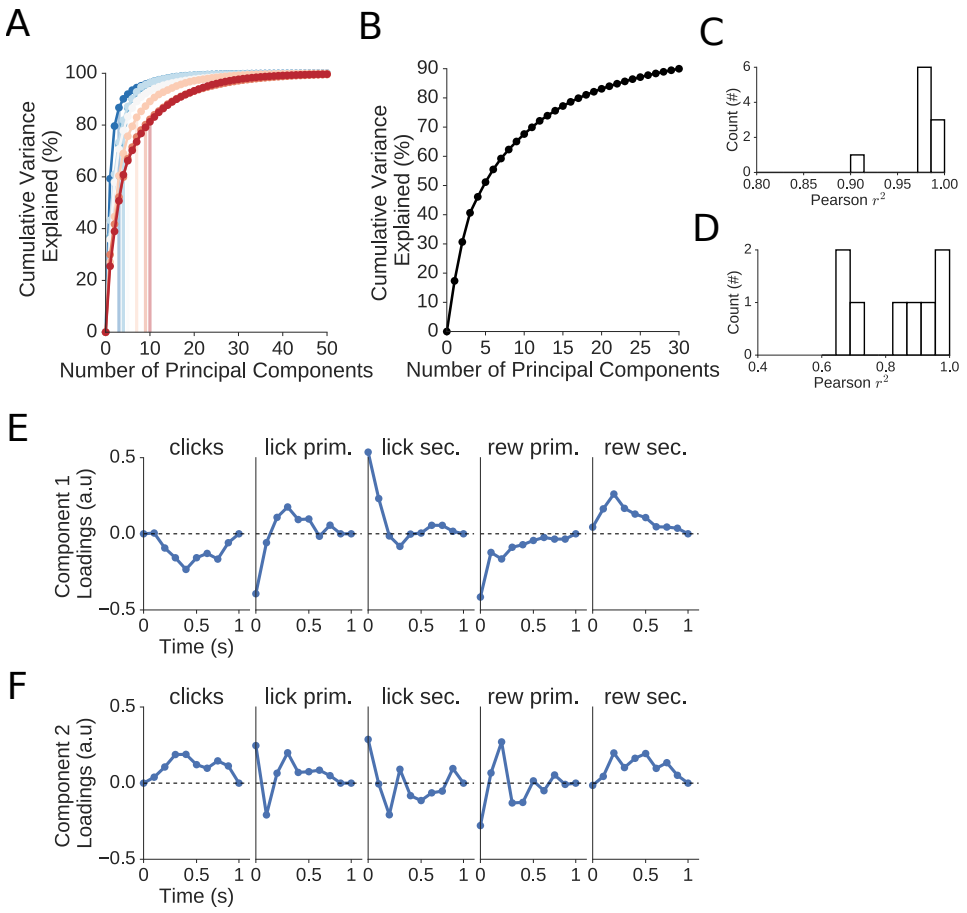


Figure 4.7: **A** Cumulative variance explained as a function of number of principal components included for each animal in our dataset. Vertical lines indicate number of principal components required to explain 80% of the variance. **B** cumulative variance explained as a function of the number of principal components included when performing PCA on the dataset of all animals. **C** accuracy in the reconstruction, based on simulated data, of the loadings of the principal components for the first principal component **D** same as **C** except for the second principal component. **E** Loadings of the first principal component separated by feature **F** same as **E** except for the second principal component.

(i.e. the loadings to the principal components). To verify that we were able to accurately reconstruct the axes along which behaviour changed, we tested the accuracy of our reconstructions using simulated data. To do so, we quantified the correlation coefficient of the loadings to each principal component based on the stimulated data. We *a priori* defined 60% of the explained variance as a cutoff below which we would consider the reconstruction inaccurate. With these criteria in place, we found that we could we could only accurately reconstruct the first principal component (r^2_{PC1}=0.91; r^2_{PC2}=0.66; r^2_{PC3}=0.42)) (Figure 4.7 C,D).

We then performed PCA on the behavioural trajectories of all animals and took the mean of the first two principal components. Despite the fact that the joint behaviour occupies a high dimensional space, we found substantial structure in the behaviour of individual mice (Figure 4.7E,F). Thus, our analysis suggests that mice efficiently learn by optimising their behaviour in a low dimensional subspace of the high-dimensional policy space.

4.4 Discussion

We developed Bayesian state-space methods for tracking changing behavioural strategies in head-fixed mice. We validated our methods on simulated data, showing that we are able to accurately reconstruct the latent, changing behavioural strategies of head-fixed mice during the performance of a simple operant task. This is true both with respect to stimulus-driven licking. Ap-

plying these methods to data from head-fixed mice learning of a sensory detection task, we showed that over the course of learning there are consistent changes in the behaviour with respect to both stimulus-evoked and stimulus-independent licking. Strikingly, although the learning problem is a priori high dimensional, our analysis showed that mice exploit the intrinsically low dimensional structure of the environment to efficiently optimise their behaviour. Together, these findings validate and show the utility of our methods for furthering understanding the structure of behaviour.

The majority of computational work on learning draws on the reinforcement learning paradigm (Sutton and Barto, 1998; Ito and Doya, 2009; Gershman, Norman, and Niv, 2015; Maia, 2009; Niv et al., 2015; Bathellier, Poh, et al., 2013). This work has largely focussed on comparing the ability of different learning rules to describe the learning process (**Leon2017**; Niv et al., 2015; Ito and Doya, 2009; Bathellier, Poh, et al., 2013). As such, it characterises behaviour from a top-down, hypothesis driven perspective. While this work led to immense progress (see Gershman and Daw, 2017 for a recent review), the use of specific models of behavioural change can artifactually induce or remove structure from the behavioural change process. This can occur because the prescription of the learning rule induces a prior on the structure of that change (Bishop, 2006; Sarkka, 2013). Additionally, validation in the application of these methods is by necessity comparative as strategies underlying behaviour are latent. This, together with the fact that most models of choice are stochastic (Sutton and Barto, 1998), means

that the ability of these models to describe behavioural data is difficult to validate unequivocally.

Our work complements this strand of investigation. By analysing behavioural change using descriptive stochastic state-space methods, we are able to describe the latent structure of behavioural strategies in a largely hypothesis free manner (Chen, 2015; Sarkka, 2013). This is complementary to the hypothesis driven reinforcement learning approach in three ways. Firstly, our analysis can serve as a reference model for reinforcement learning based models to compare against. Secondly, these models can be applied to identify animal's strategies, on a single trial basis in cases where learning rules might be unknown, e.g. during initial task learning. Finally, these models can be combined with other analytical approaches to identify structure in the learning process, which can then form the basis of hypothesis-driven investigation. Thus, our bottom-up approach to describing holds substantial promise.

Previous work has applied learning-rule-agnostic state-space models to operant task learning in experimental animals (Smith et al., 2004b; Prerau et al., 2009; Bak et al., 2016; Smith et al., n.d.). Two main conclusions have been derived from this work. Firstly, stimulus driven behaviour gradually emerges and improves with increased experience. Secondly, in addition to optimal stimulus-driven strategies, subjects engage in other strategies and show a variety of biases in their behaviour (; Bak et al., 2016). Further, over the course of learning, these strategies are dynamic and change non-monotonically (Smith et al., n.d.). Our work is consistent with this. Thus,

we show that over the course of learning an operant conditioning task, mice gradually switch from sampling-based strategies to more efficient stimulus-driven strategies. Additionally, we show that stimulus-unrelated strategies change dynamically over the course of single sessions.

Our work also extends both the reinforcement learning and state-space approaches. In particular, previous work has exclusively analysed behaviour at the choice point (**Bathelier**; **Leong2016**; 2004, b; Prerau et al., 2009; Bak et al., 2016; Smith et al., n.d.; Niv et al., 2015; Ito and Doya, 2009; Maia, 2009). However, it is increasingly clear that an animal's reward-guided behaviour is influenced by factors beyond discrete decisions in the task epoch (Skinner, 1948b; Kawai et al., 2015b; Krakauer et al., 2017).

The head-fixed paradigm provides a unique opportunity for exploring stimulus-unrelated behaviour and the way it changes. Finding a natural analytical structure for describing and analysing mice' behaviour in freely moving conditions is immensely challenging **Gomez-Marin2015**; Gomez-Marin, Paton, et al., 2014; Wiltschko et al., 2015. In contrast, this problem is greatly alleviated in the head-fixed preparation. The low cost of licking together with physical restraint lead to mice continuously interacting with the response apparatus and this interaction provides a large statistically homogenous dataset with which to characterise behaviour.

We have previously argued, in chapter 3, that in the head-fixed preparation, reward-guided, stimulus-unrelated behaviour takes the form of stochastic sampling of lick-ports. This takes place both with and outside the task

epoch. Describing the change process is *a priori* challenging due to the high dimensionality of the parameter estimation problem (Chen, 2015; Bishop, 2006). Nonetheless, we show that applying the state-space formalism allows us to accurately describe this process.

On the other hand, the head-fixed preparation has several disadvantages. Most notably, head-fixation is potentially associated with substantial stress for the animals (Weixin et al., 2017), which may qualitatively affect behaviour. Another problem affecting its use as a quantitative behaviour tool is more practical. It is well known that the distance of lick-ports from mice' snouts plays a crucial role in influencing the structure of the behaviour (Guo, Hires, et al., 2014). Without careful control, this influence might dominate measured changes in behavioural strategies across days. However, at least with respect to stimulus sensitivity, we typically observed consistency across days, indicating that this is not a major issue for the data presented here.

Finally, by analysing the fits resulting from our work, we identified that learning takes place on low-dimensional manifolds of the high-dimensional policy space. This low-dimensionality is a hallmark of efficient learning algorithms (Salimans et al., 2017; Kemp and Tenenbaum, 2009; Nesterov and Spokoiny, 2017). We have not been able to differentiate whether the low-dimensional structure derives from an efficient learning rule or from the internal structure of the world. Though we hope to clarify this with future investigation.

4.4.1 Conclusion

We have developed Bayesian-state space models for describing changing behaviour. We have shown that we are able to accurately infer changing strategies of mice with single trial-solution. Through analysing these fits we have shown that even though mice behaviour is *a priori* high dimensional, they efficiently optimise their behaviour on low-dimensional manifolds embedded in this space.

Chapter 5

Neural Correlates of Behaviour

5.1 Introduction

Perceptual decision making is an experimentally tractable cognitive behaviour (Carandini and Churchland, 2013; Parker and Newsome, 1998; Park, Meister, et al., 2014). As such it can provide insights into the mechanisms by which neural circuits give rise to function (Yuste, 2008; Carandini, 2012). Auditory cortex is a major brain region involved in sound processing (King and Nelken, 2009; Nelken, 2004; Schnupp, Nelken, and King, 2011; Nelken, Rotman, and Yosef, 1999) sound-guided behaviour (Whitfield, 1985; Winer and Schreiner, 2011; King, Bajo, et al., 2007). Previous work has demonstrated that transient and chronic lesions to auditory cortex impair performance on a variety of auditory perceptual tasks (Kuchibhotla et al., 2016a; Talwar, Musial, and Gerstein, 2001; Dykstra et al., 2012; Carcea et al., 2017;

Li, Liao, et al., 2017; Guo, Clause, et al., 2017; Nodal et al., 2010; Bajo et al., 2010) and stimulation studies causally demonstrate the involvement of auditory cortex in perceptual decision making (Znamenskiy and Zador, 2013a; Xiong, Znamenskiy, and Zador, 2015; Guo, Chambers, et al., n.d.).

Recording the activity of cortical neurons has revealed both their basic response properties and also how these are modulated by task performance. At the level of single neurons, auditory cortex has been shown to encode diverse features of task performance (Kuchibhotla et al., 2016a; Bizley, Walker, Nodal, et al., 2013; Jaramillo, Borges, and Zador, 2014; Carcea et al., 2017; Niwa et al., 2012; Downer, Niwa, and Sutter, 2015) and it has been shown that receptive fields structure can adaptively change to meet task demands (Fritz, David, et al., 2010; Fritz, Shamma, et al., 2003; Fritz, 2005). At a population level, it has been demonstrated that engagement in a sensory discrimination task leads to the suppression of cortical activity (Otazu et al., 2009).

Recent work suggests that many neuronal computations occur at the level of neural populations (Carandini, 2012; Harris and Mrsic-Flogel, 2013; Petersen and Crochet, 2013; Elsayed and Cunningham, 2017). This is both true in terms of representing sensory information, where neurons have been shown to act in concert as ensembles (Gentner and Margoliash, 2003; Bathellier, Ushakova, and Rumpel, 2012; Niessing and Friedrich, 2010). It also applies to the neural circuits that drive behaviour where it has been shown that attention (Ruff and Cohen, 2014; Herrero et al., 2008; Rabi-

nowitz et al., 2015; Harris and Thiele, 2011) as well as task demands (Cohen and Newsome, 2008) change the structure of correlations between neurons. However, the synaptic or circuit mechanisms by which these computations are implemented is still an area of active investigation (Carandini, 2012; Harris and Mrsic-Flogel, 2013).

Layer 2/3 is, in terms of connectivity, situated between the main thalamorecipient layers of the neocortex (layer 4) and its canonical main output (layer 5) and thus a major potential site of neocortical computation (Douglas and Martin, 2004). While the functional properties of the neocortical layer 2/3 are reasonably well understood, its functional role in sensory processing and behaviour remains unclear (Petersen and Crochet, 2013). One striking property of layer 2/3 is its dense, recurrent connectivity (Oviedo et al., 2010; Dantzker and Callaway, 2000; Bureau, Von Paul, and Svoboda, 2006; Levy and Reyes, 2012; Barbour and Callaway, 2008; Watkins, Kao, and Kanold, 2014), which could support local computations.

However, while the activity of neurons in layer 2/3 have been extensively studied during behaviour (see Petersen and Crochet, 2013 for a review), its properties have been largely analysed from the perspective of single neuron computations (Poort et al., 2015; Huber, Gutnisky, et al., 2012; Petersen and Crochet, 2013; Carandini, 2012; Zhang, Xu, Kamigaki, et al., 2014; Allen et al., 2017; Yamashita and Petersen, 2016; Barlow, 1972; Parker and Newsome, 1998).

Two photon imaging (Denk, Strickler, and Webb, 1990) is a powerful

tool for studying the neural mechanisms underlying perceptual decision making. It affords the ability to simultaneously monitor hundreds of genetically-defined neurons in parallel in awake behaving animals (Svoboda and Yasuda, 2006). It thereby enables the study of network properties (Mishchencko, Vogelstein, and Paninski, 2011; Cotton et al., 2013; Soudry et al., 2015) in addition to the characterisation of activity in single neurons. This method has been used by several groups to investigate the functional organisation of the auditory corte (Vasquez-Lopez et al., 2017; Bathellier, Ushakova, and Rumpel, 2012; Rothschild, Nelken, and Mizrahi, 2010; Chen, Leischner, et al., 2011; Barnstedt et al., 2015; Bandyopadhyay, Shamma, and Kanold, 2010). Additionally, recent years have seen the development of powerful statistical models capable of describing both the response properties and functional interactions in neural circuits (Paninski, 2004; Pillow, Shlens, et al., 2008).

Here we combine these recent developments to further our understanding of how activity in layer 2/3 of mouse auditory cortex drives behaviour. We imaged layer 2/3 of mouse auditory cortex while mice' perform a stimulus detection task. We apply a previously developed statical framework (Paninski, 2004; Pillow, Shlens, et al., 2008; Park, Meister, et al., 2014) for studying neural populations to fully leverage the multineuronal data provided by two-photon calcium imaging. Doing so, we show that neurons in auditory cortex encode a plethora of sensory and non-sensory features of the environment. Further, we show that neurons representing similar features tend to form

functional subnetworks. Finally, we provide evidence that these functional subnetworks act in concert to drive sound-guided behaviour.

5.2 Methods

5.2.1 Animals and Behaviour

For detailed animal welfare considerations and surgical procedures, see chapter 2. For details of the experimental apparatus, see chapter 4. The only difference to this experimental apparatus was that solenoids for reward delivery and lick detectors were extensively sound shielded such that sound produced by this apparatus was below the noise floor (attenuation > 70dB SPL).

Thirteen mice (C57B6.129S-Gt(ROSA)26Sor$^{\text{tm95.1(CAG-GCaMP6f)Hze}}$ [Jackson Laboratories; stock number 024105] x C57B6.Cg-Tg(Camk2a-cre)T29-1Stl/J [Jackson Laboratories; stock number 005359]), aged 5-7 weeks, expressing the calcium sensor GCaMP6f in excitatory neurons, were surgically implanted with a 3mm cranial window over the auditory cortex (Vasquez-Lopez et al., 2017) and a headpost for head-fixation. After a two week recovery period, during which animals were habituated to handling, animals were trained over the course of approximately 4 days to perform a simple click detection task (see chapter 3 and chapter 4 for details) while undergoing chronic two-photon imaging.

Thereafter animals underwent further behavioural testing wherein stimuli

were presented at reduced levels. Levels were heuristically chosen on the basis of session-by-session performance, determined in an initial behavioural block. Levels were chosen to cover the slope of animal's psychometric functions. The maximal level was fixed to 80 dB SPL and was presented in all sessions.

5.2.2 Window surgery

Window surgeries were performed using identical protocols to chapter 2, except that windows were unitlaterally implanted over the right auditory cortex.

5.2.3 Two-photon imaging

Two photon imaging was performed as described previously (; Barnstedt et al., 2015). Image acquisition was carried out using a commercially available two-photon laser-scanning system (B-Scope; Thorlabs). A SpectraPhysics Mai-Tai eHP laser fitted with a DeepSee prechirp unit (70fs pulse width, 80MHz repetition rate) provided the laser beam for two photon excitation. The beam was directed into a Conoptics modulator and then through the objective (16× 0.8NA water immersion objective; Nikon). The beam was scanned across the brain using an 8kHz resonance scanner (X) and a galvanometric mirror (Y). The resonance scanner was used in bidirectional mode, enabling aquisition of 512 × 512 pixels at a frame-rate of approximately 30Hz. Emitted photons were filtered (525/50) and collected and amplified

by GaAsP photomultiplier tubes (Hamamatsu). ScanImage was used to acquire data and control the microscope.

5.2.4 Analysis

5.2.4.1 Behaviour

Behavioural data were acquired as described previously chapter 4. We categorised trials as hit trials if mice licked within a 1s window following stimulus onset. Reaction times were calculated for stimulus levels including more than 20 hit trials per session. Responses faster than 100ms were not included in the calculation of reaction time data.

5.2.4.2 Two-Photon Data Preprocessing

All data were analysed using custom scripts written in Python 2.7. Two photon imaging analysis was performed using a custom-written, publically available toolbox www.github.com/yves-weissenberger/Multiphoton-Toolbox. Briefly, image-stacks were imported and converted to HDF5. Thereafter, image-stacks were motion registered using an efficient subpixel registration algorithm (manification factor 10) (Guizar-Sicairos, Thurman, and Fienup, 2008). Thereafter, image segmentation to identify neurons as regions of interest was performed by manual inspection of two-photon videos using a custom-written graphical user interface. Traces were extracted and neuropil corrected by selecting a 20 µm square region surrounding the region of interest and sub-

tracting the weighted (0.7) neuropil trace (Chen, Wardill, et al., 2013; Kerlin et al., 2010). We performed spike inference using the standard model from the c2s toolbox (Theis et al., 2016; Akerboom et al., 2012; Cotton et al., 2013; Chen, Wardill, et al., 2013). All analysis of neural data was performed on inferred spike trains, as they more accurately reflect underlying neuronal spiking than the calcium signal (Theis et al., 2016).

5.2.4.3 Triggered averages

Triggered averages were calculated by averaging neural responses to repeated events in a single session. Neurons were deemed responsive to an event type based on their responses in a ∼200ms window following event-onset. The average response over this window was compared to activity averaged over a 200ms window preceeding event-onset using a Wilcoxon Signed Rank Test ($\alpha = 0.001$).

5.2.4.4 Choice Probabilities

Choice probabilites were calculated as previously described (Britten et al., 1996). P-values of choice probabilities were determined by permutation testing (1000 shuffles).

5.2.4.5 Generalized Linear Model fitting

To determine what factors influenced neuronal activity, we fitted poisson distributed generalized linear models (GLM), with exponential nonlinearity

(Paninski, 2004) to data from individual neurons (; Pillow, Ahmadian, and Paninski, 2010; Runyan et al., 2017). Regressors used in model fitting related to both external variables (stimuli at different levels, mouse choices, licks and rewards) and, in a subset of fits (see section 5.3), the activity of other neurons. External events were binned into windows matching the acquisition rate of the two-photon microscope. Since external events can have prolonged influences on neuronal activity, we included, for each event, time lagged regressors (20 time bins of ∼30ms) for each external event. The influence of other neurons on a given neurons inferred spike rate was fixed to be instantaneous, given our temporal resolution (i.e. influence was 1 time-bin in duration). This yields the following model, predicting the spiking of neuron i at time t:

$$\begin{aligned}\lambda_i(t) = exp\big(&\beta^{\mathbf{i}}_{\mathbf{click}} \cdot \mathbf{K}x_{click}(t) + \\ &\beta^{\mathbf{i}}_{\mathbf{rew}} \cdot \mathbf{K}x_{rew}(t) + \\ &\beta^{\mathbf{i}}_{\mathbf{lickL}} \cdot \mathbf{K}x_{lickL}(t) + \\ &\beta^{\mathbf{i}}_{\mathbf{lickR}} \cdot \mathbf{K}x_{lickR}(t) + \\ &\beta^{\mathbf{i}}_{\mathbf{decision}} \cdot \mathbf{K}x_{decision}(t) + \\ &\beta^{\mathbf{i}}_{\mathbf{coupling}} \cdot \mathbf{Y}(t) + \\ &\beta_0\big)\end{aligned} \tag{5.1}$$

where β_{event} are vectors of free parameters of the model relating to a par-

ticular covariate; $\mathbf{K}$ is a toeplitz matrix projecting the binary vector $x_{event}(t)$ (except in the case of stimuli where the values of $_{click}$ vary continuously with the level of stimuli) onto β_{event} (in effect implementing the time lagged effect of regressors). $\beta_{decision}$ is a binary vector set to 1 on hit trials and 0 on miss trials. β_0 is the neurons baseline firing rate. Finally, $\mathbf{Y}$ is the vector of population activity (i.e. a 1-D vector containing the inferred spike rates of all simultaneously neurons at time-step t), excluding neuron i at the time-step under question.

GLMs were fit using a custom toolbox via Newtons method (chapter 3) using known derivatives (Paninski, 2004; McCullagh and Nelder, 1989). GLMs were regularised using an L2 penalty term of 1 for external regressors and a penalty term of 10 for coupling parameters. The larger penalty for coupling parameters ensured that models preferentially attributed variance to external events and is consistent with previous applications of these models to neuronal populations (Runyan et al., 2017; Pillow, Shlens, et al., 2008). Regularisation parameters were determined via cross-validation on training datasets.

5.2.4.6 Analysis of Generalized Linear Model Parameters

For the analysis of parameters, both those related to external event and coupling parameters, we thresholded parameters using p-values calculated using Wald's test (Pawitan, 2013) with significance threshold $\alpha = 0.05$. To ensure that our results did not emerge artifactually from our choice of threshold,

all relevant analyses were repeated with $\alpha = \{0.01, 0.005, 0.001\}$ and led to qualitatively identical conclusions (data not shown).

Stimulus drive was defined, for each neuron individually, as the sum of β weights (over time) relating the occurrence of a stimulus to activity in that neuron.

For analysis of coupling between neurons, we considered coupling from two perspectives: incoming coupling and outgoing coupling. In both cases, coupling was calculated for each neuron individually.

Our measure of incoming coupling, termed target coupling, was simply the thresholded sum of β weights relating the instantaneous activity of all other neurons to a given neuron.

$$target\ coupling = \sum_{j=1}^{N-1} \left(\beta^i_{coupling}(j)\right) \tag{5.2}$$

where N is the number of neurons in the population and $\beta^i_{coupling}(j)$ denotes the directional coupling from neuron j to neuron i.

Outgoing coupling, termed target coupling, was calculated as the thresholded sum of β weights, acoss all other neurons, relating their activity to activity in a given neuron.

$$source\ coupling = \sum_{i=1}^{N-1} \left(\beta^i_{coupling}(j)\right) \tag{5.3}$$

where now instead of summing over coupling *to* to a given neuron from other neurons (i.e. sum over j) we sum over coupling terms *from* a given

neuron to other neurons (i.e. sum over i).

5.2.4.7 Choice Decoding

Choice decoding was performed analytically as described in (Park, Meister, et al., 2014). In particular, choices were decoded by evaluating the log-likelihood ratio of the population activity under the two response options (lick or no-lick), given all other variables (including spiking of other neurons). When quantifying the performance our population decoders, we used information in a 270ms window following onset of stimuli. This window was chosen as it represents the median reaction time and excludes the majority of the animals' responses to stimuli, minimising the effects of licking and reward consumption on our estimates of choice related activity.

5.2.5 Statistical Analysis

Unless otherwise indicated, nonparametric statistical analyses were performed. Statistical tests were performed using standard Python extension modules Scipy (Oliphant, 2007) and Statsmodels (Seabold and Perktold, 2010). All reported p-values are two-tailed.

5.3 Results

We performed two-photon imaging of excitatory neurons in layer 2/3 (range: 160-240µm below cortical surface) of mouse (n=4 mice; n=8 sessions; n=2095 neurons) auditory cortex while mice performed a sound detection task, we demonstrated to be dependent on auditory cortex in chapter 2 (Figure 5.1A-D). We used the method of constant stimuli, presenting stimuli at different levels. Mice responded at reaction times consistent with previous work (O'Connor et al., 2010) (Figure 5.1E) and performed many trials per session (μ=366; s.e.m 30 trials), enabling accurate estimation of behavioural sensitivity (Figure 5.1C). Additionally, we found that stimulus intensity modulated reaction time, such that quieter stimuli were detected at longer reaction times (Figure 5.1D,E), and hence also with performance.

Previous work has shown that the location of auditory cortex, and in particular specific areas of the auditory cortex are unreliably defined stereotaxically and on the basis of surface vasculature (Stiebler et al., 1997) (though see Tsukano et al., 2016). However, the primary region of auditory cortex (A1) can be identified by its caudo-rostral tonotopic gradient which reverses at the boundary to the anterior auditory field (AAF) (; Guo, Chambers, et al., n.d.; Vasquez-Lopez et al., 2017; Bandyopadhyay, Shamma, and Kanold, 2010; Rothschild, Nelken, and Mizrahi, 2010).

Thus, prior to behavioural imaging experiments, we performed cortical two photon imaging (Figure 5.2A) under anaesthesia, to identify the primary

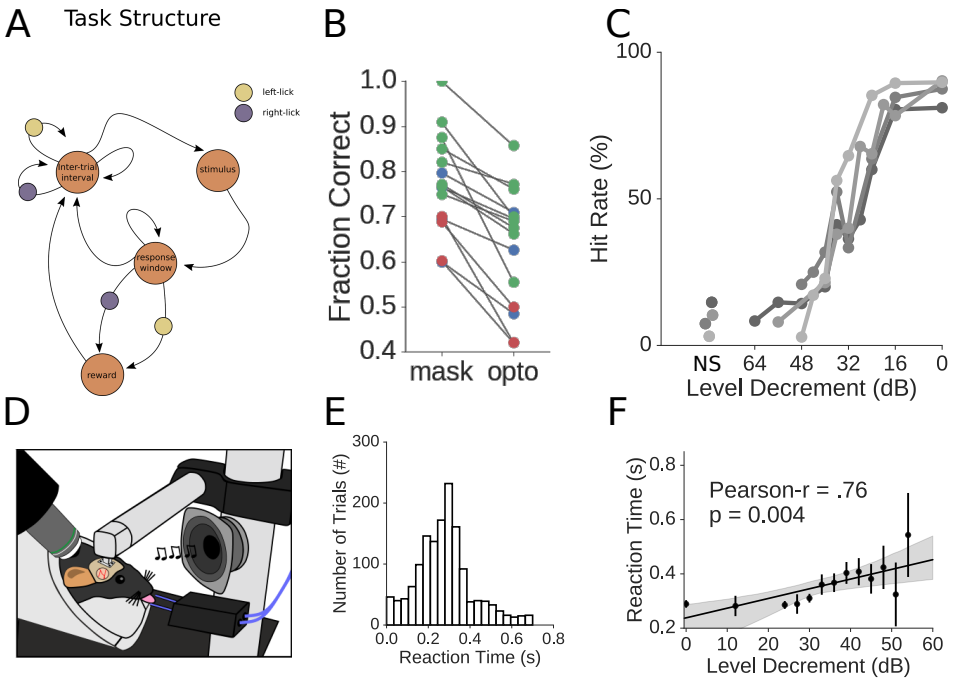


Figure 5.1: Summary of behavioural experiment **A** Schematic showing the state structure of our behavioural task. **B** Data from chapter 2 showing that optogenetic inhibition of auditory cortex impairs performance on our behavioural task. **C** Four example psychometric functions demonstating level-dependence of sound-driven behaviour. NS denotes hit rates where no stimulus was presented (catch trials) **D** Schematic of mouse performing our behavioural task **E** Distribution of reaction times **F** Linear correlation between stimulus level and mean reaction time; error-bars are 2 s.e.m.

cortex. We presented logarithmically spaced pure tones ($\frac{1}{3}$ octave spacing; 60dB SPL) and characterised neural frequency tuning curves (Figure 5.2B). We then attempted to localise auditory cortex by mapping the distribution of best-frequencies onto the cortical surface. In most experiments, we were unable to unambiguously identify auditory cortex using defining as criteria the presence of a caudorostral gradient and its reversal at its high frequency edge. Despite the fact that we were able to reliably identify a dominant caudorostral gradient (Figure 5.2C-E), we were not able to demonstrate its reversal. Thus, though our data suggests that we were imaging in A1, we lacked the certainty which could have been provided by the A1/AAF gradient reversal. Hence we refer to imaged the imaged region as putative primary cortex (pA1).

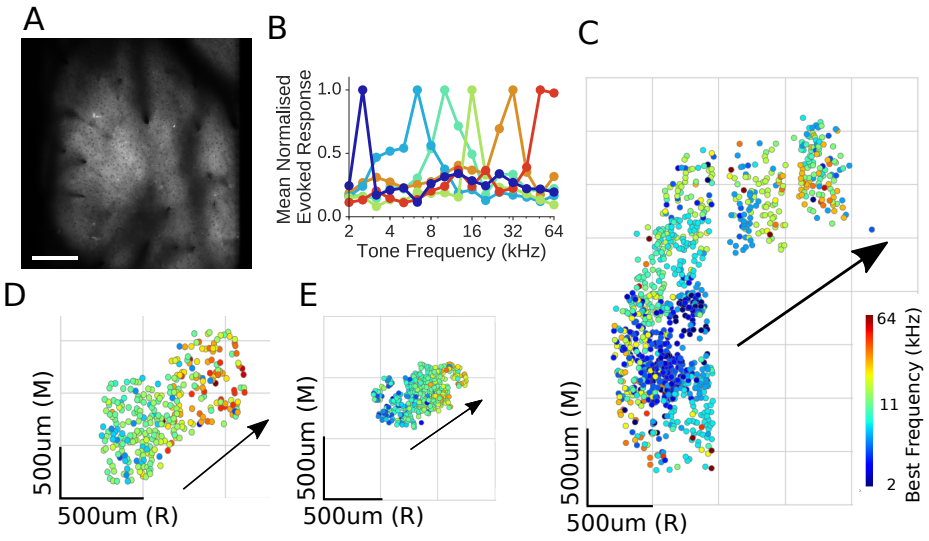


Figure 5.2: Functional identification of auditory cortex. **A** Example imaging region. Scale bar denotes 200µm **B** Tuning curves of 6 example neurons as a function of stimulus frequency. **C-E** Example tonotopic maps acquired by mapping best frequencies onto neural location. Only neurons with significant frequency tuning (one-way ANOVA $p < 0.001$ are included in the figure. M: medial; R: rostral

5.3.1 Model free analysis of neural activity

Since our optogenetic perturbation experiments (Figure 5.1B), together with previous work (Guo, Chambers, et al., n.d.; Li, Liao, et al., 2017; Talwar, Musial, and Gerstein, 2001) suggest that auditory cortex causally contributes to enactment of sound-driven behaviour, we sought to gain a better understanding of its functional role by studying the response properties of auditory cortical neurons during behaviour. To this end we performed 2-photon imaging in layer 2/3 of mouse pA1 while these same mice performed our behavioural task.

The slow timescale of GCaMP6f fluorescence (Chen, Wardill, et al., 2013) means raw fluorescence provides a poor readout of cortical spiking (Theis et al., 2016), particularly when potential causes of activity are closely spaced. To ameliorate these problems, we used a previously validated spike-inference approach (), which offered a good account of our observed fluorescence data (Figure 5.3A).

To identify neurons encoding particular events we initially used an event (e.g. reward,lick,stimulus) triggered average approach. We averaged inferred spike rates over a 200ms window preceding onset of stimulus and compared these rates with a 200ms window following event onset. We defined encoding neurons as those whose spiking was modulated by the occurrence of events (Wilcoxon Signed-Rank; $\alpha = 0.001$) (fig:5.3B-F).

We found that the application of this method yielded consistent evoked responses, both in averages across neurons tuned to specific features (Fig-

ure 5.3Bi) and in single neurons (Figure 5.3Bii) indicating that this method effectively determined neural sensitivity. In doing so, we found subpopulations of neurons whose activity robustly increased following the occurrence of measured external variables (Figure 5.3B,E,F). Consistent with the location of our imaging areas within auditory cortex, we found that the most common sensitivity was to presented auditory stimuli. Further, we found that across the entire population, the presented stimulus level had a profound effect on neural spiking, such that stimuli presented at lower levels, led to lower spike rates (Figure 5.3C,D). Finally, we found that a large fraction of neurons were sensitive to multiple events (Figure 5.3E,F). Strikingly neurons responding to reward receipt did so irrespective of the lick spout to which the reward was delivered.

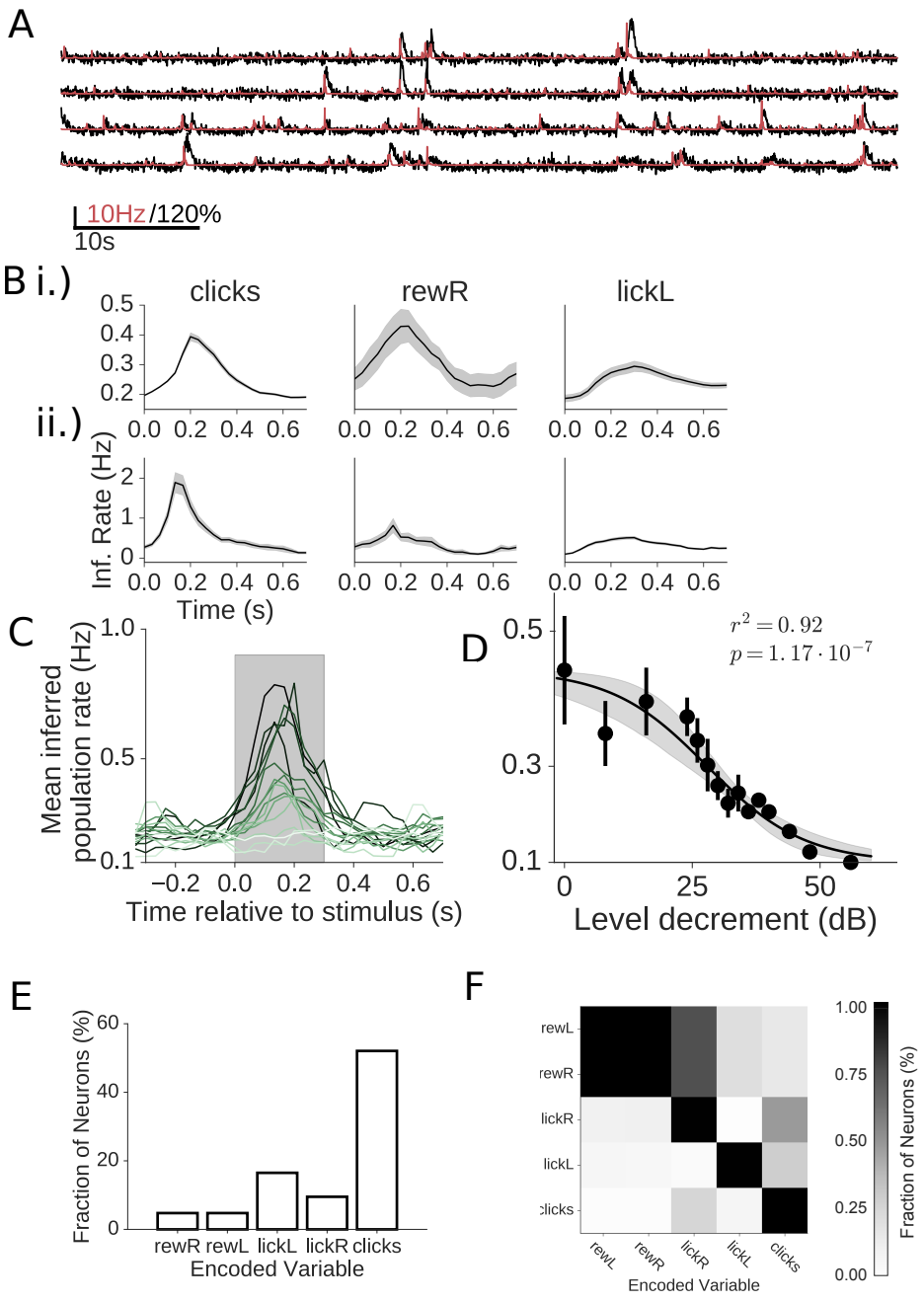


Figure 5.3: see below

Figure 5.3: Diverse responses to external events in auditory cortex. **A** Example fluorescence traces ($\Delta F/F$) in black and overlayed inferred spike rates (red). **B** Triggered average traces from all **(i)** or example **(ii)** neurons tuned to stimuli, rewards or licks. Responses are triggered to onset of events named in text above figure panels. **C** Representative example of evoked population responses as a function of stimulus level. **D** Evoked responses averaged over the grey window in C show sigmoidal level dependence of population responses. **E** Distribution of response selectivities, as assessed using the triggered average method. **F** Joint feature tuning assessed using the triggered average method. Data are normalised along rows, such that each row is the fraction of neurons active following the occurrence of a given event which also respond to other events.

In addition to sensitivity to external events, previous work has shown that neural activity encodes internal variables, such as attention (Rabinowitz et al., 2015; Ruff and Cohen, 2014; Harris and Thiele, 2011; Cohen and Newsome, 2008) as well as upcoming decisions (Poort et al., 2015; Parker and Newsome, 1998; Kwon et al., 2016; Cohen and Newsome, 2009; Bizley, Walker, Silverman, et al., 2009). Our optogenetic experiments suggest that auditory cortex makes a causal contribution to the decision process. Theoretical work has shown that, at least in principle, it is possible to infer the causal contribution of neural populations to the decision making process, based on their correlation with behavioural decisions (Britten et al., 1996; Shadlen, Britten, et al., 1996; Haefner et al., 2013; Pitkow and Angelaki, 2017). We therefore proceeded to analyse the responses of recorded neurons relative to mice decision on single trials.

To understand the relationship between neural activity and behaviour, we selected stimulus intensities at which similar rates of hit and miss trials occurred (Figure 5.1C). We selected equal numbers of hit and miss trials

from each stimulus level within this range. Given our use of a single stimulus and the fact that neurons encode the animals' licking behaviour as well as rewards, a major concern was that any differences between hit and miss trials may arise from the differences in behaviour and rewards received. To address this concern, we compared hit trials with misses, as well as with trials in which the animals' response occurred earlier than a reasonable reaction time (100ms; early trials). We reasoned that in early trials, behaviour was not driven by stimuli but was balanced with hit trials in terms of licking and reward receipt.

In doing so, we found subsets of neurons that strongly represented animals upcoming decision (Figure 5.4A,B). Across populations of neurons, we quantified animals selectivity using the choice probability metric (CP) (Britten et al., 1996). CP, which varies between 0 and 1, represents the probability that an ideal observer can correctly assign two randomly drawn neural responded to the correct category (hit vs miss). By convention, CPs above 0.5 indicate that neurons respond more strongly in hit trials, while those below 0.5 indicate larger responses during miss trials.

Consistent with previous literature we found a small population of neurons (12 % hit-vs-early 15% hit-vs-miss) whose activity covaried with behavioural choice (Figure 5.4C,D). Further, we found that the distribution of choice probabilities was significantly greater than 0 (1 sample t-test; $p < 10^{-450}$ in both cases; μ= 0.56 hit-vs-early; μ=0.52 hit-vs miss). This effect was driven by a larger fraction of neurons with significant choice probabilites

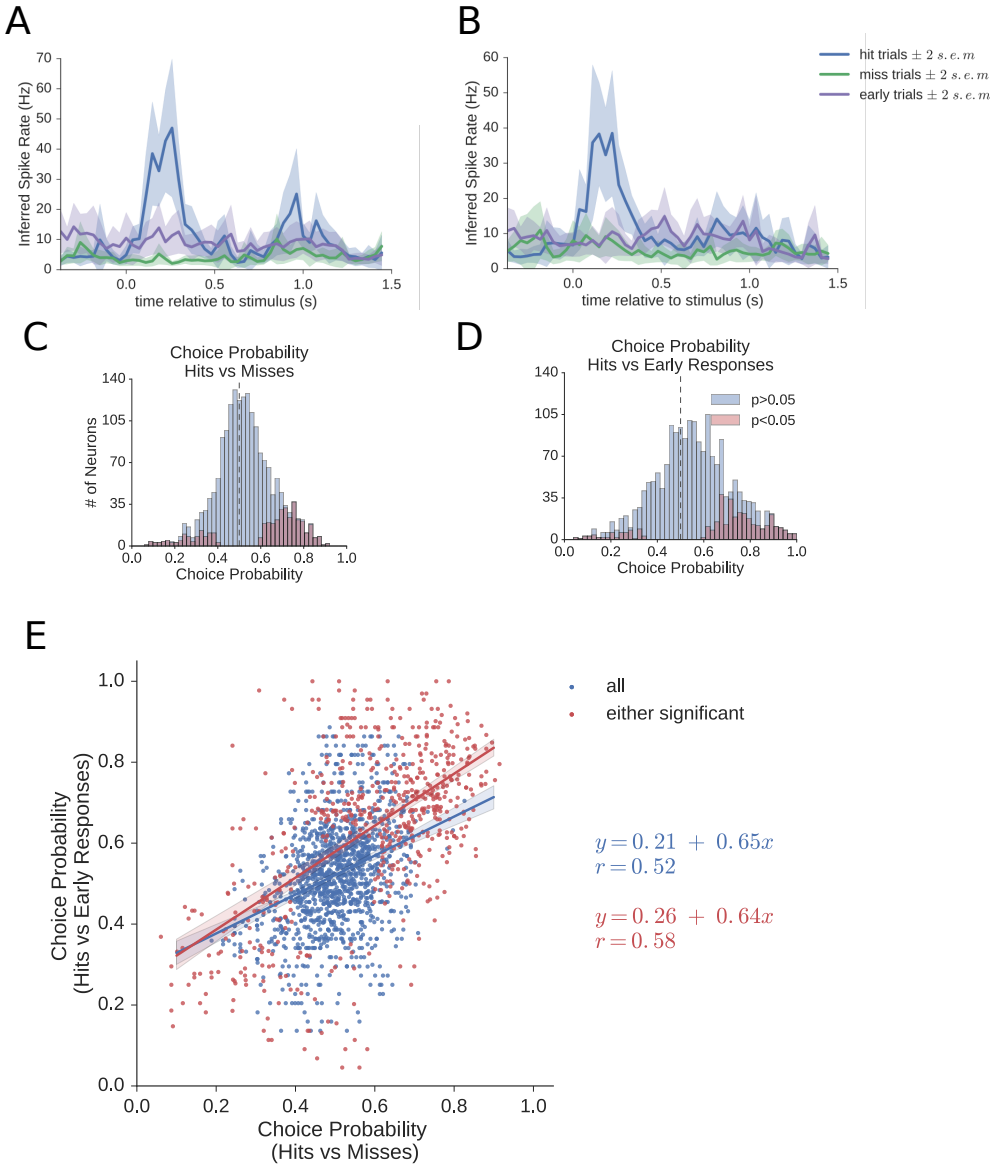


Figure 5.4: Choice related activity in primary auditory cortex. **A,B** Example inferred spike rates of two putative highly choice selective neurons. **C** Distribution of choice probabilities calculated using the first ∼250ms following stimulus onset. **D** Same as **C** except calculated comparing hit and early-response trials. **E** Correlation CP calculated based on hit-vs-miss trials and CP calculated based on hit-vs-early trials.

greater than 0.5 (χ^2; $p = 1.2 \cdot 10^{-47}$) rather than larger magnitudes of significant choice probabilities above 0.5 than below (2-sample t-test; p=0.39 hit-vs-miss, p=0.19 hit vs early).

Finally, we compared choice probabilities derived from hit vs miss trials with those derived from hit vs early trials. We reasoned that if neural responses truly encoded behavioural choice, we would expect a strong correspondence between the choice probabilities derived from hit-vs-miss and hit-vs-early trials. Consistent with this hypothesis, we observed strong correlations between the encoding of choice across the two trials types (Figure 5.4E). This was true whether we included only neurons with significant choice probabilities (in one of the two conditions) or across the entire population. Indeed, regression coefficients were highly similar between these two groups. Thus, in addition to encoding external variables, subsets of neurons in pA1 robustly encode behavioural choice.

5.3.2 Model based analysis of neural activity

While triggered averages provide a reasonable first approximation of the selectivity of neurons they are simultaneously problematic. This due to two reasons. Firstly, the structure of the behavioural task imposes correlations between the occurrence of events during a trial, which together with the slow kinetics of our sensors (Chen, Wardill, et al., 2013; Theis et al., 2016) means that selectivity may be misattributed. Secondly, mouse behaviour is variable from trial to trial such triggering to a given event will contain other variable

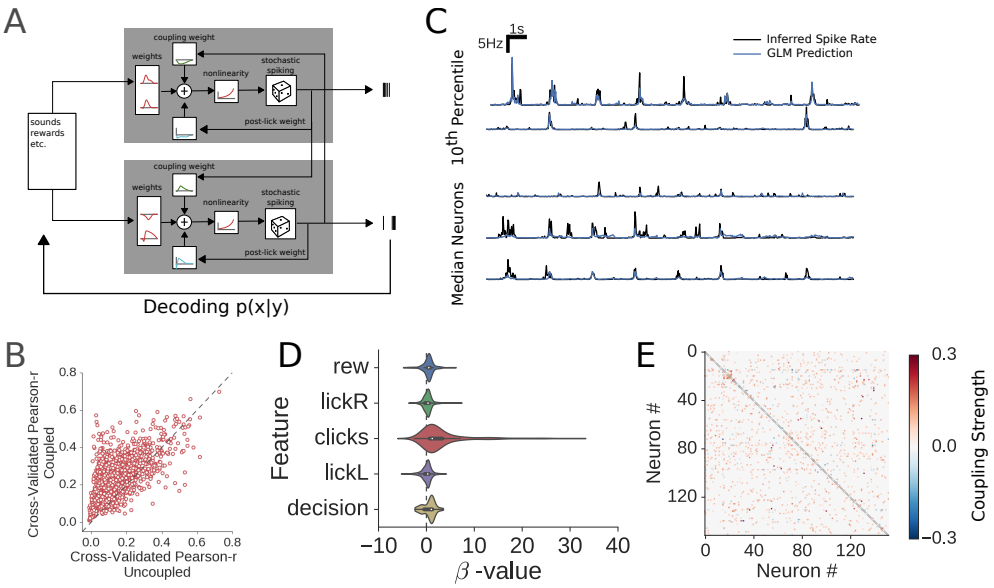


Figure 5.5: Poisson Generalized Linear Models capture neuronal firing rates. **A** Schematic of Generalized Linear Models fitted to imaged neural data. **B** Cross-validated Pearson-r of individual neurons fit with a coupled GLM or uncoupled GLM shows coupled GLM provides better fits. **C** Example GLM fits to single trial neural activity. Two randomly selected examples from the top tenth percentile of fits are shown together with three examples close to median (as determined by cross-validated pearson-r). **D** Horizontal smoothed histograms showing the marginal distributions of parameters across all imaged populations. The height of the histogram indicates the number of neurons with a level of feature preference β-values were calculated by summing thresholded β parameters of the model for each neuron. **E** Example coupling matrix showing sparse coupling of populations in pA1.

elements. Additionally, previous work has demonstrated that cortical neurons are coupled to one another (Oviedo et al., 2010; Dantzker and Callaway, 2000; Bureau, Von Paul, and Svoboda, 2006; Levy and Reyes, 2012; Barbour and Callaway, 2008; Watkins, Kao, and Kanold, 2014; Runyan et al., 2017; Zohary, Shadlen, and Newsome, 1994; Abeles and Gerstein, 1988), and that this coupling can bias the attribution of response properties, based on triggered averages, to neurons (Shadlen, Britten, et al., 1996; Pillow, Shlens, et al., 2008; Zohary, Shadlen, and Newsome, 1994).

To overcome these difficulty, we fitted poisson generalised linear model (GLMs) (McCullagh and Nelder, 1989; Paninski, 2004; Runyan et al., 2017; Pillow, Shlens, et al., 2008) to our inferred spike rates of neurons. We used poisson GLMs to our data as, consistent with previous reports (), we found that the variance of evoked responses of single neurons scaled with their mean (data not shown). Our GLMs model the function mapping from external and internal (behavioural choice) events, as well as the spiking of other neurons, to the inferred spike rates of particular neurons (Figure 5.5A).

First, we assessed whether GLMs which included the activity of other neurons as regressors (coupled-GLM) provided more accurate fits to cortical data than those without (uncoupled-GLM). Consistent with previous application of these models (; Runyan et al., 2017), as well as the comparatively dense monosynaptic connectivity in layer 2/3 of auditory cortex **??**, we found that including coupling terms substantially improved cross-validated prediction of neural activity (Figure 5.5B). Importantly, this was true both with

respect to cross-validated correlation coefficients (paired-t-test; $p < 10^{-450}$ and data log-likelihoods(paired-t-test; $p < 10^{-450}$,data now shown). Thus, we restict our subsequent analysis to fits based on the coupled GLM.

We found that the trial-by-trial activity of individual neurons was well predicted by our coupled GLM (Figure 5.5B,C) at levels similar to previous reports in cortical structures (Park, Archer, et al., 2013; Runyan et al., 2017). We therefore proceeded to analyse the model parameters by summing fit parameters belonging to a given class across neurons. Our results based on this qualitatively consistent with the analysis of neural activity by triggered averages. Consistent with our results from our analysis of raw neuronal activity, we found that the most influential regressor in the population was stimulus driven activity (Figure 5.5D), being larger than any other variable (1-way ANOVA $p < 0.05$), excluding coupling parameters (Figure 5.5E). Surprisingly, we found that the second most influential regressor was the animal's upcoming decisions. Further, similar to our triggered analysis, we found that most neurons were sensitive to more than 1 feature (μ=2; s.e.m=1) (Figure 5.6). Importantly, and consistent with previous work, we find that there is a strong positive correlation between stimulus sensitivity and the extent to which cells activity relates to mice' decisions (Figure 5.6).

5.3.3 Functional connectivity in pA1

In addition to external and behavioural parameters, neuronal activity was strongly influenced by the activity of other neurons (Figure 5.5B,E). To better

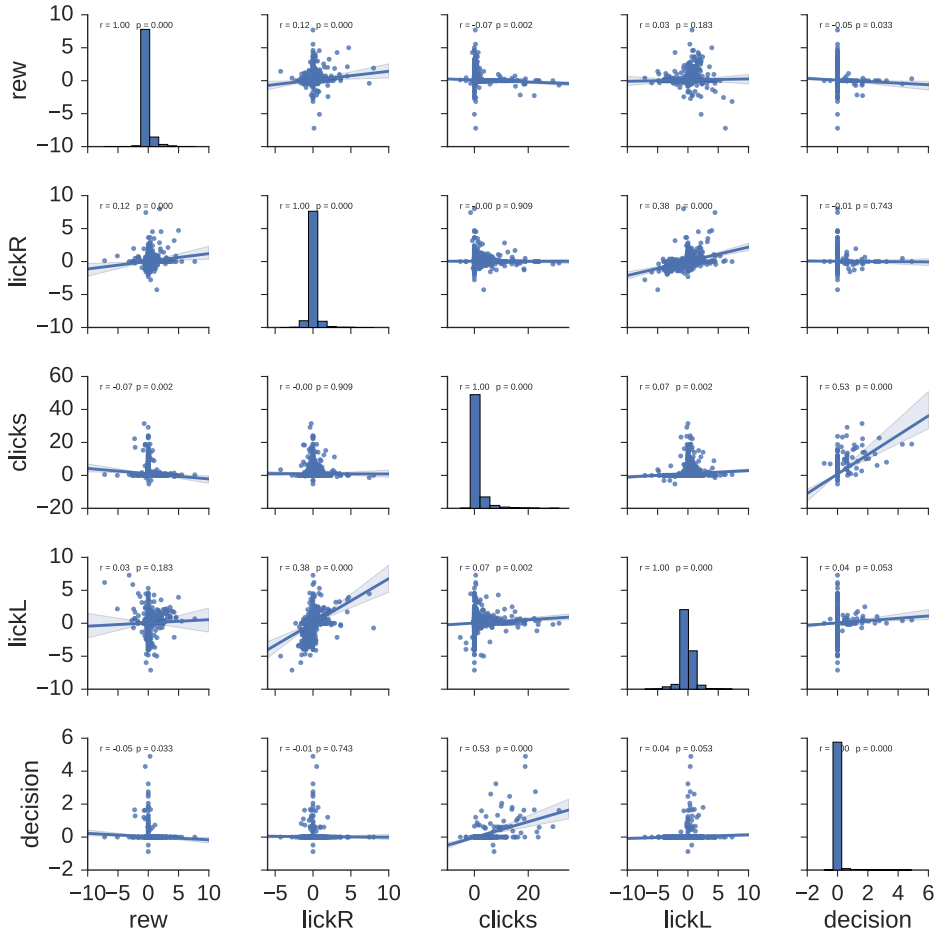


Figure 5.6: Tuning Covariance Structure mouse pA1. In off diagonal plots, each dot represents the tuning of a single neuron to pairs of features. In diagonal plots, the marginal distribution of tuning to single features. β.s were thresholded and include only β. where $p < 0.05$.

understand the functional architecture of auditory cortex, we characterised the structure of coupling between neurons. We stress that although this connectivity is modelled as interactions between individual neurons, the presence of coupling between cells does not imply that there are structural, monosynaptic connections between them (see *Discussion*). While the structure of coupling between neurons may be random (Figure 5.7A), previous work in the visual system has demonstrated that structural connectivity is well predicted by the stimulus selectivity of individual neurons (Cossell et al., 2015; Ko et al., 2013). Based on this, we hypothesised that neurons with similar event preferences would show stronger coupling (Figure 5.7B) giving rise to functional subnetworks of neurons with similar feature preferences.

To test this idea, we analysed two coupling coefficients of our model, quantifying both the strength and probability of coupling between pairs of neurons as a function of their respective feature preferences. For this analysis we defined separated neurons into disjoint groups based on the feature they were most influenced by (i.e. largest $\sum abs(\beta_{event})$).

Consistent with our predictions, and previous work (Cossell et al., 2015; Ko et al., 2013), we found large differences in coupling probability (Figure 5.7C) and strength (Figure 5.7D) depending on target source and target neurons feature preferences. Consistent with the existence of structure in the coupling between neurons in auditory cortex, we found significant differences in connectivity depending on source neuron selectivity, target neuron selectivity and the interaction between these terms (2-way ANOVA; $p < 1{\cdot}10^{-9}$, in

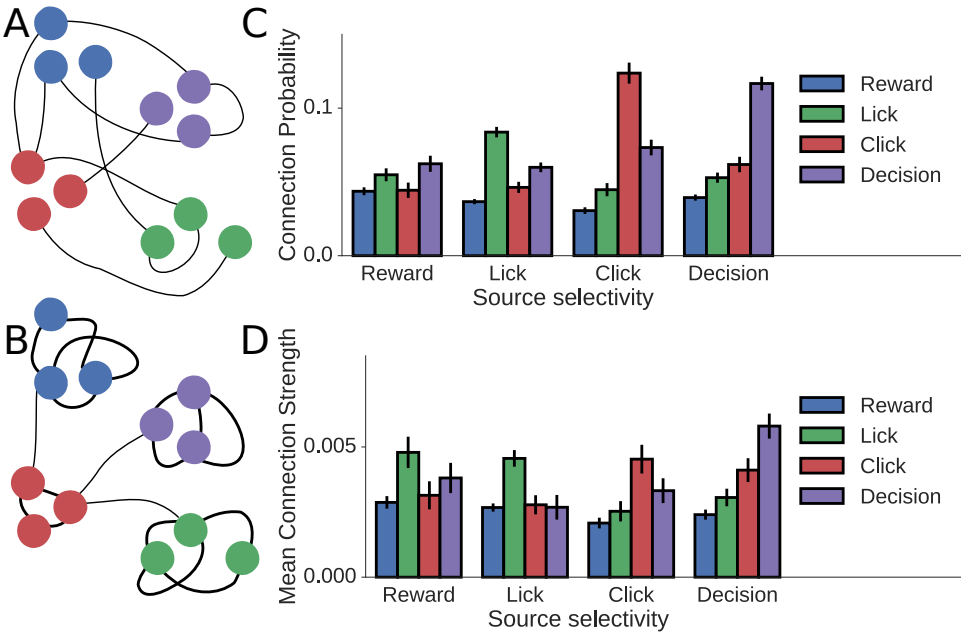


Figure 5.7: Coupling structure between neurons in pA1. **A** Schematic showing random coupling between neurons with diverse feature preferences. Colors match histograms. **B** A priroi predicted structure of coupling in pA1 populations in layer 2/3. Neurons with similar feature preference are more strongly coupling. **C** Pairwise coupling probability as a function of source and target selectivity demonstrates that neurons preferentially couple with neurons sharing the same dominant feature. **D** Same as **C** except based on coupling strength rather than probability.

each case). Further, post-hoc testing revealed that neurons encoding acoustic clicks, decisions or licks preferentially connected to neurons with the same dominant feature (Tukey HSD; $p < 0.01$, in each case). In contrast, neurons predominantly encoding rewards did not show differences in coupling. Thus, coupling in pA1 is highly structured and driven by similarity in feature encoding.

5.3.4 Model based analysis of decision related activity

Our model based analysis of neural activity corroborates our previous CP analyses, suggesting that neural activity covaries with the animals' decisions. Our GLM encoding model thus captures the relationship between external and internal events and neural activity. Importantly, this relationship can be inverted (Park, Meister, et al., 2014; Pillow, Shlens, et al., 2008) enabling the decoding of events from spiking activity. To assess our ability to predict decisions based on our model, we thus attempted to reconstruct the animals' decisions based on the complete set of observed inferred spike rates. We found that we were able to reliably decode mouse behaviour on single trials as assessed by a population choice probability (Figure 5.8A). Importantly, we found that our ability to decode mice decisions was independent of stimulus level (Pearson-r= 0.11, p=0.67). This suggests that our analysis divulges an axis through the space of population activity whose readout is not confounded by stimulus evoked responses.

The above decoding approach relies on an ideal observer model and is

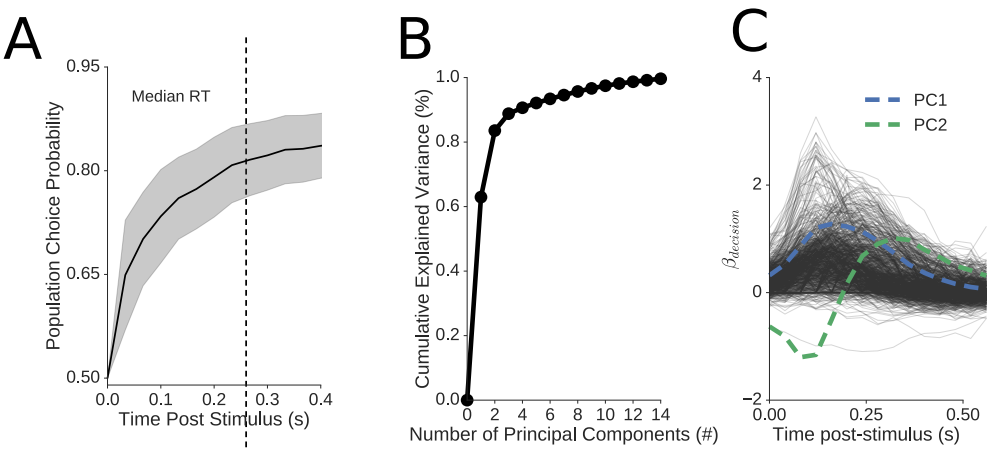


Figure 5.8: Population decoding from pA1. **A** Population choice probability as a function of time post stimulus **B** Cumulative explained variance of decoding weights accounted for by each principal component. **C** Decoding weights of individual neurons (grey) and the first two principal components

---

thus useful for determining theoretical limits on accuracy (Park, Meister, et al., 2014). However, this readout mechanism is unlikely to be neurally implementable as it required complete knowledge of activity and tuning properties of the entire population as well as all events. However, previous work has shown that contingent on low-dimensional structure of decoding weights, a neurally plausible readout mechanism can be derived, using leaky downstream integrator models ().

To test whether similar principles may apply here, we performed principal components analysis on the decoding weights of all neurons (Figure 5.8B,C). In doing so we find that we can capture most of the variance (82%) using the first two principal components. This suggests that a neurally plausible low-dimensional readout mechanism may implement decoding based on weights

identified by our model. Thus, the observed differences in activity between hit and miss trials may drive different perceptual decisions through a neurally plausible readout mechanism.

Multiple Mechanisms underlie decision related activiy

The origin of this activity remains unclear. Previous work has suggested that multiple circuit mechanisms contribute to choice related activity in sensory regions. These can be separated into two sets of mechanisms relating to separable parameters of our GLMs. The first relates to an increase in the sensory response, driven either by top-down attention like mechanisms (Nienborg and Cumming, 2009) and stochastic fluctuations in bottom-up input (Britten et al., 1996; Shadlen, Britten, et al., 1996). In our model these effects are captured by the strength of stimulus related responses. The second group of mechanisms involves to changes local functional connectivity (Cohen and Newsome, 2008; Schmitt et al., 2017) and reverberation in cortical circuits (Kwon et al., 2016). These changes might be captured in the coupling parameters of our model.

To explore the contribution of each of these factors, we separately fitted GLMs to data from hit and miss trials. We reasoned that, in addition to differences in gain of the sensory response between hit and miss trials, this would allow us to find differences in coupling structure in imaged pA1 populations. To ensure that differences in stimulus intensity would not drive differences in the fit models, we selected equal numbers of hit and miss trials from each stimulus level. We then proceeded to analyse changes in stimulus

drive (subsubsection 5.2.4.6) and coupling in neurons our previous analysis identified as decision related. We *a priori* separated neurons with negative decision weights (i.e. less expected activity in hit trials than in miss trials) (decision– neurons) from neurons with positive decision weights (decision+ neurons).

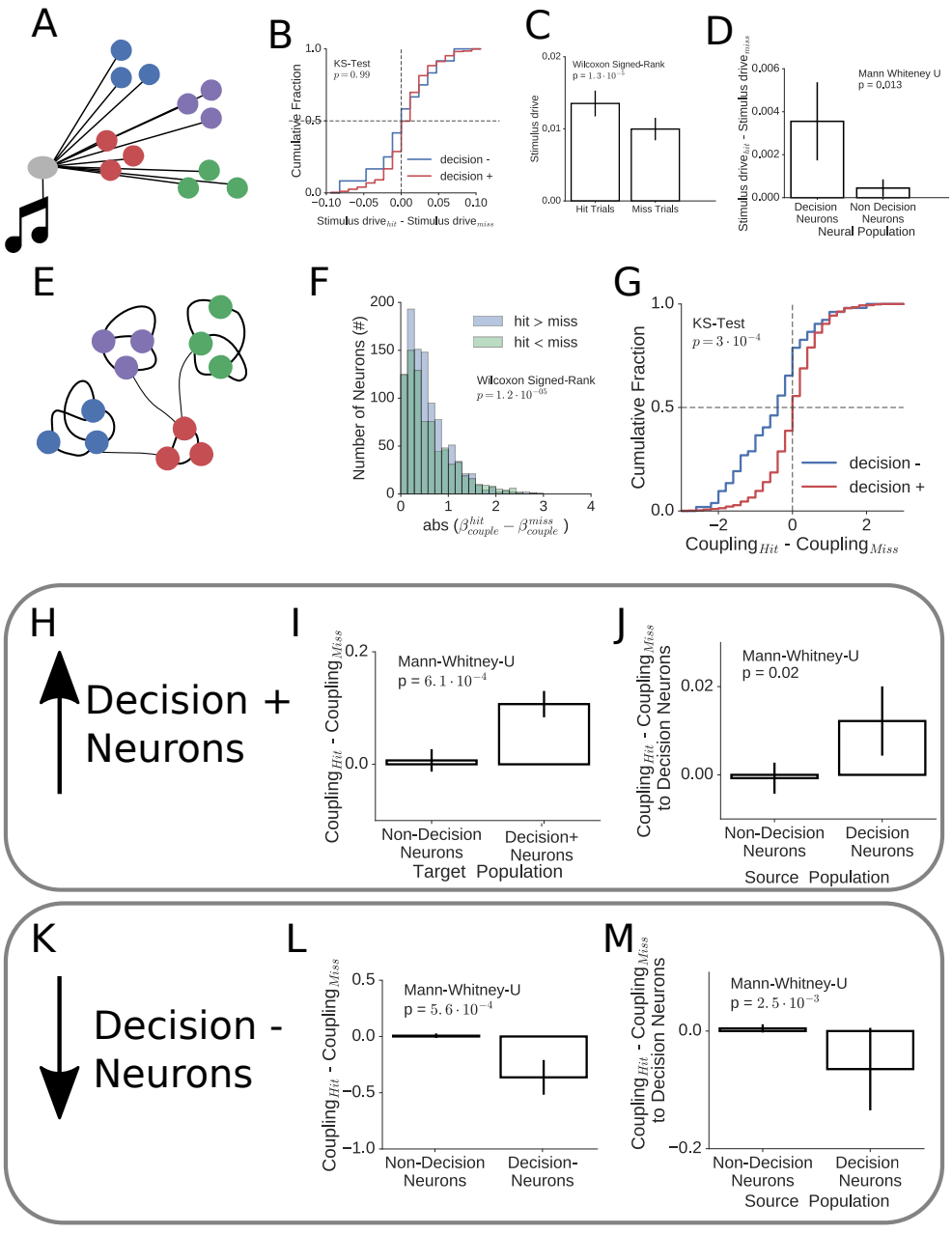


Figure 5.9: see below

Figure 5.9: Origins of choice related activity in pA1. **A** Schematic illustrating stimulus drive contributes to pA1 activity. **B** Cumulative histogram showing differences in stimulus drive between hit and miss trials does not differentiate decision+ and decision– neurons. **C** Stimulus drive to decision related neurons is larger on hit trials than on miss trials. **D** The difference in stimulus drive between hit and miss trials is larger for decision related neurons than neurons without significant decoding weights. **E** Schematic illustrating changes in functional coupling. **F** Coupling is stronger on hit trials than on miss trials across the population of decision related neurons. **G** Cumulative histgram of differences between hit trial coupling and miss trial coupling show that changes in coupling differentiate decision+ from decision– neurons. **H** Decision+ neurons increase activity on hit trials. **I** Coupling to decision+ neurons selectively increases on hit trials. **J** Selective increase of coupling from decision+ neurons to other decision+ neurons. **K-M** Same as *H-J* except for decision– neurons. Error bars represent 2 × s.e.m.

Having fitted our model, we attempted to identify differences in the gain of stimulus evoked responses between hit and miss trials (Figure 5.9A). In order to do so, we compared the stimulus drive to neurons on hit trials against the stimulus drive to neurons on miss trials. We first tested whether there were any significantly differences in the differential stimulus drive between decision+ and decision– neurons. Surprisingly, we found that decision+ and decision– neurons were not differentiable based on their stimulus drive (KS-Test; p=0.99) (Figure 5.9B). We therefore proceeded to analyse these populations together.

We found that decision related neurons received significant stimulus drive on both hit (Wilcoxon Signed Rank; $p = 2.2 \cdot 10^{-14}$) and miss trials (Wilcoxon Signed Rank; $p = 8.2 \cdot 10^{-10}$) (Figure 5.9C). Importantly, stimulus drive was highly correlated between hit and miss trials ($r^2 = 0.84$; $p < 10^{-450}$), indicating that despite reducing our data-set size we are able to accurately

infer our regression model parameters. We found that stimulus drive on hit trials was significantly larger than on miss trials on a neuron by neuron basis (Wilcoxon Signed Rank; $p = 1.3 \cdot 10^{-5}$). Further, although both stimulus drive to decision related neurons (Wilxocon Signed Rank; $p = 1.0 \cdot 10^{-5}$) and non-decision related neurons (Wilxocon Signed Rank; $p = 0.012$) was larger on hit trials, we found that this difference was larger for decision related neurons than non decision encoding neurons (Mann Whitney U; $p = 0.013$) (Figure 5.9D). Thus, we find that stimulus drive to neurons, which may arise from bottom up drive or top-down gain effects is significantly larger during hit trials than miss trials.

However, these results cannot account for the observed differences between hit and miss trials. In particular, the fact that both decision+ and decision– neurons receive a stronger stimulus drive on hit trials runs counter to the observation of reduced activity in decision– neurons on hit trials. We therefore proceeded to analyse the coupling structure of cortical neurons during hit and miss trials (Figure 5.9E).

Across recorded neural populations, we found that there was an increase in coupling onto decision related neurons (Figure 5.9F). However, in contrast to the effect of stimulus drive, this effect was differential between decision+ and decision– neurons (Kolmogorov-Smirnov; $p = 3.0 \cdot 10^{-4}$). We therefore analysed the differences in coupling properties, between hitand miss trials, separately for each subpopulation.

We found that hit-trials, as compared with miss trials, are characterised

by an increase in coupling to decision+ neurons (Wilcoxon Signed Rank; $p = 2.3 \cdot 10^{-8}$) (Figure 5.9I). Further, the change in coupling between hit and miss trials was significantly larger for decision+ neurons than non-decision neurons (Mann Whitney U; $p = 6.1 \cdot 10^{-4}$). Strikingly, we found that across the population of non-decision encoding neurons, there was no change in mean coupling strength between hit and miss trials (Wilcoxon Signed Rank; $p = 0.21$).

Next, we sought to ascertain the source of increased coupling to decision+ neurons. Based on our previous identification of a functional subnetwork of decision related neurons (Figure 5.8), we hypothesised that the observed increase in coupling on hit-trials might arise from other decision related neurons. Consistent with this hypothesis, we found a selective increase in coupling from other decision related neurons (including both decision+ and decision– neurons) (Mann Whitney U; $p = 0.020$) and no change in coupling from non-decision related neurons (Wilcoxon Signed Rank; $p = 0.43$) (Figure 5.9J). Strikingly, when we performed identical analyses on decision– neurons, we found the opposite pattern of results (Figure 5.9K-M). Specifically, we found a selective decoupling of decision– neurons from the network of decision related neurons (Figure 5.9L,M). These results show that in addition to increases in stimulus drive, differences in the coupling structure of neural populations are likely to be a major contributor to differences in neural activity that drive behaviour.

5.4 Discussion

Here we have characterised activity in mouse auditory cortex during the performance of an auditory detection task. We have shown that although activity is dominated by auditory stimuli neurons are sensitive to all features of the observed environment. Further, we have shown that it is possible to readout upcoming decisions from single neurons as well as populations in the primary auditory cortex. Extending this notion, we have shown that decisions are encoded in low-dimensional subspaces of population activity, suggesting a neurally plausible readout mechanism. Importantly this space is uncontaminated by stimulus-evoked activity. Finally, we have shown that decision related neurons form functional subnetworks in auditory cortex and that the differential activity on hit and miss trials arises both through changes in stimulus drive, as well as an increase in functional coupling within the decision encoding functional subnetworks.

Our optogenetic results, together with inactivation studies (Guo, Chambers, et al., n.d.; Talwar, Musial, and Gerstein, 2001; Li, Liao, et al., 2017) suggest that auditory cortex is involved in driving sound evoked behaviour. This is consistent with work demonstrating that activity in sensory cortex is necessary (Poort et al., 2015; Guo, Li, et al., 2014) and sufficient (**Oconnor2013**; Salzman, Britten, and Newsome, 1990; Xiong, Znamenskiy, and Zador, 2015; Znamenskiy and Zador, 2013a; Li, Liao, et al., 2017) to provide the sen sensory based decisions.

Consistent with this notion, we find neurons whose activity covaries with decisions and thus may be implicated in the decision making process. Although previous work has suggested that choice-related activity does not necessarily relate neurons to the decision process, more recent work challenges this view. In particular, previous work has suggested that it is not possible to directly infer how sensory information is read-out by downstream decision regions because choice-correlated activity in a given neuron may arise from its correlation with neurons that are directly related to choice (Zohary, Shadlen, and Newsome, 1994). However, more recently, it has been shown that only this problem is not true in general, but restricted to specific correlation patterns in the population (Pitkow and Angelaki, 2017). Further, related theoretical work (Haefner et al., 2013) shows that even in the face of correlated variability the population readout structure, if not single neuron weights, can be inferred by comparing choice related activity to neural sensitivity parameters.

Given that we see a strong correlation between decision related neurons and sound sensitivity, we surmise that the most parsimonious interpretation of our observed decision related activity is that the observed differences in firing rate we observe drive behaviour. Canonically, layer 2/3 serves as the main driving input to layer 5 of auditory cortex (Douglas and Martin, 2004), which has been previously shown to drive auditory behaviour (Znamenskiy and Zador, 2013a; Xiong, Znamenskiy, and Zador, 2015), consistent with this interpretation.

Strikingly, we observe distinct functional subnetworks of neurons in auditory cortex. In all likelihood, our observed coupling weights reflect multiple parallel processes, including direct monosynaptic connectivity between cells, polysynaptic connectivity between cells, as well as common input noise.

In the future, we hope to test the hypothesis that common noise explains the observed coupling computationally. In particular, extensions of the poisson GLM used to model population activity here have been shown to be able to discriminate common input from true coupling (Kulkarni and Paninski, 2007; Paninski et al., 2010; Vidne et al., 2012). We hope to apply these models to our data, exploring whether a single, latent top-down input, which varies from trial to trial, better explains our results than a model with changing functional connectivity.

However, several argument point to the likelihood that monosynaptic connectivity dominates the coupling observed here. Firstly, existing evidence suggests that the major cortical source of excitatory input to neurons in layer 2 are other neurons in layer 2 (Oviedo et al., 2010) and that this contribution is larger than in visual or somatosensory cortices (Dantzker and Callaway, 2000; Bureau, Von Paul, and Svoboda, 2006). Importantly, our estimate of coupling probability (5.6%) is strikingly similar to estimates of connectivity between excitatory neurons in auditory cortex (Oviedo et al., 2010; Levy and Reyes, 2012; Barbour and Callaway, 2008; Watkins, Kao, and Kanold, 2014). Secondly, previous computational work has shown that even at frame rates similar to those used in our experiments, a GLM based approach can

accurately reconstruct the connectivity of cortical circuits (Mishchencko, Vogelstein, and Paninski, 2011; Soudry et al., 2015). Additionally, previous work has demonstrated that correlations in activity assayed by two-photon imaging predict monosynaptic connectivity (Cossell et al., 2015). Finally, our findings that neurons which tend to form functional subnetworks are dominated by neurons with similar tuning, is in line with work studying layer 2/3 in the visual system as well as plasticity rules in layer 2/3 (Kim, Oh, et al., 2016; Ko et al., 2013). Thus, we contend that our observed functional coupling is dominated by local connectivity.

In light of this interpretation of our coupling data, we suggest that the increased functional connectivity observed preceding hit trials reflects local amplification of neural activity. Studies in auditory (Li, Li, et al., 2013) and other modalities (Harris and Mrsic-Flogel, 2013; Li, Li, et al., 2013; Lien and Scanziani, 2013; Douglas, Koch, et al., 1995; Reinhold, Lien, and Scanziani, 2015) have demonstrated that cortical circuits can amplify incoming signals. Further, recent work in frontal cortex has shown that cortico-thalamo-cortical circuits modulate functional connectivity in frontal cortical regions (Wimmer et al., 2015). Intriguingly, recent work in the auditory system has posited that analogous loops contribute to enhanced stimulus detection (Guo, Chambers, et al., n.d.), suggesting a possible circuit basis for the phenomena we observe. Alternatively, changes in local functional connectivity may be driven by changes in inhibitory circuits that accompany task engagement (Kuchibhotla et al., 2016a) and are recruited by top-down

mechanims (Zhang, Xu, Kamigaki, et al., 2014).

Overall our results suggest a model whereby small subpopulations of decision-related neurons, recurrently amplify activity, thereby driving robust sound-evoked behaviour.

Chapter 6

Discussion

6.1 Overview

Overall, this work represents an almost end-to-end series of experiments attempting to progress our understanding of mouse behaviour and its neural correlates. We have developed novel hardware for performing large scale behavioural experiments in mice and motivated by pilot experiments proceeded to develop novel tools for quantifying static and dynamic behaviour in the head-fixed mouse. Finally, we imaged activity in mouse auditory cortex, attempting to find functional correlates of behaviour. Since we have discussed the findings from each experiment in their respective chapters, this discussion will focus on the implications of the project and discuss ongoing work in relation to them.

6.2 Large scale behaviour

In this work I have presented an apparatus for performing large scale behavioural experiments in both head-fixed and freely moving mice. Behaviour is very complex, and is driven by a plethora of subtly interacting cognitive or neural mechanisms (Gomez-Marin, Paton, et al., 2014). Understanding behaviour therefore presents a formidable challenge and requires carefully disentangling and measuring cognitive operations through well designed behavioural experiments (see chapter 1). Due to the time consuming nature of the design, optimisation, and enactment of this process, the most effective way to perform these experiments is at scale. I therefore think that large scale platforms, such as the one developed here, will be invaluable to furthering our understanding behaviour.

In addition to accelerating experimental progress, large scale behavioural platforms also offer the ability to perform qualitatively different experiments. Specifically, they enable a systematic mapping of animalsí behavioural repertoires. While this also applies to spontaneous behaviour, disease models, genetic screens, and cognition; it is perhaps most straightforwardly applied to the study of perception. Historically, the emphasis of psychophysical experiments was the determination of the acuity of perception with a view to understanding the early visual systems and laws governing its function (Read, 2015).

However, there is now almost universally acknowledged that many chal-

lenging perceptual problems exist beyond this early processing (Marr, 1982). Indeed, the prevailing assumption is that the majority of sensory neurons in the brain are devoted to processing the complex structure of sensory stimuli (; Yamins and DiCarlo, 2016). Although several attempts have been made to apply psychophysical methods to more complex stimuli (e.g. (Houlsby et al., 2013; McDermott and Simoncelli, 2011)) none of these attempts, to the best of the my knowledge, have effectively capitalised on the ability to perform psychophysical experiments at large scales.

This kind of approach is now both possible in humans, using services such as Mechanical Turk (Crump, McDonnell, and Gureckis, 2013), or in experimental animals, using large scale platforms for animal behaviour such as the work presented herein. Interestingly, analogous large scale methods are currently being fruitfully applied in other areas of neuroscience in the form of connectomics (Schröter, Paulsen, and Bullmore, 2017; Jones, Overly, and Sunkin, 2009), genomics (Abbott, 2003; Geschwind and Konopka, 2009) and creation of brain-activity maps (Vogelstein et al., 2014). Large scale behavioural experiments may therefore similarly advance our understanding of the *complex* computational functions performed during perception and behaviour. Describing and understanding these functions will be vital and without it, we may remain in the dark about how the brain solves problems (Anderson, 1972).

Importantly, large scale behavioural paradigms and home-cage training in particular provides an opportunity to ameliorate ethical problems asso-

ciated with animal work, in addition to accelerating research. This is for two reasons; firstly, it removes the stress that animals may experience when being handled by experimenters. Secondly, in our hands, animals undergoing training and testing in home cages consume similar amounts of water to when it is available ad libitum and thus do not undergo weight loss. In spite of this, they perform hundreds or thousands of trials per day at high levels of performance.

Thus, large scale behaviour platforms, such as the one developed here hold a major promise for advancing neuroscience, by both accelerating existing research programs and enabling new ones.

6.3 Policy-based descriptions of learning

Along with increased volume of behavioural data discussed above, there is an increasing move to measuring in greater detail the structure of animals behaviour (Gomez-Marin and Mainen, 2016; Gomez-Marin, Paton, et al., 2014; Hires et al., 2015; Kawai et al., 2015a; Guo, Graves, et al., 2015; O'Connor et al., 2010). These more detailed measurements include high-speed video fo animal movement, tracking the detailed whisking patterns of mice and microphones to record their vocalisations.

The promise of more detailed measurement of behaviour are a better understanding of the substrates of behaviour and their neural bases. For example, it has recently been shown, using high speed video, that rats learn

to twice press a lever, with a 700ms delay between presses, by engaging in idiosynchatic motor patterns bridging the 700ms (Kawai et al., 2015a). This simple observation, enabled by dense measurements of behaviour, provides novel and irreplacable insights into how this task was solved. The same promise holds for neural activity, where detailed measurements of behaviour enable both testing hypotheses and evaluating the sufficiency of our accounts of neural activity. Algining neural activity from layer 4 of mouse primary somatosensory cortex with high speed recording of whisker movement, both acquired during behaviour, Hires et al., 2015 were able to explain neural activity to sub-poisson levels of variability. Thus, dense measurements of behaviour hold considerable promise for progress.

To fulfil this promise, detailed measurement must be accompanied by normative accounts of observed phenomena. Contemporary normative accounts of behaviour and learning are derived largely from value-based reinfocement learning (RL) methods. However, the road to extending these methods to such complex datasets is unclear. The main problem with their application to more complex behaviour is that to remain computationally tractable, value-based RL relies on discetising states and actions. In reality, however, both behaviour and environments are often continuous and not easily discretised into a managably sized set.

For instance, how to to discretise rats idiosyhchatic behaviour in the level pressing task described above is unclear. Similary, how to appropriately discretise a mouse's whisking is unclear, yet whisking patterns can undergo sub-

stantial adaptive changes (O'Connor et al., 2010). Thus, for many measured behaviours, and in many behavioural tasks, the value based RL paradigm is difficult to apply.

An alternative, which has guided much of the behavioural modelling done here are policy-based descriptions of behaviour. In ethological studies, policy based descriptions of behaviour are commonplace (**Krebs**), as is perhaps best exemplified by the description of patch foraging (Viswanathan et al., 1999). Further, they are commonplace in neuroscience where they often viewed as statistical models of behavioiur (Lau and Glimcher, 2005; Busse et al., 2011; Bak et al., 2016; Smith et al., 2004a; 2007; n.d.; Burgess, Lak, Steinmetz, Zatka-Haas, Bai Reddy, et al., 2017; Fründ, Wichmann, and Macke, 2014a; Seeds et al., 2014; Desrochers et al., 2010).

Here we have performed an initial exploration of normative policy based models of more densely sampled behaviour (chapter 3,chapter 4), applying these methods to the complete licking pattern of mice. Since densely sampled head-fixed behaviour is still statistically homogenous and low dimensional, I though it server as both an interesting area of application of these methods as well as a useful test-ground of their potential.

With regard to the future application of these methods, we can draw several preliminary conclusions based on the work here. Firstly, combining dense measurements of behaviour with policy based-modelling can be fruitful for instance to infer the causes of behaviour on a single trial basis. The inability to do so has been a major criticism of statical and policy based

models of behaviour (Shadlen and Kiani, 2013). In addition, we have shown that such models can be used effectively to gain a qualitative understanding of behaviour, such as the finding that mice optimise their behaviour by multiple parallel strategy changes (chapter 3). Finally, despite the comparative complexity, these models can effectively be used to describe dynamically changing behavioural strategies, at high temporal resolutions (chapter 4).

In future work, we aim to characterise in more detail the elementary properties of the learning process. In particular, we hope to infer the learning rules mice might be using to optimise behaviour in head-fixed tasks. Importantly, the learning rule can trivially be explicitly incorporated into our models of dynamic behaviour.

Altogether policy based RL methods are therefore a promising class of methods for the rich description of densely measured behaviour.

6.4 Cortical contributions to sound driven behaviour

The data and analysis presented in chapter 5 is part of an extended project in which neural activity and behaviour were tracked over the course of learning. In chapter 5, we have demonstrated the power of fitting poisson generalised linear models (GLMs) to calcium imaging data acquired during mouse behaviour. Doing so, we have demonstrated that we can generate insights into how neural population activity is read-out by downstream regions involved

in decision making. We proceeded to suggest that recurrent connectivity among decision related neurons in layer 2/3 of auditory cortex, contributes to the decision making process.

Building on the work presented in chapter 5, our main future goal is to analyse the structure of neural and behavioural change over the course of learning. I will present my perspective on this issue, together with a preliminary analysis of some of this work here.

The learning process could take place at multiple, conceptually delineated, stages of processing. In particular, learning is generally though to take place through one (or some combination of) the following mechanisms (Sutton and Barto, 1998; Gilbert, Sigman, and Crist, 2001; Fritz, David, et al., 2010; Shadlen, Britten, et al., 1996; Shadlen and Kiani, 2013; Gallistel, Fairhurst, and Balsam, 2004; Sutton and Barto, 1981; Xiong, Znamenskiy, and Zador, 2015; Schmitt et al., 2017; Leong et al., 2017; Niv et al., 2015; Law and Gold, 2008; Averbeck and Costa, 2017):

1. Plasticity of the feed-forward sensory processing
2. Plasticity of recurrent connections in relevant sensory regions
3. Plasticity in the readout of activity in sensory regions by decision related brain regions
4. Plasticity of top-down inputs to sensory regions
5. Plasticity of top-down inputs to decision realated brain regions

Despite this diversity, which is likely distributed over a panolpy of brain regions, we can classify the above changes into two groups. The changes described by (1),(2) and (4) can be grouped together under the heading of representational plasticity. The net effect of these kinds of plasticity results in changes in the encoding of behaviourally relevant stimuli in sensory areas. The other loci of plasticity, (3) and (5), can be grouped into a second category which I will refer to as readout-plasticity. In essence this involves changes in the way activity in sensory regions is read out by decision-related regions to drive motor behaviour. Formally, our heuristic model Figure 6.1 is:

$$p(response) = W_{readout} \cdot f(\mathbf{X} \cdot W_{representation}) \tag{6.1}$$

where $\mathbf{X}$ is the sensory input, $W_{\cdot}$ are weights, and f is a function representing the transformation from the physical sensory stimuli to population activity.

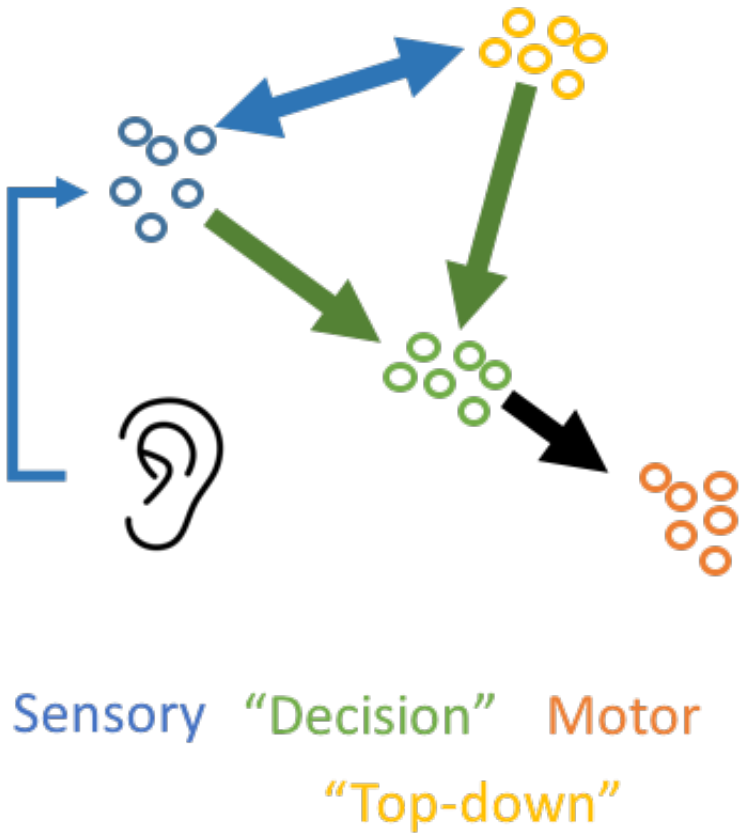


Figure 6.1: Schematic of a canonical decision making circuit . Processing stages which undergo representational plasticity are colored in blue; states where changes cause readout plasticity are coloured green.



In ongoing work, we plan to investigate the neural substrates of learning by analysing neural activity in auditory cortex over the course of learning. In the following discussion, I will argue that by fitting GLMs chapter 5 to cortical activity we can, in principle, assess the relative contributions of representational and read-out plasticity to improved performance on our operant task. I will then present some preliminary analyses of neural data over the course of learning which bears on this question.

We will illustrate this in abstract terms be-

Figure 6.2: Differential predictions of the structure of plasticity under the assumptions of representational and read-out plasticity. *(Top)* Post-learning, we can accurately read-out population activity (chapter 5) to predict animal's behaviour on single trials. We can do so by estimating the decoding weights and using them to project population activity on a 1-dimensional axis. By convention, values greater than zero along this axis predict that a given trial is a hit trial; values below 0 indicate that a given trial is a miss trial. The distance from zero corresponds with the confidence of the prediction. *(Middle)* An idealised form of representational plasticity, in which readout weights are intact from early learning. Under this assumption, neural activity on trials during initial learning mostly look like miss trials when onto the decision axis (defined post-learning). However, rare trials whose projected value exceeds 0 would result in hit trials. *(Bottom)* An idealised model of readout plasticity predicts that activity along the decision axis is distributed just as it is post-learning (i.e. no representational learning), but that activity along the decision axis, defined post learning, is unrelated to behaviour.

fore returning to our specific example. As illustrated in chapter 5, fitting population GLMs enables to to estimate how downstream decision related regions read out neural activity from auditory cortex. Specifically, for each neuron, we obtain a set of decoding weights across time, reflecting the differential activity on hit and miss trials. Combining decoding weights with neural population activity, we can predict whether a given stimulus presentation will result in a hit or a miss. Mathematically, this is implemented by projecting population activity onto a decision axis (Park, Meister, et al., 2014), which quantifies whether neural activity is more similar to activity on a prototypical hit trial (projection > 0) or miss trial (projection <0) Figure 6.2. Post-learning, this enables reliable prediction of behaviour on single trials (chapter 5).

Representational and readout plasticity make differential prections of how activity projected to this axis should look in early stages of learning. If no representational plasticity takes place over the course of learning, then learning should be mediated by altering readout weights. The prediction about activity along the decision axis is that neural activity would span this axis, but that the presence of a response to a given stimulus presentation is independent of the projection of neural activity onto the decision axis, defined post learning. Conversely if readout weights were fixed, but representations were learned, we would expect activity to always be on the 'miss-side' of the decision axis.

By inactivating auditory cortex (chapter 1 we have shown that auditory

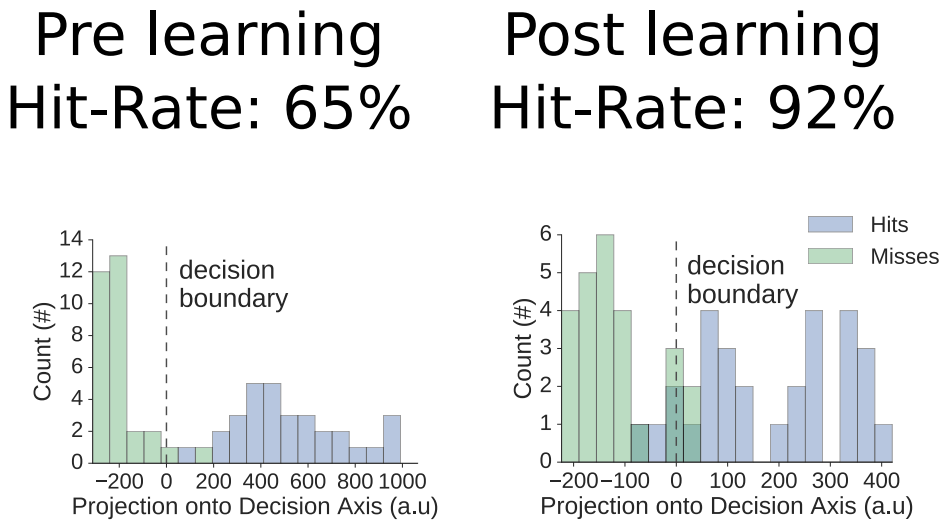
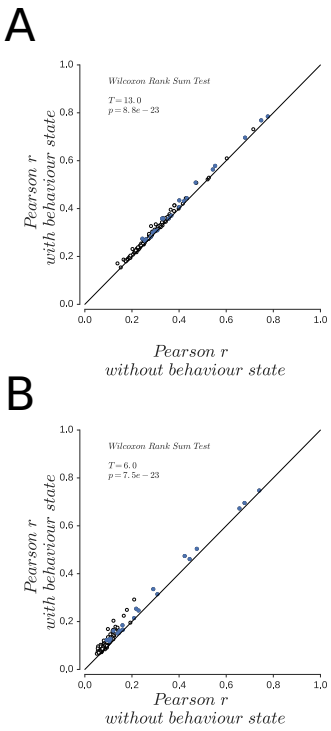


Figure 6.3: Ability to predict mice' behaviour on single trials on the first day of learning and the first day of testing in an example animal. Just as above, neural activity on trials projected to positions greater than zero along the decision axis are predicted to be hits, trials projected to below zero are predicted to be misses. A single decoder trained on post-learning data is able to perfectly predict hit and miss trials. Although there were substantially more miss trials in early learning, equal numbers of hit and miss trials on each day were selected for this analysis

---

cortex is necessary for sound detection, consistent with previous work (Guo, Clause, et al., 2017; Talwar, Musial, and Gerstein, 2001; Li, Liao, et al., 2017). Therefore, we know that information from this region is read-out by downstream decision regions. By fitting our GLMs to the behaviour of mice post-learning (chapter 5), we have been able to infer how activity in decision-regions is read-out from auditory cortex. By tracking how neural activity projects onto the decision axis, defined post-learning, across learning we aim to probe the substrates of learning.



-2cm

Figure 6.4: Cross validated prediction of neural activity including a single trial estimate stimulus sensitivity inferred using the state space GLMs described in chapter 4. Sessions correspond to the greatest change in fraction correct between the first and final 20 trials of training. In both cases this was the second day of training.

Since we used simple stimuli, our initial hypothesis was that learning is mediated by changes in how information from sensory areas is read out. Strong theoretical predictions for the dynamics of change process underlying this are defined by reinforcement learning (Sutton and Barto, 1998; Fetz, 1969; Linderman and Gershman, 2017). Specifically, temporal difference learning, as well as a variety of other reinforcement learning algorithms (Sutton and Barto, 1998), predict that readout weights should change according to a Gaussian random walk, biased by some learning rule (Linderman and Gershman, 2017).

In stark contrast to these predictions, our preliminary results (n=2 mice) suggest that read-out mechanisms are intact from early in learning. Specifically, we observe that as early as the first day of training, we can reliably decode animals behaviour using the GLM based decoder of neural activity developed in chapter 5. More strikingly, we find that we can decode mice's decisions using the

decoder trained on post-learning data. That is to say that neural activity on hit trials during the first day looks like the activity on hit trials on the final training day Figure 6.4.

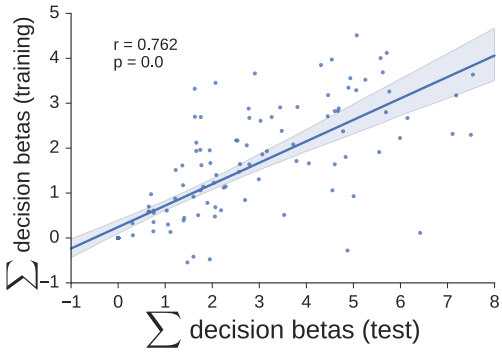


Figure 6.5: Pooled data from two animals shows that the decision readout weights are highly correlated between the first day of training and during testing

This is consistent with two complementary analyses. In the first analysis, we attempted to correlate changing behavioural stimulus sensitivity over learning with trial by trial neural activity. We did so by incorporating behavioural stimulus sensitivity, estimated trial-by-trial using the state-space GLMs described in chapter 4, into our models of neural activity. To our surprise, we found that behaviourally estimated stimulus sensitivity only marginally improved the ability of our neural models to predict activity in auditory cortex Figure 6.4. Although we, consitent with other work (Poort et al., 2015; Huber, Gutnisky, et al., 2012), find that neural activity changes over the course of learning, this demonstates that on a trial by trial basis, these changes only weakly corrrelate with changing auditory cortical activity. Secondly, we find that decision readout weights fit using our GLMs do not change substantially between the first day where behaviour is above chance Figure 6.4A and the testing phase.

Our preliminary analysis of the origins of the sensory readout structure thus suggests that learning is not mediated by iterative improvements to the read-out weights. Further, across both animals analysed the mean distance, along the decision axis, between hit and miss trials, is not significantly different before and after learning (t-test; p=0.41). This suggests that iterative improvements in sensory processing also do not account for changing behaviour.

An alternative to the concept of gradual learning exists to descirbe both learning in humans and animals. Previous work re-analysing animal behaviour in a variety of simple conditioning paradigms has shown that behavioural undergoes bi-directional step-like transitions before stabilsing at effective behaviour, rather than gradually changing (Gallistel, Fairhurst, and Balsam, 2004; Daw and Courville, 2008). Additionally, in humans step-like learning functions have been observed in the context of verbal learning (Rock, 1957; Roediger and Arnold, 2012; Wogan and Waters, 1959; Clark, Lansford, and M, 1960). Further, similar phenomena are observed during extinction (Gershman, Norman, and Niv, 2015; Bouton, 2004), where conditions behaviours are stochastically reinstated following and during behavioural extinction. Thus, previous behavioural analyses have reported step like decisions in behaviour, consistent with our findings in neural activity.

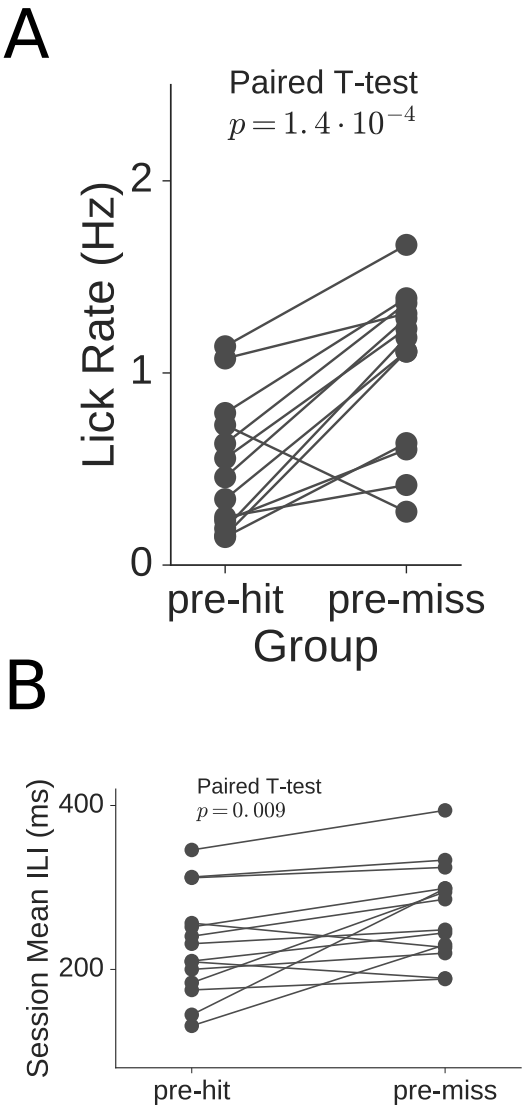


Figure 6.6: Differential behaviour pre-hit and pre-miss trials. **A** Prior to miss trials, animal's lick rates are higher. **B** Prior to miss-trials mean inter-lick-intervals are longer

An alternative hypothesis consistent with our and others findings, is that rather than learning the value of responding to the stimulus, mice learn the value of disjointly enacted behavioural policies. Plainly, in trials where mice do not respond to the stimulus, they don't do so because they were trying to get reward in another way. Based on the results in chapter 3 we would predict that the main alternative strategy would be characterised by increase bout initiation rates. Consistent with this prediction, when we analyse the behaviour of mice during learning (n=12 mice; pooling data from chapter 4, and chapter 5) we find that mice lick-rates are higher preceding miss trials than preceding hit trials (Paired-Sample T-test; $p = 1.4 \cdot 10^{-4}$). Further, this is not driven by increases in bout-lengths, as mean inter-lick intervals are even re-

duced on miss trials relative to hit trials (Paired-Sample T-test; $p = 0.009$) (this seemingly paradoxical set of findings arises from the fact that mean-inter lick intervals and lick-rates are calculated based on bounded pre-trial intervals; so for example if a mouse licked only twice in each pre-hit window with an inter-lick interval of 100ms, lick rates would be very low, and the mean inter-lick interval would also be very short).

Thus, my working hypothesis is that rather than learning the value of the stimulus, mice learn by learning the value of two different disjoint enacted behavioural policies. Future work will hopefully clarify these open questions.

6.5 Conclusion

Overall, we have I have done an end-to-end series of experiments to characterise mouse behaviour. We have developed a novel infrastructure for performing behavioural experiments and optogenetics and used it to collect behaviour. By modelling densely sampled head-fixed behaviour, we have been able to further our understanding of the structure of behaviour and allowed us to track its change. Studying neural activity has provided further insights into the process of behavioural enactment and furthered our understanding of the mechanisms underlying decision making. Future work will further bring together the neural and behavioural strands of research developed in this thesis to further eludicate the processes underlying learning simple operant behaviours.

Bibliography

Abbott, Alison (2003). "Genomics on the brain". In: *Nature* 426.6968, pp. 757–757. ISSN: 0028-0836. DOI: 10.1038/426757a.

Abeles, Moshe and George L Gerstein (1988). "Detecting spatiotemporal firing patterns among simultaneously recorded single neurons." In: *Journal of neurophysiology* 60.3, pp. 909–24. ISSN: 0022-3077.

Adrian, ED and Yngve Zotterman (1926). "The impulses produced by sensory nerve endings. Part 2. The response of a single end-organ". In: *The Journal of physiology* 61.2, pp. 151–171. ISSN: 00223751. DOI: 10.1113/jphysiol.1926.sp002281.

Ahmadian, Yashar, Jonathan W Pillow, and Liam Paninski (2011). "Efficient Markov Chain Monte Carlo Methods for Decoding Neural Spike Trains". In: *Neural Computation* 23.1, pp. 46–96. ISSN: 0899-7667. DOI: 10.1162/NECO_a_00059.

Akam, Thomas, Rui Costa, and Peter Dayan (2015). "Simple Plans or Sophisticated Habits? State, Transition and Learning Interactions in the Two-Step Task". In: *PLOS Computational Biology* 11.12. Ed. by Jean

Daunizeau, e1004648. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1004648.

Akerboom, Jasper et al. (2012). "Optimization of a GCaMP Calcium Indicator for Neural Activity Imaging". In: *Journal of Neuroscience* 32.40, pp. 13819–13840. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.2601-12.2012.

Allen, William E et al. (2017). "Global Representations of Goal-Directed Behavior in Distinct Cell Types of Mouse Neocortex". In: *Neuron* 94.4, 891–907.e6. ISSN: 10974199. DOI: 10.1016/j.neuron.2017.04.017.

Anderson, P W (1972). "More is different." In: *Science (New York, N.Y.)* 177.4047, pp. 393–396. ISSN: 0036-8075. DOI: 10.1126/science.177.4047.393. arXiv: arXiv:1011.1669v3.

Aoki, Ryo et al. (2017). "An automated platform for high-throughput mouse behavior and physiology with voluntary head-fixation". In: *Nature Communications* 8.1. ISSN: 20411723. DOI: 10.1038/s41467-017-01371-0.

Averbeck, Bruno B. and Vincent D. Costa (2017). *Motivational neural circuits underlying reinforcement learning*. DOI: 10.1038/nn.4506.

Bajo, Victoria M et al. (2010). "The descending corticocollicular pathway mediates learning-induced auditory plasticity". In: *Nature Neuroscience* 13.2, pp. 253–260. ISSN: 10976256. DOI: 10.1038/nn.2466.

Bak, Ji Hyun et al. (2016). "Adaptive optimal training of animal behavior". In: *NIPS*, pp. 1939–1947.

Bandyopadhyay, Sharba, Shihab A. Shamma, and Patrick O. Kanold (2010). "Dichotomy of functional organization in the mouse auditory cortex". In: *Nature Neuroscience* 13.3, pp. 361–368. ISSN: 10976256. DOI: 10.1038/nn.2490. arXiv: NIHMS150003.

Bannerman, David M et al. (2012). "Dissecting spatial knowledge from spatial choice by hippocampal NMDA receptor deletion". In: *Nature Neuroscience* 15.8, pp. 1153–1159. ISSN: 10976256. DOI: 10.1038/nn.3166.

Barbour, Dennis L and Edward M Callaway (2008). "Excitatory Local Connections of Superficial Neurons in Rat Auditory Cortex". In: *Journal of Neuroscience* 28.44, pp. 11174–11185. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.2093-08.2008.

Barlow, H B (1972). "Single units and sensations: a neuron doctrin for perceptual psychology?" In: *Perception* 1, pp. 371–394. ISSN: 03010066. DOI: 10.1068/p010371. arXiv: 0810.2901.

Barnstedt, O. et al. (2015). "Functional Microarchitecture of the Mouse Dorsal Inferior Colliculus Revealed through In Vivo Two-Photon Calcium Imaging". In: *Journal of Neuroscience* 35.31, pp. 10927–10939. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.0103-15.2015.

Bathellier, Brice, Sui Poh, et al. (2013). "A multiplicative reinforcement learning model capturing learning dynamics and interindividual variability in mice". In: *Proceedings of the National Academy of Sciences of the United States of America* 110.49, pp. 19950–5. ISSN: 1091-6490. DOI: 10.

1073/pnas.1312125110/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1312125110.

Bathellier, Brice, Lyubov Ushakova, and Simon Rumpel (2012). "Discrete Neocortical Dynamics Predict Behavioral Categorization of Sounds". In: *Neuron* 76.2, pp. 435–449. ISSN: 08966273. DOI: 10.1016/j.neuron.2012.07.008.

Baxter, Mark G and David Gaffan (2007). "Asymmetry of attentional set in rhesus monkeys learning colour and shape discriminations". In: *Quarterly Journal of Experimental Psychology* 60.1, pp. 1–8. ISSN: 17470218. DOI: 10.1080/17470210600971485.

Behrens, Timothy E J et al. (2007). "Learning the value of information in an uncertain world". In: *Nature Neuroscience* 10.9, pp. 1214–1221. ISSN: 1097-6256. DOI: 10.1038/nn1954. arXiv: NIHMS150003.

Bellman, R.E. (1961). "Adaptive control processes: A guided tour". In: *Princeton University Press* 28, pp. 1–19. arXiv: arXiv:1302.6677v1.

Berditchevskaia, A, R D Cazé, and S R Schultz (2016). "Performance in a GO/NOGO perceptual task reflects a balance between impulsive and instrumental components of behaviour". In: *Scientific Reports* 6. ISSN: 20452322. DOI: 10.1038/srep27389.

Bishop, Christopher M (2006). *Pattern Recognition and Machine Learning*. Vol. 4. 4, p. 738. ISBN: 9780387310732. DOI: 10.1117/1.2819119. arXiv: 0-387-31073-8.

Bizley, Jennifer K, Kerry M M Walker, Fernando R Nodal, et al. (2013). "Auditory cortex represents both pitch judgments and the corresponding acoustic cues". In: *Current Biology* 23.7, pp. 620–625. ISSN: 09609822. DOI: 10.1016/j.cub.2013.03.003.

Bizley, Jennifer K, Kerry M M Walker, Bernard W Silverman, et al. (2009). "Interdependent Encoding of Pitch, Timbre, and Spatial Location in Auditory Cortex". In: *Journal of Neuroscience* 29.7, pp. 2064–2075. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.4755-08.2009.

Blackman, Derek (1974). *Operant conditioning: an experimental analysis of behaviour*. Routledge. ISBN: 1138299839.

Bouton, Mark E (2004). "Context and Behavioral Processes in Extinction". In: *Learning & Memory* 11.5, pp. 485–494. ISSN: 1072-0502. DOI: 10.1101/lm.78804. arXiv: /www.ncbi.nlm.nih.gov/pubmed/15466298 [http:].

Boyd, Stephen SP and Lieven Vandenberghe (2004). "Convex optimization". In: *Optimization Methods and Software* 25.3, pp. 487–487. ISSN: 10556788. DOI: 10.1080/10556781003625177. arXiv: 1111.6189v1.

Boyden, Edward S et al. (2005). "Millisecond-timescale, genetically targeted optical control of neural activity". In: *Nature Neuroscience* 8.9, pp. 1263–1268. ISSN: 1097-6256. DOI: 10.1038/nn1525. arXiv: 9809069v1 [arXiv:gr-qc].

Braun, Daniel A, Carsten Mehring, and Daniel M Wolpert (2010). *Structure learning in action*. DOI: 10.1016/j.bbr.2009.08.031.

Britten, K H et al. “A relationship between behavioral choice and the visual responses of neurons in macaque MT.” In: *Visual neuroscience* 13.1, pp. 87–100. ISSN: 0952-5238.

Britten, K. H. et al. (1996). “A relationship between behavioral choice and the visual responses of neurons in macaque MT”. In: *Visual Neuroscience* 13.01, pp. 87–100. ISSN: 0952-5238. DOI: 10.1017/S095252380000715X.

Brown, Jennifer, Kathleen A. Martin, and Joshua Dudman (2016). “Behavioral evidence for feedback gain control by the inhibitory microcircuit of the substantia nigra”. In: *bioRxiv*, p. 090209. DOI: 10.1101/090209.

Brown, Verity J and Eric M Bowman (2002). “Rodent models of prefrontal cortical function.” In: *Trends in neurosciences* 25.7, pp. 340–3. ISSN: 0166-2236.

Brunton, B. W., M. M. Botvinick, and C. D. Brody (2013). “Rats and Humans Can Optimally Accumulate Evidence for Decision-Making”. In: *Science* 340.6128, pp. 95–98. ISSN: 0036-8075. DOI: 10.1126/science.1233912.

Bureau, Ingrid, Francisca Saint Von Paul, and Karel Svoboda (2006). “Interdigitated paralemniscal and lemniscal pathways in the mouse barrel cortex”. In: *PLoS Biology* 4.12, pp. 2361–2371. ISSN: 15457885. DOI: 10.1371/journal.pbio.0040382. arXiv: arXiv:1208.5792v1.

Burgess, Christopher P., Armin Lak, Nicholas A. Steinmetz, Peter Zatka-Haas, Charu Bai Reddy, et al. (2017). “High-Yield Methods for Accurate Two-Alternative Visual Psychophysics in Head-Fixed Mice”. In: *Cell Re-*

ports 20.10, pp. 2513–2524. ISSN: 22111247. DOI: 10.1016/j.celrep.2017.08.047.

Burgess, Christopher P, Armin Lak, Nicholas A Steinmetz, Peter Zatka-Haas, Kenneth D Harris, et al. (2017). "High-Yield Methods for Accurate Two-Alternative Visual Psychophysics in Head-Fixed Mice". In: *Cell Reports* 20, pp. 2513–2524. DOI: 10.1016/j.celrep.2017.08.047.

Busse, Laura et al. (2011). "The Detection of Visual Contrast in the Behaving Mouse". In: *Journal of Neuroscience* 31.31, pp. 11351–11361. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.6689-10.2011. arXiv: arXiv:1011.1669v3.

Carandini, Matteo (2012). "From circuits to behavior: A bridge too far?" In: *Nature Neuroscience* 15.4, pp. 507–509. ISSN: 10976256. DOI: 10.1038/nn.3043.

Carandini, Matteo and Anne K Churchland (2013). "Probing perceptual decisions in rodents". In: *Nature Neuroscience* 16.7, pp. 824–831. ISSN: 1097-6256. DOI: 10.1038/nn.3410.

Carcea, Ioana et al. (2017). "Dynamics of cortical activity during behavioral engagement and auditory perception". In: *Nature Communications* 8, pp. 1–12. ISSN: 2041-1723. DOI: 10.1038/ncomms14412.

Charlesworth, Jonathan D et al. (2011). "Learning the microstructure of successful behavior." In: *Nature neuroscience* 14.3, pp. 373–380. ISSN: 1097-6256. DOI: 10.1038/nn.2748.

Chen, Tsai-Wen, Trevor J Wardill, et al. (2013). "Ultrasensitive fluorescent proteins for imaging neuronal activity". In: *Nature* 499.7458, pp. 295–300. ISSN: 0028-0836. DOI: 10.1038/nature12354. arXiv: arXiv:1011.1669v3.

Chen, Xiaowei, Ulrich Leischner, et al. (2011). "Functional mapping of single spines in cortical neurons in vivo". In: *Nature* 475.7357, pp. 501–505. ISSN: 0028-0836. DOI: 10.1038/nature10193.

Chen, Z (2015). *Advanced State Space Methods for Neural and Clinical Data*, pp. 1–374. ISBN: 9781139941433. DOI: 10.1017/CBO9781139941433.

Churchland, Mark M et al. (2012). "Neural Population Dynamics During Reaching". In: *Nature* 487.7405, pp. 1–20. ISSN: 1878-5832. DOI: 10.1016/j.micinf.2011.07.011.Innate. arXiv: NIHMS150003.

Clark, L L, T G Lansford, and Dallenbach K M (1960). "Repetition and associative learning." In: *The American journal of psychology* 73, pp. 22–40. ISSN: 0002-9556.

Clause, Amanda, Tuan Nguyen, and Karl Kandler (2011). "An acoustic startle-based method of assessing frequency discrimination in mice". In: *Journal of Neuroscience Methods* 200.1, pp. 63–67. ISSN: 01650270. DOI: 10.1016/j.jneumeth.2011.05.027.

Cohen, Marlene R and William T Newsome (2008). "Context-Dependent Changes in Functional Circuitry in Visual Area MT". In: *Neuron* 60.1, pp. 162–173. ISSN: 08966273. DOI: 10.1016/j.neuron.2008.08.007.

Cohen, Marlene R and William T Newsome (2009). "Estimates of the Contribution of Single Neurons to Perception Depend on Timescale and Noise Correlation". In: *Journal of Neuroscience* 29.20, pp. 6635–6648. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.5179-08.2009.

Cossell, Lee et al. (2015). "Functional organization of excitatory synaptic strength in primary visual cortex". In: *Nature* 518. DOI: 10.1038/nature14182.

Cotton, R. James et al. (2013). "Three-dimensional mapping of microcircuit correlation structure". In: *Frontiers in Neural Circuits* 7.October, pp. 1–13. ISSN: 1662-5110. DOI: 10.3389/fncir.2013.00151.

Crump, Matthew J.C., John V. McDonnell, and Todd M Gureckis (2013). "Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research". In: *PLoS ONE* 8.3. ISSN: 19326203. DOI: 10.1371/journal.pone.0057410.

Cunningham, John P and Byron M Yu (2014). "Dimensionality reduction for large-scale neural recordings". In: *Nature Neuroscience* 17.11, pp. 1500–1509. ISSN: 1097-6256. DOI: 10.1038/nn.3776.

Dana, Hod et al. (2016). "Sensitive red protein calcium indicators for imaging neural activity". In: *eLife* 5.MARCH2016, e12727. ISSN: 2050084X. DOI: 10.7554/eLife.12727.

Dantzker, J. L. and E. M. Callaway (2000). "Laminar sources of synaptic input to cortical inhibitory interneurons and pyramidal neurons". In: *Nature Neuroscience* 3.7, pp. 701–707. ISSN: 1097-6256. DOI: 10.1038/76656.

Daw, Nathaniel D (2011). "Trial-by-trial data analysis using computational models". In: *Decision Making, Affect, and Learning: Attention and Performance XXIII*. ISBN: 9780191725623. DOI: 10.1093/acprof:oso/9780199600434.003.0001.

Daw, Nathaniel D and Aaron C Courville (2008). "The pigeon as particle filter". In: *Advances in neural information processing systems*, pp. 369–376. DOI: 10.1.1.73.3246.

Daw, Nathaniel D, Aaron C Courville, and David S. Touretzky (2006). "Representation and Timing in Theories of the Dopamine System". In: *Neural Computation* 18.7, pp. 1637–1677. ISSN: 0899-7667. DOI: 10.1162/neco.2006.18.7.1637.

Daw, Nathaniel D et al. (2011a). "Model-based influences on humans' choices and striatal prediction errors". In: *Neuron* 69.6, pp. 1204–1215. ISSN: 08966273. DOI: 10.1016/j.neuron.2011.02.027.

Daw, Nathaniel D. et al. (2011b). "Model-based influences on humans' choices and striatal prediction errors". In: *Neuron* 69.6, pp. 1204–1215. ISSN: 08966273. DOI: 10.1016/j.neuron.2011.02.027.

Dayan, Peter. "Motivated Reinforcement Learning". In:

Dayan, Peter and L F Abbott (2013). "Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems". In: *Journal of Chemical Information and Modeling* 53.9, pp. 1689–1699. ISSN: 1098-6596. DOI: 10.1017/CBO9781107415324.004. arXiv: arXiv:1011.1669v3.

Dayan, Peter, Sham Kakade, and P. Read Montague (2000). "Learning and selective attention." In: *Nature neuroscience* 3 Suppl.november, pp. 1218–1223. ISSN: 1097-6256. DOI: 10.1038/81504.

De Hoz, Livia and Israel Nelken (2014). "Frequency tuning in the behaving mouse: Different bandwidths for discrimination and generalization". In: *PLoS ONE* 9.3. ISSN: 19326203. DOI: 10.1371/journal.pone.0091676.

Deisenroth, Marc Peter (2011). "A Survey on Policy Search for Robotics". In: *Foundations and Trends in Robotics* 2.1, pp. 1–142. ISSN: 1935-8253. DOI: 10.1561/2300000021.

Deisenroth, Marc Peter, Gerhard Neumann, and Jan Peters (2011). "A Survey on Policy Search for Robotics". In: *Foundations and Trends R in Robotics* 2, pp. 1–2. DOI: 10.1561/2300000021.

Dempster, A.P., N.M. Laird, and Donald B Rubin (1977). "Maximum likelihood from incomplete data via the EM algorithm". In: *Journal of the Royal Statistical Society Series B Methodological* 39.1, pp. 1–38. ISSN: 00359246. DOI: http://dx.doi.org/10.2307/2984875. arXiv: 0710.5696v2.

Denk, W, J. Strickler, and W. Webb (1990). "Two-photon laser scanning fluorescence microscopy". In: *Science* 248.4951, pp. 73–76. ISSN: 0036-8075. DOI: 10.1126/science.2321027.

Desrochers, Theresa M et al. (2010). "Optimal habits can develop spontaneously through sensitivity to local cost". In: *Proceedings of the Na-*

tional Academy of Sciences 107.47, pp. 20512–20517. ISSN: 0027-8424. DOI: 10.1073/pnas.1013470107.

Ding, Long and Joshua I Gold (2010). "Caudate Encodes Multiple Computations for Perceptual Decisions". In: *Journal of Neuroscience* 30.47, pp. 15747–15759. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.2894-10.2010.

— (2012). "Neural correlates of perceptual decision making before, during, and after decision commitment in monkey frontal eye field". In: *Cerebral Cortex* 22.5, pp. 1052–1067. ISSN: 10473211. DOI: 10.1093/cercor/bhr178.

Django Software Foundation (2017). *Django (Version 1.5)*.

Doll, Bradley B, Dylan A Simon, and Nathaniel D Daw (2012). *The ubiquity of model-based reinforcement learning*. DOI: 10.1016/j.conb.2012.08.003. arXiv: NIHMS150003.

Douglas, R J, C Koch, et al. (1995). "Recurrent excitation in neocortical circuits." In: *Science (New York, N.Y.)* 269.5226, pp. 981–5. ISSN: 0036-8075.

Douglas, Rodney J and Kevan A.C. Martin (2004). "Neuronal Circuits of the Neocortex". In: *Annual Review of Neuroscience* 27.1, pp. 419–451. ISSN: 0147-006X. DOI: 10.1146/annurev.neuro.27.070203.144152.

Downer, Joshua D, Mamiko Niwa, and Mitchell L Sutter (2015). "Task Engagement Selectively Modulates Neural Correlations in Primary Auditory

Cortex". In: *Journal of Neuroscience* 35.19, pp. 7565–7574. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.4094-14.2015.

Durstewitz, Daniel et al. (2010). "Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning". In: *Neuron* 66.3, pp. 438–448. ISSN: 08966273. DOI: 10.1016/j.neuron.2010.03.029.

Dykstra, Andrew R et al. (2012). "Dissociation of Detection and Discrimination of Pure Tones following Bilateral Lesions of Auditory Cortex". In: *PLoS ONE* 7.9. ISSN: 19326203. DOI: 10.1371/journal.pone.0044602.

Egnor, S.E. Roian and Kristin Branson (2016). "Computational Analysis of Behavior". In: *Annual Review of Neuroscience* 39.1, pp. 217–236. ISSN: 0147-006X. DOI: 10.1146/annurev-neuro-070815-013845.

Ehret, Günter (1975). "Frequency and intensity difference limens and nonlinearities in the ear of the housemouse (Mus musculus)". In: *Journal of Comparative Physiology* 102.4, pp. 321–336. ISSN: 03407594. DOI: 10.1007/BF01464344.

Elber-Dorozko, Lotem and Yonatan Loewenstein (2017). "Striatal Action-Value Neurons Reconsidered". In: *bioRxiv*, pp. 1–34. DOI: 10.1101/087502.

Elsayed, Gamaleldin F and John P Cunningham (2017). "Structure in neural population recordings: an expected byproduct of simpler phenomena?" In: *Nature Neuroscience*. ISSN: 1097-6256. DOI: 10.1038/nn.4617.

Ernst, Marc O. and Martin S. Banks (2002). "Humans integrate visual and haptic information in a statistically optimal fashion". In: *Nature* 415.6870, pp. 429–433. ISSN: 00280836. DOI: 10.1038/415429a.

Faisal, A Aldo and Daniel M Wolpert (2009). "Near Optimal Combination of Sensory and Motor Uncertainty in Time During a Naturalistic Perception-Action Task". In: *Journal of Neurophysiology* 101.4, pp. 1901–1912. ISSN: 0022-3077. DOI: 10.1152/jn.90974.2008.

Fetz, E. E. (1969). "Operant Conditioning of Cortical Unit Activity". In: *Science* 163.3870, pp. 955–958. ISSN: 0036-8075. DOI: 10.1126/science.163.3870.955.

Fetz, Eberhard E. (1992). *Are movement parameters recognizable coded in the activity of single neurons?.pdf*. DOI: 10.1017/S0140525X00072599.

Francis, Nikolas A. and Patrick O. Kanold (2017). "Automated Operant Conditioning in the Mouse Home Cage". In: *Frontiers in Neural Circuits* 11, p. 10. ISSN: 1662-5110. DOI: 10.3389/fncir.2017.00010.

Friston, Karl J et al. (2010). "Action and behavior: A free-energy formulation". In: *Biological Cybernetics* 102.3, pp. 227–260. ISSN: 03401200. DOI: 10.1007/s00422-010-0364-z.

Fritz, J. B. (2005). "Differential Dynamic Plasticity of A1 Receptive Fields during Multiple Spectral Tasks". In: *Journal of Neuroscience* 25.33, pp. 7623–7635. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.1318-05.2005.

Fritz, Jonathan B., Stephen V. David, et al. (2010). "Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in fer-

ret frontal cortex". In: *Nature Neuroscience* 13.8, pp. 1011–1019. ISSN: 10976256. DOI: 10.1038/nn.2598.

Fritz, Jonathan, Shihab Shamma, et al. (2003). "Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex". In: *Nature Neuroscience* 6.11, pp. 1216–1223. ISSN: 10976256. DOI: 10.1038/nn1141.

Fründ, Ingo, Felix A. Wichmann, and Jakob H. Macke (2014a). "Quantifying the effect of intertrial dependence on perceptual decisions." In: *Journal of Vision* 14.7, pp. 1–16. ISSN: 1534-7362. DOI: 10.1167/14.7.9.

Fründ, Ingo, Felix A Wichmann, and Jakob H Macke (2014b). "Quantifying the effect of intertrial dependence on perceptual decisions." In: *Journal of Vision* 14.7, pp. 1–16. ISSN: 1534-7362. DOI: 10.1167/14.7.9.

Gallistel, Charles R, Stephen Fairhurst, and Peter Balsam (2004). "The learning curve: Implications of a quantitative analysis". In: *Proceedings of the National Academy of Sciences* 101.36, pp. 13124–13131. ISSN: 0027-8424. DOI: 10.1073/pnas.0404965101.

Garner, Joseph P et al. (2006). "Animal neuropsychology: Validation of the Intra-Dimensional Extra-Dimensional set shifting task for mice". In: *Behavioural Brain Research* 173.1, pp. 53–61. ISSN: 01664328. DOI: 10.1016/j.bbr.2006.06.002.

Gelman, Andrew, John B Carlin, et al. (2004). *Bayesian Data Analysis*, p. 696. ISBN: 978-1-4398-9820-8. DOI: 10.1007/s13398-014-0173-7.2. arXiv: arXiv:1011.1669v3.

Gelman, Andrew and Cosma Rohilla Shalizi (2013). "Philosophy and the practice of Bayesian statistics". In: *British Journal of Mathematical and Statistical Psychology* 66.1, pp. 8–38. ISSN: 00071102. DOI: 10.1111/j.2044-8317.2011.02037.x. arXiv: 1006.3868.

Gentner, Timothy Q. and Daniel Margoliash (2003). "Neuronal populations and single cells representing learned auditory objects". In: *Nature* 424.6949, pp. 669–674. ISSN: 00280836. DOI: 10.1038/nature01731.

Geramifard, Alborz et al. (2013). "A Tutorial on Linear Function Approximators for Dynamic Programming and Reinforcement Learning". In: *Foundations and Trends in Machine Learning* 6.4, pp. 375–454. DOI: 10.1561/2200000042.

Gershman, Samuel J and Nathaniel D Daw (2017). "Reinforcement Learning and Episodic Memory in Humans and Animals: An Integrative Framework". In: *Annual Review of Psychology* 68.1, pp. 101–128. ISSN: 0066-4308. DOI: 10.1146/annurev-psych-122414-033625.

Gershman, Samuel J and Yael Niv (2010). *Learning latent structure: Carving nature at its joints*. DOI: 10.1016/j.conb.2010.02.008.

Gershman, Samuel J, Kenneth A Norman, and Yael Niv (2015). "Discovering latent causes in reinforcement learning". In: *Current Opinion in Behavioral Sciences* 5, pp. 43–50. ISSN: 23521546. DOI: 10.1016/j.cobeha.2015.07.007. arXiv: NIHMS150003.

Geschwind, Daniel H and Genevieve Konopka (2009). *Neuroscience in the era of functional genomics and systems biology*. DOI: 10.1038/nature08537. arXiv: NIHMS150003.

Gess, Austen et al. (2011). "Automated auditory recognition training and testing". In: *Animal Behaviour* 82.2, pp. 285–293. ISSN: 00033472. DOI: 10.1016/j.anbehav.2011.05.003. arXiv: NIHMS150003.

Gilbert, Cd, Mariano Sigman, and Re Crist (2001). "The neural basis of perceptual learning". In: *Neuron* 31, pp. 681–697. ISSN: 0896-6273. DOI: 10.1016/S0896-6273(01)00424-X.

Gimenez, Tyler L, Maja Lorenc, and Santiago Jaramillo (2015). "Adaptive categorization of sound frequency does not require the auditory cortex in rats". In: *Journal of Neurophysiology* 114.2, pp. 1137–1145. ISSN: 0022-3077. DOI: 10.1152/jn.00124.2015.

Gold, Joshua I and Michael N Shadlen (2007). "The Neural Basis of Decision Making". In: *Annual Review of Neuroscience* 30.1, pp. 535–574. ISSN: 0147-006X. DOI: 10.1146/annurev.neuro.29.051605.113038. arXiv: NIHMS150003.

Gomez-Marin, Alex and Zachary F Mainen (2016). *Expanding perspectives on cognition in humans, animals, and machines*. DOI: 10.1016/j.conb.2016.01.011.

Gomez-Marin, Alex, Joseph J Paton, et al. (2014). "Big behavioral data: psychology, ethology and the foundations of neuroscience." In: *Nature*

neuroscience 17.11, pp. 1455–1462. ISSN: 1546-1726. DOI: 10.1038/nn.3812.

Green, David M. and John A. Swets (1966). "Signal detection theory and psychophysics". In: *Society* 1, p. 521. ISSN: 00225002. DOI: 10.1901/jeab.1969.12-475.

Guizar-Sicairos, Manuel, Samuel T. Thurman, and James R. Fienup (2008). "Efficient subpixel image registration algorithms". In: *Optics Letters* 33.2, p. 156. ISSN: 0146-9592. DOI: 10.1364/OL.33.000156. arXiv: 78 [0146-9592].

Guo, Jian Zhong, Austin R. Graves, et al. (2015). "Cortex commands the performance of skilled movement". In: *eLife* 4.DECEMBER2015, pp. 1–18. ISSN: 2050084X. DOI: 10.7554/eLife.10774. arXiv: arXiv:1011.1669v3.

Guo, Wei, Anna R Chambers, et al. "Behavioral/Systems/Cognitive Robustness of Cortical Topography across Fields, Laminae, Anesthetic States, and Neurophysiological Signal Types". In: DOI: 10.1523/JNEUROSCI.0065-12.2012.

Guo, Wei, Amanda R. Clause, et al. (2017). "A Corticothalamic Circuit for Dynamic Switching between Feature Detection and Discrimination". In: *Neuron* 95.1, 180–194.e5. ISSN: 08966273. DOI: 10.1016/j.neuron.2017.05.019.

Guo, Zengcai V, S Andrew Hires, et al. (2014). "Procedures for Behavioral Experiments in Head-Fixed Mice". In: *PLoS ONE* 9.2. Ed. by Sidney

Arthur Simon, e88678. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0088678.

Guo, ZengcaiV, Nuo Li, et al. (2014). "Flow of cortical activity underlying a tactile decision in mice". In: *Neuron* 81.1, pp. 179–194. ISSN: 08966273. DOI: 10.1016/j.neuron.2013.10.020. arXiv: 1512.01655.

Haefner, Ralf M et al. (2013). "Inferring decoding strategies from choice probabilities in the presence of correlated variability". In: *Nature Neuroscience* 16.2, pp. 235–242. ISSN: 1097-6256. DOI: 10.1038/nn.3309.

Harding, Luke and Leonard Barden (1997). *Deep Blue win a giant step for computerkind*.

Harlow, Harry F. (1949). "The formation of learning sets." In: *Psychological Review* 56.1, pp. 51–65. ISSN: 0033-295X. DOI: 10.1037/h0062474.

Harris, Kenneth D. and Thomas D. Mrsic-Flogel (2013). "Cortical connectivity and sensory coding". In: *Nature* 503.7474, pp. 51–58. ISSN: 00280836. DOI: 10.1038/nature12654.

Harris, Kenneth D and Alexander Thiele (2011). "Cortical state and attention". In: *Nature Reviews Neuroscience* 12.9, pp. 509–523. ISSN: 1471-003X. DOI: 10.1038/nrn3084.

Hartline, HK (1938). "The response of single optic nerve fibers of the vertebrate eye to illumination of the retina." In: *American Journal of Physiology* 121, pp. 400–415. ISSN: 0002-9394. DOI: 10.1234/12345678.

Hassabis, Demis et al. (2017). *Neuroscience-Inspired Artificial Intelligence*. DOI: 10.1016/j.neuron.2017.06.011. arXiv: 1404.7282.

He, Kaiming et al. (2015). "Deep Residual Learning for Image Recognition". In: *Multimedia Tools and Applications*, pp. 1–17. ISSN: 1380-7501. DOI: 10.1007/s11042-017-4440-4. arXiv: 1512.03385.

Heffner, Henry and Bruce Masterton (1980). "Hearing in Glires: Domestic rabbit, cotton rat, feral house mouse, and kangaroo rat". In: *The Journal of the Acoustical Society of America* 68.6, pp. 1584–1599. ISSN: 0001-4966. DOI: 10.1121/1.385213.

Heitz, Richard P and Jeffrey D Schall (2012). "Neural Mechanisms of Speed-Accuracy Tradeoff". In: *Neuron* 76.3, pp. 616–628. ISSN: 08966273. DOI: 10.1016/j.neuron.2012.08.030. arXiv: NIHMS150003.

Herrero, Jose L et al. (2008). "Acetylcholine cpntributes to attentional modulation in V1 by muscarinic receptors (Supplementary Materials)". In: *Nature* 454.7208, pp. 1–22. ISSN: 00085472. DOI: 10.1038/nature. arXiv: NIHMS150003.

Herrnstein, R J (1961). "Relative and absolute strength of response as a function of frequency of reinforcement, 12". In: *Journal of the Experimental Analysis of Behavior* 4.3, pp. 267–272. ISSN: 0022-5002. DOI: 10.1901/jeab.1961.4-267.

Hinton, G. E. (2006). "Reducing the Dimensionality of Data with Neural Networks". In: *Science* 313.5786, pp. 504–507. ISSN: 0036-8075. DOI: 10.1126/science.1127647.

Hinton, Geoffrey (2014). *Geoffrey Hinton AMA*.

Hires, Samuel Andrew et al. (2015). "Low-noise encoding of active touch by layer 4 in the somatosensory cortex". In: *eLife* 4.AUGUST2015, pp. 14921–14926. ISSN: 2050084X. DOI: 10.7554/eLife.06619.

Hopfield, J J (1982). "Neural networks and physical systems with emergent collective computational abilities." In: *Proceedings of the National Academy of Sciences* 79.8, pp. 2554–2558. ISSN: 0027-8424. DOI: 10.1073/pnas.79.8.2554.

Houlsby, Neil M.T. et al. (2013). "Cognitive Tomography Reveals Complex, Task-Independent Mental Representations". In: *Current Biology* 23.21, pp. 2169–2175. ISSN: 09609822. DOI: 10.1016/j.cub.2013.09.012.

Howe, Mark W et al. (2013). "Prolonged dopamine signalling in striatum signals proximity and value of distant rewards". In: *Nature* 500.7464, pp. 575–579. ISSN: 0028-0836. DOI: 10.1038/nature12475.

Huang, Wei et al. (2013). "MTORC2 controls actin polymerization required for consolidation of long-term memory". In: *Nature Neuroscience* 16.4, pp. 441–448. ISSN: 10976256. DOI: 10.1038/nn.3351. arXiv: NIHMS150003.

Hubel, D. H. and T. N. Wiesel (1962). "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex". In: *The Journal of Physiology* 160.1, pp. 106–154. ISSN: 00223751. DOI: 10.1113/jphysiol.1962.sp006837. arXiv: arXiv:1011.1669v3.

Hubel, David H and Torsten N Wiesel (1965). "Receptive Fields and Functional Architecture in Two Nonstriate Visual Areas (18 and 19) of the Cat." In: *Journal of neurophysiology* 28, pp. 229–289. ISSN: 0022-3077.

Huber, D, D A Gutnisky, et al. (2012). "Multiple dynamic representations in the motor cortex during sensorimotor learning". In: *Nature* 484.7395, pp. 473–478. ISSN: 0028-0836. DOI: 10.1038/nature11039.

Huber, Daniel, Leopoldo Petreanu, et al. (2008). "Sparse optical microstimulation in barrel cortex drives learned behaviour in freely moving mice". In: *Nature* 451.7174, pp. 61–64. ISSN: 14764687. DOI: 10.1038/nature06445.

Huh, Namjung et al. (2009). "Model-based reinforcement learning under concurrent schedules of reinforcement in rodents". In: *Learning & Memory* 16.5, pp. 315–323. ISSN: 1072-0502. DOI: 10.1101/lm.1295509.

Hunter, John D. (2007). "Matplotlib: A 2D graphics environment". In: *Computing in Science and Engineering* 9.3, pp. 99–104. ISSN: 15219615. DOI: 10.1109/MCSE.2007.55. arXiv: 0402594v3 [arXiv:cond-mat].

Iigaya, Kiyohito et al. (2017). "Learning Fast And Slow: Deviations From The Matching Law Can Reflect An Optimal Strategy Under Uncertainty". In: *bioRxiv*. DOI: 10.1101/141309.

Inagaki, Hidehiko K et al. (2017). "Discrete attractor dynamics underlying selective persistent activity in frontal cortex". In: *BioRxiv*. DOI: 10.1101/203448. arXiv: /doi.org/10.1101/203448 [https:].

Ison, James R, Paul D Allen, and William E. O'Neill (2007). "Age-related hearing loss in C57BL/6J mice has both frequency-specific and non-frequency-specific components that produce a hyperacusis-like exaggeration of the acoustic startle reflex". In: *JARO - Journal of the Association*

for Research in Otolaryngology 8.4, pp. 539–550. ISSN: 15253961. DOI: 10.1007/s10162-007-0098-3.

Ito, Makoto and Kenji Doya (2009). "Validation of Decision-Making Models and Analysis of Decision Variables in the Rat Basal Ganglia". In: *Journal of Neuroscience* 29.31, pp. 9861–9874. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.6157-08.2009.

Jaramillo, Santiago, Katharine Borges, and Anthony M Zador (2014). "Auditory Thalamus and Auditory Cortex Are Equally Modulated by Context during Flexible Categorization of Sounds". In: *Journal of Neuroscience* 34.15, pp. 5291–5301. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.4888-13.2014.

Jaramillo, Santiago and Anthony M. Zador (2014). "Mice and rats achieve similar levels of performance in an adaptive decision-making task". In: *Frontiers in Systems Neuroscience* 8.September, pp. 1–11. ISSN: 1662-5137. DOI: 10.3389/fnsys.2014.00173. arXiv: NIHMS150003.

Jhuang, Hueihan et al. (2010). "Automated home-cage behavioural phenotyping of mice". In: *Nature Communications* 1.6. ISSN: 20411723. DOI: 10.1038/ncomms1064.

Ji, Xu-Ying et al. "Thalamocortical Innervation Pattern in Mouse Auditory and Visual Cortex: Laminar and Cell-Type Specificity". In: DOI: 10.1093/cercor/bhv099.

Jones, Allan R, Caroline C Overly, and Susan M Sunkin (2009). *The allen brain atlas: 5 years and beyond*. DOI: 10.1038/nrn2722.

Jun, James J. et al. (2017). "Fully Integrated Silicon Probes for High-Density Recording of Neural Activity". In: *Nature* in press.7679, pp. 232–236. ISSN: 0028-0836. DOI: 10.1038/nature24636.

Kaas, J.H. (2007). "Evolution of Nervous Systems". In: *Evolution of Nervous Systems* 1985, pp. 499–516. ISSN: 01662236. DOI: 10.1016/B0-12-370878-8/00089-6.

Kao, Jonathan C et al. (2015). "Single-trial dynamics of motor cortex and their applications to brain-machine interfaces". In: *Nature Communications* 6, p. 7759. ISSN: 2041-1723. DOI: 10.1038/ncomms8759.

Katz, Leor et al. (2016). "Dissociated functional significance of choice-related activity across the primate dorsal stream". In: *Nature* 535.7611, Salt Lake City USA. ISSN: 0028-0836. DOI: 10.1038/nature18617. arXiv: 15334406.

Kawai, Risa et al. (2015a). "Motor Cortex Is Required for Learning but Not for Executing a Motor Skill". In: *Neuron* 86.3, pp. 800–812. ISSN: 10974199. DOI: 10.1016/j.neuron.2015.03.024.

— (2015b). "Motor Cortex Is Required for Learning but Not for Executing a Motor Skill". In: *Neuron* 86.3, pp. 800–812. ISSN: 10974199. DOI: 10.1016/j.neuron.2015.03.024.

Kemp, Charles and Joshua B Tenenbaum (2009). "Structured statistical models of inductive reasoning." In: *Psychological Review* 116.1, pp. 20–58. ISSN: 1939-1471. DOI: 10.1037/a0014282.

Kepecs, Adam et al. (2008). "Neural correlates, computation and behavioural impact of decision confidence". In: *Nature* 455.7210, pp. 227–231. ISSN: 00280836. DOI: 10.1038/nature07200. arXiv: NIHMS150003.

Kerlin, Aaron M et al. (2010). "Broadly Tuned Response Properties of Diverse Inhibitory Neuron Subtypes in Mouse Visual Cortex". In: *Neuron* 67.5, pp. 858–871. ISSN: 08966273. DOI: 10.1016/j.neuron.2010.08.002.

Kile, Benjamin T. and Douglas J. Hilton (2005). "The art and design of genetic screens: mouse". In: *Nature Reviews Genetics* 6.7, pp. 557–567. ISSN: 1471-0056. DOI: 10.1038/nrg1636.

Kim, Sung Soo, Hervé Rouault, et al. (2017). "Ring attractor dynamics in the Drosophila central brain". In: *Science* 356.6340, pp. 849–853. ISSN: 0036-8075. DOI: 10.1126/science.aal4835.

Kim, Taekeun, Won Chan Oh, et al. (2016). "Emergence of functional subnetworks in layer 2/3 cortex induced by sequential spikes in vivo". In: *Proceedings of the National Academy of Sciences* 113.10, E1372–E1381. ISSN: 0027-8424. DOI: 10.1073/pnas.1513410113.

King, Andrew J, Victoria M Bajo, et al. (2007). "Physiological and behavioral studies of spatial coding in the auditory cortex". In: *Hearing Research* 229.1-2, pp. 106–115. ISSN: 03785955. DOI: 10.1016/j.heares.2007.01.001.

King, Andrew J and Israel Nelken (2009). *Unraveling the principles of auditory cortical processing: Can we learn from the visual system?* DOI: 10.1038/nn.2308.

Kirkpatrick, James et al. (2017). "Overcoming catastrophic forgetting in neural networks". In: *Proceedings of the National Academy of Sciences* 114.13, pp. 3521–3526. ISSN: 0027-8424. DOI: 10.1073/pnas.1611835114. arXiv: 1612.00796.

Klein-Flugge, M. C. et al. (2016). "Neural Signatures of Value Comparison in Human Cingulate Cortex during Decisions Requiring an Effort-Reward Trade-off". In: *Journal of Neuroscience* 36.39, pp. 10002–10015. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.0292-16.2016.

Kluyver, Thomas et al. (2016). *Jupyter Notebooks—a publishing format for reproducible computational workflows*, pp. 87–90. ISBN: 9781614996491. DOI: 10.3233/978-1-61499-649-1-87.

Ko, Ho et al. (2013). "The emergence of functional microcircuits in visual cortex". In: *Nature* 496.7443, pp. 96–100. ISSN: 0028-0836. DOI: 10.1038/nature12015. arXiv: arXiv:1011.1669v3.

Koay, S et al. (2015). "A mouse navigation and accumulation of visual evidence task". In: *Society for Neuroscience Annual Meeting*.

Kolling, Nils and Thomas Akam (2017). *(Reinforcement?) Learning to forage optimally*. DOI: 10.1016/j.conb.2017.08.008.

Krakauer, John W et al. (2017). "Neuroscience Needs Behavior: Correcting a Reductionist Bias". In: *Neuron* 93.3, pp. 480–490. ISSN: 08966273. DOI: 10.1016/j.neuron.2016.12.041.

Kuchibhotla, Kishore V et al. (2016a). "Parallel processing by cortical inhibition enables context-dependent behavior". In: *Nature Neuroscience* 20.1, pp. 62–71. ISSN: 1097-6256. DOI: 10.1038/nn.4436.

— (2016b). "Parallel processing by cortical inhibition enables context-dependent behavior". In: *Nature Neuroscience* 20.1, pp. 62–71. ISSN: 1097-6256. DOI: 10.1038/nn.4436.

Kulkarni, Jayant E. and Liam Paninski (2007). "Common-input models for multiple neural spike-train data". In: *Network: Computation in Neural Systems* 18.4, pp. 375–407. ISSN: 0954-898X. DOI: 10.1080/09548980701625173.

Kwon, Sung Eun et al. (2016). "Sensory and decision-related activity propagate in a cortical feedback loop during touch perception". In: *Nature Neuroscience* 19.9, pp. 1243–1249. ISSN: 1097-6256. DOI: 10.1038/nn.4356.

Lake, Brenden M, Ruslan Salakhutdinov, and Joshua B Tenenbaum (2015). "Human-level concept learning through probabilistic program induction". In: *Science* 350.6266, pp. 1332–1338. ISSN: 0036-8075. DOI: 10.1126/science.aab3050. arXiv: science.aab3050 [10.1126].

Latimer, Kenneth W et al. (2015). "Single-trial spike trains in parietal cortex reveal discrete steps during decision-making". In: *Science* 349.6244, pp. 184–187. ISSN: 0036-8075. DOI: 10.1126/science.aaa4056. arXiv: arXiv:1402.7245v1.

Lau, Brian and Paul W Glimcher (2005). "Dynamic Response-by-Response Models of Matching Behavior in Rhesus Monkeys". In: *Journal of the Experimental Analysis of Behavior* 84.3, pp. 555–579. ISSN: 0022-5002. DOI: 10.1901/jeab.2005.110-04.

Law, Chi-Tat and Joshua I Gold (2008). "Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area". In: *Nature Neuroscience* 11.4, pp. 505–513. ISSN: 1097-6256. DOI: 10.1038/nn2070.

LeCun, Yann et al. (2015). "Deep learning". In: *Nature* 521.7553, pp. 436–444. ISSN: 0028-0836. DOI: 10.1038/nature14539. arXiv: arXiv:1312.6184v5.

LeDoux, Joseph E., Claudia R Farb, and Lizabeth M Romanski (1991). "Overlapping projections to the amygdala and striatum from auditory processing areas of the thalamus and cortex". In: *Neuroscience Letters* 134.1, pp. 139–144. ISSN: 03043940. DOI: 10.1016/0304-3940(91)90526-Y.

Leong, Yuan Chang et al. (2017). "Dynamic Interaction between Reinforcement Learning and Attention in Multidimensional Environments". In: *Neuron* 93.2, pp. 451–463. ISSN: 08966273. DOI: 10.1016/j.neuron.2016.12.040.

Levy, Robert B and Alex D Reyes (2012). "Spatial Profile of Excitatory and Inhibitory Synaptic Connectivity in Mouse Primary Auditory Cortex". In: *Journal of Neuroscience* 32.16, pp. 5609–5619. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.5158-11.2012.

Lewi, Jeremy, Robert Butera, and Liam Paninski (2009). "Sequential Optimal Design of Neurophysiology Experiments". In: *Neural Computation* 21.3, pp. 619–687. ISSN: 0899-7667. DOI: 10.1162/neco.2008.08-07-594.

Lewi, Jeremy, David M Schneider, et al. (2011). "Automating the design of informative sequences of sensory stimuli". In: *Journal of Computational Neuroscience* 30.1, pp. 181–200. ISSN: 09295313. DOI: 10.1007/s10827-010-0248-1.

Li, Jingcheng, Xiang Liao, et al. (2017). "Primary Auditory Cortex is Required for Anticipatory Motor Response". In: *Cerebral Cortex* 27.6, pp. 3254–3271. ISSN: 14602199. DOI: 10.1093/cercor/bhx079.

Li, Ling Yun, Ya Tang Li, et al. (2013). "Intracortical multiplication of thalamocortical signals in mouse auditory cortex". In: *Nature Neuroscience* 16.9, pp. 1179–1181. ISSN: 10976256. DOI: 10.1038/nn.3493. arXiv: NIHMS150003.

Li, Nuo, Kayvon Daie, et al. (2016). "Robust neuronal dynamics in premotor cortex during motor planning". In: *Nature* 532.7600, pp. 459–464. ISSN: 0028-0836. DOI: 10.1038/nature17643.

Li, Yi and Joshua Tate Dudman (2013). "Mice infer probabilistic models for timing". In: *Proceedings of the National Academy of Sciences* 110.42, pp. 17154–17159. ISSN: 0027-8424. DOI: 10.1073/pnas.1310666110.

Li, Yuxi (2017). "Deep Reinforcement Learning: An Overview". In: pp. 1–30. arXiv: arXiv:1701.07274v5.

Lien, Anthony D. and Massimo Scanziani (2013). "Tuned thalamic excitation is amplified by visual cortical circuits". In: *Nature Neuroscience* 16.9, pp. 1315–1323. ISSN: 10976256. DOI: 10.1038/nn.3488.

Lima, Susana Q and Gero Miesenböck (2005). "Remote control of behavior through genetically targeted photostimulation of neurons". In: *Cell* 121.1, pp. 141–152. ISSN: 00928674. DOI: 10.1016/j.cell.2005.02.004.

Lin, John Y. et al. (2009). "Characterization of engineered channelrhodopsin variants with improved properties and kinetics". In: *Biophysical Journal* 96.5, pp. 1803–1814. ISSN: 00063495. DOI: 10.1016/j.bpj.2008.11.034. arXiv: arXiv:1011.1669v3.

Linderman, Scott W and Samuel J Gershman (2017). "Using computational theory to constrain statistical models of neural data A Theory-driven Pipeline for Neural Data Analysis". In: pp. 1–14. DOI: 10.1101/104737.

Liu, Haoyang, Keqin Liu, and Qing Zhao (2013). "Learning in a changing world: Restless multiarmed bandit with unknown dynamics". In: *IEEE Transactions on Information Theory* 59.3, pp. 1902–1916. ISSN: 00189448. DOI: 10.1109/TIT.2012.2230215. arXiv: 1011.4969.

Luck, Steven J and Edward K Vogel (2013). *Visual working memory capacity: From psychophysics and neurobiology to individual differences*. DOI: 10.1016/j.tics.2013.06.006. arXiv: arXiv:1011.1669v3.

Macke, Jakob H et al. (2011). "Empirical models of spiking in neural populations". In: *Advances in Neural Information Processing Systems 25*, pp. 1–8.

Maia, Tiago V. (2009). "Reinforcement learning, conditioning, and the brain: Successes and challenges". In: *Cognitive, Affective, & Behavioral Neuroscience* 9.4, pp. 343–364. ISSN: 1530-7026. DOI: 10.3758/CABN.9.4.343.

Mao, Junwen and Laurel H Carney (2015). "Tone-in-Noise Detection Using Envelope Cues: Comparison of Signal-Processing-Based and Physiological Models". In: *JARO - Journal of the Association for Research in Otolaryngology* 16.1, pp. 121–133. ISSN: 14387573. DOI: 10.1007/s10162-014-0489-1.

Mao, Junwen, Azadeh Vosoughi, and Laurel H Carney (2013). "Predictions of diotic tone-in-noise detection based on a nonlinear optimal combination of energy, envelope, and fine-structure cues." In: *The Journal of the Acoustical Society of America* 134.1, pp. 396–406. ISSN: 1520-8524. DOI: 10.1121/1.4807815.

Marder, Eve, Marie L Goeritz, and Adriane G Otopalik (2015). *Robust circuit rhythms in small circuits arise from variable circuit components and mechanisms*. DOI: 10.1016/j.conb.2014.10.012.

Marr, David (1982). "Vision". In: *book*.

Mathis, Mackenzie Weygandt, Alexander Mathis, and Naoshige Uchida (2017). "Somatosensory Cortex Plays an Essential Role in Forelimb Motor Adaptation in Mice". In: *Neuron* 93.6, 1493–1503.e6. ISSN: 10974199. DOI: 10.1016/j.neuron.2017.02.049.

McCullagh, P.N.J.A. and J.A. Nelder (1989). *Generalized Linear Models*, pp. 1–514. ISBN: 978-0412317606. DOI: 10.1007/978-1-4899-3242-6.

McDermott, Josh H. and Eero P Simoncelli (2011). "Sound texture perception via statistics of the auditory periphery: Evidence from sound synthesis". In: *Neuron* 71.5, pp. 926–940. ISSN: 08966273. DOI: 10.1016/j.neuron.2011.06.032.

McKenzie, Sam et al. (2014). "Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas". In: *Neuron* 83.1, pp. 202–215. ISSN: 10974199. DOI: 10.1016/j.neuron.2014.05.019. arXiv: NIHMS150003.

Mechan, Annis O et al. (2009). "A comparison of learning and memory characteristics of young and middle-aged wild-type mice in the IntelliCage". In: *Journal of Neuroscience Methods* 180.1, pp. 43–51. ISSN: 01650270. DOI: 10.1016/j.jneumeth.2009.02.018.

Mishchencko, Yuriy, Joshua T Vogelstein, and Liam Paninski (2011). "A BAYESIAN APPROACH FOR INFERRING NEURONAL CONNECTIVITY FROM CALCIUM FLUORESCENT IMAGING DATA". In: *The Annals of Applied Statistics* 5.2B, pp. 1229–1261. DOI: 10.1214/09-AOAS303. arXiv: arXiv:1107.4228v1.

ML, Sheffield (2012). *GPy*. Tech. rep.

Mnih, Volodymyr et al. (2015a). "Human-level control through deep reinforcement learning". In: *Nature* 518.7540, pp. 529–533. ISSN: 0028-0836. DOI: 10.1038/nature14236. arXiv: 1312.5602.

Mnih, Volodymyr et al. (2015b). "Human-level control through deep reinforcement learning". In: *Nature* 518.7540, pp. 529–533. ISSN: 0028-0836. DOI: 10.1038/nature14236. arXiv: 1312.5602.

Montague, P R, Peter Dayan, and T J Sejnowski (1996). "A framework for mesencephalic dopamine systems based on predictive Hebbian learning." In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 16.5, pp. 1936–1947. ISSN: 0270-6474. DOI: 10.1.1.156.635.

Nelken, Israel (2004). *Processing of complex stimuli and natural scenes in the auditory cortex*. DOI: 10.1016/j.conb.2004.06.005.

Nelken, Israel, Yaron Rotman, and Omer Bar Yosef (1999). "Responses of auditory-cortex neurons to structural features of natural sounds". In: *Nature* 397.6715, pp. 154–157. ISSN: 00280836. DOI: 10.1038/16456.

Nesterov, Yurii and Vladimir Spokoiny (2017). "Random Gradient-Free Minimization of Convex Functions". In: *Foundations of Computational Mathematics* 17.2, pp. 527–566. ISSN: 16153383. DOI: 10.1007/s10208-015-9296-2.

Nevin, John A. (1969). "Signal detection theory and operant behaviorA review of David M. Green and John A. Swets' Signal detection theory and psychophysics1". In: *Journal of the Experimental Analysis of Behavior* 12.3, pp. 475–480. ISSN: 0022-5002. DOI: 10.1901/jeab.1969.12-475.

Nienborg, Hendrikje and Bruce G Cumming (2009). "Decision-related activity in sensory neurons reflects more than a neuron ' s causal effect". In: *Na-*

ture 459.7243, pp. 89–92. ISSN: 0028-0836. DOI: 10.1038/nature07821. arXiv: NIHMS150003.

Niessing, J??rn and Rainer W Friedrich (2010). "Olfactory pattern classification by discrete neuronal network states". In: *Nature* 465.7294, pp. 47–52. ISSN: 00280836. DOI: 10.1038/nature08961.

Niv, Y. et al. (2015). "Reinforcement Learning in Multidimensional Environments Relies on Attention Mechanisms". In: *Journal of Neuroscience* 35.21, pp. 8145–8157. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.2978-14.2015.

Niwa, Mamiko et al. (2012). "Active Engagement Improves Primary Auditory Cortical Neurons' Ability to Discriminate Temporal Modulation". In: *Journal of Neuroscience* 32.27, pp. 9323–9334. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.5832-11.2012.

Nodal, Fernando R et al. (2010). "Lesions of the auditory cortex impair azimuthal sound localization and its recalibration in ferrets." In: *Journal of Neurophysiology* 103.3, pp. 1209–1225. ISSN: 0022-3077. DOI: 10.1152/jn.00991.2009.

Obeso, Ignacio et al. (2014). "The subthalamic nucleus and inhibitory control: Impact of subthalamotomy in Parkinson's disease". In: *Brain* 137.5, pp. 1470–1480. ISSN: 14602156. DOI: 10.1093/brain/awu058.

O'Connor, D. H. et al. (2010). "Vibrissa-Based Object Localization in Head-Fixed Mice". In: *Journal of Neuroscience* 30.5, pp. 1947–1967. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.3762-09.2010.

Ohl, Frank W et al. (1999). "Bilateral ablation of auditory cortex in Mongolian gerbil affects discrimination of frequency modulated tones but not of pure tones". In: *Learning & Memory* 6.4, pp. 347–362. ISSN: 1072-0502. DOI: 10.1101/lm.6.4.347.

Oikonomidis, Lydia et al. (2017). *A dimensional approach to modeling symptoms of neuropsychiatric disorders in the marmoset monkey*. DOI: 10.1002/dneu.22446.

O'Keefe, J. and J. Dostrovsky (1971). "The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat". In: *Brain Research* 34.1, pp. 171–175. ISSN: 00068993. DOI: 10.1016/0006-8993(71)90358-1.

Oliphant, Travis E (2007). "SciPy: Open source scientific tools for Python". In: *Computing in Science and Engineering* 9, pp. 10–20. ISSN: 1521-9615. DOI: 10.1109/MCSE.2007.58.

Otazu, Gonzalo H et al. (2009). "Engaging in an auditory task suppresses responses in auditory cortex". In: *Nature Neuroscience* 12.5, pp. 646–654. ISSN: 1097-6256. DOI: 10.1038/nn.2306. arXiv: NIHMS150003.

Oviedo, Hysell V et al. (2010). "The functional asymmetry of auditory cortex is reflected in the organization of local cortical circuits". In: *Nature Neuroscience* 13.11, pp. 1413–1420. ISSN: 10976256. DOI: 10.1038/nn.2659.

Paninski, Liam (2004). "Maximum likelihood estimation of cascade point-process neural encoding models". In: *Network: Computation in Neural*

Systems 15.4, pp. 243–262. ISSN: 0954-898X. DOI: 10.1088/0954-898X_15_4_002.

Paninski, Liam et al. (2010). *A new look at state-space models for neural data*. DOI: 10.1007/s10827-009-0179-x.

Park, Hame, Jan-Matthis Lueckmann, et al. (2016). "Spatiotemporal dynamics of random stimuli account for trial-to-trial variability in perceptual decision making". In: *Scientific Reports* 6.1, p. 18832. ISSN: 2045-2322. DOI: 10.1038/srep18832.

Park, Il Memming, Evan W. Archer, et al. (2013). "Spectral methods for neural characterization using generalized quadratic models". In: *Advances in Neural Information Processing Systems* 26, pp. 2454–2462. ISSN: 10495258.

Park, Il Memming, Miriam L R Meister, et al. (2014). "Encoding and decoding in parietal cortex during sensorimotor decision-making." In: *Nat Neurosci* 17.August, pp. 1395–1403. ISSN: 1546-1726. DOI: 10.1038/nn.3800.

Park, Mijung, Gergo Bohner, and Jakob H Macke (2015). "Unlocking neural population non-stationarities using hierarchical dynamics models". In: *Advances in Neural Information Processing Systems*, pp. 145–153.

Park, Mijung and Jonathan W Pillow (2011). "Receptive field inference with localized priors". In: *PLoS Computational Biology* 7.10. ISSN: 1553734X. DOI: 10.1371/journal.pcbi.1002219.

Parker, A J and W T Newsome (1998). "SENSE AND THE SINGLE NEURON: Probing the Physiology of Perception". In: *Annual Review of Neu-*

roscience 21.1, pp. 227–277. ISSN: 0147-006X. DOI: 10.1146/annurev.neuro.21.1.227.

Pawitan, Yudi. (2013). "In All Likelihood: Statistical Modelling and Inference Using Likelihood". In: p. 544. ISSN: 0021-9665.

Peters, Jan (2010). "Policy gradient methods". In: *Scholarpedia* 5.11, p. 3698. ISSN: 1941-6016. DOI: 10.4249/scholarpedia.3698.

Peters, Jan and Stefan Schaal (2008). "Reinforcement learning of motor skills with policy gradients". In: *Neural Networks* 21.4, pp. 682–697. ISSN: 08936080. DOI: 10.1016/j.neunet.2008.02.003.

Petersen, Carl C H and Sylvain Crochet (2013). *Synaptic Computation and Sensory Processing in Neocortical Layer 2/3*. DOI: 10.1016/j.neuron.2013.03.020.

Pillow, Jonathan W (2005). "Prediction and Decoding of Retinal Ganglion Cell Responses with a Probabilistic Spiking Model". In: *Journal of Neuroscience* 25.47, pp. 11003–11013. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.3305-05.2005.

Pillow, Jonathan W, Yashar Ahmadian, and Liam Paninski (2010). "Model-Based Decoding, Information Estimation, and Change-Point Detection Techniques for Multineuron Spike Trains." In: *Neural computation* 45.1, pp. 1–45. ISSN: 1530-888X. DOI: 10.1162/NECO_a_00058.

Pillow, Jonathan W, Jonathon Shlens, et al. (2008). "Spatio-temporal correlations and visual signalling in a complete neuronal population". In: *Na-*

ture 454.7207, pp. 995–999. ISSN: 0028-0836. DOI: 10.1038/nature07140. arXiv: NIHMS150003.

Pineau, J et al. (2015). "Bayesian Reinforcement Learning: A Survey". In: *Foundations and Trends R in Machine Learning* 8.5, pp. 6–359. DOI: 10.1561/2200000049.

Pitkow, Xaq and Dora E Angelaki (2017). *Inference in the Brain: Statistics Flowing in Redundant Population Codes*. DOI: 10.1016/j.neuron.2017.05.028. arXiv: 1702.03492.

Pitkow, Xaq, Sheng Liu, et al. (2015). "How Can Single Sensory Neurons Predict Behavior?" In: *Neuron* 87.2, pp. 411–424. ISSN: 10974199. DOI: 10.1016/j.neuron.2015.06.033. arXiv: 15334406.

Poddar, Rajesh, Risa Kawai, and Bence P Ölveczky (2013). "A fully automated high-throughput training system for rodents". In: *PLoS ONE* 8.12. ISSN: 19326203. DOI: 10.1371/journal.pone.0083171.

Poort, Jasper et al. (2015). "Learning Enhances Sensory and Multiple Non-sensory Representations in Primary Visual Cortex". In: *Neuron* 86.6, pp. 1478–1490. ISSN: 10974199. DOI: 10.1016/j.neuron.2015.05.037.

Prerau, M J et al. (2009). "Characterizing Learning by Simultaneous Analysis of Continuous and Binary Measures of Performance". In: *Journal of Neurophysiology* 102.5, pp. 3060–3072. ISSN: 0022-3077. DOI: 10.1152/jn.91251.2008.

Rabinowitz, Neil C. et al. (2015). "Attention stabilizes the shared gain of V4 populations". In: *eLife* 4.NOVEMBER2015, pp. 1–24. ISSN: 2050084X. DOI: 10.7554/eLife.08998. arXiv: arXiv:1011.1669v3.

Rasmussen, Edward and K I Williams (2006). *Gaussian processes for machine learning*, p. 248. ISBN: 0-262-18253-X.

Read, J. C.A. (2015). *The place of human psychophysics in modern neuroscience*. DOI: 10.1016/j.neuroscience.2014.05.036.

Reinhold, Kimberly, Anthony D. Lien, and Massimo Scanziani (2015). "Distinct recurrent versus afferent dynamics in cortical visual processing". In: *Nature Neuroscience* 18.12, pp. 1789–1797. ISSN: 15461726. DOI: 10.1038/nn.4153.

Rippere, Vicky (2015). *Methods in behavioral research*. Vol. 17. 5, p. 527. ISBN: 9780078035159. DOI: 10.1016/0005-7967(79)90091-3. arXiv: arXiv:1011.1669v3.

Robie, Alice A et al. (2017). "Mapping the Neural Substrates of Behavior". In: *Cell* 170.2, 393–406.e28. ISSN: 10974172. DOI: 10.1016/j.cell.2017.06.032.

Rock, I (1957). "The role of repetition in associative learning." In: *The American journal of psychology* 70.2, pp. 186–93. ISSN: 0002-9556.

Roediger, Henry L. and Kathleen M Arnold (2012). *The one-trial learning controversy and its aftermath: remembering Rock (1957)*. DOI: 10.5406/amerjpsyc.125.2.0127.

Roitman, Jamie D and Michael N Shadlen (2002). "Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task." In: *Journal of Neuroscience* 22.21, pp. 9475–89. ISSN: 1529-2401. DOI: 10.1016/S0377-2217(02)00363-6.

Rosenthal, Nadia and Steve Brown (2007). "The mouse ascending: perspectives for human-disease models." In: *Nature cell biology* 9.9, pp. 993–9. ISSN: 1465-7392. DOI: 10.1038/ncb437.

Rossant, Cyrille et al. (2016). "Spike sorting for large, dense electrode arrays". In: *Nature Neuroscience* 19.4, pp. 634–641. ISSN: 1097-6256. DOI: 10.1038/nn.4268.

Rothschild, Gideon, Israel Nelken, and Adi Mizrahi (2010). "Functional organization and population dynamics in the mouse primary auditory cortex". In: *Nature Neuroscience* 13.3, pp. 353–360. ISSN: 10976256. DOI: 10.1038/nn.2484. arXiv: NIHMS150003.

Roweis, S and Z Ghahramani (1999). "A unifying review of linear gaussian models." In: *Neural computation* 11.2, pp. 305–45. ISSN: 0899-7667.

Rudebeck, Peter H et al. (2006). "Separate neural pathways process different decision costs". In: *Nature Neuroscience* 9.9, pp. 1161–1168. ISSN: 10976256. DOI: 10.1038/nn1756.

Ruff, Douglas A. and Marlene R. Cohen (2014). "Attention can either increase or decrease spike count correlations in visual cortex". In: *Nature Neuroscience* 17.11, pp. 1591–1597. ISSN: 15461726. DOI: 10.1038/nn.3835.

Runyan, Caroline A et al. (2017). "Distinct timescales of population coding across cortex". In: *Nature Publishing Group* 548.7665, pp. 92–96. ISSN: 0028-0836. DOI: 10.1038/nature23020.

Rushworth, Matthew F.S. et al. (2011). *Frontal Cortex and Reward-Guided Learning and Decision-Making*. DOI: 10.1016/j.neuron.2011.05.014.

Sahani, Maneesh and Jf Linden (2003a). "Evidence optimization techniques for estimating stimulus-response functions". In: *Advances in neural information processing ...* Pp. 317–324. ISSN: 10495258.

— (2003b). "Evidence optimization techniques for estimating stimulus-response functions". In: *Advances in Neural Information Processing Systems*, pp. 317–324. ISSN: 10495258.

Salimans, Tim et al. (2017). "Evolution Strategies as a Scalable Alternative to Reinforcement Learning". In: *arXiv*. ISSN: 1744-4292. DOI: 10.1.1.51.6328. arXiv: 1703.03864.

Salzman, C. Daniel, Kenneth H. Britten, and William T. Newsome (1990). "Cortical microstimulation influences perceptual judgements of motion direction". In: *Nature* 346.6280, pp. 174–177. ISSN: 0028-0836. DOI: 10.1038/346174a0. arXiv: NIHMS150003.

Samejima, K. (2005). "Representation of Action-Specific Reward Values in the Striatum". In: *Science* 310.5752, pp. 1337–1340. ISSN: 0036-8075. DOI: 10.1126/science.1115270.

Sanborn, Adam N, Thomas L Griffiths, and Richard M Shiffrin (2010). "Uncovering mental representations with Markov chain Monte Carlo". In:

Cognitive Psychology 60.2, pp. 63–106. ISSN: 00100285. DOI: 10.1016/j.cogpsych.2009.07.001.

Sarkka, Simo (2013). *Bayesian Filtering and Smoothing*. Cambridge University Press. ISBN: 9781139344203. DOI: 10.1017/CBO9781139344203.

Schaefer, Andreas T and Adam Claridge-Chang (2012). *The surveillance state of behavioral automation*. DOI: 10.1016/j.conb.2011.11.004.

Scheggia, Diego and Francesco Papaleo (2016). "An Operant Intra-/Extradimensional Set-shift Task for Mice". In: *Journal of Visualized Experiments* 5350337915.107. ISSN: 1940-087X. DOI: 10.3791/53503.

Schmidt, Robert et al. (2013). "Canceling actions involves a race between basal ganglia pathways". In: *Nature Neuroscience* 16.8, pp. 1118–1124. ISSN: 1097-6256. DOI: 10.1038/nn.3456.

Schmitt, L Ian et al. (2017). "Thalamic amplification of cortical connectivity sustains attentional control". In: *Nature* 545.7653, pp. 219–223. ISSN: 0028-0836. DOI: 10.1038/nature22073.

Schnupp, Jan, Israel Nelken, and Andrew King (2011). *Auditory Neuroscience*, pp. 1–50. ISBN: 9780262113182. DOI: 10.1017/CBO9781107415324.004. arXiv: arXiv:1011.1669v3.

Schröter, Manuel, Ole Paulsen, and Edward T Bullmore (2017). *Micro-connectomics: Probing the organization of neuronal networks at the cellular scale*. DOI: 10.1038/nrn.2016.182.

Schultz, Wolfram (1998). "Predictive reward signal of dopamine neurons." In: *Journal of neurophysiology* 80.1, pp. 1–27. ISSN: 0022-3077. DOI: 10.1016/j.jss.2013.10.029. arXiv: 1407.4831.

— (2013). *Updating dopamine reward signals*. DOI: 10.1016/j.conb.2012.11.012.

Schultz, Wolfram, Paul Apicella, and T Ljungberg (1993). "Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task." In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 13.3, pp. 900–13. ISSN: 0270-6474. DOI: 8441015.

Scott, Stephen H. (2008). "Inconvenient Truths about neural processing in primary motor cortex". In: *The Journal of Physiology* 586.5, pp. 1217–1224. ISSN: 00223751. DOI: 10.1113/jphysiol.2007.146068.

Seabold, Skipper and Josef Perktold (2010). "Statsmodels: econometric and statistical modeling with Python". In: *Proceedings of the 9th Python in Science Conference*, pp. 57–61.

Seeds, Andrew M et al. (2014). "A suppression hierarchy among competing motor programs drives sequential grooming in Drosophila". In: *eLife* 3, e02951. ISSN: 2050084X. DOI: 10.7554/eLife.02951.

Sejnowski, Terrence J (2010). "Learning optimal strategies in complex environments". In: *Proceedings of the National Academy of Sciences* 107.47, pp. 20151–20152. ISSN: 0027-8424. DOI: 10.1073/pnas.1014954107.

Sereno, M I and J M Allman (1991). *Cortical visual areas in mammals*.

Seung, H S (1996). "How the brain keeps the eyes still". In: *Proceedings of the National Academy of Sciences* 93.23, pp. 13339–13344. ISSN: 0027-8424. DOI: 10.1073/pnas.93.23.13339.

Shadlen, Michael N, KH Britten, et al. (1996). "A computational analysis of the relationship between neuronal and behavioral responses to visual motion". In: *J. Neurosci.* 16.4, pp. 1486–1510.

Shadlen, Michael N and Roozbeh Kiani (2013). *Decision making as a window on cognition*. DOI: 10.1016/j.neuron.2013.10.047. arXiv: NIHMS150003.

Silva, Alcino et al. (2015). "Long-Term Memory Underlying Hippocampus-Dependent Social Recognition in Mice Long-Term Memory Underlying Hippocampus-Dependent Social Recognition in Mice". In: 1063.November, pp. 47–56. DOI: 10.1002/(SICI)1098-1063(2000)10.

Silver, David, Aja Huang, et al. (2016). "Mastering the game of Go with deep neural networks and tree search". In: *Nature* 529.7587, pp. 484–489. ISSN: 0028-0836. DOI: 10.1038/nature16961. arXiv: 1610.00633.

Silver, David, Julian Schrittwieser, et al. (2017). "Mastering the game of Go without human knowledge". In: *Nature* 550.7676, pp. 354–359. ISSN: 0028-0836. DOI: 10.1038/nature24270.

Silver, L. M. (1995). "Mouse genetics: concepts and applications." In: *Mouse genetics: concepts and applications.*

Siniscalchi, Michael J et al. (2016). "Fast and slow transitions in frontal ensemble activity during flexible sensorimotor behavior". In: *Nature Neuroscience* 19.9, pp. 1234–1242. ISSN: 1097-6256. DOI: 10.1038/nn.4342.

Skinner, B. F. (1938). "The Behavior of Organisms: An experimental analysis". In: *The Psychological Record*, p. 486. ISSN: 0066-4308. DOI: 10.1037/h0052216. arXiv: arXiv:1011.1669v3.

— (1948a). "'Superstition' in the pigeon." In: *Journal of Experimental Psychology* 38.2, pp. 168–172. ISSN: 0022-1015. DOI: 10.1037/h0055873.

Skinner, B F (1948b). "Superstition in the pigeon." In: *Journal of experimental psychology* 38.2, pp. 168–172. ISSN: 0022-1015. DOI: 10.1037/h0055873.

Smith, Anne C et al. "Bayesian Analysis of Interleaved Learning and Response Bias in Behavioral Experiments". In: DOI: 10.1152/jn.00946.2006.

Smith, Anne C et al. (2004a). "Dynamic Analysis of Learning in Behavioral Experiments". In: *Journal of Neuroscience* 24.2, pp. 447–461. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.2908-03.2004.

— (2004b). "Dynamic Analysis of Learning in Behavioral Experiments". In: *J. Neurosci.* 24.2, pp. 447–461. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.2908-03.2004.

Smith, Anne C et al. (2007). "Bayesian Analysis of Interleaved Learning and Response Bias in Behavioral Experiments". In: *Journal of Neurophysiology* 97.3, pp. 2516–2524. ISSN: 0022-3077. DOI: 10.1152/jn.00946.2006.

Sofroniew, Nicholas James et al. (2016). "A large field of view two-photon mesoscope with subcellular resolution for in vivo imaging". In: *eLife* 5.JUN2016, e14472. ISSN: 2050084X. DOI: 10.7554/eLife.14472. arXiv: arXiv:1011.1669v3.

Soudry, Daniel et al. (2015). "Efficient "Shotgun" Inference of Neural Connectivity from Highly Sub-sampled Activity Data". In: *PLoS Computational Biology* 11.10. ISSN: 15537358. DOI: 10.1371/journal.pcbi.1004464.

Srivastava, Nitish et al. (2014). "Dropout: A Simple Way to Prevent Neural Networks from Overfitting". In: *Journal of Machine Learning Research* 15, pp. 1929–1958. ISSN: 15337928. DOI: 10.1214/12-AOS1000. arXiv: 1102.4807.

Stangor, Charles (2011). *Research methods for the behavioral sciences*, p. 467. ISBN: 0-8400-3197-1.

Starkweather, Clara Kwon et al. (2017). "Dopamine reward prediction errors reflect hidden-state inference across time". In: *Nature Neuroscience* 20.4, pp. 581–589. ISSN: 1097-6256. DOI: 10.1038/nn.4520. arXiv: NIHMS150003.

Stephens, D W and J R Krebs (1986). *Foraging Theory*. Vol. 121, p. 247. ISBN: 0691084424. DOI: 10.2307/2409049.

Stephens, Greg J, Bethany Johnson-Kerner, et al. (2008). "Dimensionality and dynamics in the behavior of C. elegans". In: *PLoS Computational*

Biology 4.4. ISSN: 1553734X. DOI: 10.1371/journal.pcbi.1000028. arXiv: arXiv:0705.1548v1.

Stevenson, Ian H and Konrad P Kording (2011). "How advances in neural recording affect data analysis". In: *Nature Neuroscience* 14.2, pp. 139–142. ISSN: 1097-6256. DOI: 10.1038/nn.2731. arXiv: NIHMS150003.

Stiebler, I. et al. (1997). "The auditory cortex of the house mouse: Left-right differences, tonotopic organization and quantitative analysis of frequency representation". In: *Journal of Comparative Physiology - A Sensory, Neural, and Behavioral Physiology* 181.6, pp. 559–571. ISSN: 03407594. DOI: 10.1007/s003590050140.

Sutton, Richard S and Andrew G Barto (1981). "Toward a modern theory of adaptive networks: Expectation and prediction." In: *Psychological Review* 88.2, pp. 135–170. ISSN: 1939-1471. DOI: 10.1037/0033-295X.88.2.135. arXiv: arXiv:1011.1669v3.

Sutton, Richard S. and Andrew G. Barto (1998). *Reinforcement Learning: An Introduction*, p. 331. ISBN: 0262193981.

Sutton, Richard S et al. (1999a). "Policy Gradient Methods for Reinforcement Learning with Function Approximation". In: *In Advances in Neural Information Processing Systems 12*, pp. 1057–1063. ISSN: 0047-2875. DOI: 10.1.1.37.9714.

— (1999b). "Policy Gradient Methods for Reinforcement Learning with Function Approximation". In: *Advances in Neural Information Processing Systems 12*, pp. 1057–1063. ISSN: 0047-2875. DOI: 10.1.1.37.9714.

Svoboda, Karel and Ryohei Yasuda (2006). *Principles of Two-Photon Excitation Microscopy and Its Applications to Neuroscience*. DOI: 10.1016/j.neuron.2006.05.019.

Talwar, Sanjiv K, Pawel G Musial, and George L Gerstein (2001). "Role of mammalian auditory cortex in the perception of elementary sound properties". In: *Journal of neurophysiology* 85.6, pp. 2350–2358. ISSN: 0022-3077.

Tervo, D. Gowanlock R, Joshua B Tenenbaum, and Samuel J Gershman (2016). *Toward the neural implementation of structure learning*. DOI: 10.1016/j.conb.2016.01.014.

Tervo, Dougal G R, Mikhail Proskurin, et al. (2014). "Behavioral variability through stochastic choice and its gating by anterior cingulate cortex". In: *Cell* 159.1, pp. 21–32. ISSN: 10974172. DOI: 10.1016/j.cell.2014.08.037.

Theis, Lucas et al. (2016). "Benchmarking Spike Rate Inference in Population Calcium Imaging". In: *Neuron* 90.3, pp. 471–482. ISSN: 10974199. DOI: 10.1016/j.neuron.2016.04.014.

Tremblay, Robin, Soohyun Lee, and Bernardo Rudy (2016). *GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits*. DOI: 10.1016/j.neuron.2016.06.033.

Truccolo, Wilson (2004). "A Point Process Framework for Relating Neural Spiking Activity to Spiking History, Neural Ensemble, and Extrinsic Co-

variate Effects". In: *Journal of Neurophysiology* 93.2, pp. 1074–1089. ISSN: 0022-3077. DOI: 10.1152/jn.00697.2004. arXiv: NIHMS150003.

Tsividis, Pedro A et al. (2017). "Human Learning in Atari". In: {*AAAI Spring Symposium on Science of Intelligence: Computational Principles of Natural and Artificial Intelligence*}.

Tsukano, Hiroaki et al. (2016). "Quantitative map of multiple auditory cortical regions with a stereotaxic fine-scale atlas of the mouse brain". In: *Scientific Reports* 6.October 2015, pp. 1–12. ISSN: 20452322. DOI: 10.1038/srep22315.

Van Der Walt, Stéfan, S. Chris Colbert, and Gaël Varoquaux (2011). "The NumPy array: A structure for efficient numerical computation". In: *Computing in Science and Engineering* 13.2, pp. 22–30. ISSN: 15219615. DOI: 10.1109/MCSE.2011.37. arXiv: 1102.1523.

Van Essen, David C and John H R Maunsell (1983). "Hierarchical organization and functional streams in the visual cortex". In: *Trends in Neurosciences* 6.C, pp. 370–375. ISSN: 01662236. DOI: 10.1016/0166-2236(83)90167-4.

Vasquez-Lopez, Sebastian A et al. (2017). "Thalamic input to auditory cortex is locally heterogeneous but globally tonotopic". In: *eLife* 6, e25141. ISSN: 2050-084X. DOI: 10.7554/elife.25141.

Vidne, Michael et al. (2012). "Modeling the impact of common noise inputs on the network activity of retinal ganglion cells". In: *Journal of Compu-*

tational Neuroscience 33.1, pp. 97–121. ISSN: 09295313. DOI: 10.1007/s10827-011-0376-2.

Vinyals, Oriol, Charles Blundell, and Timothy Lillicrap (2016). "Matching networks for one shot learning". In: *Advances In Neural*.

Viswanathan, G M et al. (1999). "Optimizing the success of random searches". In: *Nature* 401.6756, pp. 911–914. ISSN: 0028-0836. DOI: 10.1038/44831.

Vogelstein, J. T. et al. (2014). "Discovery of Brainwide Neural-Behavioral Maps via Multiscale Unsupervised Structure Learning". In: *Science* 344.6182, pp. 386–392. ISSN: 0036-8075. DOI: 10.1126/science.1250298. arXiv: NIHMS150003.

Walker, Kerry M.M. et al. (2011). "Cortical encoding of pitch: Recent results and open questions". In: *Hearing Research* 271.1-2, pp. 74–87. ISSN: 03785955. DOI: 10.1016/j.heares.2010.04.015.

Walton, M E, S W Kennerley, et al. (2006). "Weighing up the benefits of work: Behavioral and neural analyses of effort-related decision making". In: *Neural Networks* 19.8, pp. 1302–1314. ISSN: 08936080. DOI: 10.1016/j.neunet.2006.03.005.

Walton, Mark E, David M Bannerman, et al. (2003). "Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions." In: *The Journal of neuroscience* 23.16, pp. 6475–6479. ISSN: 1529-2401. DOI: 23/16/6475[pii].

Waterston, R H et al. (2002). “Initial sequencing and comparative analysis of the mouse genome Mouse Genome Sequencing Consortium*”. In: *Nature* 420, pp. 520–562.

Watkins, Christopher J. C. H. and Peter Dayan (1992). “Q-learning”. In: *Machine Learning* 8.3-4, pp. 279–292. ISSN: 0885-6125. DOI: 10.1007/BF00992698.

Watkins, Paul V., Joseph P. Y. Kao, and Patrick O. Kanold (2014). “Spatial pattern of intra-laminar connectivity in supragranular mouse auditory cortex”. In: *Frontiers in Neural Circuits* 8.March, pp. 1–18. ISSN: 1662-5110. DOI: 10.3389/fncir.2014.00015.

Wearden, J H (1983). “Undermatching and overmatching as deviations from the matching law”. In: *Journal of the experimental analysis of behavior* 3.3, pp. 332–340. ISSN: 19383711. DOI: 10.1901/jeab.1983.40-333.

Weber, Bruce (1997). *Swift and Slashing, Computer Topples Kasparov - The New York Times*.

Weber, Theophane et al. (2015). “Reinforced Variational Inference”. In: *NIPS* 2, pp. 1–9.

Weixin, Zhong et al. (2017). “Learning and Stress Shape the Reward Response Patterns of Serotonin Neurons”. In: *The Journal of Neuroscience*, pp. 1181–17. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.1181-17.2017.

Werren, J. H. et al. (2010). “Functional and Evolutionary Insights from the Genomes of Three Parasitoid Nasonia Species”. In: *Science* 327.5963,

pp. 343–348. ISSN: 0036-8075. DOI: 10.1126/science.1178028. arXiv: NIHMS150003.

Whitfield, I. C. (1985). "The role of auditory cortex in behavior." In: *Cerebral Cortex.* Vol. 4, Associa. Springer, Boston, MA, pp. 329–351. DOI: 10.1007/978-1-4757-9619-3_8.

Wichmann, F A and N J Hill (2001). "The psychometric function: I. Fitting, sampling, and goodness of fit." In: *Perception & psychophysics* 63.8, pp. 1293–313. ISSN: 0031-5117.

Williams, Ronald J. (1992). "Simple statistical gradient-following algorithms for connectionist reinforcement learning". In: *Machine Learning* 8.3-4, pp. 229–256. ISSN: 0885-6125. DOI: 10.1007/BF00992696.

Wiltschko, Alexander B. et al. (2015). "Mapping Sub-Second Structure in Mouse Behavior". In: *Neuron* 88.6, pp. 1121–1135. ISSN: 10974199. DOI: 10.1016/j.neuron.2015.11.031.

Wimmer, Klaus et al. (2015). "Sensory integration dynamics in a hierarchical network explains choice probabilities in cortical area MT". In: *Nature Communications* 6, p. 6177. ISSN: 2041-1723. DOI: 10.1038/ncomms7177.

Winer, Jeffery A. and Christoph. Schreiner (2011). *The Auditory Cortex*. Ed. by Jeffery A. Winer and Christoph E. Schreiner. Boston, MA: Springer US, p. 715. ISBN: 978-1-4419-0073-9. DOI: 10.1007/978-1-4419-0074-6.

Wogan, M and R H Waters (1959). "The role of repetition in learning." In: *The American journal of psychology* 72, pp. 612–3. ISSN: 0002-9556.

Xiong, Qiaojie, Petr Znamenskiy, and Anthony M Zador (2015). "Selective corticostriatal plasticity during acquisition of an auditory discrimination task". In: *Nature* 521.7552, pp. 348–351. ISSN: 0028-0836. DOI: 10.1038/nature14225. arXiv: NIHMS150003.

Yamashita, Takayuki and Carl C.H. Petersen (2016). "Target-specific membrane potential dynamics of neocortical projection neurons during goal-directed behavior". In: *eLife* 5.JUN2016, pp. 1–11. ISSN: 2050084X. DOI: 10.7554/eLife.15798.

Yamins, Daniel L.K. and James J DiCarlo (2016). *Using goal-driven deep learning models to understand sensory cortex*. DOI: 10.1038/nn.4244. arXiv: science.aab3050 [10.1126].

Yang, Hongdian, Sung E Kwon, et al. (2015). "Origins of choice-related activity in mouse somatosensory cortex". In: *Nature Neuroscience* 19.1, pp. 127–134. ISSN: 1097-6256. DOI: 10.1038/nn.4183. arXiv: 15334406.

Yang, Tianming and Michael N Shadlen (2007). "Probabilistic reasoning by neurons". In: *Nature* 447.7148, pp. 1075–1080. ISSN: 0028-0836. DOI: 10.1038/nature05852. arXiv: NIHMS150003.

Yizhar, Ofer et al. (2011). *Optogenetics in Neural Systems*. DOI: 10.1016/j.neuron.2011.06.004. arXiv: 9605103 [cs].

Yu, Byron M et al. (2009). "Gaussian-Process Factor Analysis for Low-Dimensional Single-Trial Analysis of Neural Population Activity". In: *Journal of Neurophysiology* 102.1, pp. 614–635. ISSN: 0022-3077. DOI: 10.1152/jn.90941.2008.

Yuste, Rafael (2008). "Circuit neuroscience: the road ahead". In: *Frontiers in Neuroscience* 2.1, pp. 6–9. ISSN: 16624548. DOI: 10.3389/neuro.01.017.2008.

Zemelman, Boris V et al. (2003). "Photochemical gating of heterologous ion channels: Remote control over genetically designated populations of neurons". In: *Proceedings of the National Academy of Sciences* 100.3, pp. 1352–1357. ISSN: 0027-8424. DOI: 10.1073/pnas.242738899.

Zhang, Han, Tao Xu, Hongsheng Li, et al. (2016). "StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks". In: *arXiv*. arXiv: 1612.03242.

Zhang, S., M. Xu, T. Kamigaki, et al. (2014). "Long-range and local circuits for top-down modulation of visual cortex processing". In: *Science* 345.6197, pp. 660–665. ISSN: 0036-8075. DOI: 10.1126/science.1254126. arXiv: 1408.2756.

Znamenskiy, Petr and Anthony M Zador (2013a). "Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination". In: *Nature* 497.7450, pp. 482–485. ISSN: 0028-0836. DOI: 10.1038/nature12077.

— (2013b). "Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination". In: *Nature* 497.7450, pp. 482–485. ISSN: 0028-0836. DOI: 10.1038/nature12077.

Zohary, Ehud, Michael N. Shadlen, and William T. Newsome (1994). "Correlated neuronal discharge rate and its implications for psychophysical

performance". In: *Nature* 370.6485, pp. 140–143. ISSN: 0028-0836. DOI: 10.1038/370140a0.

This page is intentionally left blank

Appendix A

Behaviour Apparatus

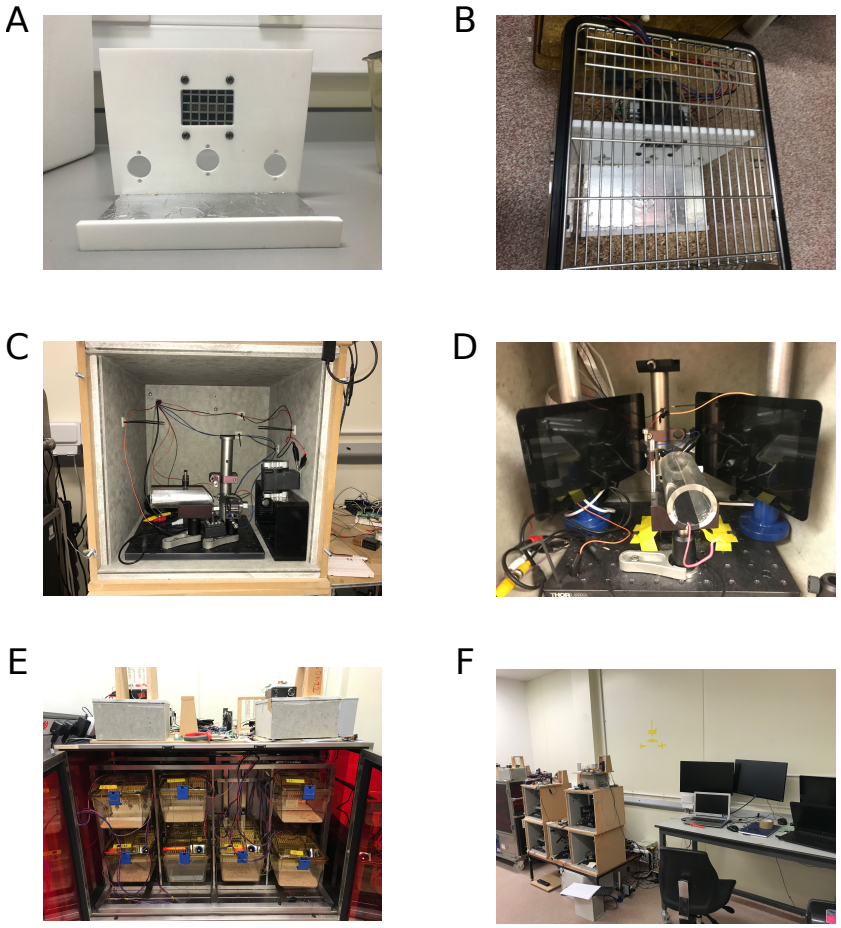


Figure A.1: Behaviour apparatus. **A** Example of three-port behavioural apparatus. Central speaker is protected by a grill. **B** Behavioural apparatus inset into an animals home-cage. **C** Example of a head-fixed apparatus for auditory experiments, in sound shielded enclosure. **D** Example of a head-fixed apparatus used for visual discrimination experiment in which mice performed an orientation discrimination task. **E** Rack of individually ventilated cages with behaviour apparatus inset and electronics on top. Image illustrates the small footprint of our apparatus. **F** Room for behavioural experiments with 5 enclosures for head-fixed experiments visible.

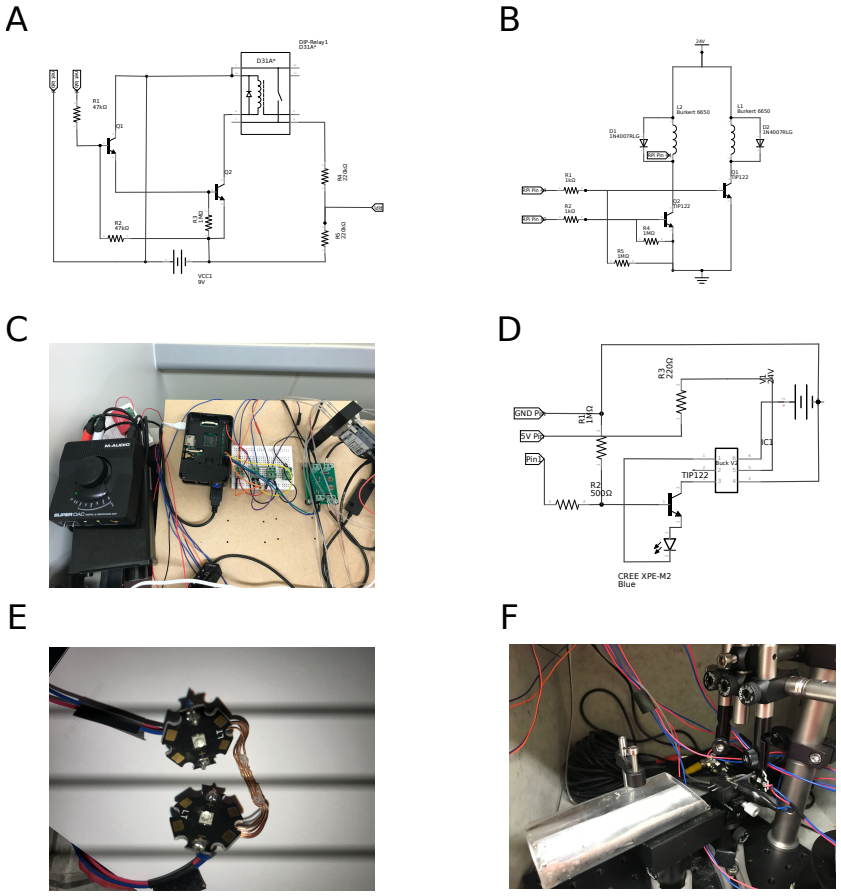


Figure A.2: Circuit schematics and apparatus for optogenetic experiments. **A** Circuit used for lick detection. **B** Circuit for reward delivery. **C** Printed circuit board of circuit shown in B together with connectors, shows small footprint of apparatus. **D** Constant current LED driver circuit used for optogenetics. **E** Head-cap used for head-fixed optogenetic experiments. **F** Head-fixed optogenetic apparatus showing masking lights.