

The Psychophysics of Reward: Empirical Studies and Modeling of Performance  
for Medial Forebrain Electrical Stimulation in the Rat

A Thesis  
In the Department  
of  
Psychology

Presented in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy at

Concordia University

Montreal, Quebec, Canada

June, 2014

© Rebecca Brana Solomon, 2014

**CONCORDIA UNIVERSITY  
SCHOOL OF GRADUATE STUDIES**

This is to certify that the thesis prepared

By: Rebecca Brana Solomon

Entitled: The Psychophysics of Reward: Empirical Studies and Modeling of  
Performance for Medial Forebrain Electrical Stimulation in the Rat

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Psychology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

\_\_\_\_\_ Chair  
Dr. C. DeWolf

\_\_\_\_\_ External Examiner  
Dr. M. Waraczynski

\_\_\_\_\_ External to Program  
Dr. R. Courtemanche

\_\_\_\_\_ Examiner  
Dr. A. Arvanitogiannis

\_\_\_\_\_ Examiner  
Dr. W. Brake

\_\_\_\_\_ Thesis Supervisor  
Dr. P. Shizgal

Approved by: \_\_\_\_\_  
Dr. A. Arvanitogiannis, Graduate Program Director

September 12, 2014 \_\_\_\_\_  
Dr. A. Roy, Dean, Faculty of Arts & Science

## ABSTRACT

The Psychophysics of Reward: Empirical Studies and Modeling of Performance for Medial Forebrain Electrical Stimulation in the Rat

Rebecca Brana Solomon, Ph.D.  
Concordia University, 2014

Brain stimulation reward (BSR), the effect of electrical stimulation that animals seek to reinitiate, is a useful tool to investigate reward-seeking behaviour and its neural underpinnings. The experiments in this thesis pursue this approach by applying the "reward-mountain" model of performance for BSR. This model provides a framework for describing the computational processes that link the induced neural activity to reward-seeking behaviour. The data to which the model is fit are obtained by measuring operant performance for BSR (time spent pressing a lever) as a function of subjective intensity of the stimulation (controlled by pulse frequency) and opportunity cost (work time required to earn a reward). Determining the stage of neural circuitry responsible for the behavioural impact of physiological manipulations is among the principal goals of this strategy.

At the core of the model is the subject's computation of "payoff" via the integration of reward intensity and costs. An important initial stage, often overlooked in neuroscientific studies of decision-making, is the transformation of the objective into subjective variables. The formal relationship between these variables (termed psychophysical functions) is often non-linear: what is experienced is not necessarily a direct reflection of the external world. An analysis of these transformations is important for the full understanding of cost-benefit decision-making. A central goal of the

experiments in this thesis is to estimate the psychophysical functions of reward-seeking variables.

Chapter 1 reviews the BSR literature and describes the reward-mountain model. The experiment described in Chapter 2 concerns the valuation of time: the translation of the experimenter-set opportunity cost (the objective price) into the equivalent subjective domain (subjective price). The experiment described in Chapter 3 estimates the frequency-response function of the directly stimulated neurons subserving the rewarding effect. This function translates the experimenter-set pulse frequency (the inducing stimulus) into the firing frequency of the neuron (the induced physiological response). Chapter 4 describes a proof-of-principle study: the ability of the reward-mountain paradigm to detect the effect of a lesion challenge on pursuit of BSR and to link this effect to one or more stages of processing. Chapter 5 concludes with a general discussion.

## ACKNOWLEDGEMENTS

Completion of this doctoral dissertation would not have been possible without the help and support of many special people.

First, I feel very privileged to have worked with and learned from Dr. Peter Shizgal. His immense knowledge, insight, and consistent encouragement make him a wonderful mentor.

I am extremely grateful to Dr. Kent Conover who developed the many statistical tools used in this dissertation; I thank him for all of the data analysis help he has provided me.

I wish to thank my thesis committee members, Dr. Andreas Arvanitogiannis, Dr. Wayne Brake, Dr. Richard Courtemanche, and Dr. Meg Waraczynski for their comments and input.

My sincere thanks go to Dave Munro and Steve Cabilio for their generous hardware and software technical assistance.

I would like to thank my lab mates for making the lab a fun and enjoyable place to do research: Dr. Giovanni Hernandez, Marie-Pierre Cossette, Brian Dunn. Special thanks to Dr. Yannick Breton for advancing the programming required to conduct the experiments, Ivan Trujillo-Pisanty for contributing one of his subjects for a couple of the experiments in this thesis, and Sarah Nolan-Poupart who proofread and edited parts of this thesis.

I also gratefully acknowledge *Le Fonds québécois de la recherche sur la nature et les technologies* for their funding.

Finally, heartfelt thanks to my parents Jane and Ben, family Jonathan and Jessica, and grandparents Joshua and Dorothy who supported me in every way possible to see the completion of this work.

# TABLE OF CONTENTS

<b>LIST OF FIGURES .....</b>	<b>xi</b>
<b>LIST OF TABLES .....</b>	<b>xiv</b>
<b>Chapter 1: General introduction .....</b>	<b>1</b>
1.1. Brain stimulation reward and natural rewards.....	2
1.2. Physiological properties of the substrate for BSR .....	10
1.2.1. Experimenter-controlled electrical stimulation parameters .....	11
1.2.2. The counter model .....	11
1.2.3. Electrophysiological properties of the substrate .....	12
1.2.4. The role of dopamine in BSR .....	19
1.2.5. Input to VTA dopamine neurons: proposed neural candidates.....	22
1.2.5.1. Cholinergic neurons of the pedunculopontine tegmental nucleus (PPTg) and lateral tegmental nucleus (LDTg).....	22
1.2.5.2. Glutamatergic input to the VTA neurons .....	23
1.2.6. A role for the central extended amygdala in BSR.....	25
1.2.7. Lesion effects .....	27
1.3. The measurement of intracranial self-stimulation.....	29
1.3.1. The 1-dimensional measurement approach.....	29
1.3.2. The 2-dimensional measurement approach.....	30
1.3.3. The 3-dimensional reward mountain model .....	35
1.3.4. The stages of the minimal model .....	41
1.3.4.1 The counter model.....	43
1.3.4.2. Reward intensity function: reward intensity as a function of pulse frequency and train duration.....	44
1.3.4.3. Subjective estimation of external values and computation of payoff.....	47
1.3.4.4 Transformation of payoff into behaviour .....	48
1.3.4.4.1 How the “payoff” of everything else is specified in terms of rewarding electrical stimulation .....	50
1.3.4.4.2. Time allocation expressed in terms of pulse frequency and price .....	51
1.3.5. Displacements of the mountain can reveal at what stages perturbations act .....	52
1.3.6. Overview of reward-mountain methodology .....	54
1.4. Thesis aim: the transformation of objective variables into their subjective equivalents.....	59
<b>Chapter 2: Estimation of subjective opportunity costs.....</b>	<b>64</b>
<b>Introduction.....</b>	<b>66</b>
2.1. Time as an opportunity cost .....	66
2.2. Schedules of reinforcement .....	69
2.3. The translation of objective opportunity cost into subjective opportunity cost .....	71
2.4. Plausible forms of the subjective price function .....	74
2.4.1. The objective price function .....	74

2.4.2. Shizgal’s subjective price function .....	75
2.4.3 Opportunity cost treated as a delay to the reward delivery .....	79
2.4.3.1. The inverted hyperbolic delay discounting subjective price function .....	81
2.4.3.2. The inverted exponential delay discounting subjective price function .....	85
2.5. How to measure the subjective price function .....	87
<b>Materials &amp; Methods .....</b>	<b>93</b>
2.6. Subjects.....	93
2.7. Apparatus & Materials .....	93
2.8. Surgical procedure.....	94
2.9. Experimental procedure.....	95
2.9.1. Screening: preliminary testing .....	95
2.9.2. Schedules of reinforcement.....	95
2.9.3. General experimental structure .....	97
2.9.4. Training: frequency-sweeps .....	98
2.9.5. Training: price-sweeps .....	101
2.9.6. Mountain conditions .....	102
2.9.6.1. Radial-sampling matrix .....	102
2.9.6.2. The structure of the mountain testing paradigm .....	102
2.9.6.3. Training: 3-sampling matrix mountain condition.....	103
2.9.6.4. Experimental subjects.....	107
2.9.6.5. 9-sampling matrix mountain condition.....	107
<b>Statistics and Results .....</b>	<b>112</b>
2.10. Raw data .....	112
2.11. Psychometric functions: determining the $F_{M50}$ and $OP_{M50}$ values.....	113
2.12. The pulse frequency-objective price trade-off functions.....	119
2.13. Fitting the pulse frequency-objective price trade-off functions to the data.....	126
2.14. Subjective price as a function of objective price.....	127
2.15. Model comparisons based on AIC values .....	134
<b>Discussion .....</b>	<b>139</b>
2.16. Model comparisons based on the theoretical validity of predicted subjective prices .....	140
2.17. Model comparisons based on the credibility of reward mountain and delay discounting parameters .....	143
2.18. Conclusion: inferring the best model using statistical and theoretical criteria .....	145
2.19. Implications .....	146
<b>Appendix A: The experimenter-set price versus the actual price .....</b>	<b>150</b>
<b>Appendix B: The derivation for the expression of the contour line at half-maximal time allocation of the reward mountain model: F as a function of OP for each of the 4 proposed subjective price functions.....</b>	<b>151</b>

<b>Appendix C:</b> Reward-mountain parameters for all models for each rat .....	155
<b>Appendix D:</b> Pulse frequency trade-off functions and corresponding subjective price functions for all models, for each rat.....	162
<b>Appendix E:</b> The derivation of the expression for $SP$ as a function $F$ at half-maximal time allocation .....	171
<b>Appendix F:</b> AIC tables comparing the subjective price function and objective price function for each rat.....	174

**Chapter 3: Psychophysical inference of frequency-following fidelity in the neural substrate for brain stimulation reward ..... 177**

<b>Introduction.....</b>	<b>179</b>
3.1. The proposed frequency-response function.....	179
3.2. The counter model.....	182
3.3. Illustrative example of the current-pulse frequency trade-off relationship .....	186
3.4. Previous current-pulse frequency trade-off studies .....	188
3.5. The 4-dimensional reward model .....	191
3.6. The current-pulse frequency trade-off function derived from the 4-dimensional model .....	199
<b>Methods &amp; statistical analyses.....</b>	<b>200</b>
3.7. Subjects.....	200
3.8. Apparatus and materials & surgical procedure .....	200
3.9. Training: current sweeps .....	200
3.10. 4-dimensional testing paradigm .....	201
3.11. Data analysis .....	202
<b>Results.....</b>	<b>203</b>
<b>Discussion .....</b>	<b>208</b>
3.12. The counter model was supported .....	208
3.13. Frequency-following fidelity broke down at high pulse frequencies .....	209
3.14. Why does frequency following fail at high pulse frequencies? .....	210
3.15. Reward intensity saturation is not caused by frequency following failure .....	210
3.16. Implications: properties of the reward-relevant neurons .....	211
3.17. A direct application of the frequency following function: behavioural-neural activity correspondence .....	213



3.18. Further implications: first-stage neurons and glutamate .....	217
3.19. Implications for reward-mountain methodology .....	219
<b>Appendix A:</b> Derivation of the frequency-response function.....	223
<b>Appendix B:</b> The derivation for the expression of the contour line at half-maximal time allocation of the reward mountain model: F as a function of OP for each of the 4 proposed subjective price functions.....	226
<b>Appendix C:</b> The relationship between current and the number of stimulated neurons .....	229
<b>Appendix D:</b> The resampling technique.....	234
<b>Appendix E:</b> The 3-dimensional sections of the 4-dimensional model for all rats.....	242
<b>Appendix F:</b> The parameter values of the 4-dimensional model for all rats.....	255
<b>Chapter 4: Three-dimensional analysis of the effects of a lesion challenge to the rewarding effect of electrical brain stimulation .....</b>	<b>258</b>
<b>Introduction.....</b>	<b>260</b>
4.1. Measuring and interpreting lesion effects .....	262
4.2. The effect lesions to forebrain and hypothalamic regions on ICSS .....	262
4.3. The effect of lesions to the mid and hindbrain on ICSS .....	265
4.4. Summary of lesion-induced effects .....	267
4.5. Validity of the rate-frequency curve method to assess lesion-induced effects .....	268
4.6. Validity of the reward-mountain model to assess lesion-induced effects .....	268
<b>Materials and Methods.....</b>	<b>273</b>
4.7. Subjects.....	273
4.8. Baseline condition .....	273
4.9. Electrolytic lesion and post-lesion condition .....	274
4.10. Histology .....	274
<b>Statistics &amp; results.....</b>	<b>276</b>
<b>Discussion .....</b>	<b>291</b>
4.11. Time course analyses.....	292
4.12. The control group and stability.....	296
4.13. The dual integrator model and implications for lesion-induced effects .....	296
4.14. Dopamine and $OP_e$ changes from baseline .....	301

4.15. Further implications.....	302
<b>Appendix A:</b> The meaning of common logarithmic unit changes using rate frequency curve methodology .....	304
<b>Appendix B:</b> Example of how the 1000 bootstrapped difference value estimates of a parameter is obtained .....	305
<b>Chapter 5: General discussion and conclusion.....</b>	<b>306</b>
5.1. Psychophysical functions of reward-seeking variables .....	309
5.1.1. Psychophysical functions of reward intensity: previous studies .....	309
5.1.2. The psychophysical function of opportunity cost: Chapter 2 summary .....	310
5.1.3. The psychophysical function of the frequency-following response .....	311
5.2. Implications: psychophysical functions and neural signals.....	316
5.3. Implications: the psychophysical functions bolster the reward-mountain model .....	317
5.4. Implications: the necessity for specifying psychophysical functions in the reward-mountain model ..	321
5.5. Summary.....	322
<b>REFERENCES.....</b>	<b>324</b>

# LIST OF FIGURES

## Chapter 1: General introduction

Figure 1 .....	31
Figure 2 .....	34
Figure 3 .....	36
Figure 4 .....	38
Figure 5 .....	42
Figure 6 .....	55
Figure 7 .....	57
Figure 8 .....	63

## Chapter 2: Estimation of subjective opportunity costs

Figure 1 .....	68
Figure 2 .....	76, 77
Figure 3 .....	80
Figure 4 .....	84
Figure 5 .....	86
Figure 6 .....	91
Figure 7 .....	96
Figure 8 .....	100
Figure 9 .....	105, 106
Figure 10 .....	109, 110
Figure 11 .....	111
Figure 12 .....	115, 116
Figure 13 .....	120
Figure 14 .....	123, 124
Figure 15 .....	129, 130
Figure 16 .....	132, 133
Figure ApD.1 .....	163, 164
Figure ApD.2 .....	165, 166
Figure ApD.3 .....	167, 168

Figure ApD.4 .....	169, 170
--------------------	----------

**Chapter 3: Psychophysical inference of frequency-following fidelity in the neural substrate for brain stimulation reward**

Figure 1 .....	181
Figure 2 .....	187
Figure 3 .....	197
Figure 4 .....	198
Figure 5 .....	204
Figure 6 .....	205
Figure 7 .....	206
Figure 8 .....	222
Figure ApA.1 .....	225
Figure ApD.1 .....	239
Figure ApD.2 .....	241
Figure ApE.1 .....	243
Figure ApE.2 .....	244
Figure ApE.3 .....	245
Figure ApE.4 .....	246
Figure ApE.5 .....	247
Figure ApE.6 .....	248
Figure ApE.7 .....	249
Figure ApE.8 .....	250
Figure ApE.9 .....	251
Figure ApE.10 .....	252
Figure ApE.11 .....	253
Figure ApE.12 .....	254

**Chapter 4: Three-dimensional analysis of the effects of a lesion challenge to the rewarding effect of electrical brain stimulation**

Figure 1 .....	271
Figure 2 .....	277
Figure 3 .....	279, 281

Figure 4 .....	283
Figure 5 .....	284
Figure 6 .....	286
Figure 7 .....	289
Figure 8 .....	290
Figure 9 .....	295

**Chapter 5: General discussion and conclusion**

Figure 1 .....	314
Figure 2 .....	320

## LIST OF TABLES

### **Chapter 2: Estimation of subjective opportunity costs**

Table 1 .....	117
Table 2 .....	135
Table 3 .....	135
Table 4 .....	135
Table 5 .....	135
Table 6 .....	136
Table 7 .....	136
Table ApA.1 .....	150
Table ApC.1 .....	156
Table ApC.2 .....	157
Table ApC.3 .....	158
Table ApC.4 .....	159
Table ApC.5 .....	160
Table ApC.6 .....	161
Table ApF.1 .....	175
Table ApF.2 .....	175
Table ApF.3 .....	175
Table ApF.4 .....	175
Table ApF.5 .....	176
Table ApF.6 .....	176

### **Chapter 3: Psychophysical inference of frequency-following fidelity in the neural substrate for brain stimulation reward**

Table ApF.1 .....	256
Table ApF.2 .....	256
Table ApF.3 .....	256
Table ApF.4 .....	256
Table ApF.5 .....	257
Table ApF.6 .....	257
Table ApF.7 .....	257

# *Chapter 1*

## *General introduction*

### ***1.1. Brain stimulation reward and natural rewards***

Brain stimulation reward (BSR), the phenomenon that an animal will seek out trains of electrical brain stimulation, was an unexpected discovery made by scientists James Olds and Peter Milner in 1953. With the goal of studying avoidance behaviour using aversive effects of electrical brain stimulation, electrodes were aimed at the reticular activating system. Ironically, what they observed was the opposite of what they had initially set out to study: the rat would go to locations in the testing apparatus where it had received the electrical stimulation. What they learned subsequently was that the electrode tip was situated, not in the reticular activation system, but instead in the septal area. In further tests, the experimenters demonstrated that rats lever-pressed for electrical stimulation, not only of the septal area, but also for areas such as the tegmentum, subthalamus, and cingulate gyrus of the cortex (Olds & Milner, 1954). This behaviour has been termed intracranial self-stimulation (ICSS); it reveals that electrical stimulation of particular brain areas can serve as a reinforcer. This exciting discovery led to optimism that perhaps motivated behaviours, reinforcement, and learning could be understood on a neural level. In addition, studies conducted soon after the discovery of BSR demonstrated that dependence-inducing drugs exert some of their influence on ICSS (Wise, 1996). Thus, working out the chemical neuroanatomy of the neural circuitry subserving BSR could also shed light on the mechanisms of actions of specific drugs and issues significant to psychopharmacological research.

A rat will vigorously press a lever or run down an alley to obtain rewarding electrical stimulation of many brain areas, to the point of exhaustion (Olds, 1958) but what exactly is the meaning or nature of this signal? Does the electrical signal mimic



some of those that are normally triggered by a natural occurring goal objects or does it exert its influence in some aberrant fashion? Early studies suggested a relationship between natural goal objects and rewarding electrical brain stimulation (e.g., Balagura & Hoebel, 1967; Hoebel and Titelbaum, 1962; Hoebel & Thompson, 1969; Margules & Olds, 1962; Mogenson & Morgan, 1967; Routtenberg & Lindy, 1965; Hoebel, 1965; Hoebel, 1968; Hoebel, 1969). For instance, it has long been known that electrical stimulation of the lateral hypothalamus (LH) elicits eating; in accordance, lesions of this area cause complete termination of eating (Glickman & Schiff, 1967; Valenstein, Cox, & Kakolewski, 1970). It was later shown that the LH electrode that elicits feeding also supports robust self-stimulation (Hoebel & Teitelbaum, 1962; Margules & Olds, 1962) which suggests the possibility that the rewarding electrical brain stimulation is encoding some aspect of the food reward. Furthermore, studies have shown that the LH stimulation causes consummatory behaviours not only when water or food is present but also causes goal-directed responses when food or water is absent (Miller, 1957; Andersson, 1953). In terms of sexual reward, the rats will self-stimulate for olfactory-midbrain pathway activation while often displaying sexual responses in response to this stimulation (Herberg, 1963). Thus, studies in this vein pointed to a structural and functional overlap of natural rewards and rewarding electrical brain stimulation.

The early studies suggest that electrical brain stimulation may be representing certain elements of natural goal objects, yet there were no precise quantitative tests of this hypothesis. Furthermore, it was not clear what aspects of the goal object were being represented. The most compelling evidence that BSR is linked to natural rewards comes from a series of more recent, refined experiments using a forced-choice testing paradigm

that investigates the relationship between rewarding electrical brain stimulation and gustatory stimuli (Conover and Shizgal 1994a, 1994b; Conover, Woodside, & Shizgal, 1994). The authors first point out obvious important differences between a natural food reward and the electrical reward. For example, in an operant paradigm, unlike the electrical reward, a natural goal object can be seen, smelled, heard, and touched. Also, for a food reward, the instrumental response precedes the consummatory response; there is a delay between them, and the two responses are typically performed in different locations within the test cage. In contrast, to obtain rewarding stimulation, a single response is sufficient to both procure and consume the reward in an immediate manner. Additionally, rewarding electrical stimulation does not lead to the accumulation of substance in the gut, unlike food and water which lead to postingestive feedback signals that influence behaviour. Take together, the electrical signal seems to be mimicking some feature(s) of natural rewards, but not all. They reasoned that in order to accurately compare natural and electrical rewards, they must control for the differences as much as possible. The forced-choice paradigm described below offers such control.

In Conover and Shizgal's forced-choice preparations, the rat's task is to choose between the electrical reward and the sucrose reward. The rat has a stimulation electrode aimed at the lateral hypothalamus (LH), an intra-oral catheter, and intragastric cannula attached to a drain tube such that the sucrose solution drips out from the gut thereby eliminating postingestive effects. Touching an empty drinking spout triggers the delivery of a concentrated drop of sucrose solution to the intra-oral catheter via a connected pump. Touching a second spout delivers electrical brain stimulation. In both cases, a single response is sufficient to procure and consume the reward. At the end of a 2 minute trial,

the number and kind of each reward obtained was recorded. The “standard” sucrose reward was held at a fixed concentration across trials. The “alternate” reward was an electrical reward that varied in strength (number of pulses in a stimulation train) across trials. When the strength of the electrical stimulation was relatively low (the stimulation contained few pulses), the rat chose the sucrose reward almost exclusively. However, when the strength of the stimulation train was increased, the rat reversed its preference, choosing the electrical reward instead. Also, a key finding was that moderate strength stimulation trains that the rat had worked for previously without the option of the sucrose reward, when presented alongside the sucrose choice, were now foregone. The experimenters’ interpretation was that if two different rewards can compete with each other, then the rewards must be subject to an evaluation on a common measurement scale and therefore must share a common property.

A further experiment using the forced-choice paradigm (Conover & Shizgal, 1994a) demonstrated that the rats could combine the value of both rewards, underscoring the notion that both rewards share a common characteristic. The standard reward that was held constant across trials was a compound consisting of both the sucrose reward and the electrical reward. The alternate electrical reward was varied from trial to trial. As compared to the sucrose standard alone, the rat assigned a higher value to the compound electrical stimulation-sucrose reward: reward summation between electrical stimulation and sucrose occurred. For summation of two different rewards to occur, the two objects must share a common property.

Early studies of BSR focused a great deal on the internal state of the animal. For instance, at some brain sites such as the LH, a reduction in hunger can reduce the rate of self-stimulation (Hoebel, 1969; Hoebel, 1976). Consequently, it was suggested that because a reduction in hunger decreases responding for both a natural food reward and self-stimulation, then the electrical stimulation could be mimicking the effects of a food reward. However, this proposal is problematic because a reduction in responding does not unambiguously imply that the rewarding effects of the stimulation have been blunted; alternatively, response capacity may have been affected. To address the role of physiological state and BSR, the forced-choice method outlined above was employed to test the effects of the rat's physiological state on competition and summation of a gustatory reward and rewarding electrical stimulation (Conover, Shizgal, & Woodside, 1994; Conover & Shizgal, 1994b). Conover, Woodside and Shizgal (1994b) perturbed the rat's sodium balance as the physiological manipulation and presented the rat with a saline reward and LH electrical stimulation. They showed that inducing sodium depletion increased the value of saline without altering the value of the electrical stimulation. A similar study by Conover, and Shizgal (1994b) modified the rat's postingestive feedback such that the solution was allowed to accumulate in its gut via closure of the gastric cannula thereby degrading the sucrose value over time. Using this preparation, in a summation paradigm, a compound reward consisting of sucrose and LH electrical stimulation was available while the electrical stimulation was presented without the sucrose. They showed that closing the gastric cannula decreased the choice of the compound reward across trials but not the choice of electrical stimulation alone: postingestive feedback preferentially affected the value of sucrose but not the electrical

stimulation. Taken together, these studies suggest that the physiological state of the subject (in the context of sodium depletion and postingestive feedback) does not affect the evaluation of the electrical stimulation at the stimulation sites tested in those studies.

In contrast, an important later study implied that this interpretation of physiological states and reward processing may depend on the nature of the physiological manipulation. Fulton, Woodside and Shizgal (2000) demonstrated that chronic food restriction and body weight loss can enhance the performance for BSR at some sites in the LH but not others. Importantly, this finding suggests that there may be at least two functionally different subpopulations of LH neurons, one of which is altered by *long-term* energy signals as opposed to short-term signals induced by acute manipulations in the forced-choice experiments.

The forced-choice experiments showed that the effects of LH stimulation and sucrose can be evaluated on a common measurement scale and be combined agonistically. In addition, the studies demonstrated that the physiological state affected the evaluation of the natural rewards while it did not affect the subjective value of the electrical stimulation. On the other hand, Fulton and colleagues (2000) showed that long-term energy signals had an effect on the subjective valuation of the rewarding electrical signal, but only for a subset of subjects. Taken together, what can be inferred concerning the nature of the rewarding electrical signal? According to one account, the neural signal subserving BSR is proposed to represent a natural goal object's subjective reward magnitude or "reward intensity" (Shizgal, 1997; Shizgal, 2012) on a unidimensional scale, allowing for qualitatively different goals to be compared. In keeping with this framework, in a natural context, a subject assigns a reward intensity value to all

competing goal objects. For goal selection to occur, such as deciding whether to hunt, nest-build, or mate, there must be a way of representing the value of the competing goals that are qualitatively different from each other via a common currency. When all of the competing resources can be evaluated on a common scale, comparisons can be made such that the most optimal goal or resource is selected. It is important to note that the assignment of subjective reward intensity to a goal object is different than the assignment of a subjective sensory value. For instance, in reward processing, a sucrose concentration is initially assigned a subjective sensory value of “sweetness.” In contrast, the assignment of subjective reward intensity conveys the subjective values of the strength of how “rewarding” or “good” or “valuable” the goal object is on a general scale. In line with this interpretation of goal selection, sensory processing is at an early stage of reward processing, influenced by the subject’s physiological state as shown in the forced-choice studies. In a subsequent downstream stage, assignment of a reward intensity value to a goal object occurs. In addition, as revealed by Fulton and colleagues (2000), there may be some populations of neurons at this downstream stage of reward representation that are influenced by long-term physiological states. In summary, when a rat receives a rewarding train of electrical stimulation, the reward-processing stage that is being tapped into is at the level of unidimensional reward evaluation. There are other computations associated with goal selection such as the accompanying costs that influence the subject’s decisions; these will be discussed later on (Section 1.3).

Given that the electrical signal mimics some properties of natural rewards, rewarding electrical stimulation is an appropriate reinforcer to use in experiments that investigate motivated behaviours. Although there are several differences between

responding for the electrical reward and natural reward as pointed out above, the advantage of the electrical reward is that it offers a direct way to probe reward-seeking behaviour at a single stage of reward processing. That is, the stage investigated is at a purely unidimensional, “reward” level without the confounding effects of response variables such as the handling of food or physiological feedback. As a result, a subject’s behaviour is more stable with electrical rewards than with natural rewards within experimental sessions lasting many hours over many months. Thus, the use of BSR allows us to collect large amounts of data which are necessary for scaling and psychophysical experiments. As well, the electrical stimulation is easily manipulated by adjusting the electrical stimulation parameters such as current which sets the effective radius of excitation and pulse frequency which sets the firing rate (Gallistel, Shizgal, & Yeomans, 1981). This ease of manipulation provides the experimenter with precise control over the intensity as well as the timing of the reward delivery. Also, at a well positioned electrode location, the strength of the rewarding effect is powerful and the rats will work to obtain it even when the response cost or work requirement is high; this potency permits the study of response costs (as will be described in Chapter 2). Finally, the electrode can be aimed at particular brain sites of interest allowing for the investigator to better understand the relative contributions of brain areas to the complex circuitry underlying reward-seeking behaviours.

## ***1.2. Physiological properties of the substrate for BSR***

Numerous studies have shown that discrete brain regions supporting intracranial self-stimulation are extensive throughout the brain (Routtenberg & Malsbury, 1969). These sites are distributed across the diencephalic and telencephalic structures (Olds, 1963), but are also found in brain stem structures such as the nucleus of the solitary tract (Carter & Phillips, 1975) and the cerebellum (Corbett, Fox, & Milner, 1982). However, the most extensively studied self-stimulation sites are found in the medial forebrain bundle (MFB), particularly at the level of the lateral hypothalamus (LH) and posterior hypothalamus, as well as at the ventral tegmental area (VTA). The MFB is defined as the central pathway linking the forebrain to the midbrain, extending from the olfactory tubercle to the midbrain tegmentum (Nieuwenhuys, Geeraedts, & Veening, 1982). This pathway is composed of an ascending and descending heterogeneous dense bundle of 50 projection systems and components encompassing differing origin and termination sites through the forebrain (Nieuwenhuys, Geeraedts, & Veening 1982). Due to the complexity of this substrate, identifying the nature and origin of the neurons that are activated and are responsible for the rewarding effect of the electrical stimulation has been challenging. Although the field of neuroscience has experienced a substantial growth in sophisticated experimental techniques over the 60 years since the discovery of BSR, the exact identity of the directly stimulated neurons that give rise to the rewarding effect remains unknown. Nonetheless, advances have been made in identifying the properties of these neurons as well as possible sites of origin, narrowing down potential candidates. This identification has been a central goal in studies of brain stimulation reward. Several approaches have been adopted, including: electrophysiological characterization using psychophysical



techniques, pharmacological studies, neuronal tracers, neurochemical recordings, lesions, and refined behavioural measurement strategies. The characteristics of the neurons and pertinent studies will be reviewed.

### ***1.2.1. Experimenter-controlled electrical stimulation parameters***

Before describing the properties of the substrate implicated in BSR, it is important to specify the parameters of the electrical pulse trains. The pulses are rectangular, cathodal, monophasic, constant-current, usually of short duration (0.1 ms). They are separated by interpulse intervals (also termed the period which is the reciprocal of the pulse frequency, in ms). The pulse frequency is the number of pulses per second (Hz). A train of electrical stimulation is defined as the length of time from the onset of the first pulse to the offset of the last pulse: typically a short train is employed (in the present experiments, 0.5 s). Depending on the experiment, the various parameters can be manipulated.

### ***1.2.2. The counter model***

The rate of firing is set by the pulse frequency (within limits); the number of stimulated neurons is roughly proportional to the current (this principle is developed in Chapter 3). Gallistel (1978) and Gallistel, Shizgal, and Yeomans (1981) demonstrated that an x-fold increase in pulse frequency has the same effect on reward intensity as an x-fold increase in current intensity. This finding implies that firing 10 neurons, 100 times produces the same rewarding effect as firing 100 neurons, 10 times. This demonstration led to the “counter model” which describes how the postsynaptic effects of the induced

neural firing is integrated over time and space. Specifically, the targets of the directly activated neurons act in a simple manner, as a spike counter, as if to count the aggregate number of spikes conveyed by the directly stimulated neurons within a given time window (Gallistel, Shizgal, & Yeomans, 1981).

Thus, the rewarding effect is a function of aggregate spike rate. Following suit, increasing either the pulse frequency or the current increases the rewarding effect. The experiments presented below hold the current constant while varying the pulse frequency to vary the strength of the stimulation. (The experiment in Chapter 3 manipulates the current as well.)

The formal descriptions of the rat's processing of the reward signal in stages subsequent to the initial spatial and integration of the signal will be discussed at length in section 1.3.

### ***1.2.3. Electrophysiological properties of the substrate***

The identity of the directly stimulated (“first-stage”) reward-relevant neurons has been narrowed down by assessing their electrophysiological properties, such as their refractory period and conduction velocity. Electrical stimulation of the MFB activates multiple neural populations due to the heterogeneity of that particular structure, yet, many of the activated populations play no role in the rewarding effect. Therefore, a purely electrophysiological approach does not provide much insight into the properties of the reward-relevant neurons. Instead, by employing psychophysical techniques, the quantitative neurophysiological and anatomical properties of the neurons responsible for reward-seeking behaviour can be revealed (Deutch, 1964; Gallistel, 1975; Gallistel,

Shizgal, & Yeomans, 1981). The neuronal properties obtained from the psychophysical techniques can then be compared to the identified neurons with known electrophysiological characteristics that are obtained with conventional electrophysiological techniques.

To estimate the physiological properties of the neurons using behavioural techniques, early studies (Deutsch, 1964) used changes in behavioural output (such as lever pressing) to infer changes in reward intensity. This methodology was eventually recognized as flawed (Yeomans, 1975; Gallistel, 1975) because the relationship between the reward intensity and performance is unknown and most likely non-linear. For instance, a given manipulation could cause a change in reward intensity yet might not be detected in changes in lever pressing rates, resulting in erroneous conclusions. Instead, subsequent studies used behavioural “trade-off” methodology which allows the experimenter to “see through” the behaviour such that the analysis is unconfounded by performance variables (Gallistel, Shizgal, & Yeomans, 1981). Two stimulus variables are used rather than one. The effectiveness of one stimulus variable is assessed not by its effect on behaviour but by its effect on the second variable. A fixed response level (behavioural output) is chosen as the behavioural index (such as half-maximal responding); setting a fixed behavioural index implies setting a fixed rewarding effect. The various combinations of two stimulation variables that produce a criterion level of responding reveal the physiological properties of the neurons that carry the reward signal.

In particular, refractory periods, conduction velocity, and the continuity of reward-related fibers between two stimulation sites have been inferred by paired-pulse experiments that use behavioural trade-off methodology. To estimate recovery from

refractoriness, pulse pairs are delivered by a single electrode in a bundle of axons. The first pulse in each pair, referred to as the C (conditioning) pulse, evokes an action potential. The second pulse, referred to as the T (test) pulse, only produces an action potential if the C-T interval is longer than the refractory period of the axons. The number of action potentials doubles when the C-T interval is longer than the refractory period of the axons. In the behavioural version of the refractory C-T experiment, the trade-off stimuli are the C-T interval and number of C-T pulse pairs in a stimulation train. To maintain a given level of behaviour, different combinations of C-T intervals and number of pulse pairs within a stimulation train are assessed. When the T-pulse falls within the absolute refractory period, the number of pulse pairs must be increased to compensate for the failure of the T-pulse to fire the axons. The estimated range of C-T intervals wherein the required number of pulse pairs decline indicates the recovery of the axons that are fired by the T-pulse, and thus the refractory periods of the elements of the activated substrate. The refractory periods have been shown to be 0.4 to 1.2 ms (Yeomans, 1975; Yeomans, 1979; Shizgal, Schindler, & Rompre, 1989; Murray & Shizgal, 1996). The finding that recovery occurs gradually (over a range of refractory period durations) and not abruptly reflects the notion that different reward-relevant neural populations are activated by the stimulation (Murray & Shizgal, 1994). That is, the stimulation activates fibers of different calibers with refractory periods of different durations. Hence, the contribution of these different neural types is reflected in the measured *range* of refractory periods.

Furthermore, recovery from refractoriness also depends on the phases of the post-stimulation excitability cycle and the distance from the electrode to the activated neuron.

In the first post-stimulation recovery phase (the absolute refractory period) the neuron cannot fire, no matter how intense the T-pulse. In the second phase (the relative refractory period), the neuron can fire, but only if the intensity of the T-pulse is well above resting threshold. In the context of paired pulse behavioural trade-off experiments, the C-T estimates are the consequences of the excitability phases and how far an activated neuron is from the electrode. Namely, there are two areas of interest within the brain region fired by the C-pulse (the area closest to the electrode, the area further away from the electrode). The density of the current producing excitation declines in magnitude with distance from the electrode. (The inverse square law indicates that the current intensity is inversely proportional to the square of the distance from the tip.) The neurons closest to the electrode tip will be exposed to current densities considerably higher than their resting threshold. In this region, the T-pulse will produce firings as soon as the absolute refractory periods of the neurons are exceeded. However, neurons further away from the tip will fire again only after their relative refractory period has ended. Thus, the post-synaptic excitability of the neurons in both regions is reflected in the estimated refractory period duration range.

The collision technique is used to infer whether the axons of reward-relevant neurons directly link two self-stimulation sites. Also, if two sites are linked, then the conduction velocity of the axon can be determined. In this method, two electrodes are used: the first electrode delivering the C-pulses, and second electrode delivering the T-pulses. When stimulating an axon with one electrode, two action potentials are always evoked: one orthodromically (toward the terminals), and one antidromically (toward the cell body). If two electrodes placed on a longitudinal plane simultaneously stimulate the

same axon, then a total of four action potentials are evoked. However, the spikes traveling towards each other will collide and cancel each other out. In order for them not to collide, there must be a sufficient amount of time between the induction of the C and T pulse so that the first action potential induced (by the C pulse) and the refractory period propagating in its wake can pass beyond the second electrode (that induces the T pulse). Only if the C-T interval is greater than the sum of the inter-electrode conduction time and corresponding refractory period, will both spikes triggered by the C and T pulse reach the synaptic terminals. The “collision interval” is the C-T interval at which the collision effect is just avoided and is the measure of conduction time between the two electrodes. In the behavioural version of this technique, the C-T interval is traded-off against the number of pulse pairs within a stimulation train. At short C-T intervals, action potentials lost due to collision must be replaced by an increase in the number of pulse pairs. As the C-T interval is increased, there is an abrupt decline in the required number of pulse pairs needed to maintain the behavioural criterion because the lost action potentials are recovering.

If there is a synapse between two electrodes, a simple collision effect is not seen. However, evidence of a collision effect shows that the two electrodes are stimulating the same axon. Using this method several studies showed (Shizgal, Bielajew, Corbett, Skelton, & Yeomans, 1980; Bielajew & Shizgal, 1986; Shizgal & Murray, 1996) that reward relevant neurons course between the lateral hypothalamus (LH) and ventral tegmental area (VTA). Linkage of the reward relevant neurons has also been demonstrated from the lateral preoptic area (LPO is a neighbouring site anterior to the LH) to the lateral hypothalamus (LH) (Bielajew, Thrasher, & Fouriez, 1987; Bielajew,

Konkle, Fouriez, Boucher-Thrasher, Schindler, 2001) as well as between the posterior mesencephalon and the VTA (Boye & Rompre 1996), and from the LPO to the VTA (Bielajew, Bushnick, Konkle, & Schindler, 2000).

Conduction velocity is determined by dividing the interelectrode difference (mm) by the sum of the collision interval and refractory period. In the LH-VTA pathway, conduction velocities are in the range of 1 to 8 m/s which suggest that the fibers responsible for the effect are myelinated, with diameters of 0.5 to 2  $\mu\text{m}$  (Bielajew & Shizgal, 1982; Bielajew & Shizgal, 1986; Murray & Shizgal, 1996).

The behavioural collision test was adapted to demonstrate the direction in which the LH-VTA neurons project (Bielajew & Shizgal, 1986). Anodal hyperpolarization block was employed to temporarily disrupt the conduction between a self-stimulation electrode and the terminals that transmit reward information to the next stage. If the anodal block is interposed between the soma and electrode, reward effectiveness will not be altered. However, if the anodal block is interposed between the electrode and terminals, reward effectiveness will be reduced. After determining linkage between two electrodes of the LH-VTA path, this hypothesis was tested by setting the LH electrode as the cathode and VTA electrode as the anode, and subsequently reversing the order. The direction of the pathway could be inferred by comparing the behavioural effectiveness of stimulation delivered with either polarity. The results are consistent with rostral to caudal conduction, from the LH to the VTA, which implies that at least some of the first-stage neurons project in this direction.

To summarize, most of the recovery from refractoriness in the LH-VTA pathway occurs between 0.4 and 1.2 ms as revealed through behavioural C-T experiments.

Conduction velocities are in the range of 1 to 8 m/s which suggest fine myelinated fibers, 0.5 to 2  $\mu\text{m}$  in diameter; at least some of the first-stage neurons project in the rostral-caudal direction. The experiment presented in Chapter 3 adds another physiological characteristic of the first-stage neurons using psychophysical techniques that employ trade-off methodology: the maximum firing rate of the neurons of the LH.

These psychophysically derived characteristics have been compared with the electrophysiological characteristics of neurons observed via unit recordings. Early studies demonstrated that refractory periods of the MFB derived from behaviourally derived psychophysical studies (Deutsch, 1964) matched with the electrophysiological properties of the thalamus, midbrain reticular formation, pontine tegmentum (Gallistel, Rolls, & Green, 1969; Rolls, 1971). However, given the finding that some of the reward-relevant neurons project rostral-caudally (Bielajew & Shizgal, 1986), later studies shifted the focus to the forebrain structures. Indeed, the behaviourally psychophysical characteristics have been shown to overlap with the electrophysiological properties of many forebrain areas such as lateral septum, medial septum, diagonal band of Broca, bed nucleus of the stria terminalis, ventral pallidum, lateral and medial preoptic areas, magnocellular preoptic nucleus, olfactory tubercle, substantia nigra (substantia nigra pars reticulata), interstitial nucleus of the stria medullaris (Rompere & Shizgal, 1986; Shizgal, Schindler, & Rompere, 1989; Murray & Shizgal, 1996b).

The behaviourally derived properties of the substrate provide crucial information not readily obtained from unit recordings on their own. However, the limitation to this approach is that even if a given neuronal type corresponds to the physiological and anatomical properties inferred from psychophysical data, there is no guarantee that this



proposed neuronal type constitutes the first-stage neurons. It is possible that a neuronal type simply resembles the first-stage neurons and plays no role in reward-processing. Thus, other complementary approaches are necessary for the assessment neurons as candidates for the directly stimulated stage. Nonetheless, on the basis of these properties, one can confidently rule out the kinds of neurons that do not match these characteristics such as the dopamine neurons which have figured prominently in the BSR and reward circuitry literature. The role of dopaminergic neurons is described in the next section.

#### ***1.2.4. The role of dopamine in BSR***

An early hypothesis regarding the nature of the first-stage neurons is that the dopamine (DA) neurons are the directly activated neurons in the path and carry the reward signal (Wise, 1978; Stein, 1962; German & Bowden, 1974; Corbett & Wise, 1980). Initially, the correlation between the brain regions supporting self-stimulation and the location of the catecholamine pathways contributed significantly to this proposal (Ungerstedt, 1971; Corbet & Wise, 1980). In addition, pharmacological studies that have shown an involvement of DA in BSR have largely influenced this hypothesis. For instance, using the curve-shift paradigm, DA antagonists have been shown to decrease rewarding impact of the electrical stimulation (eg, Franklin, 1978; Gallistel, Boytim, Gomita, & Klebanoff, 1982) whereas DA agonists have been shown to increase its impact (eg, Crowe, 1970; Rompré & Bauco, 1990).

The DA neurons that have been implicated in reward processes originate in the ventral tegmental area (VTA) and substantia nigra (SN) and project rostrally via the MFB to the nucleus accumbens (NAC) and medial prefrontal cortex (review, Ikemoto, 2010).

Two active states of DA have been described (Schultz, 2002): sustained low frequency firing tonic firing and intermittent high frequency bursting. Self-stimulation of the MFB causes increases in DA of the NAC in both these states (tonic release: Hernandez et al., 2007; Hernandez & Hoebel, 1988; Nakahara, Ozaki, Miura, Miura, Nagatsu, 1989; You, Chen, & Wise, 2001; You, Tzschentke, Brodin & Wise, R. A. 1998; phasic release: Beyene, Carelli, Wightman, 2010; Cheer et al., 2007). The two different DA states have been suggested to play different roles in BSR. Initially, it was proposed that phasic DA may act as a learning signal, involved in the acquisition of the instrumental response. The hypothesis was based on the application of temporal difference learning to electrophysiological data from Schultz's group (Montague, Dayan, & Sejnowski, 1996; Schultz, Dayan, & Montague, 1997). In addition, this proposal was supported by the finding that the DA signal fell below the detection limit within a minute of the onset of self-stimulation (Garris, Kilpatrick, Bunin, Michael, & Walker, 1999). However, a more recent voltammetry study showed that a DA phasic response was recorded in the NAC after every train when a 10 s post reward time-out was imposed (Cheer et al., 2007), thus complicating the initial learning phasic signal hypothesis. The tonic state has been suggested to be implicated in maintaining the BSR instrumental response (Hernandez et al, 2006) and response vigour (Niv, Daw, & Dayan, 2007). What functions these different DA states in reward processes play continues to be determined.

Despite the pharmacological and neurochemical evidence, on the basis of physiological properties, it is highly unlikely that DA neurons compose the first-stage (Gallistel, Shizgal, & Yeomans, 1981). DA neurons are unmyelinated, have too slow conduction velocities (0.3 - 1.5 m/s) and too long refractory period periods (1.8 - 20 ms).

In contrast, the psychophysically derived properties of the reward-relevant neurons suggest myelination, have faster conduction velocities (1 – 8 m/sec), and shorter refractory periods (0.4 - 1.2 ms). As well, the direction of DA neurons along the MFB is caudal-rostral whereas the behaviourally derived data show that at least some of the reward-relevant neural projections are rostral-caudal.

Alternate functions have been proposed for mesolimbic DA: it may invigorate (Salamone, Correa, Farrar, & Mingote, 2007; Niv et al., 2007) and enable animals to exert effort to execute the required instrumental task (Cousins & Salamone, 1994; Aberman & Salmone, 1999; Denk et al., 2005), initiate goal-directed responses (Nicola, 2010) and increase the willingness to pay for the reward (Hernandez, Breton, Conover, & Shizgal, 2010). However, the DA neurons' role in transmitting a reward signal should not be discounted: although the DA neurons of the LH-VTA pathway are probably not the first-stage neurons, they may receive input from them and carry the reward signal in some capacity (Wise, 1980; Shizgal 1989; Moisan & Rompre, 1998). This hypothesis is supported in recent studies using optogenetic techniques which enables the selective activation of a neuronal type using transgenetics and light gated channels. Witten and colleagues (2011) showed that direct activation of DA VTA neurons was sufficient for the rats to acquire and sustain self-stimulation (a nose-poking response) in rats. It is difficult to argue against the proposed involvement of VTA DA at some stage in the reward neural circuitry given that a rat will self-stimulate for its activation.

### ***1.2.5. Input to VTA dopamine neurons: proposed neural candidates***

Given that rewarding electrical stimulating may depend on transynaptic involvement of VTA dopamine neurons, what neurons may provide input to the VTA DA neurons?

#### ***1.2.5.1. Cholinergic neurons of the pedunculopontine tegmental nucleus (PPTg) and lateral tegmental nucleus (LDTg)***

Numerous lines of evidence suggest that acetylcholine input into the DA neurons of the VTA may play an important role in ICSS. Older studies suggested a link between BSR and acetylcholine: acetylcholine or carbachol infused into the posterior hypothalamus or VTA increased bar pressing rates (Olds, Yuwiler, Olds, & Yun, 1964; Redgrave & Horrell, 1976) suggesting a rewarding effect for this neurotransmitter. Injection of atropine (a muscarinic blocker) into the VTA raised the ICSS threshold (reduced the rewarding impact) for LH rewarding electrical stimulation by over 300% (Yeomans, Kofman, & McFarlane, 1995). Also, the inhibition by oligonucleotides of the M5 muscarinic receptor mRNA in the VTA produced an increase in the threshold for ICSS responding (Yeomans, et al., 2000; Yeomans, Forster, & Blaha, 2001). In terms of anatomy, cholinergic neurons of the pedunculopontine tegmental nucleus (PPTg) and neighbouring lateral tegmental nucleus (LDTg) of the brainstem send ascending axons to the VTA and neighbouring substantia nigra. These cells make contact with the DA somata of the VTA (Woolf, 1991).

Yeomans, Mathur, and Tampakeras (1993) assessed whether the PPTg neurons are responsible for the attenuating effects of muscarinic blockers on ICSS. They showed

that manipulating PPT neurons by autoreceptor activation or inactivation modulates ICSS: behavioural thresholds increased by over 400% with autoreceptor activation and decreased by 20-80% with deactivation. In a later study, extracellular acetylcholine was increased in the VTA during self-stimulation of the (perifornical) LH and infusion of atropine into the VTA via the microdialysis probe stopped self-stimulation (Rada, Mark, Yeomans, & Hoebel, 2000). Taken together, they proposed that the (unidentified) descending myelinated axons activate PPTg neurons which in turn ascend to activate VTA dopamine neurons which then ascend to project to the NAC and other structures. It is also possible that the descending MFB fibers may not directly project to the PPTg neurons, but instead could go first to other sites such as the nearby parabrachial nucleus which is implicated in BSR (Routtenburg & Malsbury, 1969; Arvanitogiannis, Flores, & Shizgal, 1997).

Another view regarding these cholinergic neurons from the PPTg (Wise 1998) is that given that these neurons project rostrally along the MFB (Woolf & Butcher, 1986) the first-stage neurons might actually be the PPTg and LDTg neurons that are activated antidromically. Electrical stimulation of the MFB could trigger antidromic action potentials that may activate collaterals that project to the VTA DA neurons.

#### ***1.2.5.2. Glutamatergic input to the VTA DA neurons***

Several lines of evidence suggest that glutamatergic input to the VTA is implicated in reward processes. Glutamate agonists injected into the VTA increase the firing rate of VTA DA neurons (Chergui et al., 1993; Johnson, Seutin, & North, 1992) and cause DA release in the NAC (Kalivas, Duffy, & Barrow, 1989). LH stimulation

causes extracellular increases of glutamate in the VTA along with simultaneous release of DA in the NAC (You, Chen, & Wise, 2001). Together, these studies suggest that glutamatergic activation may be activating the DA neurons of the VTA that ascend to the NAC.

Anatomically, by combining retrograde tracer with in situ hybridization of vesicular glutamate transports, Geisler, Derst, Veh and Zahm 2007 demonstrated the glutamatergic afferents into the VTA along the MFB originates in many descending structures: prefrontal cortex, prelimbic cortex, lateral hypothalamic and preoptic areas, medial, ventral pallidum, medial preoptic area, medial septum/diagonal band complex, subthalamic extended amygdala, and lateral habenula. Ascending glutamatergic inputs were also revealed: periaqueductal and central gray, mesencephalic and pontine reticular formation, pedunculopontine, laterodorsal tegmental nuclei, parabrachial and cuneiform nuclei, and median raphe.

Thus, there are many candidates that may provide glutamatergic input to the VTA and influence reward processing. For example, the lateral habenula which supports ICSS (Sutherland & Nakajima, 1981) sends glutamatergic inputs along with DA and GABA projections to the VTA (Omelchenko, Bell, & Sesack, 2009). DA neurons of the VTA are inhibited by stimulation of the lateral habenula (Christophe, Leonzio, & Wilcox, 1986), which suggests some sort of modulatory role for this region.

Another link to glutamatergic transmission in ICSS comes from a recent optogenetic study (Kempadoo, Tourino, Cho, Magnani, Leininger, Stuber, Zhang, Myers, Deisseroth, Lecea, Bonci, 2013). In this study, it was first directly demonstrated that the specific projection from the LH to VTA promotes ICSS of the VTA in the

mouse. This finding is particularly important because while previous studies had only suggested, on the basis of collision studies and tracer studies, that the LH-VTA pathway underlies ICSS, the specificity of the employed technique allowed a direct confirmation of this hypothesis. This investigation focused on the neuropeptide neurotensin (NT) that has previously been shown to be found in one third of the LH neurons projecting to the VTA, and known to interact with DA neurons, such as by increasing their firing rate. Behaviourally, they showed that blocking a neurotensin receptor Nts1 in the VTA attenuated optogenetic self-stimulation of the LH-VTA pathway. Next they showed that blocking the NMDA receptor (glutamate receptor) in the VTA also attenuated optogenetic self-stimulation of this pathway. Given that both glutamate and neurotensin plays a role in ICSS, they investigated their relationship to each other and to VTA DA neurons. By whole cell patch-clamp, they demonstrated that Nts1 activation potentiated NMDA (N-methyl-D-aspartate) mediated-current in VTA DA neurons. In summary, this study suggests that NT is an important LH peptide that directly mediates ICSS by enhancing glutamate transmission in VTA DA neurons.

Taken together, it is possible that the directly stimulated neurons are glutamatergic neurons arising from various forebrain structures that project to VTA DA neurons. However, on the basis of these studies, we cannot rule out that the glutamatergic inputs could be transynaptically activated by the first-stage neurons.

#### ***1.2.6. A role for the central extended amygdala in BSR***

A recent proposal concerns the role of the central extended amygdala network of the basal forebrain in the subject's valuation of reward intensity. Waraczynski (2006)

argues that this structure, particularly the central sublenticular extended amygdala (SLEAc), is anatomically in an ideal position to receive motivationally salient stimuli input from the basolateral amygdala and frontal cortex as well as from sensory input from the brainstem. The central extended amygdala in turn sends efferents to brainstem structures that are implicated in autonomic, hormonal visceromotor responses. It also sends efferents to the striatopallidum, the structure that has traditionally been linked to reward and goal-directed behaviours. These efferents and afferents course through the MFB. In terms of reward processing, it is proposed that the central extended amygdala may be more important than the nucleus accumbens (NAC). That is, the NAC may play in learning and directing responses towards the reward. In contrast the SLEAc's role may be more direct in reward processing, such as the transformation of sensory input into survival utility and subsequent informing of response mechanisms. Several studies have demonstrated effects of pharmacological manipulations of the SLEAc on performance for BSR (Waraczynski, 2008; Waraczynski, Salaeme, & Farral, 2010; Waraczynski, Zwifelhofer, & Kuehn, 2012). For instance GABA receptor agonists of the SLEAc increased BSR thresholds (Waraczynski, 2008). DA receptor manipulations of this area had only modest effects (Waraczynski, Salaeme, & Farral, 2010) whereas dopamine-glutamate manipulations influenced BSR thresholds, but counter to the predicted direction (Waraczynski, Zwifelhofer, & Kuehn, 2012). The exact role of the SLEAc in BSR, whether it is the first-stage or downstream stage and its relationship to other brain areas and BSR are avenues to be investigated.



### *1.2.7. Lesion effects*

Another method used to assess the origin of the reward-relevant neurons are lesion techniques. The destruction of a candidate brain area and subsequent assessment of behaviour under study (such as vigour of lever pressing) tests the necessity of the destroyed brain area for the relevant behaviour. The model of reward circuitry that has guided a large part of the lesions studies is that of the “descending path hypothesis” which proposes that the directly stimulated reward fibers of the medial forebrain bundle (MFB) originate in the basal forebrain and descend through the MFB. The descending path hypothesis is based on Bielajew and Shizgal’s (1986) finding that at least some of the reward-relevant neurons project in the rostral-caudal direction, and that electrophysiological properties of the neurons match those of the forebrain. It is also based on the proposal that VTA DA neurons are transynaptically activated by the electrical stimulation (Hernandez et al., 2010; Moisan & Rompré, 1998), and that many of the forebrain areas were activated following electrical stimulation (Flores, Arvanitogiannis, & Shizgal, 1997; Arvanitogiannis, Flores, & Shizgal, 1997; Arvanitogiannis, Flores, Pfaus, & Shizgal, 1996). Other lesions studies investigated the mid and hindbrain areas as well.

Lesions of several brain areas have been shown to reduce the reward efficacy of the electrical stimulation. However, the effects of lesions to the forebrain areas have often been variable within and across studies, not long lasting or not substantial, thus casting doubt on the strength of the descending path hypothesis. In some cases, the brain damage that was unique to the group of rats showing a behavioural effect was not evident. In addition, there are methodological differences between studies such as the sort of lesion

employed as well as the electrical stimulation parameters, complicating comparisons across studies. Chapter 4 provides a detailed overview of the different brain areas tested and the magnitude and meaning of the effects.

In terms of the forebrain, lesions to the anterior LH (Shizgal & Murray, 1996; Gallistel, Leon, Sim, Lim, & Waraczynski, 1996), sublenticular extended amygdala (Arvanitogiannis, Waraczynski, & Shizgal, 1996; Waraczynski, 2003), and lateral preoptic area (Arvanitogiannis, Waraczynski, & Shizgal, 1996) have demonstrated relatively large and long-lasting lesion effects on LH and VTA stimulation. In terms of more posterior areas, lesions to the lateral habenula of the midbrain have been particularly effective (Morissette & Boye, 2008).

If a sole region was the origin of the first-stage neurons, then one would expect more dramatic lesion effects than what has characteristically been seen. These non-substantial lesion effects may suggest a diffuse reward circuitry system in which several distinct regions give rise to the rewarding effect of electrical stimulation. A related perspective is a model that can account for small effects when a large proportion of directly stimulated substrate is destroyed (Arvanitogiannis, Waraczynski, & Shizgal, 1996), further discussed in Chapter 4.

### ***1.3. The measurement of intracranial self-stimulation***

Following the discovery that rats will work for electrical stimulation of various brain areas, it was evident that measurement strategies were needed to quantify the subjective strength of its rewarding effect. Assessing the subjective strength of the rewarding effect produced by the electrical stimulation before and subsequent to changes in stimulation parameters or physiological manipulations such as pharmacological agents or lesions is essential for understanding the underlying neural circuit.

#### ***1.3.1. The 1-dimensional measurement approach***

To infer the subjective strength of the rewarding effect of the stimulation, early studies employed a 1-dimensional response-rate method that measures the vigour of responding (such as the rate of bar pressing) for a stimulation train of fixed parameters. The subjective strength of its rewarding effect was presumed to be reflected in the rate at which the subject performs the required task to attain the electrical stimulation (Olds, 1958). If a drug caused rates of responding for the electrical stimulation reward to increase, then it was inferred that the drug increased the rewarding efficacy of the electrical stimulation. If rates of responding decreased, it was inferred that the drug attenuated the rewarding efficacy of electrical stimulation. Similarly, effects of lesions to particular brain areas on response rates were commonly used to infer the role of a given brain area (e.g., Olds & Olds, 1968; Boyd & Gardner, 1967; Keesey & Powley, 1973; Ward, 1960). The limitation to this method is that if the average response changed after a manipulation, one cannot be certain whether manipulation affected the rewarding efficacy of the stimulation or alternatively, the performance capacity of the subject

(Valenstein, 1964; Edmonds & Gallistel, 1974; Gallistel, Stellar, & Bubis, 1974). A further limitation of this response rate method is the dependence on one arbitrary chosen train of fixed stimulation parameters that may mask effects. For instance, if a train of fixed stimulation parameters that sustains maximal or near maximal responding is used, then drugs proposed to increase the rewarding effect of the electrical stimulation will not have the ability to further increase responding, thus resulting in faulty inferences. However, if the chosen fixed stimulation parameters sustain a low level of responding, then that same drug dose will produce a large increase in response rate.

### ***1.3.2. The 2-dimensional measurement approach***

The 2-dimensional rate-frequency curve shift method (Edmonds and Gallistel, 1974; Gallistel, Stellar, & Bubis, 1974; Miliareisis, Rompre, Laviolet, Philippe, Coulombe, 1986) was developed to differentiate between the subjective strength of the stimulation's rewarding effect and the rat's ability to perform the task. A further advantage is that this method does not make inferences based solely on the effects of one arbitrarily chosen train of fixed stimulation parameters. The rate-frequency curve plots performance as a function of the pulse frequency of a stimulation train while all other stimulation parameters such as current stays constant. The dependent performance variable is usually rate of bar pressing or as illustrated in Figure 1, "proportion of rewards harvested," the proportion of rewards that the rat reaped within the amount of time available. A high proportion of rewards harvested implies that the rat is attending to the lever a great deal because the stimulation is very rewarding; a low proportion implies that

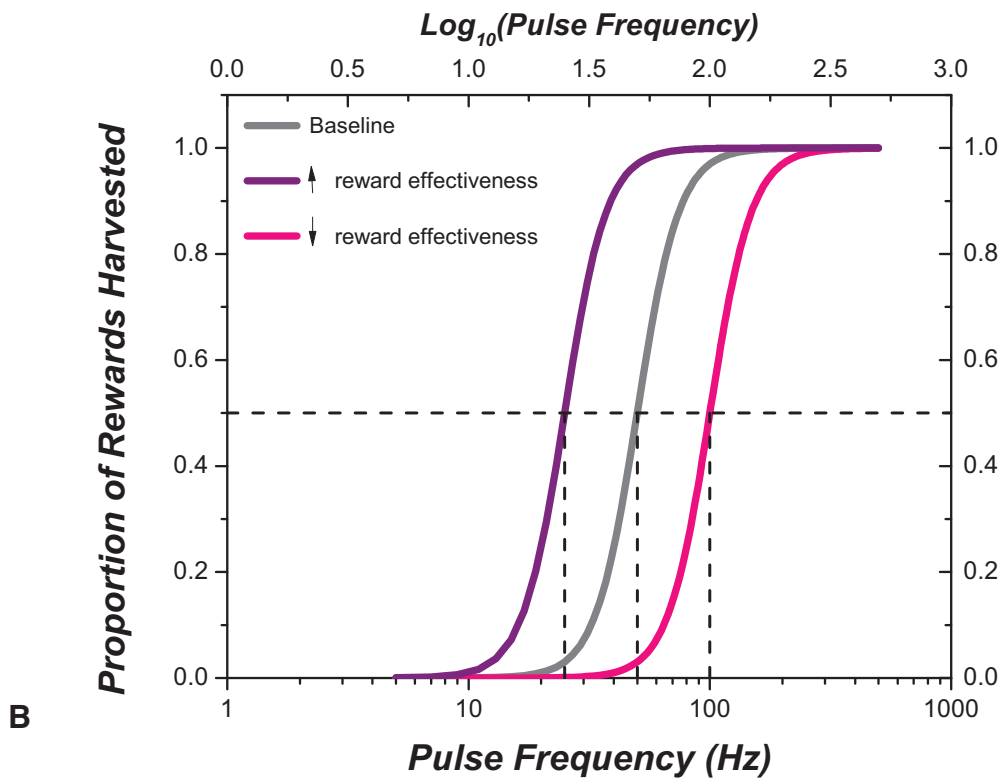
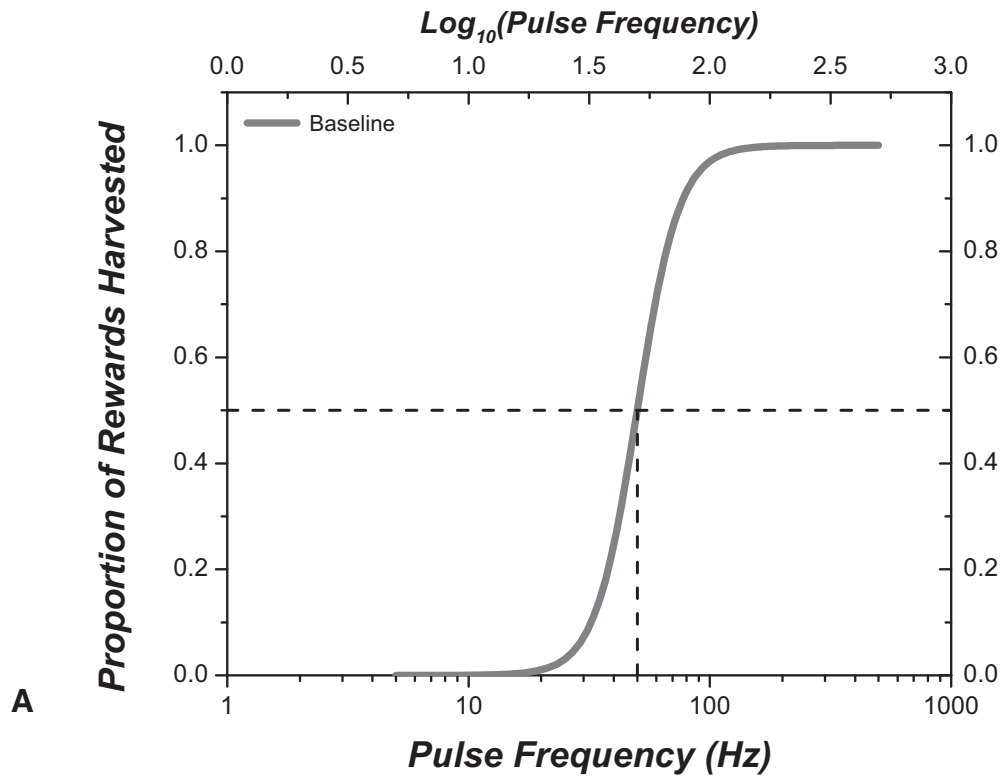


Figure 1. A. An example of a rate frequency curve. As pulse frequency is increased, the proportion of rewards harvested is initially low but increases rapidly over a relatively small pulse frequency range and eventually levels-off. The index of measurement is the pulse frequency required to maintain a given level of behaviour: in this example, the criterion level of behaviour is half-maximal performance denoted by the dotted vertical line (referred to as  $F_{M50}$ ). B. Lateral shifts of the curve from baseline. If a drug potentiates the rewarding effect of the electrical stimulation, the curve will shift leftward from baseline ( $F_{M50}$  is reduced) as denoted by the purple curve. Less electrical stimulation is required to maintain half-maximal performance because the rewarding efficacy of each train of electrical stimulation has been boosted by the drug. On the other hand, the curve will shift rightward if a drug attenuates the rewarding effect of the stimulation ( $F_{M50}$  is increased) as denoted by the pink curve. In this case, to maintain half-maximal performance, a higher pulse frequency train of electrical stimulation is needed relative to baseline because the rewarding efficacy of the stimulation has been attenuated by the drug.

the rat is not attending to the lever because the stimulation is not very rewarding. Figure 1A illustrates the rate frequency curve in semi-logarithmic space: as pulse frequency is increased, proportion of reward harvested is initially low but increases rapidly over a relatively small pulse frequency range and eventually levels-off.

The index of measurement is the pulse frequency required to maintain a given level of behaviour, usually half-maximal performance. This index is referred to as the pulse frequency threshold or  $F_{M50}$ . The effect of a physiological manipulation is reflected in the lateral shifts of this curve (Figure 1B). If a drug potentiates the rewarding effect of the electrical stimulation, the curve will shift leftward from baseline ( $F_{M50}$  is reduced). Less electrical stimulation is required to maintain a given level of performance because the rewarding impact of each train of electrical stimulation has been boosted by the drug. On the other hand, the curve will shift rightward if a drug attenuates the rewarding effect of the stimulation ( $F_{M50}$  is increased). In this case, to maintain a given level of performance, a higher pulse frequency train of electrical stimulation is needed relative to baseline because the rewarding impact of the stimulation train has been attenuated by the drug. Manipulations affecting the rewarding efficacy of the electrical stimulation are said to affect the “sensitivity” of the reward substrate.

Manipulations that affect the subject’s capacity to perform the task had been traditionally understood to affect the curve’s upper asymptote but not the pulse frequency threshold. However, it was later shown that manipulating the response cost such as effort (energetic exertion) can indeed cause lateral shifts as well (Fouriez, Bielajew & Pagatto 1990; Frank & Williams 1985). Figure 2 shows the lateral shifts that occur when increasing or decreasing the cost. When the stimulation train is made more “expensive”

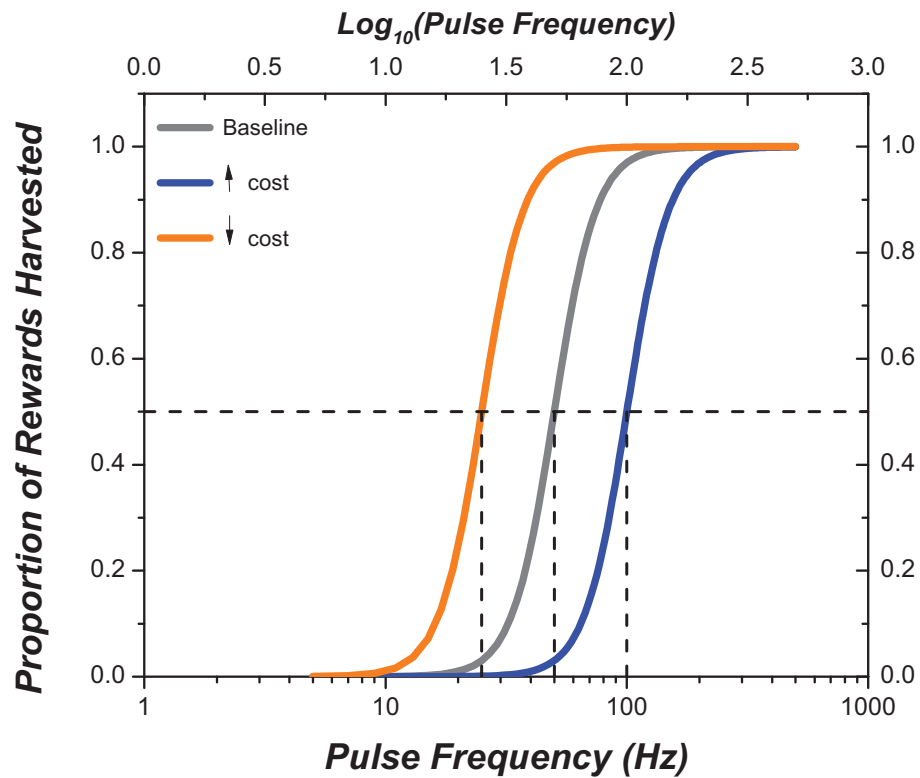


Figure 2. Lateral shifts of the rate frequency curve occur when increasing or decreasing the cost. When the stimulation train is made more “expensive” (e.g., energetic costs are increased), to maintain a given level of responding, (such as half-maximal performance) the required pulse frequency is increased and the curve shifts rightward. When the stimulation is made “cheaper” (e.g., energetic costs are reduced) the pulse frequency required to maintain a given level is reduced and the curve shifts leftward.



energetic costs are increased), the curve shifts rightward. That is, to maintain a given level of responding, an increase in pulse frequency is required. When the stimulation is made “cheaper” (energetic costs are reduced), the curve shifts leftward; to maintain a given level of responding, a reduction in pulse frequency is required.

The lateral shifts that occur due to the manipulation of the reward costs indicate that using this 2-dimensional measurement strategy to infer changes in reward intensity following a physiological manipulation is inherently flawed. One cannot distinguish whether a physiological manipulation to the reward circuitry affects the intensity of the reward or another variable implicated in reward seeking, such as the perceived cost of the reward. Thus, indistinguishable curve-shifts on a 2-dimensional plot may result from physiological manipulations that affect different stages of processing. Figure 3 demonstrates that decreasing the reward effectiveness of the stimulation can have the same effect on the rate frequency curve as increasing the cost of the stimulation: the two curves are superimposable. Consequently, incorrect conclusions can result from the use of this 2-dimensional method.

### ***1.3.3. The 3-dimensional reward mountain model***

To eliminate this ambiguity, Arvanitogiannis and Shizgal (2008) and Hernandez and colleagues (2010) proposed a 3-dimensional model termed the “reward-mountain model” that measures performance as a function of both reward strength and cost. The reward strength of the stimulation is experimenter-set by the pulse frequency (Hz). The experimenter-imposed cost, also referred to as the “price,” is an “opportunity cost”: the time (in seconds) that is required for the subject to hold down the lever to reap the reward

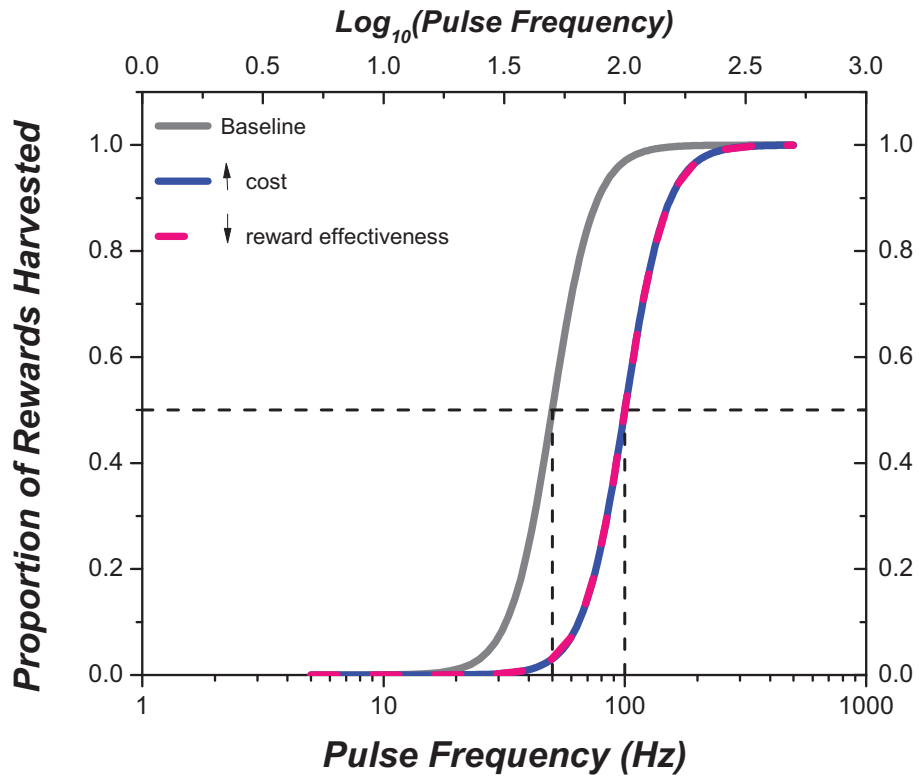


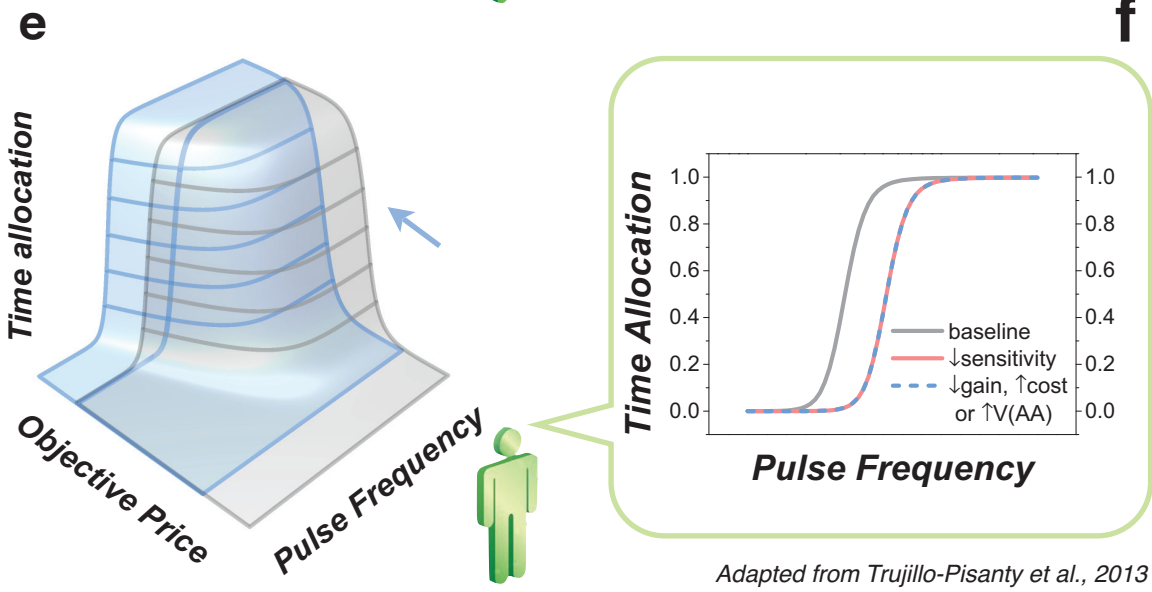
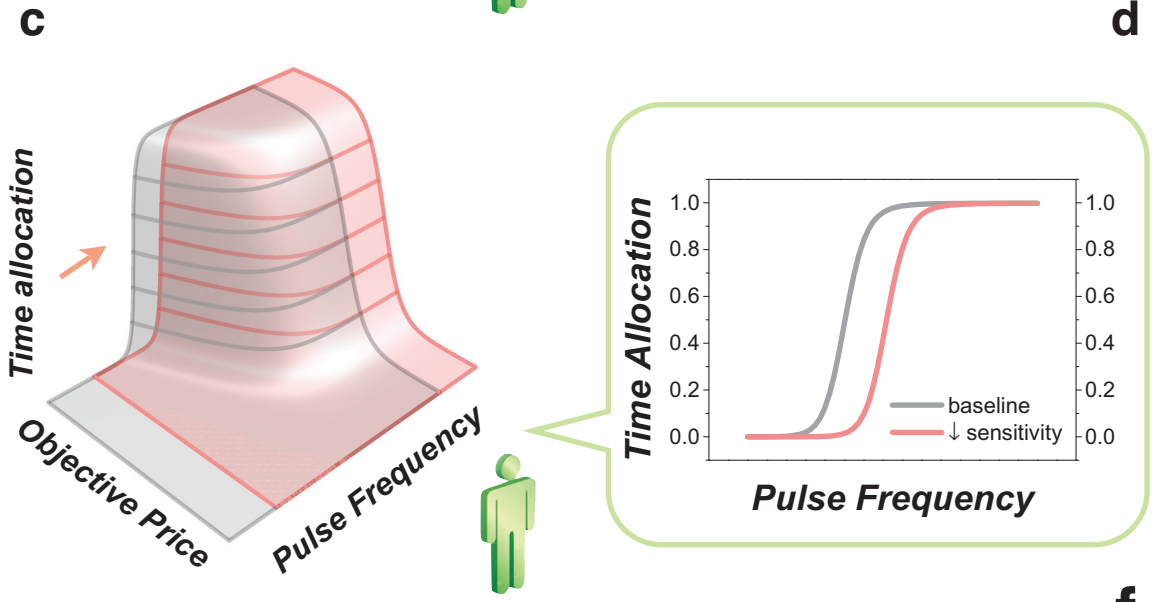
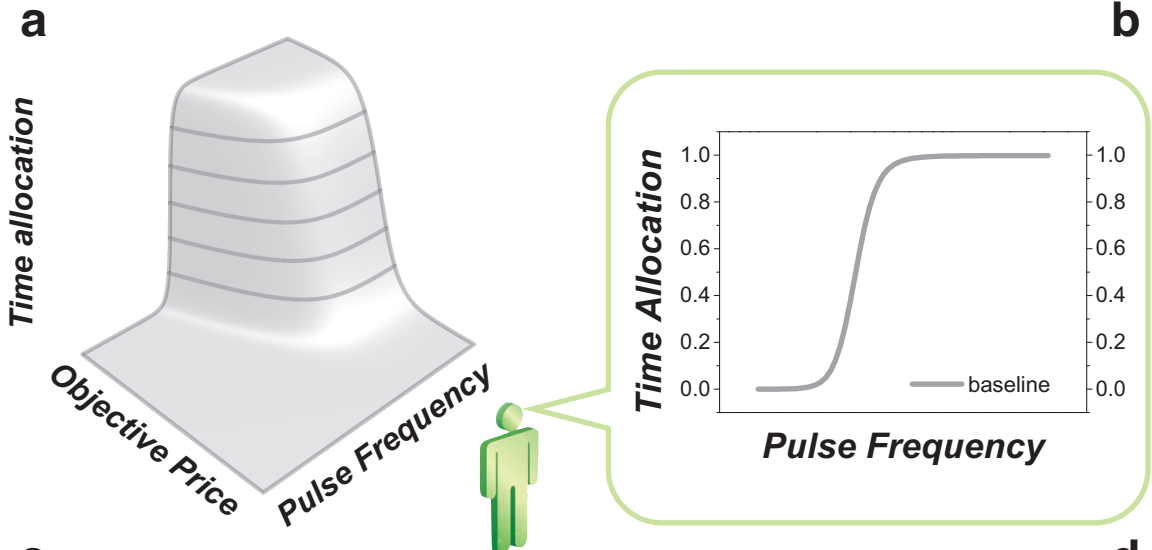
Figure 3. Decreasing the reward effectiveness or increasing the cost of the stimulation can have identical effects on the displacement of the rate frequency curve. The indistinguishable curve shifts imply that this measurement technique is inherently flawed: one cannot distinguish whether a physiological manipulation to the reward circuitry affects the subjective intensity of the reward or another variable implicated in reward seeking, such as the perceived cost of the reward.

at the expense of other “opportunities” or activities it could be engaging in such as grooming or exploring in the operant box. The schedule of reinforcement used in the present studies is that of the “fixed cumulative handling time schedule” (FCHT) which tightly controls opportunity costs so as to preclude engagement in alternate activities while the subject works for the reward. Specifically, this schedule requires that the subject hold down the lever for the cumulative required time (termed the price). For instance, for a 10 s price, the rat could hold-down the lever, release the lever for any given amount of time, and resume the hold-down task for the rest of the imposed time interval to obtain the reward.

The dependent variable is time allocation, the amount of time that the rat holds down the lever (“working”) as a proportion of the total time available. Time allocation increases as the pulse frequency is increased or when the price is decreased (see Figure 6 A).

We distinguish between the “objective” price and “subjective” price. The objective price (*OP*) is the experimenter-set price while the subjective price (*SP*) is the rat’s interpretation of the objective price. The two variables may not necessarily be identical along the whole range. For now we assume that objective price equals the subjective price but will revisit this relationship in Section 1.4.

Figure 4 shows again the ambiguity of the 2-dimensional curve shifts as in Figure 3 but also demonstrates how these shifts are explained by the 3-dimensional representation and are distinguished clearly in that view. On the basis of the limited perspective of the 2-dimensional representations, one cannot ascertain the direction in which the corresponding 3-dimensional structure has moved. The green figure views the



Adapted from Trujillo-Pisanty et al., 2013

Figure 4. Displacements of the 3-dimensional reward mountain representation (left column) are ambiguous in the corresponding 2-dimensional representation (right column). The green figure facing the pulse frequency axis perceives the world in 2 dimensions. The price axis is hidden to him and he sees the corresponding 3-dimensional structure as a 2-dimensional silhouette, the *left* outline of the 3-dimensional structure. Panels b,d,f show the left outlines of the silhouettes perceived by the green figure. In panel f, the dashed blue line that is due to a shift of the 3-dimensional structure (panel e) along the price axis is superimposed on the solid pink outline of the mountain that has shifted along the pulse frequency axis (panel c). The superimposed curves demonstrate that the 2-dimensional projections from both of these different effects are essentially identical, yet, a 3-dimensional representation can clearly distinguish between these effects.

world in only 2 dimensions. From his perspective (Figure 4a), facing the pulse frequency axis, the mountain is collapsed onto its 2-dimensional silhouette, in the plane defined by time allocation as a function of pulse frequency. The left outline of the silhouette of the reward mountain in grey (Figure 4a) is represented by the grey curve seen in Figure 4b of the 2-dimensional plot. If a manipulation such as a lesion caused a reduction in rewarding efficacy of the stimulation, a shift of the mountain structure along the pulse frequency axis (Figure 4c) occurs. This would be viewed in 2 dimensions as a rightward shift of the silhouette along the pulse frequency axis (Figure 4d). If the perceived cost of the stimulation were increased by some manipulation, from the 3-dimensional perspective, the mountain would shift along the price axis (Figure 4e). The critical consequence of this shift along the price axis and argument for the necessity of the 3-dimensional measurement strategy is that this shift drags the mountain such that the silhouette of the 2-dimensional perspective of time allocation as a function of pulse frequency is also shifted rightward (Figure 4f).

The 2-dimensional projections from both of these effects (on reward strength and cost) are essentially identical, yet, as shown in Figure 4, a 3-dimensional representation can clearly distinguish between them. The reward mountain eliminates this ambiguity. A further advantage of this measurement strategy is that shifts in the position of the mountain model can be related to different variables and stages of reward processing. This link can be made because the reward mountain model is based on the “minimal model,” the sequence of stages that ties the neural signal through a multistage circuitry (Gallistel, Shizgal, & Yeomans, 1981) to behaviour. The following section describes

these stages of reward processing and the associated displacements of the 3-dimensional reward mountain structure that may occur after perturbations to the described stages.

#### ***1.3.4. The stages of the minimal model***

The minimal model depicted in Figure 5 describes the stages that link the electrical signal to the behavioural output of the subject. Briefly, the directly stimulated neurons (Figure 5A-1) provide input to an “integrator” (Figure 5A-2) that combines the effect of incoming action potentials over time (duration of the pulse train) and space (the number of neurons that are activated). The output of this integrator is the rat’s subjective evaluation of the neural signal produced by the integrator, termed “the subjective reward intensity.” This output is passed through a peak detector to measure the peak reward intensity of a given stimulation train (Figure 5A-3). Next, the rat makes subjective estimates of the associated costs of obtaining the reward (Figure 5A-4). Then, the subject weighs the peak subjective reward intensity of the electrical signal with the subjective costs of obtaining the reward to compute a “payoff” of the electrical stimulation; this computation is likened to the weighing of the benefits to costs of a good (Figure 5A-5). The subject also computes a payoff of the alternate activities it may engage in during the experiment when not working for the reward such as exploring or grooming (Figure 5A-6). These two payoffs are compared; subsequently, the subject makes a choice between either allocating its time to performing the task to obtain the reward (“working”) or engaging in other activities. This choice is reflected in its behaviour and measured as time allocated to performing the task as a function of both reward strength and cost

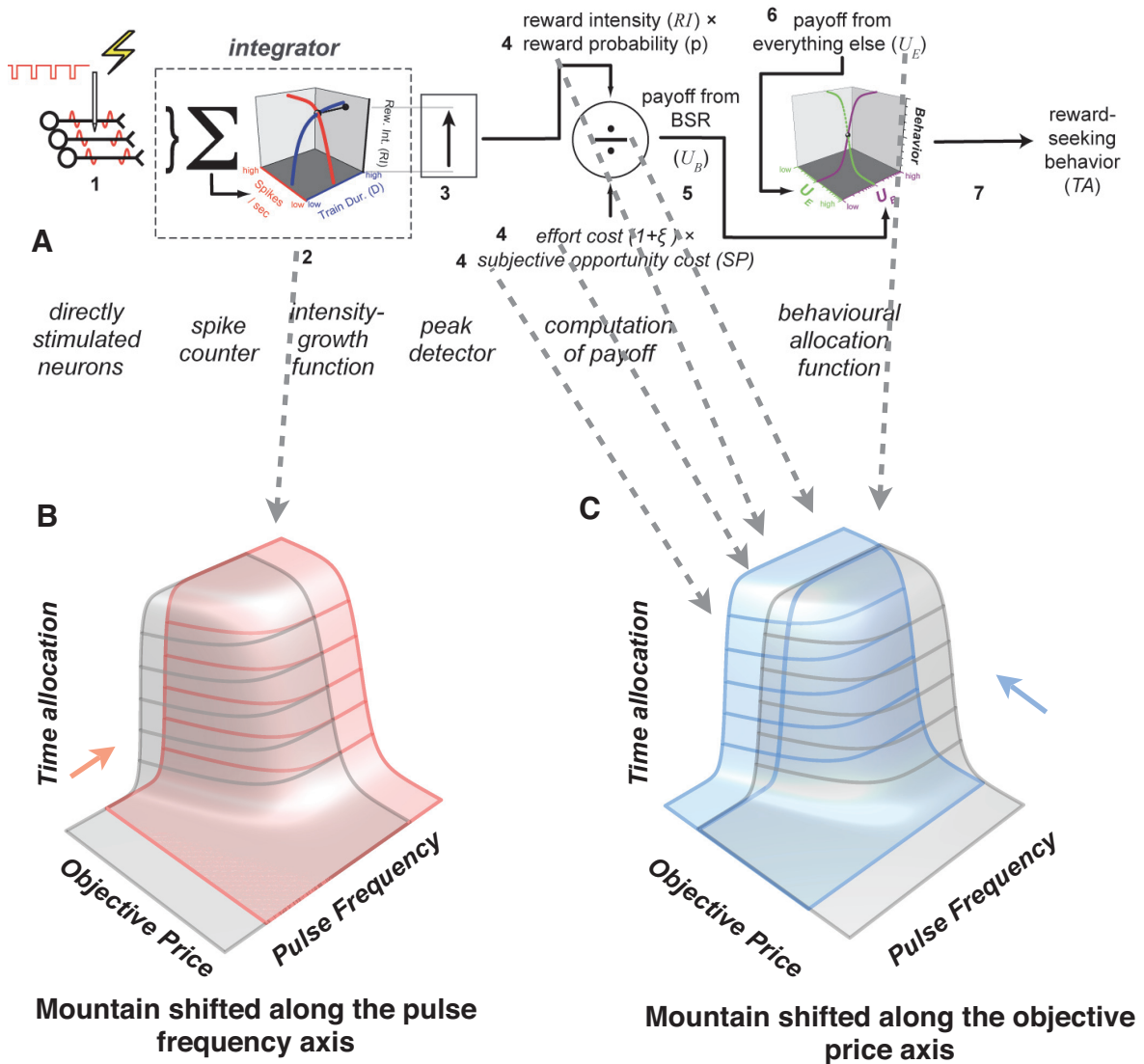


Figure 5. Graphical representation and summary of the reward-mountain model. The action potentials produced by the electrical stimulation (1) are integrated temporally and spatially and transformed into a signal representing subjective reward intensity (2). The peak reward intensity (3) is combined in a scalar manner with subjective estimates of probability and cost (4) so as to compute a payoff ( $U_B$ ) (5). The rat computes a payoff of “everything else” ( $U_E$ ) (6). It compares the payoff of BSR with the payoff of everything else and allocates its behaviour accordingly (7). B. Time allocation is represented as a function of reward strength (pulse frequency) and objective opportunity cost (objective price) of the stimulation. Perturbations acting prior to the output of the integrator shift the 3-dimensional structure along the pulse frequency axis. These shifts are measured as changes in parameter  $F_{hm}$  that locates the structure along the pulse-frequency axis. C. Perturbations acting subsequently to the output of the integrator displaces the structure along the price axis as measured by changes in parameter  $OP_e$  that locates the structure along the price axis.

Adapted from Breton et al., 2013



(Figure 5A-7). A more in depth and formal description of the stages of the model is outlined out below.

#### ***1.3.4.1. The counter model***

A train of electrical stimulation excites the neurons surrounding the electrode tip, dubbed the first-stage neurons. The subsequent stage(s) of processing is termed the “integrator.” Within the integrator, the targets of the directly activated neurons have been shown to act in a simple manner, as a spike counter as if to count the aggregate number of spikes conveyed by the directly stimulated neurons within a given time window (Gallistel, Shizgal, & Yeomans, 1981). This counting is represented by  $\Sigma$  in the illustration in Figure 5A.

The current sets the effective radius of neural activation, consequently, the number of neurons activated. The experimenter-applied pulse frequency sets the spike rate. (It is assumed for the moment that each neuron fires once per induced pulse, this will be revisited in Chapter 3). Thus, the counter model specifies that holding the stimulation train duration constant, a high frequency, low current train will produce the same rewarding effects as a low frequency, high current train. For example, firing 10 neurons, 100 times produces the same rewarding effect as firing 100 neurons, 10 times.

### ***1.3.4.2 Reward intensity function: reward intensity as a function of pulse frequency and train duration***

The aggregate spike rate is transformed into a reward intensity by a 3-dimensional function. Reward intensity grows non-linearly and approaches asymptote as the aggregate spike rate is increased or as the duration of the stimulation train is increased. Holding current constant, the spike rate can be expressed as the experimenter-set pulse frequency  $F$ . The reward intensity is a logistic function that initially grows steeply and levels-off (Leon & Gallistel, 1992; Simmons & Gallistel, 1994; Conover & Shizgal, 2005; Sonnenschein, Conover, & Shizgal; 2003):

$$RI(D, F) = RI_{\max} \times \frac{F^g}{F^g + [F_{hm}(D)]^g}$$

Rearranging terms such that a relative reward intensity value is computed:

$$RI_{rel}(D, F) = \frac{F^g}{F^g + [F_{hm}(D)]^g}$$

where,

$RI$  = reward intensity

$RI_{rel}$  = relative reward intensity,  $RI/RI_{\max}$  which varies from 0 to 1

$RI_{\max}$  = the maximal reward intensity

$D$  = duration of the stimulation train in seconds (held constant at 0.5 s in these experiments)

$F$  = pulse frequency (Hz)

$F_{hm}$  = pulse frequency that produces half-maximal reward intensity (Hz)

$g$  = the intensity growth exponent, controls the steepness at which  $RI_{rel}$  grows

Embedded in the reward intensity function described above is the hyperbolic strength duration function that measures  $F_{hm}(D)$  to account for the effect of train duration on reward intensity. In the strength duration function,  $F_{hm}$ , the frequency required to drive reward intensity to its half-maximal value, is a rapidly declining rectangular hyperbolic function of train duration (Gallistel, 1978; Sonnenschein et al. 2003; Hernandez et al. 2010):

$$F_{hm}(D) = F_{hm_R} \times \left(1 + \frac{C}{D}\right)$$

where,

$F_{hm_R}$  = the rheobase: the pulse frequency required to produce a reward of half-maximal intensity at an infinitely long duration, the horizontal asymptote

$C$  = the chronaxie; the train duration at which  $F_{hm}(D)$  is twice  $F_{hm_R}$ , determines the curvature of the function

Taken together, the subjective reward intensity as a function of pulse frequency is represented by the blue curve in Figure 5A-2: reward intensity as a function of spike rate grows then levels-off. No further increases in pulse frequency cause increases in subjective reward intensity. Also as represented by red curve in Figure 5A, the reward

intensity grows as a function of train duration (the duration of a stimulation train in seconds) and levels-off. In other words, prolonging the duration of the stimulation train does not alter performance once the duration exceeds some critical value (Shizgal & Matthews, 1977; Gallistel, 1978, Sonnenschein et al., 2003).

The stored value of reward intensity is assumed to reflect the peak reward intensity value. It is proposed that the output of the intensity-growth function is passed through a “peak detector” en route to memory such that the peak (or maximum subjective reward intensity) of a given train is stored (Gallistel, 1978; Sonnenschein, Conover, & Shizgal, 2003). This peak detector is represented by the vertical upward arrow in the Figure 5A-3. The peak detector can be thought of as a running tab or gauge throughout the duration  $D$  of the stimulation train. At each point in time ( $t$ ) during the duration of the train, a detector assesses the reward intensity magnitude. The peak reward intensity is the maximal reward intensity that has been detected by the end of the stimulation train. At the end of the stimulation train, this peak reward intensity ( $RI_{peak}$ ) is recorded and stored in memory. This stage is expressed as:

$$RI_{peak}(t, F) = RI(D, F) \\ 0 \leq t \leq D$$

where,

$RI_{peak}$  = peak reward intensity registered during the stimulation train

$D$  = the total duration of stimulation train in seconds

$t$  = the time during the stimulation train

Thus, in the above formulation, peak reward intensity ( $RI_{peak}$ ) is a function of pulse frequency ( $F$ ) and time ( $t$ ) of the entire train duration  $D$ . The maximal reward intensity of a stimulation train of given pulse frequency  $F$  and train duration  $D$  is assessed.

#### ***1.3.4.3. Subjective estimation of external values and computation of payoff***

After the properties of the electrical stimulation are “integrated” and the peak reward intensity is stored in memory, as described in the previous steps, the subjective reward intensity is then combined with “external” variables relating to the reward such as subjective estimates of subjective reward probability, effort cost (rate of exertion to meet a response requirement such as the number of bar presses or the amount of force required to hold down a lever) and opportunity cost. Thus, the rat makes the estimates of these external values, transforming the objective values into subjective ones (Figure 5A-4).

On the basis of the matching law and its generalizations that specify that different dimensions of reinforcement are combined multiplicatively (Miller, 1976; Killeen, 1972, Baum & Rachlin, 1969; Gallistel & Leon, 1991; Neuringer, 1977; Rachling 1989), the stored reward intensity values are proposed to be combined with these “external” values in a scalar (multiplicative) fashion.

This combination is named the “payoff” represented by  $U_B$ , an index of how “worthwhile” it is to the rat to work for the electrical stimulation. (This is represented in Figure 5A-5). Colloquially, we can describe the subject as computing the index of payoff as a ratio of benefits (the reward intensity of the electrical stimulation) to the costs of obtaining the electrical stimulation (the product of the effort and opportunity cost).

$$U_B(p, P, RI_{peak}, \xi) = \frac{RI_{peak} \times p}{(1 + \xi) \times OP}$$

where,

$RI_{peak}$  = peak reward intensity registered during the stimulation train

$U_B$  = net payoff of BSR

$p$  = subjective probability that BSR will be delivered once the work requirement has been satisfied

$OP$  = objective price, the opportunity cost in seconds set by the experimenter, assuming for now that the experimenter-set objective price equals the subjective price

$\xi$  = the subjective rate of exertion required to hold down the lever (kcal)

(The addition of 1 in the denominator prevents the extreme growth of  $U_B$  as  $\xi$  becomes very small.)

Next, the rat will evaluate the payoff deriving from activities other than the operant task that it can engage in such as resting, exploring, and grooming. The payoff derived from “everything else” is represented by  $U_E$ .

#### ***1.3.4.4. Transformation of payoff into behaviour***

The rat will then compare the payoff derived from working for electrical stimulation to the payoff derived from everything else and allocate its time to these options accordingly. Therefore, the last stage is the translation of the comparison of

payoffs into behaviour. Specifically, the comparison that the rat makes is represented by the expression related to McDowell's (2005) generalization of Herrnstein's single operant matching law (Herrnstein, 1970; Herrnstein, 1974; Heyman, 1988). (Figure 5A-7)

$$TA(U_B, U_E) = TA_{\min} + \left[ (TA_{\max} - TA_{\min}) \times \frac{U_B^a}{U_B^a + U_E^a} \right]$$

where,

$a$  = the payoff-sensitivity exponent, represents how sensitive the rat is to the price of the reward; it accounts for over or under-matching

$TA_{\max}$  = the maximum time allocation

$TA_{\min}$  = minimum time allocation

$U_E$  = payoff from "everything else" or activities other than the operant task

$U_B$  = payoff from electrical stimulation

(The maximal and minimal time allocation ( $TA_{\max}$  and  $TA_{\min}$ ) in the equation scales the time allocation.)

The payoff from electrical stimulation ( $U_B$ ) is expanded in terms of the stimulation parameters as described previously in Section 1.3.4.2. and 1.3.4.3. The payoff of everything else ( $U_E$ ) can also be expressed in terms of electrical stimulation parameters. How this is possible is described below.

**1.3.4.4.1. How the payoff of “everything else” is specified in terms of rewarding electrical stimulation variables**

The payoff from “everything else”,  $U_E$  can actually be expressed in terms reward intensity deriving from the electrical stimulation and the associated costs at specific experimental parameters. This way of expressing  $U_E$  is possible because the payoff of BSR ( $U_B$ ) is equal to the payoff of everything else ( $U_E$ ) when the rat allocates half of its time to working and half of its time to everything else at a specific reward intensity and cost of the stimulation. Specifically, the price at which time allocation is half-maximal ( $TA = 0.5$ ) at the maximal reward intensity ( $RI_{max}$ ) can be measured and is referred to as parameter  $OP_e$ . Therefore, the payoff from the electrical stimulation at which the parameters (reward intensity and cost) are set to values such that the rat spends half of its time working for the stimulation (thus half of its time engaging in other activities, ‘everything else’) is denoted as  $U_{midBSR}$  and can be expressed as:

$$U_{midBSR} = \frac{RI_{max} \times p}{(1 + \xi) \times OP_e}$$

Because the payoff of working for the stimulation for half of the available time ( $U_{midBSR}$ ) is equal to the payoff of ‘everything else’ ( $U_E$ ),

$$U_E = \frac{RI_{max} \times p}{(1 + \xi) \times OP_e}$$



### 1.3.4.4.2. Time allocation expressed in terms of pulse frequency and price

Given that,

$$RI(D, F) = RI_{\max} \times \frac{F^g}{F^g + [F_{hm}(D)]^g}$$

and,

$$RI_{peak}(t, F) = RI(D, F) \quad 0 \leq t \leq D$$

and,

$$U_B(p, P, RI_{peak}, \xi) = \frac{RI_{peak} \times p}{(1 + \xi) \times OP}$$

and,

$$U_E = \frac{RI_{max} \times p}{(1 + \xi) \times OP_e}$$

and,

$$TA(U_B, U_E) = TA_{\min} + \left[ (TA_{\max} - TA_{\min}) \times \frac{U_B^a}{U_B^a + U_E^a} \right]$$

Then, time allocation as a function of pulse frequency and price and can be expressed:

$$TA(D, F, P, \xi) = TA_{\min} + \left[ (TA_{\max} - TA_{\min}) \times \frac{\left( \frac{RI_{\max} \times p}{(1 + \xi) \times OP} \times \frac{F^g}{F^g + [F_{hm}(D)]^g} \right)^a}{\left( \frac{RI_{\max} \times p}{(1 + \xi) \times OP} \times \frac{F^g}{F^g + [F_{hm}(D)]^g} \right)^a + \left( \frac{RI_{\max} \times p}{(1 + \xi) \times OP_e} \right)^a} \right]$$

Multiplying the payoff expressions by  $\left( \frac{(1 + \xi) \times OP}{RI_{\max} \times p} \right)$  in order to simplify the expression:

$$TA(D, F, P) = TA_{\min} + \left[ (TA_{\max} - TA_{\min}) \times \frac{\left( \frac{F^g}{F^g + [F_{hm}(D)]^g} \right)^a}{\left( \frac{F^g}{F^g + [F_{hm}(D)]^g} \right)^a + \left( \frac{OP}{OP_e} \right)^a} \right]$$

Thus, time allocation as a function of pulse frequency and objective price is defined by the above expression. The wire mesh in Figure 6 represents this function.

### ***1.3.5. Displacements of the mountain can reveal at what stages perturbations act***

Displacements of the reward-mountain model along the pulse frequency axis (Figure 5B) occur due to manipulations that act on stages prior to the output of the reward integrator, before reward intensity is processed. For instance, these perturbations may include lesions that reduce the number of neurons that carry the reward signal or drugs

that alter the output of these neurons. Alterations that occur before the output of the integrator are said to cause changes in “sensitivity” of the reward substrate.

Displacements of the reward mountain along the price axis are caused by influences acting on later stages of the minimal model. There are several stages following reward integration that can be affected. For instance, alterations can influence the perceived cost of the reward, or alternatively, they can alter the payoff derived from alternative activities that the subject can engage in. In addition, perturbations could influence the scaling of the integrator’s output; this is referred to as a changing the substrate’s “gain.” An increase in gain would increase by a common multiplicative factor the reward intensity produced by each pulse frequency.

Within the framework of the minimal model, shifts along the price axis do not reveal the exact stage that is affected, only that it occurred at a stage subsequent to the output of the integrator. Nonetheless, whether the physiological manipulation acted on the circuitry implicated in the stages that occur before or after the integrator provides important insight into the circuitry under investigation. For instance, a long-standing view regarding cocaine (and by extension, dopamine) is that it increases the sensitivity of the reward substrate. Conventional methods using 2-dimensional representations cannot directly and unambiguously evaluate this hypothesis. The 3-dimensional methodology demonstrated that the effect of cocaine on rewarding electrical stimulation caused the reward mountain to shift exclusively along the price axis, challenging the conventional view (Hernandez, Breton, Conover, and Shizgal, 2010). Additional studies involving dopamine agonists and antagonists (Hernandez, Trujillo-Pisanty, Cossette, Conover, & Shizgal 2012; Trujillo-Pisanty, Conover, and Shizgal, 2013) show that this methodology

detects displacements exclusively along one axis (price axis). Taken together, this methodology provided important insights into the role of dopamine in reward-seeking behaviour. It has also been used to investigate the role of cannabinoid receptor blockade (Trujillo-Pisanty, Hernandez, Moreau-Debord, Cossette, Conover, Cheer, & Shizgal, 2011) in which the displacement of the reward mountain exclusively along the price axis was also demonstrated. As well, several studies have further validated this methodology (Arvanitiogiannis & Shizgal, 2008; Breton, Conover, & Shizgal, 2013).

### ***1.3.6. Overview of the reward mountain model methodology***

Figure 6 shows an example of a 3-dimensional reward mountain with simulated parameters (panel A and B). Its corresponding contour graph in panel C is seen from an aerial view (B). Each contour curve is horizontal a slice through the mountain structure and corresponds to a given time allocation. The two “location parameters” are  $F_{hm}$  (denoted by the red line), which locates the reward mountain along the pulse frequency axis and  $OP_e$  (denoted by the blue line), which locates the structure along the objective price axis. Specifically,  $F_{hm}$  is the pulse frequency axis at half-maximal reward intensity and  $OP_e$  is the price at which the subject allocates half-maximal time to working for the reward when the reward intensity is set to a maximal value). These parameters are the indices used to assess the extent of changes along the pulse frequency axis ( $F_{hm}$ ) and price axis ( $OP_e$ ) before and after a manipulation.

To obtain this reward mountain, conventionally, 3 sampling vectors are employed (Figure 7A). A “*pulse frequency sampling matrix*” is comprised of 14 points (red) in which the price is held constant and pulse frequency varies. A “*price sampling matrix*” is

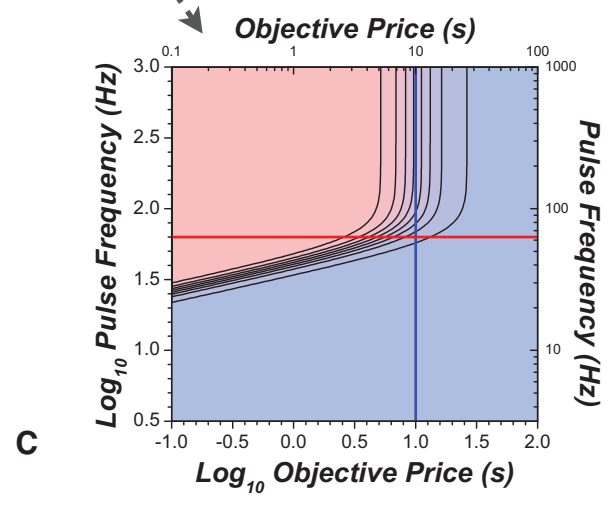
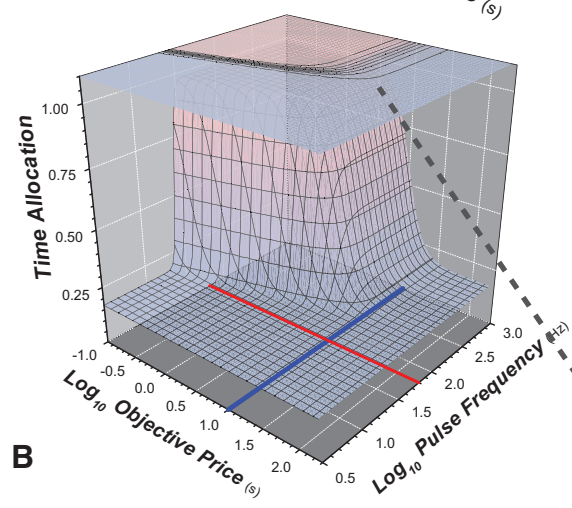
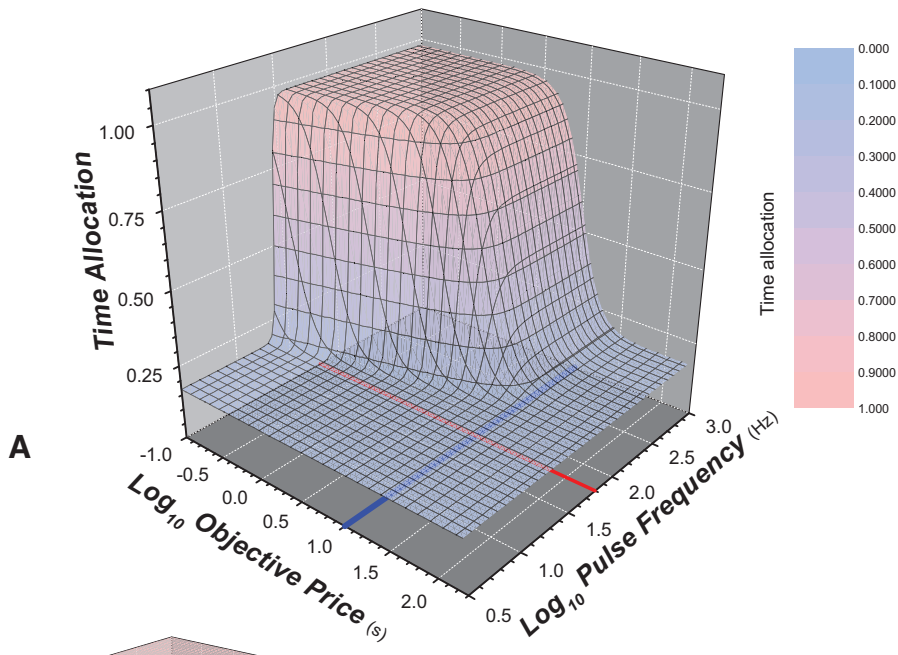
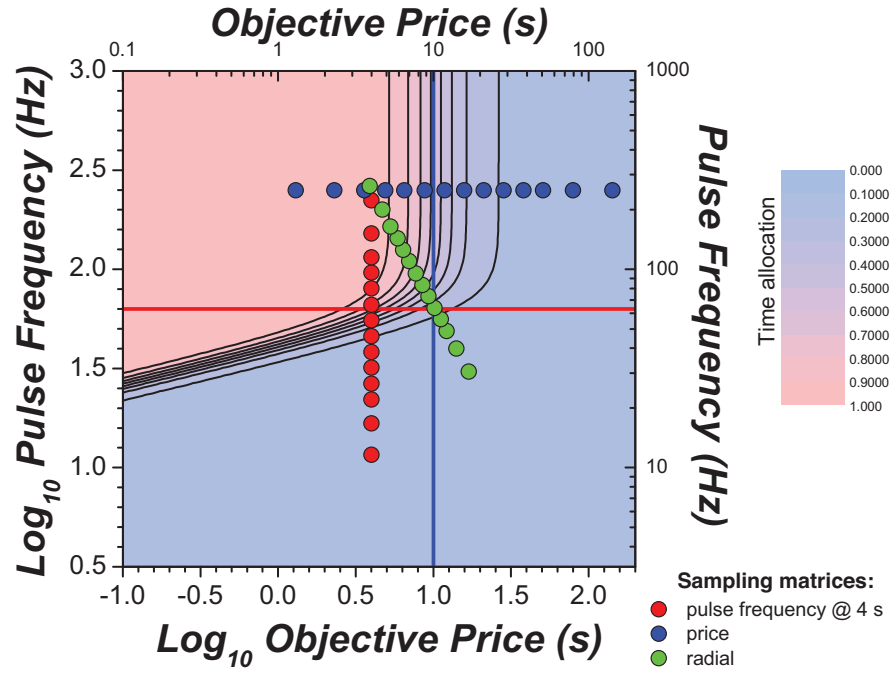
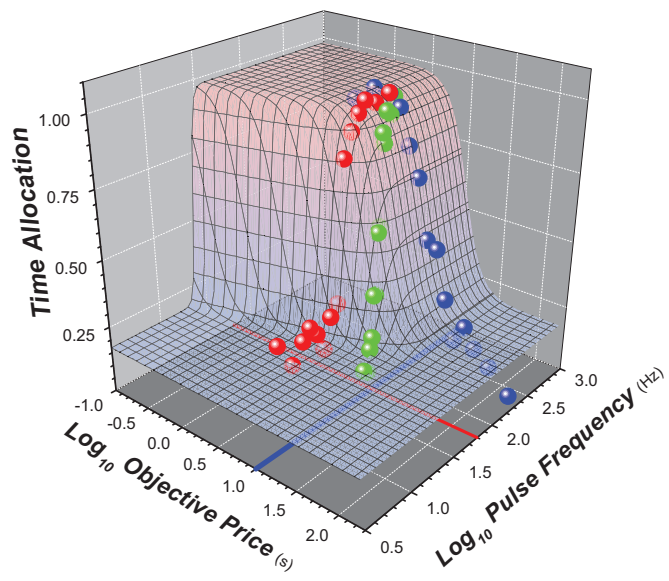


Figure 6. An example of a 3-dimensional reward mountain with simulated parameters (panel A and B). The colour gradient indicates that time allocation is highest at pink shades and lowest at blue shades. Its corresponding contour graph in panel C is seen from an aerial view (B). Each contour curve is a horizontal slice through the mountain structure and corresponds to a given time allocation. The two “location parameters” are  $F_{hm}$  that locates the reward mountain along the pulse frequency denoted by the red line and  $OP_e$  that locates the structure along the price axis denoted by the blue line. These parameters are the indices used to assess the extent of changes along the pulse frequency axis ( $F_{hm}$ ) and price axis ( $OP_e$ ) before and after an experimental manipulation.



A



B

Figure 7. To obtain the reward mountain, conventionally, 3 sampling matrices are employed (Figure 7A). A pulse frequency sampling matrix is comprised of 14 points (red) in which the price is held constant. A price sampling matrix is comprised of 14 points (blue) in which the price varies but the pulse frequency is constant. A radial sampling matrix is comprised of 14 points (green) in which the objective price is increased while the pulse frequency is decreased. These three sampling matrices provide sufficient information to fit the 3-dimensional time allocation function represented by the wire mesh in panel B. Figure 7 shows the sampling vectors used to fit the mountain on the contour plot (A) and the actual points in the 3-dimensional space (B).



comprised of 14 elements (points in blue) in which the price varies but the pulse frequency is constant. A “*radial sampling matrix*” is comprised of 14 points (green) in which the price is increased while the pulse frequency is decreased. These three sampling matrices provide sufficient information to fit the 3-dimensional time allocation function. Figure 7 shows the sampling vector used to fit the mountain on the contour plot (A) and the actual data points (time allocation as a function of pulse frequency and objective price) in the 3-dimensional space (B).

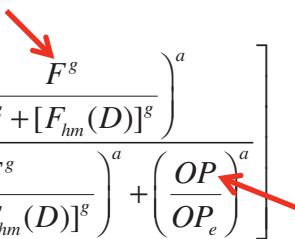
#### ***1.4. Thesis aim: the transformation of objective variables into their subjective equivalents***

Within the stages of the minimal model, certain inherent, unrealistic assumptions have been made regarding the independent variables. The first assumption is that of the experimenter-imposed opportunity cost (in seconds). The experimenter manipulates the opportunity cost, and the subject combines the subjective reward intensity to estimate a payoff. However, before this step, the subject must evaluate the experimenter-imposed opportunity cost (termed the objective opportunity cost or objective price, *OP*) and translate it into the subjective equivalent. In the model, we assume that objective opportunity costs and subjective opportunity costs are equal to each other along the entire range of prices. However, this is an unrealistic assumption at very low costs. For instance, reducing the cost from 0.2 to 0.1 s does not allow the animal to perform additional “alternate activities” (e.g., grooming, resting) of any significance, therefore, low opportunity costs are most likely subjectively equivalent along the low objective

price range. The experiment in Chapter 2 develops the psychophysical function that describes the translation of objective opportunity costs into their subjective equivalents.

The other assumption concerns that of the second independent variable, pulse frequency. This assumption entails the physiological properties of the neurons. In the model, it is assumed that each neuron fires once per experimenter-induced pulse. However, this assumption must break down as the pulse frequency becomes high; there must be a limit to the maximum firing rate of any axon due to physiological properties such as synaptic blocking or fatigue as suggested by Gallistel (1978). The goal of the experiment in Chapter 3 was to model frequency following and its progressive failure as the pulse frequency is increased. By this method, the maximal firing frequency for the reward-relevant neurons of the lateral hypothalamus is determined.

In summary, the objective values of pulse frequency ( $F$ ) and price ( $OP$ ) have been used in previous studies using the reward mountain methodology in the time allocation function presented below. It is unrealistic to assume that the objective values are equivalent to their subjective (in the case of price) or physiological (in the case of pulse frequency) values. In the present experiments, the psychophysical functions describing the subjective equivalents of pulse frequency ( $F$ ) and objective price ( $OP$ ) will be developed (Chapter 2 and Chapter 3).

$$TA(D, F, P) = TA_{\min} + \left[ (TA_{\max} - TA_{\min}) \times \frac{\left( \frac{F^g}{F^g + [F_{hm}(D)]^g} \right)^a}{\left( \frac{F^g}{F^g + [F_{hm}(D)]^g} \right)^a + \left( \frac{OP}{OP_e} \right)^a} \right]$$


Once these psychophysical functions are developed, they will be incorporated into the reward mountain model for future experiments. Independent variable  $F$  (pulse frequency) will be replaced with the psychophysical function describing  $FF$ , the firing frequency (the actual induced firing frequency). Independent variable  $OP$  (objective price) will be replaced with the psychophysical function describing  $SP$  (the subjective price). Accordingly, the parameters  $F_{hm}$  and  $SP_e$  will be replaced with  $FF_{hm}$  and  $SP_e$ :

$$TA(D, F, P) = TA_{\min} + \left[ (TA_{\max} - TA_{\min}) \times \frac{\left( \frac{FF^g}{FF^g + [FF_{hm}(D)]^g} \right)^a}{\left( \frac{FF^g}{FF^g + [FF_{hm}(D)]^g} \right)^a + \left( \frac{SP}{SP_e} \right)^a} \right]$$

Figure 8 describes the stages of the reward mountain model that will be formally described by psychophysical transformations. The frequency-following response function is represented by the pink question mark and the opportunity-cost function is represented by the blue question mark.

The experiment in Chapter 4 uses the reward mountain model methodology to test the effects of a manipulation, electrolytic lesions on operant behaviour. To date, experiments employing pharmacological and electrical manipulations have used and validated the reward mountain methodology. In the present thesis, the use of this methodology to assess effects of electrolytic lesions on reward-seeking behaviour will further validate this approach and in particular validate this methodology to be used

alongside more specific lesioning techniques such as optogenetic mediated silencing or excitotoxic lesions in future studies.

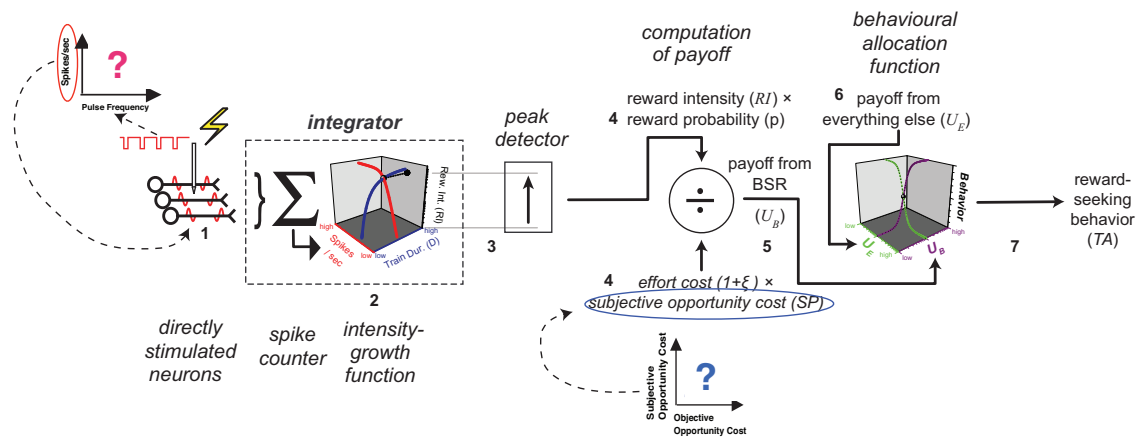


Figure 8. Graphical representation of the reward-mountain model. A. The psychophysical functions estimated in this thesis are denoted by the question marks: (i) the frequency-following response function represented in pink and (ii) the opportunity-cost function represented in blue.

## *Chapter 2*

### *Estimation of subjective opportunity costs*

#### **Abstract**

Time is a precious commodity to a forager because pursuit of one objective usually precludes simultaneous pursuit of another. Time spent searching for food reduces the time available for mating, nest building, and other essential activities. Thus, each exclusive activity entails what economists call an “opportunity cost:” the forgone benefits from competing activities that were eschewed in favor of the one that was pursued. Opportunity costs are an important factor in decision-making and include components related to travel, search, procurement, handling, and consumption of prey items. The present experiment derives the function mapping objective opportunity costs into subjective ones. The experiment was carried out in a simplified environment, an operant chamber. Operant performance was rewarded according to a “cumulative handling-time” schedule of reinforcement: rewarding electrical stimulation was delivered via a medial forebrain bundle electrode when the cumulative time a lever had been depressed by the rat reached a criterion duration (defined as the price or cost).

This function is particularly important in the context of the reward-mountain paradigm because the opportunity cost is a variable that is experimentally manipulated and is used to scale reward intensity. Previous versions of the reward-mountain model

incorporated the assumption that objective opportunity costs are equal to subjective opportunity costs along the entire tested range. Yet, psychophysical functions often differ from such identity functions; what we experience is not necessarily a direct reflection of the external world. We proposed that the opportunity-cost function is a non-identity function in the form of a hockey stick. The present results supported this form of the function. When opportunity costs were very low (e.g., working for only a small fraction of a second triggers delivery of a reward), the subjective cost varied little as a function of the objective cost. Specifically, although the rat may be capable of detecting the difference between 0.1 s and 0.2 s, it did not appear to increment its opportunity-cost function over such intervals. This range is represented by the flat “blade” of the function. The finding that low costs are subjectively equivalent makes sense: there may be no beneficial activities that the rat could substitute for work over such short intervals. As the objective opportunity cost rose, the subjective cost began to rise as well, and beyond values of  $\sim 2$  s, we estimated the subjective cost to rise at the same rate as the objective cost. This portion is represented by the “handle” of the function.

Furthermore, the subjective cost function we proposed was compared to ones derived from conventional hyperbolic and exponential models of delay discounting. The findings suggest the possibility that a different principle of intertemporal choice is involved in evaluating time spent working for reward and time spent waiting for reward after work requirements have been satisfied.

## ***Introduction***

### ***2.1. Time as an opportunity cost***

Time can be thought of as a cost. Humans and animals invest their valuable time acquiring, producing, and maintaining goods and services. The time that is invested reaping those particular benefits, however, is also time that could have been spent engaging in another activity. For example, the time spent studying for a good grade could have been time spent working to earn money. Economists call this interpretation of a cost an “opportunity cost:” the foregone benefits of the next best competing activity that were rejected in favour of the one that was pursued. Opportunity costs are especially evident in a foraging context: time spent searching for food reduces time available for other essential activities for survival (Stephens & Krebs, 1986).

As described by the reward-mountain model, the subject integrates the many variables that concern the reward and associated costs, computes the goal’s overall payoff, and decides how to allocate its time. However, before the subject computes the overall payoff, one of the crucial, initial stages of reward processing is the transformation of all of the objective variables into their subjective equivalents. The mathematical relationship describing the transformation of a physical stimulus into its subjective strength is termed a psychophysical function. Traditionally, psychophysical functions have described sensory experience such as the transformation of the physical intensity of light into perceived brightness. However, studies of reward seeking, in particular BSR, have extended psychophysical scaling to the realm of valuation, motivation and decision-making. For example, several studies have described the transformation of pulse frequency into subjective reward intensity (Hamilton, Stellar, & Hart, 1985; Gallistel &



Leon, 1991; Leon & Gallistel, 1992; Mark & Gallistel, 1993; Simmons & Gallistel, 1994). Similarly to the psychophysical transformation of pulse frequency into reward intensity, the subject must transform the opportunity cost variable into its subjective equivalent. Although the psychophysical function that computes reward intensity (the “reward-intensity” function) has been formally described, the psychophysical function that computes subjective costs (the “subjective price” function) has not. This is a salient lacuna. The present experiment estimates this function.

The *objective opportunity cost* refers to the actual, measurable, time that the animal could spend engaging in other activities, a value that everyone would agree on. The *subjective opportunity cost* refers to the animal’s estimate of the cost, which takes into account the personal significance of the time that could be spent engaging in other activities. For instance, in a natural setting, a forager may have to decide if it is worth spending three minutes hunting for its prey at the expense of other activities. In an operant box, the subject has to decide if it is worth spending a given amount of time depressing the lever (“working”) for a reward at the expense of time away from other available activities such as resting, grooming, or exploring.

The subjective price function is particularly important in the context of the reward-mountain paradigm. The reward mountain model has previously assumed that the subjective price function is an identity function (objective opportunity costs are equal to subjective opportunity costs along the whole tested cost range). Yet, psychophysical functions are often non-identity functions. The relationship that has been assumed hitherto (an identity function) is illustrated in Figure 1 and is referred to as the “objective

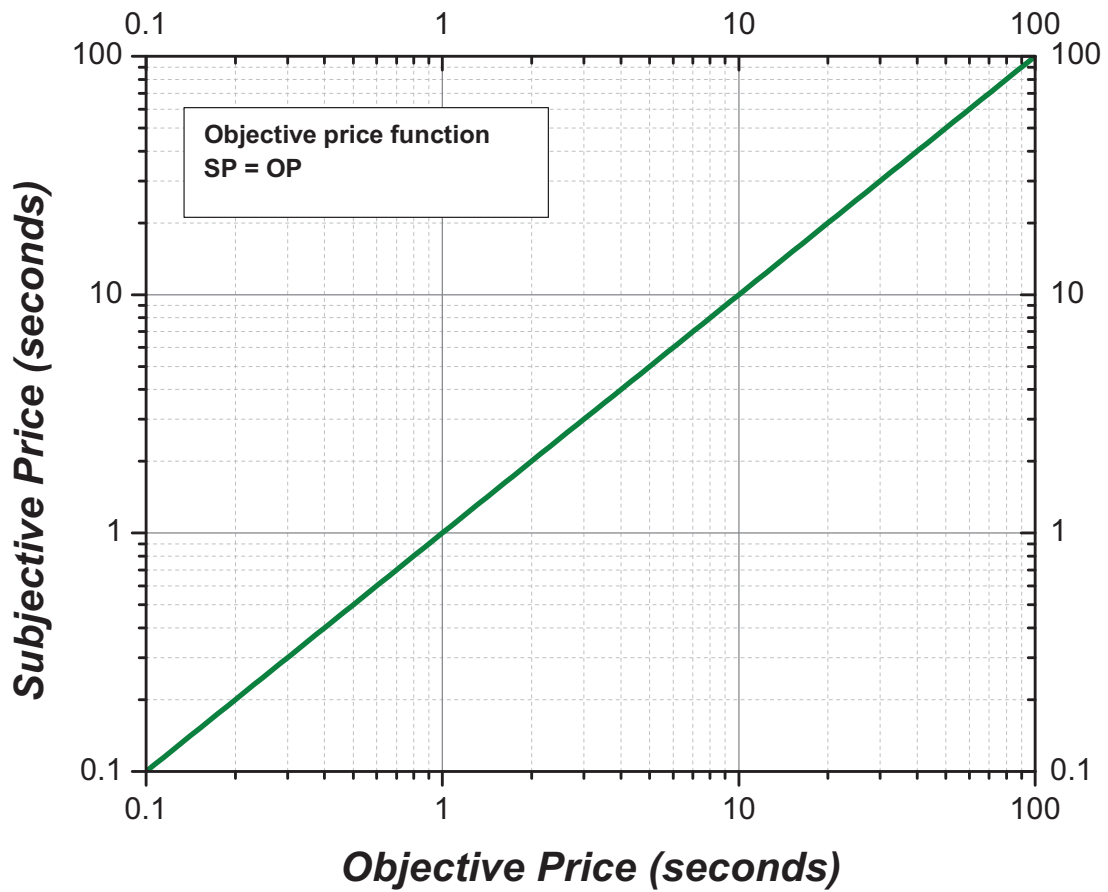


Figure 1. “The objective price function.” The relationship between subjective price and objective price is scalar along the full range of objective prices. This relationship has been assumed in previous studies using the reward mountain model.

price function.” In the present thesis, the subjective price function is hypothesized to be a non-identity function, the reasons for which will be elaborated on later (section 2.3).

## ***2.2. Schedules of reinforcement***

Before describing the proposed opportunity cost transformation, it is important to understand how objective opportunity costs are manipulated in our experiments. In traditional operant experiments, the subject’s response requirement is controlled by “ratio” or “interval” schedules of reinforcement. Shizgal and Conover (2005) argue that both of these schedules do not stringently control the opportunity cost because the subject has partial control of it. For instance, on a traditional ratio schedule, the subject must emit the experimenter-set response, such as a required number of bar presses, in order to earn a reward. However, although the work requirement is determined by the experimenter, the subject controls the time between rewards (the minimum inter-reward interval); the faster the subject works, the shorter the time required to obtain the reward. The other commonly used schedule of reinforcement, the interval schedule, does not control both types of costs. On an interval schedule, the subject is reinforced for the first response it makes after an experimenter-set amount of time. Consequently, the maximum rate at which rewards can be earned is controlled by the experimenter-set minimum inter-reward interval. The subject works during a large portion of the time interval; a characteristic scalloped response pattern is seen. However, the opportunity cost is not strictly set by this interval: although the subject works during most of the interval, it does not respond over its entirety. Specifically, at the beginning of the interval, response rates are low. During this time, the subject can be engaged in activities other than the

operant task, such as resting or exploring. Therefore, the subject has partial control of the opportunity cost. Furthermore, the subject chooses the rate at which it responds and thus has some control over the effort cost. Therefore, taken together, both the ratio and interval schedules do not strictly control opportunity costs.

To control opportunity costs so as to preclude engagement in alternate activities while the subject works for the reward, we used the fixed cumulative handling-time schedule (FCHT) in which both the work requirement and the minimum inter-reward interval are experimenter-set (Breton, Marcus, & Shizgal, 2009). The opportunity cost (also referred to as the “price”) as controlled by the FCHT schedule requires that the subject hold down the lever for a cumulative experimenter-set time interval to earn a reward. That is, if the opportunity cost were set to 10 s, the subject could pause in between bouts of work (lever holding) and would be rewarded for depressing the lever for the total cumulative time of 10 s.

Electrical brain stimulation reward used in combination with the FCHT schedule provides a simplified environment to study the contribution of opportunity costs to reward-seeking behaviour. For a foraging animal in a natural setting, opportunity costs entail multiple components: searching, procuring, handling (e.g., opening a nutshell), and consumption. By substituting electrical brain stimulation reward for a natural reward, (e.g., food and water) and employing the FCHT schedule, components of the opportunity cost are collapsed into one: the work time required to trigger the delivery of rewarding electrical stimulation. These tightly controlled experimental variables (reward intensity and opportunity costs) afford the precise manipulation required to estimate the subjective price function (method described in section 2.5).

### ***2.3. The translation of objective opportunity cost into subjective opportunity cost***

The FCHT schedule sets the objective opportunity cost; the subject translates these objective costs into subjective costs, but how? The price the subject pays is time; therefore the estimation of time intervals contributes to the estimate of the subjective price. Importantly, the functions describing the estimation of time intervals and the estimation of prices are not identical. For instance, a subject may be able to discriminate very small time intervals such as 0.1 s and 0.2 s, yet when these intervals are set as an opportunity cost, they may be interpreted as equivalent. That is, reducing the cost from 0.2 to 0.1 s does not allow the animal to perform additional “alternate activities” of any significance. For this reason, low opportunity costs are proposed to be subjectively equivalent along a range of low objective prices. Estimating the psychophysical cost function will test whether this hypothesis holds, and if so, reveal the range at which objective prices are subjectively equal to each other.

Before developing the forms of the subjective price functions to be evaluated, it is first important to recognize that embedded within the subjective price function is the *subjective time interval function*. How can the subjective time interval function be expressed? We propose that the subject estimates the subjective time interval ( $st$ ) from the experimenter-set cumulative time required to depress the lever (objective price,  $op$ ), as described by the function  $f$ . The subjective time interval function has been extensively measured by the use of different methods and schedules of reinforcement. In this experiment, the exact form of the subjective time interval function is not of great importance; the crucial characteristic of this function is that most leading theories of

interval timing studies would agree that the estimates of subjective time are scalar over a wide range of intervals (Gibbon, 1977). Thus,

$$st = f(op)$$

where,

$st$  = the subjective time estimate (s)

$op$  = the objective price (s)

$f$  = the function translating the set objective price to a subjective time estimate

Once this transformation occurs, the subjective time is then converted into subjective price by a second function, called  $g$ :

$$sp = g(st)$$

Combining both of the transformations above, we see that  $f$  is embedded within  $g$ , therefore the subjective price function can be expressed as:

$$sp = g(f(op))$$

The embedding of  $f$  in  $g$  can be referred to as the “subjective price function” and we can simplify this and name it  $h$  which incorporates the transformation of objective time intervals into subjective time and subjective time into subjective price:

$$sp = h(op)$$

What is the predicted form of the subjective price function? It seems rational that the subjective cost should mirror the objective cost; failure of the two to match would cause distortions from reality, compromising an animal's fitness. For instance, if subjective cost rose faster than objective cost, an animal would perceive costly options as being more expensive than they actually are. This particular animal would avoid valuable, but costly options, perceiving them as "too expensive," while a competitor that more accurately perceives the real cost pursued them instead. On the other hand, if subjective cost rose more slowly than objective cost, the animal would perceive costly but valuable options as being less expensive than they actually are. This particular animal may pursue many options with the erroneous perception that the options are inexpensive and forego other important options.

As described above, it is reasonable to predict that objective costs and subjective costs should match over time intervals likely to be encountered in a foraging context. In other words, the relationship is predicted to be scalar, with a slope of 1. However, this direct match may break down at low costs because as described above, the subjective opportunity costs associated with very low objective prices are predicted to be equivalent over the low objective cost range. Thus, the shape of the function is proposed to resemble a hockey stick. It should have a flat "blade" that extends over the range of initial, low costs. This "blade" extends into upward-curving portion where the animals begin to discriminate opportunity costs to a straight "handle" which extends over the range of higher costs over which the objective and subjective costs mirror each other. The

opportunity cost function indicates at what point the relationship between the objective and subjective opportunity costs are scalar and at what point this relationship breaks down. (See Figures 2, 4, 5).

## ***2.4. Plausible forms of the subjective price function***

### ***2.4.1. The objective price function***

The most parsimonious subjective price function which has been assumed in previous uses of the reward-mountain model is referred to as the “objective price function.” The scaling of objective price to subjective price is 1 over the entire range of objective prices. It is an identity function (a function that always returns the same value that was used in the argument).

$$SP = OP$$

where ,

$SP$  = subjective price (s)

$OP$  = objective price (s)

This function has a slope of 1 and passes through the origin, illustrated in Figure 1 described in section 2.1.



#### 2.4.2. Shizgal's subjective price function

Another plausible formal account of the subjective price function is “Shizgal’s subjective price function.” The form of this function is the integral of a sigmoid which gives rise to the proposed hockey stick shape, described as follows:

$$SP = SP_{\min} + (SP_{bnd} \times \ln(1 + e^{\frac{OP - SP_{\min}}{SP_{bnd}}}))$$

where,

$SP_{\min}$  = minimum subjective price (s)

$SP_{bnd}$  = controls the abruptness of the transition from the “blade” to “handle”

$OP$  = objective price (s)

$SP$  = subjective price (s)

This function is described in Figure 2. The shape of the function at different possible values of  $SP_{\min}$  and  $SP_{bnd}$  is illustrated in Figure 2B and Figure 2C.

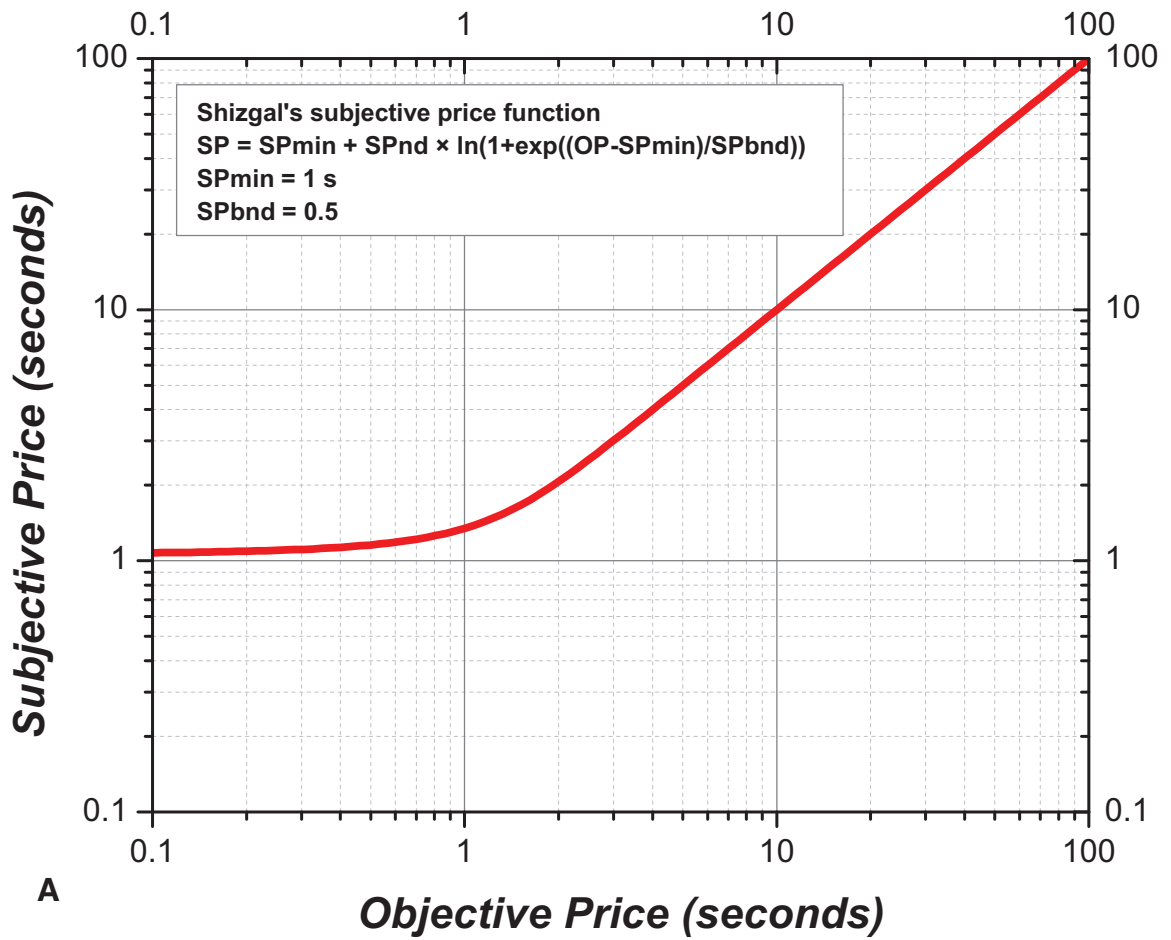


Figure 2. A. “Shizgal’s subjective price function.” The relationship between subjective price and objective price is scalar along a range of objective prices but breaks down at low objective prices.

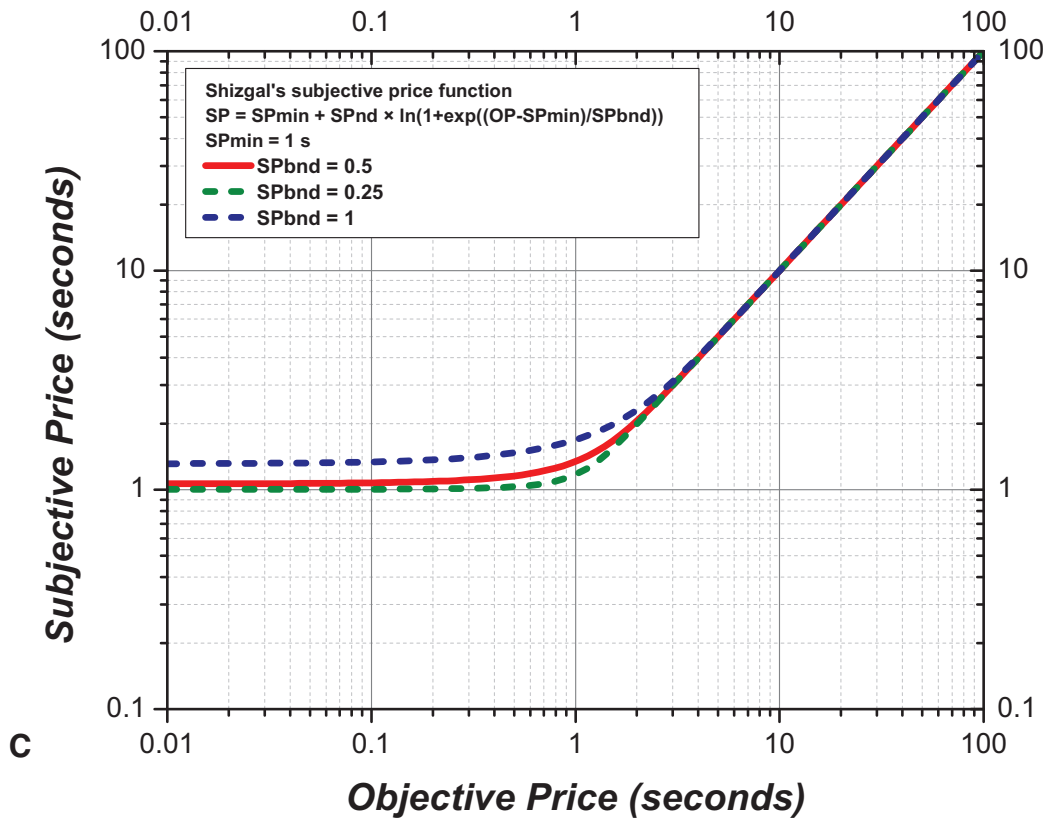
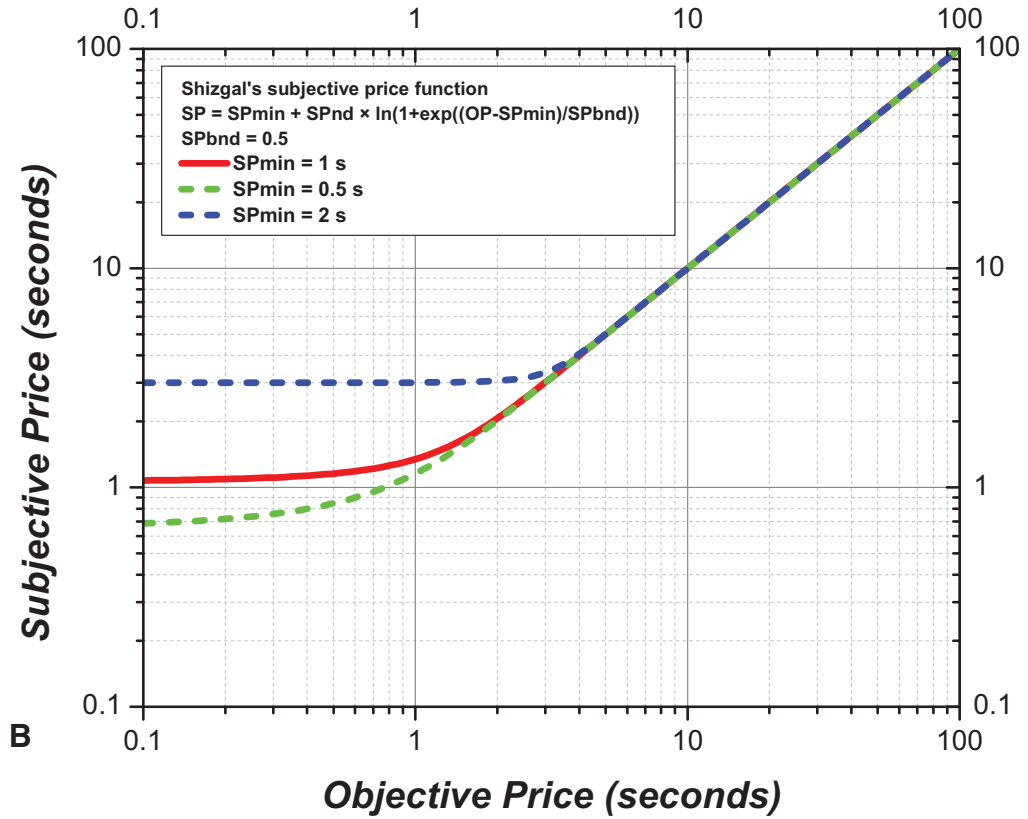


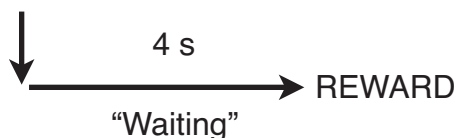
Figure 2. A. Shizgal's subjective price function at various parameter values of  $SP_{\min}$ : 0.5 s, 1 s, 2 s. Parameter  $SP_{\min}$  sets the horizontal starting value of the dependent scale of the function. B. Shizgal's subjective price function at various parameter values of  $SP_{\text{bnd}}$ : 0.25, 0.5, 1. The smaller the bend value, the steeper the bend from "handle" to "blade".

### ***2.4.3 Opportunity cost treated as a delay to the reward delivery***

Up until now, we have assumed that the subject treats the time interval required to hold down the lever as an opportunity cost, but what if it is treated as a different, but related, psychological variable instead: *delay* to the reward delivery? In traditional studies of reward delays, there is an experimenter-imposed delay until reward delivery that occurs after the subject makes only a brief response on the lever. In contrast, in studies using the reward-mountain methodology, the subject must depress the lever for a required amount of time in order to reap the reward. Yet, when setting an opportunity cost, there is an intrinsic delay to reward delivery. For instance, if the price is set to 4 s, the delay to reward delivery is 4 s as well. Could the subject be treating this experimenter-set opportunity cost as it does a simple delay, or are these two variables treated differently? In other words, two distinct processes may be responsible for the evaluation of time spent waiting for the reward and time spent working for the reward, or, the same process may govern these two evaluations. An illustration contrasting the two different paradigms, one in which a delay is imposed and one in which an opportunity cost is imposed, is presented in Figure 3.

### Delay set to 4 s

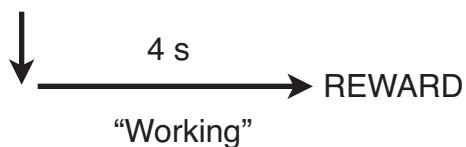
*Work requirement has been met (1 lever press)*



(the subject is in operant box but not depressing the lever because it has been retracted after work requirement has been met)

### Opportunity cost set to 4 s

*Subject begins working (depressing the lever)*



(depressing the lever for the entire 4 s interval)

Figure 3. Two different experimental paradigms in which the delay to reward or the opportunity cost is manipulated are contrasted. In a delay discounting experiment, a delay to reward is imposed after a simple work requirement (1 lever press) has been met. In this example, the imposed delay is 4 s. In an experiment manipulating opportunity costs, the subject is required to depress a lever for the duration of an experimenter-set amount of time (opportunity cost), 4 s in this example. However inherent in the opportunity cost is a delay to the reward delivery. The experiment in Chapter 2 investigates whether two different evaluation processes are responsible for time spent working and time spent waiting, or if the same process is responsible for these transformations.

### 2.4.3.1. *The (inverted) hyperbolic subjective price function*

Thus, one possible form of the subjective price function,  $h$ , is a variant of Mazur's (1987) hyperbolic delay discount function. Mazur's (1987) delay discounting function describes how a future *reward* is discounted: as the delay to the delivery of the reward is increased, the reward's present value is discounted hyperbolically:

$$Value = \frac{Scale}{1 + (k \times Delay)}$$

where,

$Value$  = discounted subjective value of the future reward

$Scale$  = proportional to reward amount; controls the vertical scale of the function

$k$  = determines how abruptly the value of the function declines as the delay grows

$Delay$  = the delay in seconds from the completion of the response requirement to the reward

If we extend Mazur's definition of value to our definition of payoff:

$Value$  = in economic terms, the *subjective* "payoff" of the reward which is the ratio of reward intensity to subjective price ( $RI/SP$ )

$Scale$  = reward amount is equivalent to the subjective reward intensity ( $RI$ ) in BSR studies

$k$  = determines how abruptly the value of the function declines as the delay grows

$Delay$  = can be considered the objective price ( $OP$ ) if the subject is treating the cost as a delay

Thus, rewriting Mazur's hyperbolic delay discounting function in terms of the parameters used in the reward mountain model:

$$Payoff = \frac{RI}{1 + (k \times OP)}$$

Thus, the payoff declines hyperbolically as the *OP* (objective price) is increased.

Expanding the definition of payoff:

$$\frac{RI}{SP} = \frac{RI}{1 + (k \times OP)}$$

According to this interpretation, the reward intensity is the factor being discounted by the subjective price. However, we can easily re-plot this as a subjective price function. If we invert the function, holding reward intensity constant, as we do in our experiments, the discounted factor (numerator) is now subjective price. In this form, the discount now increases with the independent variable (*OP*) instead of decreasing, giving rise to a hockey-stick shape. Thus, the subjective price function in accordance with Mazur's hyperbolic function is:

$$\frac{SP}{RI} = \frac{1 + (k \times OP)}{RI}$$



*RI* is canceled out. Thus,

$$SP = 1 + (k \times OP)$$

where,

*SP* = subjective price (s)

*OP* = subjective price (s)

*k* = parameter controlling the rate at which the function rises

Figure 4 illustrates this function with different values of *k* and is referred to as the “(inverted) hyperbolic subjective price function.” A notable difference between Shizgal’s subjective price function (Figure 2) and Mazur’s form (Figure 4) is the sharpness of the bend from blade to handle as well as the very slow rise over the 1 to 10 s objective price domain. Evaluating these two models would shed light on whether the price is treated psychologically as an actual delay or as an entirely different psychological variable, subjective opportunity cost.

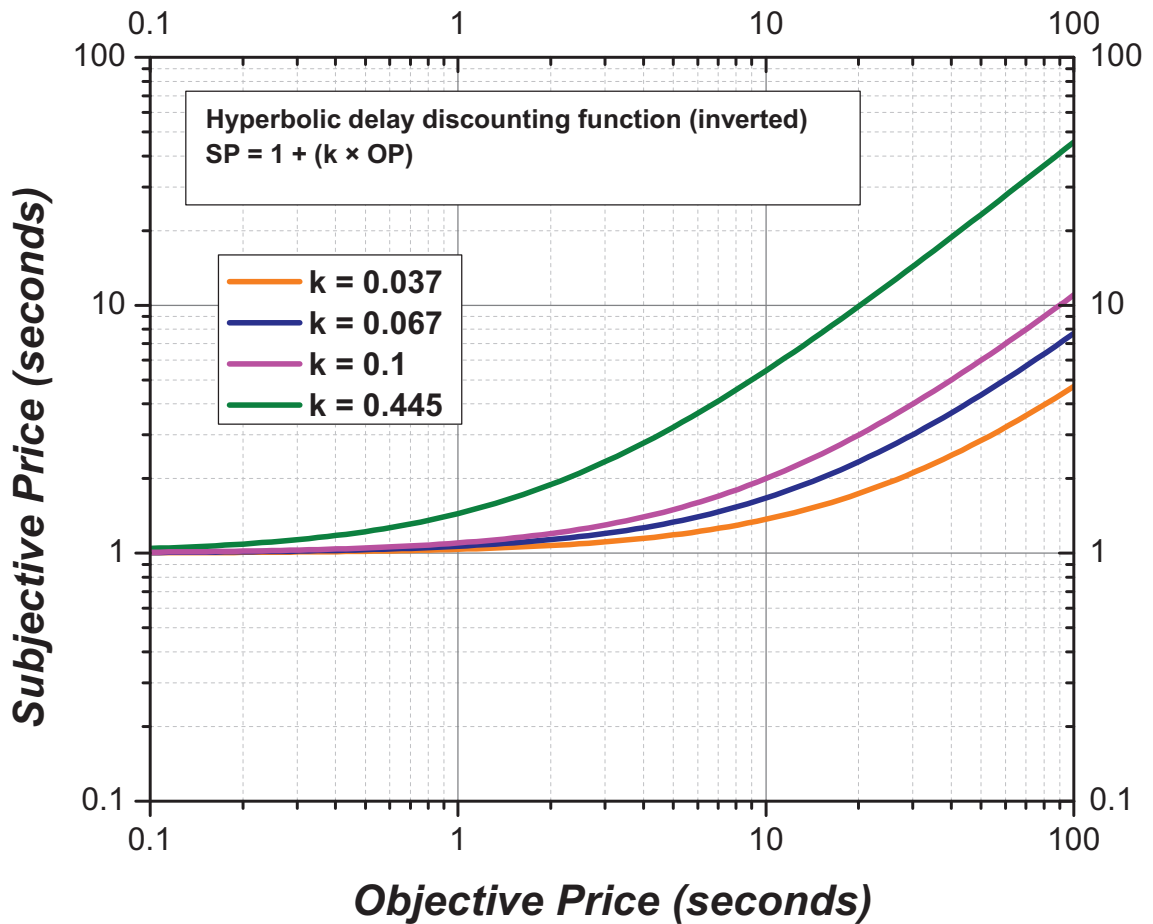


Figure 4. “The hyperbolic delay discounting function (inverted).” The form of the function is of a hockey-stick shape. As the objective price is increased, the function is flat along a range and then rises. The  $k$  values are from those derived in BSR and delay studies (Mazur, Stellar, Waraczynski, 1987; Fourzeios & Randall, 1997). The greater the  $k$ , the faster the function rises.

### ***2.4.3.2. The (inverted) exponential subjective price function***

A second delay discounting function, the exponential delay discounting function described by Samuelson (1937) is also tested. Its inverted form is expressed as:

$$SP = e^{k \times OP}$$

where,

$k$  = controls the rate at which the function grows

The inverted exponential delay discounting subjective price function is described in Figure 5 at various  $k$  values.

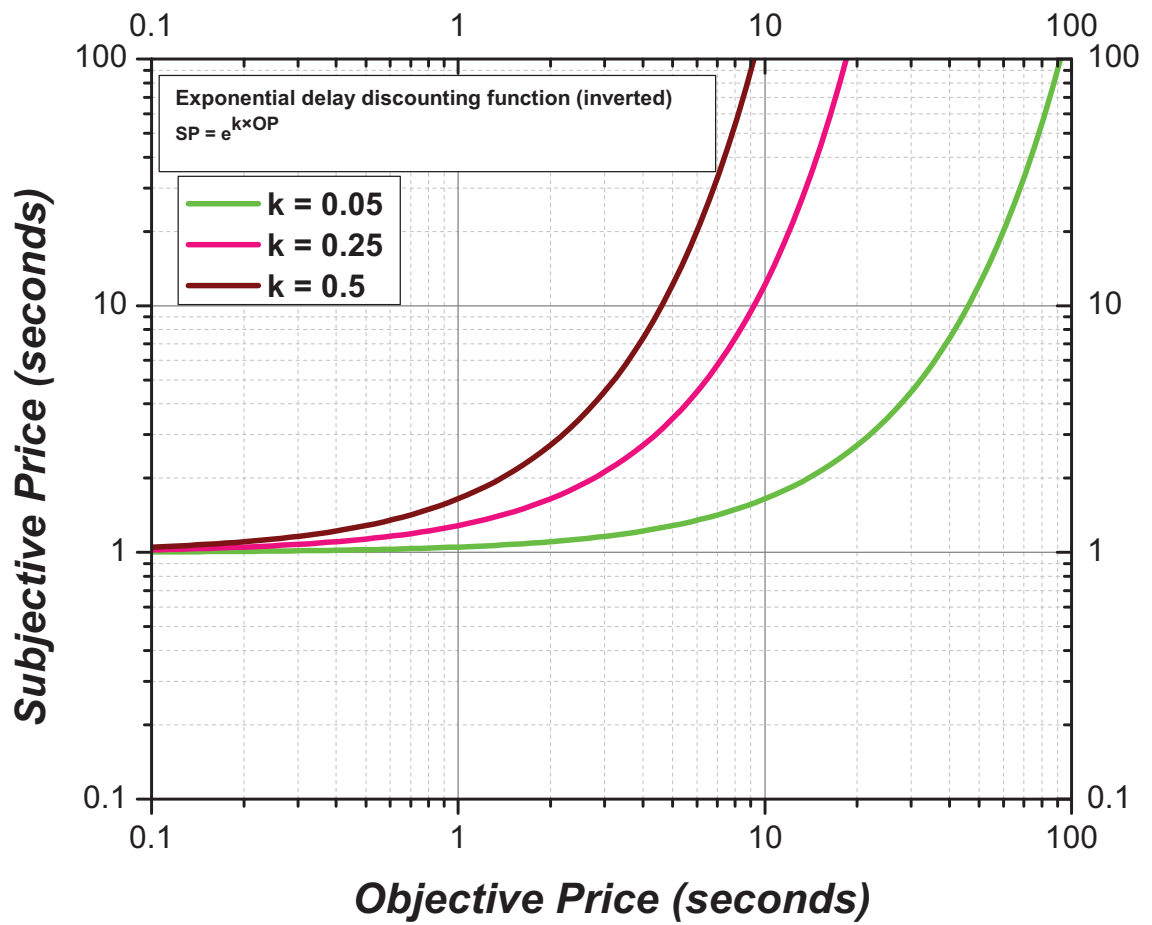


Figure 5. “The exponential delay discounting function (inverted)”. The form of the function at different values of  $k$ . The larger the  $k$  value, the more rapidly the function rises.

## 2.5. How to measure the subjective price function

Trade-off methodology is used to measure the subjective price function. This methodology assesses the effectiveness of one variable by its effect on a second variable. The two variables trade off against each other so as to hold behaviour constant. A fixed response level is chosen as the behavioural index (half-maximal responding). The various combinations of the two variables that produce the criterion level of responding reveal an underlying function, the subjective interpretation of the variables. The two trade-off variables used in the present experiment are *pulse frequency* and *objective price*. Analogous to traditional accounts of the matching law, payoff from the electrical reward is expressed as:

$$U_B(p, SP, RI_{peak}, \xi) = \frac{RI_{peak} \times p}{(1 + \xi) \times SP}$$

where,

$RI_{peak}$  = reward intensity at its peak value

$SP$  = subjective price (s)

$p$  = the subjective probability of obtaining the reward after the work requirement has been met

$\xi$  = the subjective rate of exertion required to hold down the lever

If reward delivery is certain upon meeting the response requirement and the effort entailed is minimal, the function can be approximated by:

$$U_B(SP, RI_{peak}) \approx \frac{RI_{peak}}{SP}$$

Reward intensity (set by the pulse frequency) and subjective price (set by the objective price) are related in a scalar (multiplicative) fashion. Accordingly, an increase in subjective price should be offset by an equivalent proportional increase in reward intensity along the whole range of tested variables. On a logarithmic scale, *reward intensity* as a function of *subjective price*, would be represented by a straight line with a slope of 1, similar to the shape of the function in Figure 1.

However, it is the objective variables that are directly manipulated: pulse frequency and objective price. The scalar relationship of reward intensity and subjective price may not hold in their respective objective domain (pulse frequency and objective price) if the relationship between the objective variables and their corresponding subjective variables is not scalar. Specifically, according to the proposed subjective price models, very low objective prices are subjectively equal to each other (the relationship between subjective price and objective price is not scalar along the low range of prices). Thus, any increase in the objective price along this low price range will not require any offset in reward intensity (and thus pulse frequency) to maintain the behavioural criterion. On a common logarithmic scale of *pulse frequency* as a function of *objective price*, this relationship is represented by a horizontal line at low objective prices that gradually increases, reaching a slope of 1, similar to the shape of the function in Figure 2.

The procedure outlined below describes how the trade-off function is acquired, as well as the predicted form of this trade-off function, as illustrated in Figure 6. The dependent measure is time allocated to holding down the lever. Rats are presented with a

series of “pulse frequency sweeps,” each conducted at a different price. A sweep is a series of trials over which the pulse frequency is varied from one trial to the next. Figure 6A shows the corresponding behavioural allocation curves, time allocation as a function of pulse frequency. Each curve is denoted by a different colour representing a different price. The behavioural criterion is half-maximal time allocation ( $F_{M50}$ ), denoted by the vertical dashed line for every curve. As the prices are halved within the one-to-one subjective-objective price range, the pulse frequency curves shift leftward by the same magnitude, one half. However, as the objective price approaches the price at which the scalar subjective-objective relationship breaks down, each halving of the price produces a smaller magnitude in the shift of the curves. For instance, at the lowest prices, 0.125 and 0.25 s, the curves overlap, indicating that these prices are subjectively equivalent because no pulse frequency offset was required.

Figure 6B describes the corresponding trade-off plot: required pulse frequency at the behavioural criterion (half-maximal time allocation) as a function of objective price. Along the low objective price range, the required pulse frequencies are equivalent to each other as indicated by the flat portion of the curve. As subjective and objective prices become roughly equal, the required pulse frequencies increase at the same rate that objective prices increase (indicated by the rising portion of the curve). After the points on the trade-off plot are acquired, the different forms of the subjective price functions are fit to the data points. The best fitting subjective price function is assessed.

The rationale laid out above and in Figure 6 is the general foundation as to how the subjective price function will be estimated. The trade-off functions fit to the data will

incorporate the proposed subjective price functions as well as the reward mountain parameters described in Chapter 1.



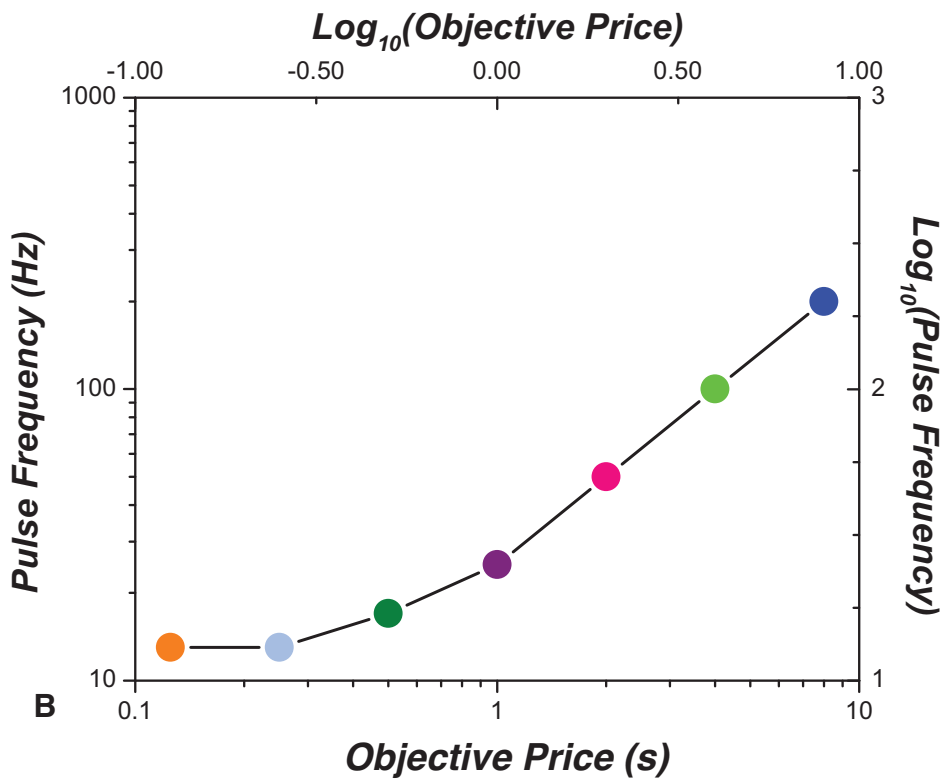
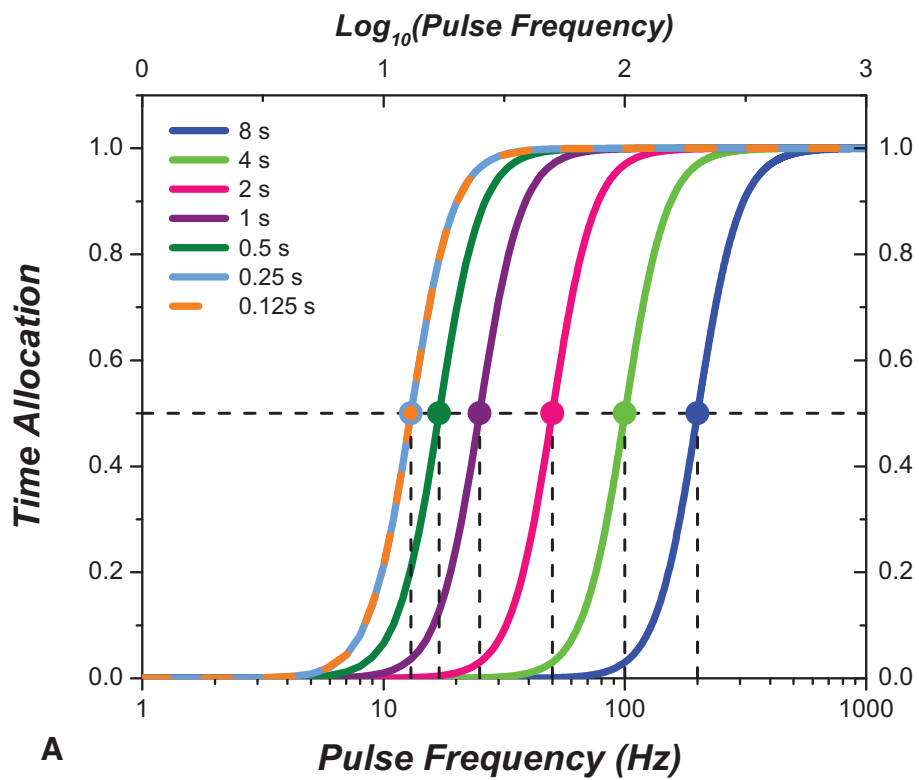


Figure 6.

A. A simulation of a family of behavioural time allocation functions relating time allocation to pulse frequency. In the experiment, rats are presented with a series of “pulse frequency sweeps” in which pulse frequency is varied throughout a session along a pulse frequency range. Each pulse frequency sweep is conducted at a different price. Each curve illustrates the behaviour obtained in a pulse frequency-sweep condition. The vertical dotted line represents the pulse frequency at half-maximal time allocation ( $F_{M50}$ ), the index of measurement. Along the veridical subjective-objective price range, as the price is reduced by one half, the curves shift leftward by that magnitude. However, as the objective-subjective price scalar relationship breaks down, each halving of the price produces a smaller magnitude of the shift of the curves.

B. The pulse frequency-objective price trade-off plot. The  $F_{M50}$  values for each of the curves are plotted. The flat portion of the curve represents the flat part of the subjective price function. At higher prices, the trade-off curve increases, representing the portion of the subjective price function in which objective prices are distinguishable from each other. In this relationship, the relationship is scalar at a higher range: as the prices are increased by a doubling, the required pulse frequency is increased by a doubling.

## ***Materials & Methods***

### ***2.6. Subjects***

The subjects were 6 male Long Evans rats (Charles River, St. Constant, Quebec, Canada) that weighed between 450 g and 600 g at the time of surgery. They were housed individually in plastic ‘shoebox’ cages and had unlimited access to food (Purina Rat Chow) and water. A reverse light cycle was in effect (lights off from 0800 to 2000).

### ***2.7. Apparatus and materials***

The test boxes had the following dimensions: 34 cm x 23 cm x 60 cm. The boxes had four Plexiglas walls with a hinged Plexiglas front door. Two retractable levers (1.5 cm x 5 cm) (ENV-112B, MED Associates, St. Albans, Vermont) were located in the center of the right and left walls, 10 cm above the wire-mesh floor; only the right lever was used in the present experiment. Cue lights were located 2 cm over the levers and were illuminated when the lever was depressed. A house light on the back wall, 35 cm from the floor, flashed between trials.

Monopolar stimulating electrodes (0.25 mm diameter) were constructed from 000 stainless steel insect pins. The insect pins were insulated with Formvar to within 0.5 mm of the tip. An insulated wire was soldered to the middle of the insect pin and terminated in a gold-plated male amphenol pin. The insect pin served as the cathode. The anode consisted of two skull screws which were connected by a wire that was crimped to a male amphenol pin. The amphenol pins were inserted into an externally threaded, nine-pin connector (Scientific Technology Centre, Carleton University, Ottawa, Ontario, Canada) which was attached to the skull with dental acrylic and anchored by 6 jeweller’s screws

embedded in the frontal and parietal bones. Dental acrylic was used to secure the electrode and connector to the screw anchors and skull.

A mating connector terminated the stimulation leads, which were attached to a slip-ring assembly to allow the rat to circle without twisting the leads. A second cable linked the slip-ring assembly to the output of the constant-current stimulator.

### ***2.8. Surgical procedure***

An injection of atropine sulfate (0.05 mg/kg sc) was given prior to surgery in order to reduce bronchial secretions. Ten minutes later, ketamine-xylazine (100mg/kg ip) was administered to induce anesthesia; to confirm that the level of anesthesia was sufficiently deep, the tail was pinched 5 min after administration. Before the rat was maintained on isoflurane for the rest of the surgery, buprenorphine (0.17 ml/kg sc) was administered as an analgesic and penicillin-g was administered to prevent infections. The anesthetized rat was mounted into a stereotaxic frame. Pilot holes were drilled for the six jeweller's screws that served as anchors for the electrode assembly. The electrode was aimed at the lateral hypothalamus: a hole was drilled in the skull over the stimulation targets, which are 2.8 mm posterior to bregma, 1.7 mm lateral to the midline and 9.0 mm ventral from the skull according to the Paxinos and Watson (2007) atlas. The stimulation electrodes were lowered into place using standard stereotaxic manipulators and secured with dental acrylic. Rats were given a 1-week recovery period after surgery to allow healing before preliminary testing began.

## ***2.9. Experimental procedure***

The description of the conditions of the experiment is described below; the sequence of the conditions is illustrated in Figure 7.

### ***2.9.1. Screening: preliminary testing***

The preliminary testing allowed the experimenter to determine whether the stimulation electrode was in the correct location. Throughout the experiment, stimulation consisted of 0.5 s trains of cathodal, constant-current pulses, square wave in shape, 0.1 ms in duration. The subject was connected to the stimulator by a cable and could move freely around the test cage. Using manually operated stimulators, the rat was initially given a low-intensity train of electrical stimulation (a low current intensity and pulse frequency) and the stimulation intensity was increased if the rat failed to approach the lever and no signs of aversion or motor-effects were observed. The rat was trained to press the lever using standard shaping techniques. Once the rat learned to press the lever, the current and pulse frequency were gradually increased to determine the parameters supporting maximal response rates for the rat. Once successfully shaped, subjects were transferred to an automated operant set-up controlled by a customized program named “PREF” developed by Steve Cabillio.

### ***2.9.2. Schedules of reinforcement***

In the automated setup, the schedule of reinforcement was manipulated in all phases of the experiment. The schedule of reinforcement sets the “price,” which refers to the cumulative amount of time the rat is required to depress the lever to harvest the

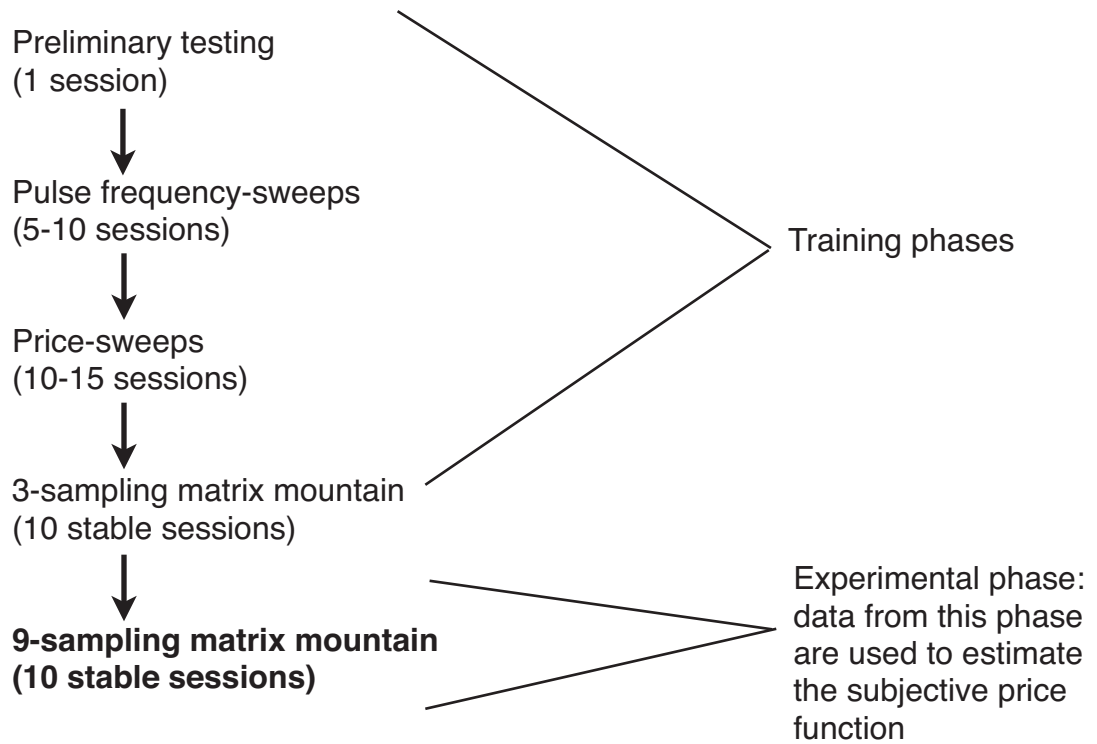


Figure 7. Sequence of conditions of the subjective price experiment. It is the data from the 9 sampling vector mountain condition that are used to assess the best fitting subjective price function.

reward. The subjects performed on a fixed cumulative handling time (FCHT) reinforcement schedule, which requires the subject to be hold down the lever for a cumulative amount of experimenter-set time in order to harvest the reward (Breton et al., 2009; Hernandez et al., 2010).

### ***2.9.3. General experimental structure***

The present experiment measured time allocated to holding down the lever as a function of the two independent variables, pulse frequency and price. The levels of these two variables can be represented in the experimental parameter space [ $\text{Log}_{10}(F)$  vs.  $\text{Log}_{10}(OP)$ ]. Thus, an individual point in the space is a function of both a pulse frequency value and an objective price value. The structure of the experiment was arranged such that the pulse frequency and objective price of a given point in the space was set on a given “trial.” A daily session was divided into several relatively brief trials; the trial time was set such that the rat could harvest a maximum of 20 rewards (e.g., 1 s price times 20 rewards = trial time of 20 s, 2 s price times 20 rewards = 40 s, etc.). The points in the parameter space that were tested depended on the experimental condition (frequency sweeps, price sweeps, mountain conditions).

The lever was extended at the beginning of the trial and was withdrawn for 2-4 s immediately after the rat is rewarded. The duration of this black-out delay was dependent on the subject. A low time allocation may be due to aversive or motor effects. Therefore, if the time allocation was 0.6 or less, the black-out delay was increased by 1 s to drive maximal performance to at least 0.7 to 0.8. (The use of black-out delays is elaborated on in the next section.)

When the lever was depressed, the cue light was illuminated to aid learning. In between trials, the house light flashed for 10 s to signal the beginning of the next trial. Before every trial, during the last second of the flashing house-light period, the rat received 1 train of priming stimulation. For the training “sweep” conditions, the pulse frequency of the priming train was set to the same value as in the train that would be triggered by the lever on that particular trial. In the actual “mountain” experimental condition, the priming stimulation was a high pulse frequency value, constant across all of trials. (Because of the PC clock system constraints, the actual computer-generated price is not exactly the experimenter-set price, but very close. The error is somewhat larger at the lower prices, but is inconsequential given that the models predict a flat subjective price range along these very low prices. Appendix A demonstrates the differences between the experimenter-set prices and computer-generated prices at 0.3 log unit intervals, from 0.125 s to 8 s. Prices will be described in this chapter by the experimenter-set price.)

#### ***2.9.4. Training: frequency sweeps***

The first phase of training in the automated setup consisted of daily sessions of “frequency sweeps.” A “sweep” in this context colloquially refers to an ascending or descending sequence of experimental parameter values. Nine points of the sample space were tested: in a frequency sweep, the price remained fixed while the pulse frequency decreased by equal logarithmic steps from trial to trial. For example, if the logarithmic step size was set to 0.1, then the pulse frequencies were: 200 Hz, 159 Hz, 126 Hz, 100 Hz, 80 Hz, 63 Hz, 50 Hz, 40 Hz, 32 Hz. Because it was the pulse frequency that changed



in this condition, this sampling matrix is termed the “*pulse-frequency sampling matrix.*” Each trial was defined by an element (point) of the pulse-frequency sampling matrix.

An individual sweep consisted of a total of 10 trials ranging from highest pulse frequency value of the pulse-frequency sampling matrix to the lowest pulse frequency value. The first trial in the set served as a warm-up trial; this pulse frequency was the same as that of the second trial. The data from the warm-up trial were discarded in the analysis. One daily session consisted of a total of 10 sweeps. The data from the first sweep were discarded in the analysis. Time allocation varies in sigmoidal fashion as a function of pulse frequency. The range of the pulse frequencies tested was set so as to drive performance from its minimal to maximal values and to include both lower and upper asymptotes (Figure 6A). The pulse frequency-sweep structure is illustrated in Figure 8A.

In the first frequency-sweep session, the price was set to 1 s. If the maximal time allocation was at least 0.8, the price was raised to 2 s on the next session. If the maximal time allocation was less than 0.8, the current and pulse frequencies were adjusted to drive performance to a time allocation of at least 0.8. The price was increased until it reached 4 s. The range of pulse frequencies was adjusted such that there were several points along each of the upper and lower asymptotes and the sloping portion of the psychometric curve relating time allocation curve to pulse frequency. Next, a frequency sweep was conducted at a price of 0.125 s to determine whether the time allocation to the lever would be significantly compromised due to the potential motor or aversive side effects that may be manifested at a high rate of reward delivery. If the time allocation of this frequency sweep was less than 0.6, then the black-out delay was increased in order to

**A Pulse frequency sweep @ 4s**  
 (trial time = 4 s × 20 = 80 s)

**B Price sweeps @ high pulse frequency (200 Hz)**  
 (trial time = price (s) × 20)

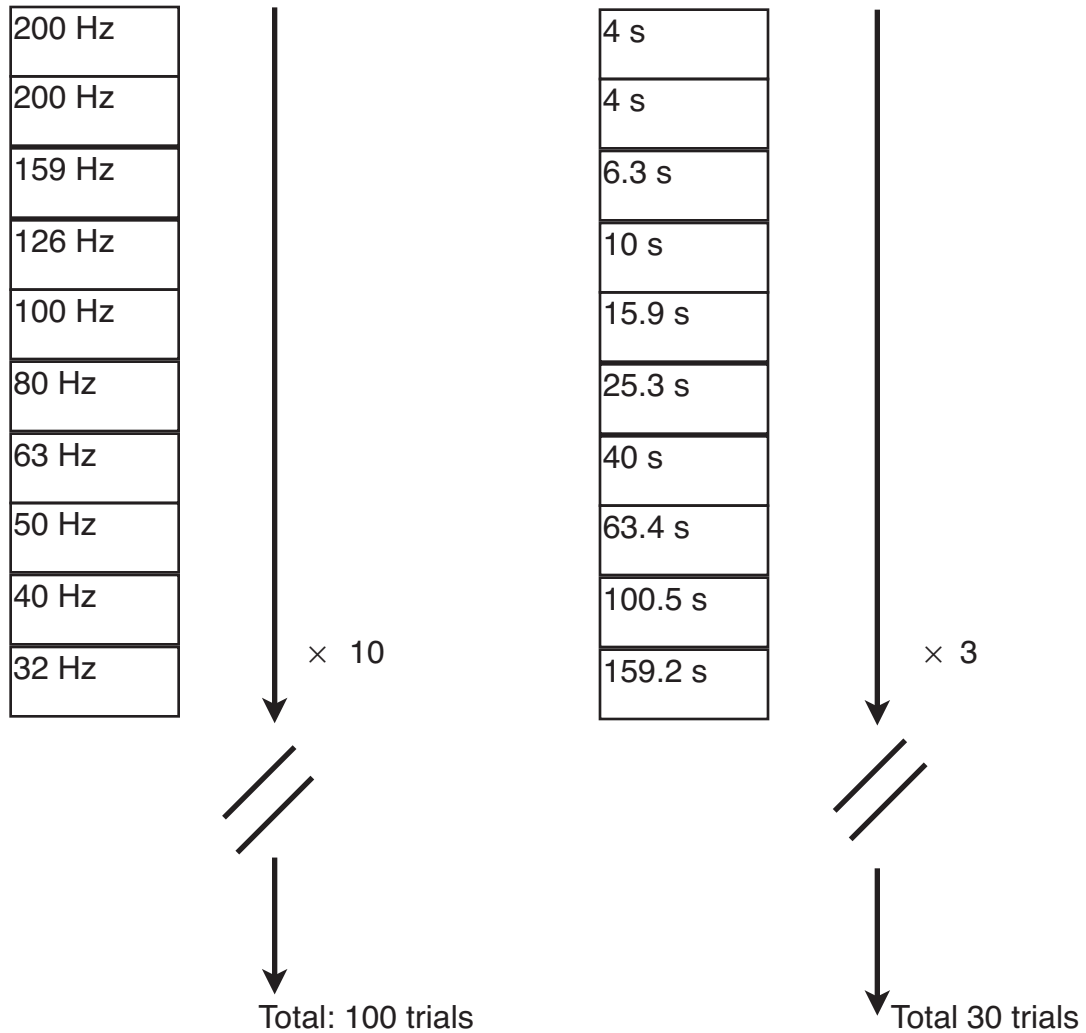


Figure 8. A. An example of a pulse frequency-sweep training session. Each box represents a trial. The pulse frequency is reduced across trials while the price stays constant. Ten trials are repeated 10 times throughout the session. B. An example of a price-sweep training session. The price is increased across 10 trials while the pulse frequency is set to a constant high value. These 10 trials are repeated 3 times.

drive maximal time allocation back to the level achieved at the higher prices. If there was a need for adjustment, the black-out delay determined in this condition was used for the rest of the experiment. This training phase was approximately 5-10 daily sessions.

### ***2.9.5. Training: price sweeps***

The second training phase entailed the price sweeps: the pulse frequency remained fixed while the prices were spaced by equal logarithmic steps. For example, at a step size of 0.2 log units, the tested prices were: 4 s, 6.3 s, 10 s, 15.9 s, 25.3 s, 40 s, 63.4 s, 100.5, 159.2 s. Because it was the price that changes in this condition, this sampling vector is termed the “*price-sampling matrix*.” The pulse frequency remained constant, set to the maximum value used in the previous condition. Each trial in the condition was defined by an element of the price-sampling matrix.

An individual sweep consisted of a total of 10 trials. The first trial in the set served as a warm-up: the first trial of the sweep was a warm-up, set to the same price as the one in effect on the second trial. The range and starting price were adjusted such that there were several points on the upper and lower asymptote and on the sloping portion of the time allocation versus price plot. The trial times were set as above (the price  $\times$  20 rewards). Because the rat required long trial time duration at high prices, there were only 2-3 sweeps conducted per daily session. In total, the rats trained in this condition for approximately 10-15 daily sessions. The price-sweep structure is illustrated in Figure 8B.

## **2.9.6. Mountain conditions**

### **2.9.6.1. Radial-sampling matrix**

The third sampling matrix is the “*radial-sampling matrix*” which is incorporated into the mountain experimental paradigm described below. In the radial-sampling matrix, both price and pulse frequency are varied simultaneously by equal logarithmic steps. Specifically, the price is increased while the pulse frequency is decreased such that the trajectory of the radial-sampling matrix in the parameter space [ $\text{Log}_{10}(F)$  vs.  $\text{Log}_{10}(OP)$ ] passes through or near the point defined by the fitted position parameters, [ $\text{Log}_{10}(OP_e)$ ,  $\text{Log}_{10}(F_{hm})$ ]. (This trajectory ensures there is sufficient data around these important parameter values to provide the most accurate fit possible of the mountain model.) The range of the experimental parameters (pulse frequency and price) for this matrix was established using data from the frequency sweep and price sweep, which were entered in a simulator developed by Yannick Breton in Matlab (The Mathworks, Natick, MA). The green segment in Figure 9B represents the radial-sampling matrix for one rat.

### **2.9.6.2. The structure of the mountain testing paradigm**

In the mountain condition, points were sampled at random and without replacement from the sampling matrices. The “standard” mountain paradigm (used in previous experiments, Hernandez et al., 2010) employs 3 sampling matrices: the pulse-frequency sampling matrix @ a 4 s price, the price-sampling matrix @ a high pulse frequency, and the radial-sampling vector. As described above, a trial is defined by a point from a sampling matrix. The trial time was set such that the rat could harvest a

maximum of 25 rewards (i.e., 4 s price times 25 rewards = trial time of 100 s). The range of pulse frequency and price values of the sampling matrices were chosen based on the behaviour that the rat displayed in the pulse frequency and price-sweep training conditions such that the chosen values capture the sigmoidal nature of the behaviour (time allocation versus pulse frequency or time allocation versus price). A *survey* is defined as a test of all of the points of the 3 sampling matrices.

Each trial was presented in a “triad” fashion, in between two “bracketing trials.” The first bracketing trial (the leading trial) consisted of the maximum pulse frequency the rat would encounter in the experiment, at a price of 1 s. The second bracketing trial (the trailing trial) consisted of a very low pulse frequency at a price of 1 s. Due to the possibility that the randomized nature of the sampling could confuse the subject, the bracketing trials were employed to give the subject a frame with which to “compare” the pulse frequency (of the test trial) to two extremes. The priming stimulation did not change from trial to trial, but was always set to the pulse frequency employed on the first bracketing trial. (See Figure 11 for an illustration of how the trials were structured).

### ***2.9.6.3. Training: 3-sampling matrix mountain condition***

The rats were trained in the experimental mountain paradigm described above. After 5 sessions in this mountain condition, the reward mountain model was fit to the data. If the radial-sampling matrix did not pass through, or close to the crosshair defined by the two location parameters [ $\text{Log}_{10}(OP_e)$ ,  $\text{Log}_{10}(F_{hm})$ ], then the vector was adjusted using the mountain simulator. After 5 sessions, the data were analyzed again, and if the radial sampling vector passed through the location parameters, 5 more sessions were

conducted.

In total, for this condition condition, there are 126 trials: 42 test trials defined by the points of the sampling vectors and 84 bracketing trials. Each daily session was approximately 6 hours in duration.

These 3 sampling matrices were needed to estimate the parameters of the 3-dimensional mountain model: time allocation to depressing the lever as a function of stimulation pulse frequency and price. However, note that this is a training condition for the rats and was used to determine which experimental parameters should be used in the next “*9-sampling matrix mountain testing paradigm.*” The structure of this condition is summarized in Figure 9A. Figure 9B illustrates an example of the sampling matrices used in this condition for one rat. (The statistical analysis for the data from this condition is described in Chapter 3, Appendix D.)

## **The structure of the 3-sampling matrix mountain condition**

There are 3 sampling matrices in this training condition: 1 pulse frequency sampling matrix, 1 price sampling matrix, and 1 radial sampling matrix.

Each sampling matrix has 14\* elements (points). An element is a unique pulse frequency-price combination.

A trial is defined by a point of a sampling matrix.

The names of the sampling vectors are in bold:

### **pulse-frequency sampling matrix @ 4 s**

the pulse frequency varies across 14 points by common logarithmic steps while the price remains constant

### **price-sampling matrix @ high pulse frequency**

The price varies across 14 points while the pulse frequency remains constant, set to a high value

### **radial-sampling matrix**

The price is increased while the pulse frequency is decreased across 14 points

Figure 9A. Experimental structure of the 3-sampling matrix mountain condition. (\* For two rats, F3 and F9 there were 9 points in a sampling matrix.) This is a training condition. The data obtained in this condition is used to guide the parameter values to be set in the next condition, the 9 sampling matrix mountain condition.

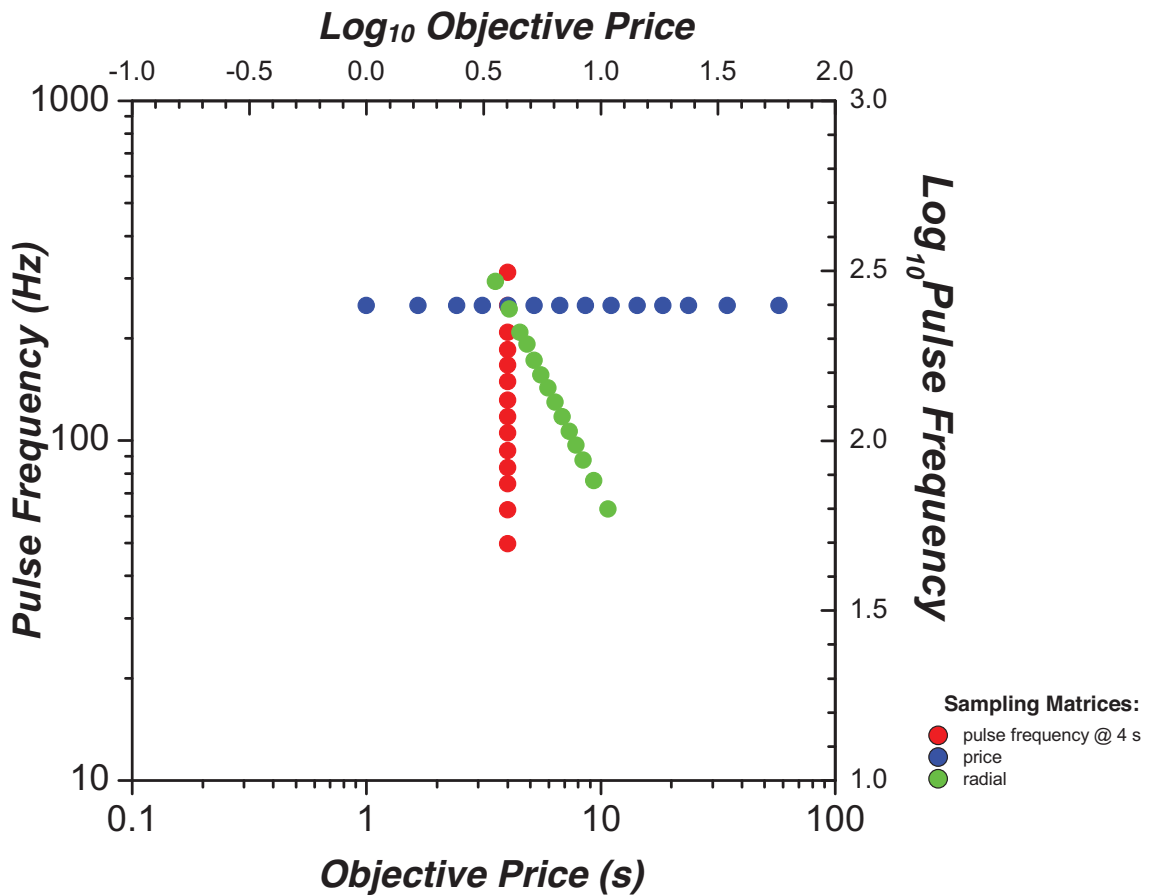


Figure 9B. An example of the sampling matrices used in a 3-sampling matrix mountain paradigm. Each colour represents a distinct sampling matrix. Each point is a unique combination of pulse frequency and price. A trial in the session is defined by a point of a sampling matrix. A point of the sampling matrix is sampled once, without replacement within a session.



#### ***2.9.6.4. Experimental subjects***

Note that the number of elements (points) of the sampling matrices and the distribution of the experimental trials in the experiment differed slightly across two groups of rats. In group 1 (rats F3, F9) each sampling matrix consisted of 9 points. The 3rd to the 7<sup>th</sup> points in all three sampling spaces were tested twice per survey. These points were tested twice to increase the amount of sampling data along the steep portions of the psychometric curves, which are the portions that most influence the values of the position and slope parameters of the mountain model. These additional data shrink the confidence intervals surrounding the estimates of these parameters. Nine points were tested per sampling matrix; there were a total of 14 trials for each sampling matrix.

For the second group of rats (rats F12, F16, F17, F18), each sampling matrix consisted of 14 points, instead of the 9 points in group 1 (rats F3, F9). The intervals separating the 10 central points were half as large as those separating the points on the extremes. Thus, the steep portions of the psychometric curves were sampled more densely than the less critical, flatter regions at either end.

#### ***2.9.6.5. 9-sampling matrix mountain condition***

In this condition, 9 sampling matrices were employed. In addition to the 3 sampling matrices in the previous condition (pulse-frequency sampling matrix @ 4s, price-sampling matrix, radial-sampling matrix), 6 more pulse frequency sampling matrices were added, at prices of 0.125 s, 0.25 s, 0.5 s, 1 s, 2 s, and 8 s. (The pulse frequencies within the matrix were chosen based on the behaviour displayed in the 3-sampling matrix mountain condition.) As above, the sequence of test trials was

randomized, and test trials were presented in between 2 bracketing trials. The psychometric curves were adjusted such that there were several points along both asymptotes as well as along the slope. In total, there were 378 trials, 126 experimental trials, and 252 bracketing trials. Each daily session lasted 7 to 8 hours.

Once all of the pulse frequency curves were adjusted and the radial-sampling matrix passed through the location parameters, as described above, 10 sessions were collected. *It is the data obtained from this condition that are used for the subjective price model analyses.* The structure of this condition is summarized in Figure 10A and example of the sampling matrices used is described in Figure 10B. The structure of the trials within a session is summarized in Figure 11.

## **The structure of the 9-sampling matrix mountain condition**

There are 9 sampling matrices in the experimental condition: 7 pulse frequency sampling matrices, 1 price sampling matrix, and 1 radial sampling matrix.

Each sampling matrix has 14\* elements (points). An element is a unique pulse frequency-price combination.

A trial is defined by a point of a sampling matrix.

The names of the sampling vectors are in bold:

**pulse-frequency sampling matrix @ 0.125 s**

**pulse-frequency sampling matrix @ 0.25 s**

**pulse-frequency sampling matrix @ 0.5 s**

**pulse-frequency sampling matrix @ 1s**

**pulse-frequency sampling matrix @ 2 s**

**pulse-frequency sampling matrix @ 4 s**

**pulse-frequency sampling matrix @ 8 s**

Pulse frequency matrices: the pulse frequency varies across 14 points by common logarithmic steps while the price remains the same

**price-sampling matrix @ high pulse frequency**

The price varies across 14 points while the pulse frequency remains constant, set to a high value

**radial-sampling matrix**

The price is increased while the pulse frequency is decreased across 14 points

Figure 10A. Experimental structure of the 9 sampling matrix mountain condition. \*two rats had 9 points in the sampling matrix instead of 14. Data obtained from this condition is used to estimate the subjective price function.

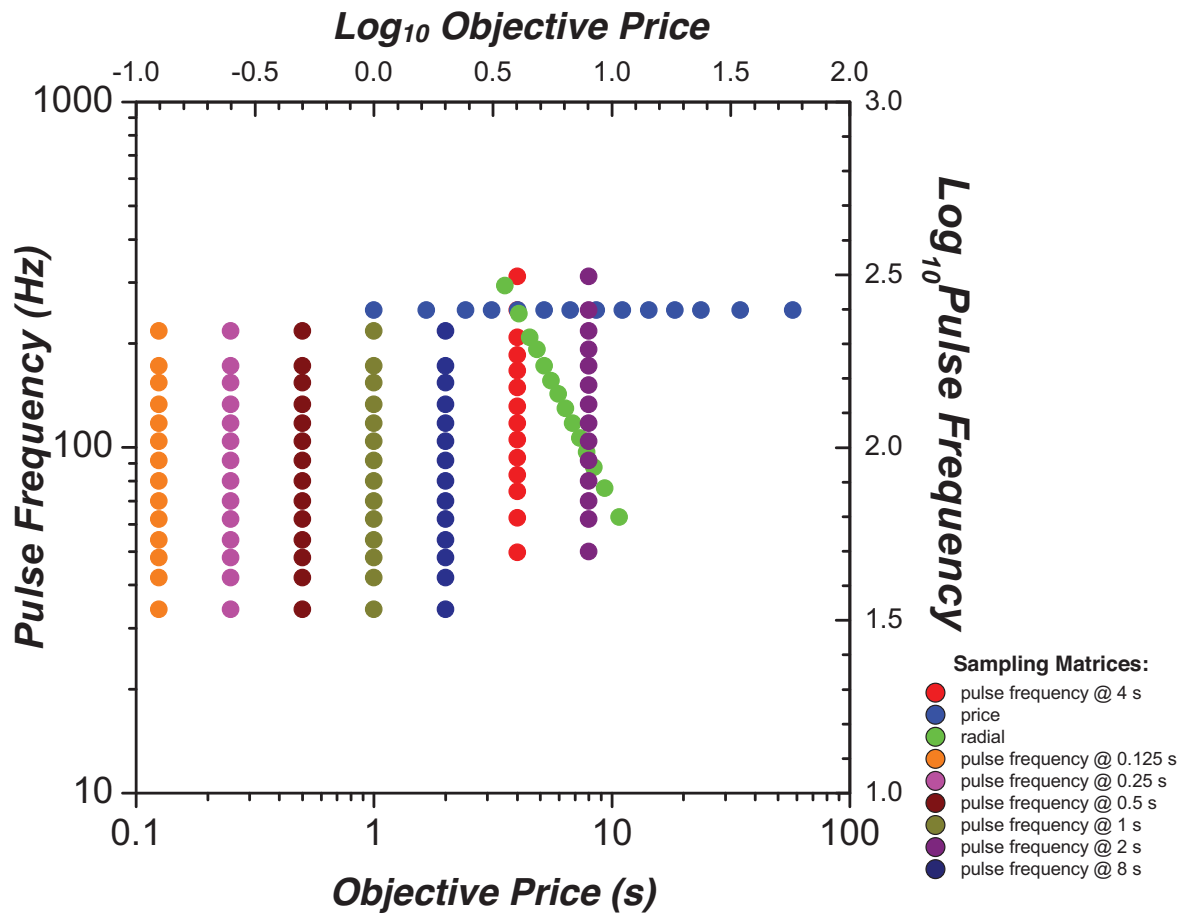


Figure 10B. An example of the sampling matrices used in a 9-sampling matrix mountain paradigm. Each colour represents a distinct sampling matrix. Each point is a unique combination of pulse frequency and price. A trial in the session is defined by a point of a sampling matrix. A point of the sampling matrix is sampled once without replacement within a session.

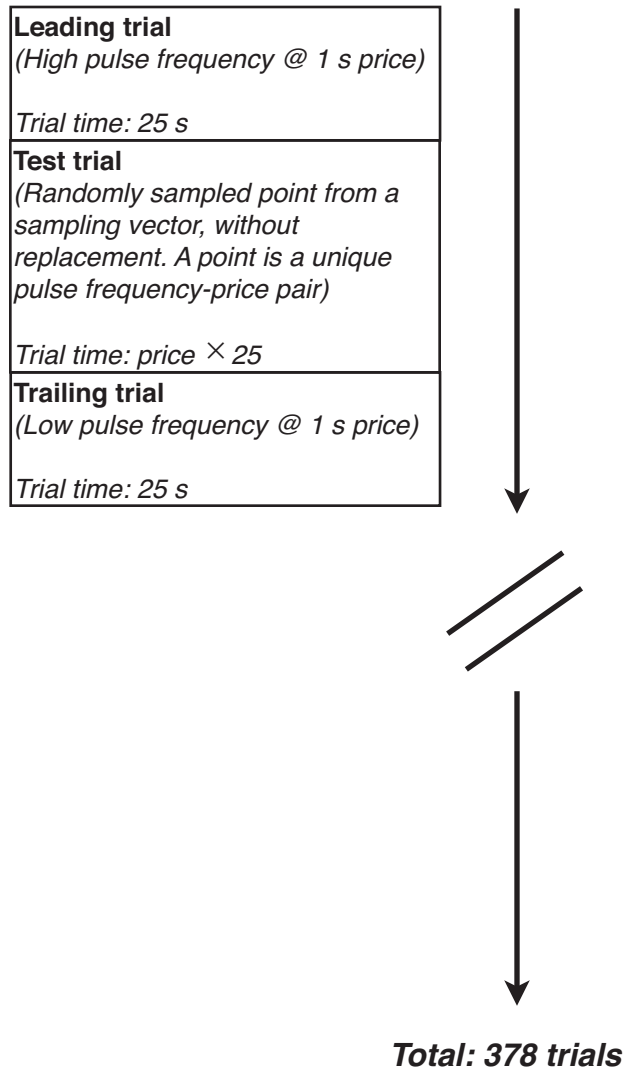


Figure 11. Sequence of trials in the 9-sampling matrix mountain condition. The rectangles represent a trial in the experiment. The trials within the condition are presented in a “triad” form. The test trial is always between two bracketing trials: the leading, (high reward) trial, and the trailing (low reward trial). The test trial is defined by a point from one of the 9 sampling matrices and is sampled without replacement. In total the experimental session has 378 trials.

## *Statistics and Results*

### *2.10. Raw data*

In the operant box, the rat is faced with a choice. It can choose to “work” for the reward, that is, hold down the lever for the cumulative experimenter-set length of time (referred to as the “price”). Alternatively, it can choose to engage in activities other than working, such as grooming, resting, and exploring. These are referred to as “everything else” or “alternate activities.”

The raw data obtained were the distribution of “holds” (intervals during which the lever was depressed by the rat) and “releases” (intervals during which the lever was extended but not depressed by the rat). The times during the trial that the rat held and released the lever and the duration of these holds and releases were recorded. The goal was to sum the cumulative amount of time the subject held the lever (total work time) during a given trial. This total hold (work) time was then converted into a proportion by dividing the total hold time by the total trial length. This proportion of time allocated to holding the lever is the dependent measure, referred to as the time allocation.

Before the hold time was transformed into the time allocation measure, two adjustments were made to the raw data. The first adjustment accounts for a rat’s occasional tendency to interrupt a depression of the lever with a very brief, apparently involuntary, release, which we call a “tap” (Breton, Marcus & Shizgal, 2009). Specifically, for the present experiment, short releases were classified as brief taps and were not considered to be representative of alternate activities in the operant box because these brief intervals are too short for the rat to be engaged in alternate activities. During these very brief releases, the rat is usually standing with its paw over the lever and is not

engaged in other activities. Therefore, a correction was implemented to treat releases of less than 1 s as work; these short release intervals were included in the measure of time allocation.

A second correction was made to the raw data. At the start of any given randomized trial, the rat does not know the parameters of the reward (reward intensity and cost). Thus, at the beginning of a new trial, the rat shows persistence in holding the lever, enabling it to discover the reward strength and cost as quickly as possible. Previous data (Breton, 2013b) has shown that the rat needs only one encounter with the reward to learn the values of both reward parameters on a given trial. In the present study, we are interested in the rat's behaviour once the reward parameters have been learned. A subject's behaviour before the first encounter (i.e. before it has learned the reward intensity and cost) is not representative of how it allocates its time to obtaining a given reward with known parameters. Thus, we discarded from the analysis the data obtained before the first reward encounter.

After the two adjustments were made to the raw data, the “adjusted” amount of time the rat holds the lever as a proportion of total trial time, was calculated. This “adjusted” time allocation will be referred to simply as “time allocation” in the analyses described below.

### ***2.11. Psychometric functions: determining the $F_{M50}$ and $OP_{M50}$ values***

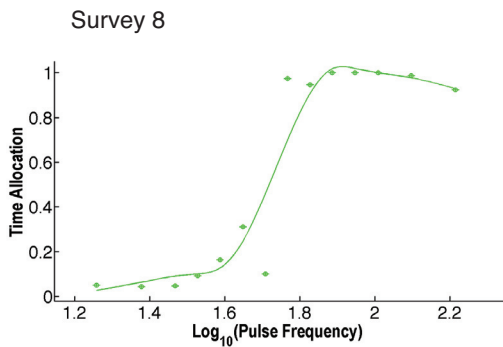
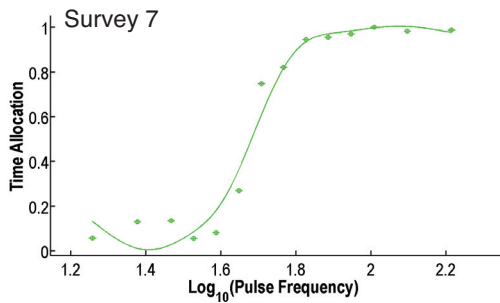
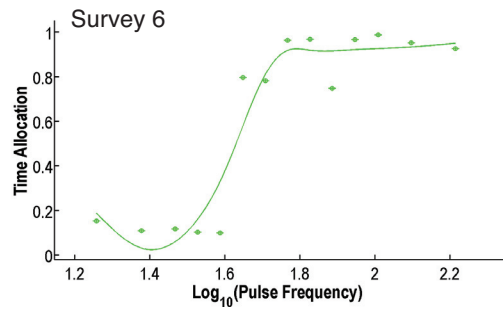
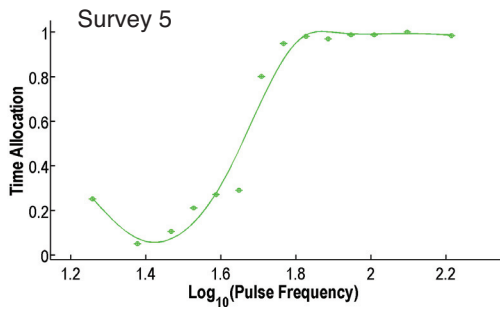
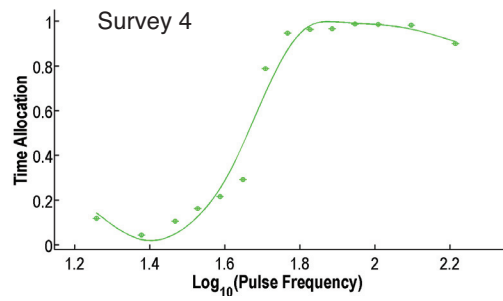
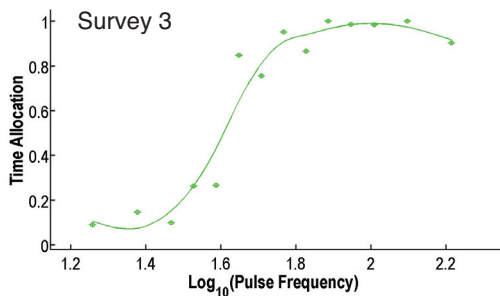
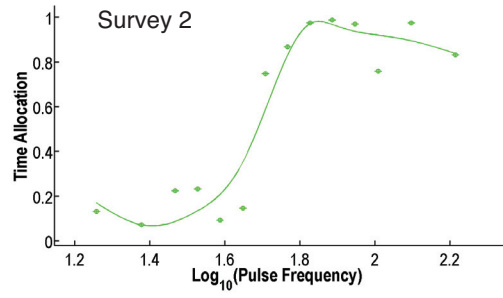
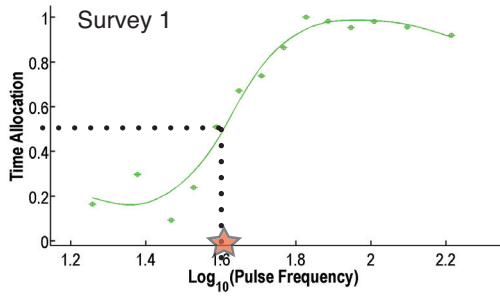
In total, 9 sampling matrices were tested: 7 pulse-frequency sampling matrices, 1 price-sampling matrix, 1 radial-sampling matrix. Overall, 10 surveys (10 daily tests of the 9 sampling matrices) were collected for each rat. Thus, after completion of the

experiment, for one subject, there were altogether 90, 2-dimensional curves relating time allocation to pulse frequency or to price that can be plotted. (However, for rat F3 a total of 11 surveys were collected, rat F16 a total of 8 surveys, and rat F17 a total of 9 surveys.). Spline functions (polynomials joined together by fixed points or knots) were fit to each 2-dimensional psychometric curve (using the Matlab spline function routine). For each of the 90 individual psychometric curves (9 sampling matrices  $\times$  10 tests), the pulse frequency at half-maximal responding, denoted by  $F_{M50}$  (for the psychometric curves corresponding to the *pulse-frequency* sampling matrices and *radial-sampling* matrices), or objective price at half-maximal responding, denoted by  $OP_{M50}$  (for the psychometric curves corresponding to the *price* and *radial-sampling* matrices), were determined.

Figure 12 illustrates an example of a set of psychometric curves with fit spline functions, obtained after 10 surveys, for the tests of the pulse-frequency sampling matrix @ 4 s for rat F18. The first graph of Figure 12 (for Survey 1) demonstrates how the  $F_{M50}$  was determined: the half-maximal time allocation is calculated (represented by the dotted horizontal line), and the pulse frequency corresponding to the half-maximal value was measured, represented by the dotted vertical line and red star on the x-axis. The  $F_{M50}$  values were derived in this manner for all of the psychometric curves.

An example of the distribution of half-maximal values collected after completion of the experiment corresponding to the tests of all sampling vectors are illustrated in Table 1 for rat F18. There were 10 tests of each of the 9 sampling matrices, therefore 10 corresponding psychometric curves (as in Figure 12). This scheme implies a distribution of 10  $F_{M50}$  or  $OP_{M50}$  values. Note that for the test of the price-sampling matrix, the  $OP_{M50}$





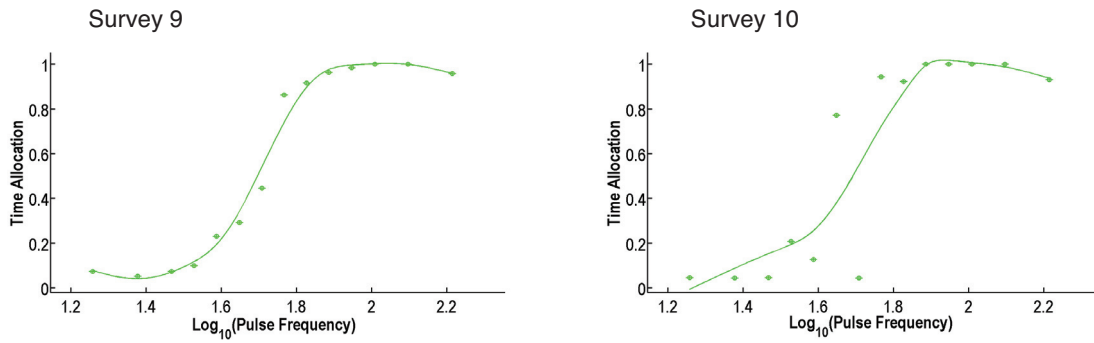


Figure 12. A set of psychometric curves with fit spline functions, obtained after 10 surveys (thus 10 days), for the tests of the pulse frequency sampling matrix @ 4 s for rat F18. The first graph (for Survey 1) demonstrates how the  $F_{M50}$  is determined: the half-maximal time allocation is calculated (represented by the dotted horizontal line), and the pulse frequency corresponding to the half-maximal value is measured, represented by the dotted vertical line and red star on the x-axis. The  $F_{M50}$  values are derived in this manner for all of the psychometric curves (vertical and horizontal lines not shown for the rest of the curves).

**Pulse frequency sampling matrix at 0.125 s (-0.903 log units)**

Survey #	LogF <sub>M50</sub>	LogOP
Survey 1	1.528	-0.903
Survey 2	1.577	-0.903
Survey 3	1.591	-0.903
Survey 4	1.584	-0.903
Survey 5	1.551	-0.903
Survey 6	1.568	-0.903
Survey 7	1.556	-0.903
Survey 8	1.457	-0.903
Survey 9	1.386	-0.903
Survey 10	1.597	-0.903

**Pulse frequency sampling matrix at 0.25 s (-0.602 log units)**

Survey #	LogF <sub>M50</sub>	LogOP
Survey 1	1.573	-0.602
Survey 2	1.576	-0.602
Survey 3	1.505	-0.602
Survey 4	1.457	-0.602
Survey 5	1.443	-0.602
Survey 6	1.487	-0.602
Survey 7	1.498	-0.602
Survey 8	1.464	-0.602
Survey 9	1.541	-0.602
Survey 10	1.578	-0.602

**Pulse frequency sampling matrix at 0.5 s (-0.301 log units)**

Survey #	LogF <sub>M50</sub>	LogOP
Survey 1	1.565	-0.301
Survey 2	1.479	-0.301
Survey 3	1.503	-0.301
Survey 4	1.404	-0.301
Survey 5	1.440	-0.301
Survey 6	1.460	-0.301
Survey 7	1.498	-0.301
Survey 8	1.453	-0.301
Survey 9	1.389	-0.301
Survey 10	1.410	-0.301

**Pulse frequency sampling matrix at 1 s (0 log units)**

Survey #	LogF <sub>M50</sub>	LogOP
Survey 1	1.497	0
Survey 2	1.737	0
Survey 3	1.525	0
Survey 4	1.509	0
Survey 5	1.505	0
Survey 6	1.440	0
Survey 7	1.459	0
Survey 8	1.470	0
Survey 9	1.402	0
Survey 10	1.540	0

**Pulse frequency sampling matrix at 2 s (0.301 log units)**

Survey #	LogF <sub>M50</sub>	LogOP
Survey 1	1.542	0.301
Survey 2	1.544	0.301
Survey 3	1.547	0.301
Survey 4	1.546	0.301
Survey 5	1.563	0.301
Survey 6	1.569	0.301
Survey 7	1.571	0.301
Survey 8	1.599	0.301
Survey 9	1.597	0.301
Survey 10	1.598	0.301

**Pulse frequency sampling matrix at 4 s (0.602 log units)**

Survey #	LogF <sub>M50</sub>	LogOP
Survey 1	1.600	0.602
Survey 2	1.664	0.602
Survey 3	1.637	0.602
Survey 4	1.641	0.602
Survey 5	1.663	0.602
Survey 6	1.599	0.602
Survey 7	1.684	0.602
Survey 8	1.746	0.602
Survey 9	1.699	0.602
Survey 10	1.646	0.602

*The spline functions corresponding to this data set are illustrated in Figure 12*

**Pulse frequency sampling matrix at 8 s (0.903 log units)**

Survey #	LogF <sub>M50</sub>	LogOP
Survey 1	1.830	0.903
Survey 2	1.779	0.903
Survey 3	1.758	0.903
Survey 4	1.836	0.903
Survey 5	1.706	0.903
Survey 6	1.771	0.903
Survey 7	1.854	0.903
Survey 8	1.786	0.903
Survey 9	1.906	0.903
Survey 10	1.764	0.903

**Price sampling matrix at 164 Hz (2.215 log units)**

Survey #	LogF	LogOP <sub>M50</sub>
Survey 1	2.215	0.986
Survey 2	2.215	0.976
Survey 3	2.215	1.059
Survey 4	2.215	1.012
Survey 5	2.215	1.108
Survey 6	2.215	1.127
Survey 7	2.215	1.124
Survey 8	2.215	1.236
Survey 9	2.215	1.077
Survey 10	2.215	1.079

**Radial sampling matrix**

Survey #	LogF <sub>M50</sub>	LogOP <sub>M50</sub>
Survey 1	1.860	0.983
Survey 2	1.830	1.015
Survey 3	1.886	0.955
Survey 4	1.886	0.955
Survey 5	1.835	1.010
Survey 6	1.888	0.952
Survey 7	1.889	0.952
Survey 8	1.909	0.930
Survey 9	1.859	0.984
Survey 10	1.883	0.958

Table 1. The distribution of  $F_{M50}$  values and  $OP_{M50}$  values for Rat 18, across the complete experiment, 10 surveys (sessions). Values are expressed in logarithmic units. For the tests of the pulse frequency sampling matrix,  $F_{M50}$  is derived as illustrated in Figure 12. For the tests of the price-sampling matrix,  $OP_{M50}$  is derived. For the tests of the radial sampling vector, both  $F_{M50}$  and  $OP_{M50}$  are derived. These values are plotted in  $[Log_{10}(F) \text{ vs. } Log_{10}(OP)]$  space. For example, for the pulse frequency sampling matrix @ 0.125 s, the first point plotted in space is for survey 1: coordinates (-0.903, 1.528), the second point, survey 2: (-0.903, 1.5770), etc. All of the values in the table are plotted in this manner in the  $[Log_{10}(F) \text{ vs. } Log_{10}(OP)]$  space. The pulse frequency-objective price trade-off functions can then be fit to the points in space.

was measured because it was the objective price that varied across the sampling matrix. Also note that for the test of the radial-sampling matrix,  $F_{M50}$  and  $OP_{M50}$  values were obtained for each psychometric curve because both the pulse frequency and objective price varied. This distribution of  $F_{M50}$  and  $OP_{M50}$  values are plotted in  $[\text{Log}_{10}(F) \text{ vs. } \text{Log}_{10}(OP)]$  space, illustrated in Figure 13.

### ***2.12. The pulse frequency-objective price trade-off functions***

The pulse frequency-objective price trade-off functions in the  $[\text{Log}_{10}(F) \text{ vs. } \text{Log}_{10}(OP)]$  space are then fit to the distribution of the  $F_{M50}$  and  $OP_{M50}$  values for each rat. There are 4 pulse frequency-objective price trade-off functions: each one incorporates one of the proposed subjective price functions as well as the reward mountain parameters.

The form of the pulse frequency-objective price trade-off functions is the contour line of the 3-dimensional reward mountain at half-maximal time allocation. This trade-off function is collapsed into two dimensions in the space  $[\text{Log}_{10}(F) \text{ vs. } \text{Log}_{10}(OP)]$  as illustrated in Figure 14A and B. Derivation of the 4 forms of the pulse frequency-objective price trade-off functions from the general  $TA$  reward mountain equation are described in Appendix B.

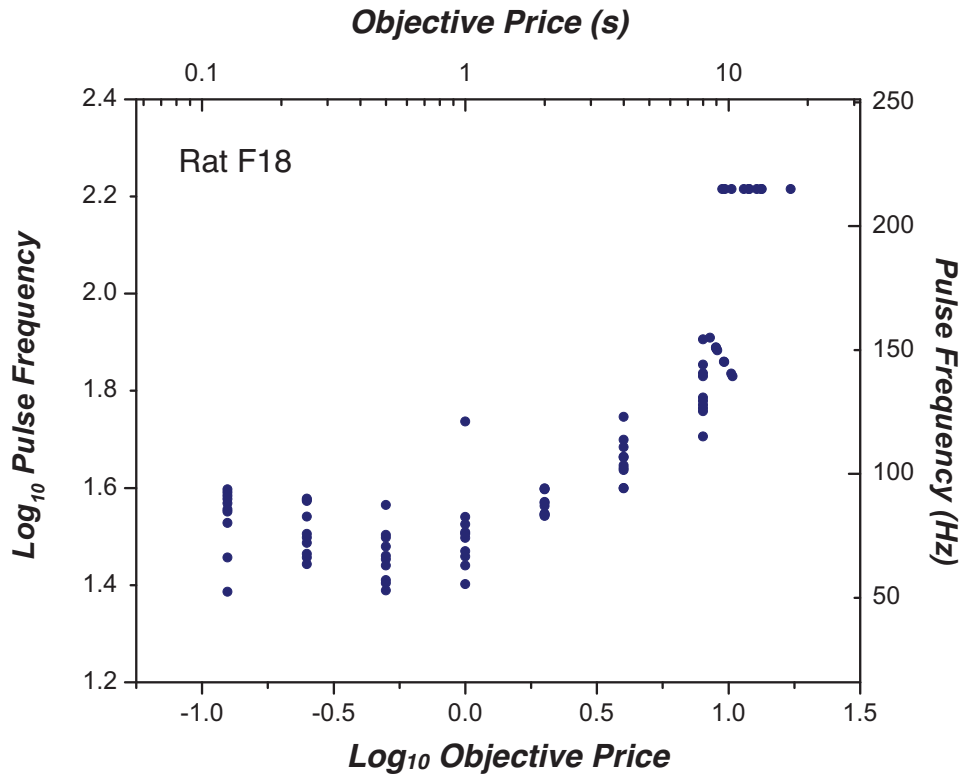


Figure 13. Distribution of all values from Table 1 for Rat F18 after completion of the experiment. It is to this distribution that the proposed pulse frequency-objective price trade-off functions are fit.

The four pulse frequency-objective price trade-off functions fit to the  $F_{M50}$  and  $OP_{M50}$  data points are:

**(1). The objective price trade-off function:**

*(This is the form of the subjective price function that has been assumed in previous studies using the reward mountain model)*

$$\text{Log}F_{M50} = \text{Log}(F_{hm}) + \frac{1}{g} \times \text{Log} \left( \frac{SP}{SP_e - SP} \right)$$

where,

$$SP = OP$$

*(with 3 parameters:  $F_{hm}$ ,  $SP_e$ ,  $g$ )*

**(2). Shizgal's subjective price trade-off function:**

$$\text{Log}F_{M50} = \text{Log}(F_{hm}) + \frac{1}{g} \times \text{Log} \left( \frac{SP}{SP_e - SP} \right)$$

where,

$$SP = SP_{\min} + (SP_{bnd} \times \ln(1 + e^{\frac{OP - SP_{\min}}{SP_{bnd}}}))$$

*(with 4 parameters  $F_{hm}$ ,  $SP_e$ ,  $g$ ,  $SP_{\min}$ ,  $SP_{bnd}$ )*

**(3). The inverted hyperbolic discounting trade-off function:**

$$\text{Log}F_{M50} = \text{Log}(F_{hm}) + \frac{1}{g} \times \text{Log} \left( \frac{SP}{SP_e - SP} \right)$$

where,

$$SP = 1 + (k \times OP)$$

(with 4 parameters:  $F_{hm}$ ,  $SP_e$ ,  $g$ ,  $k$ )

**(4). The inverted exponential discounting trade-off function:**

$$\text{Log}F_{M50} = \text{Log}(F_{hm}) + \frac{1}{g} \times \text{Log} \left( \frac{SP}{SP_e - SP} \right)$$

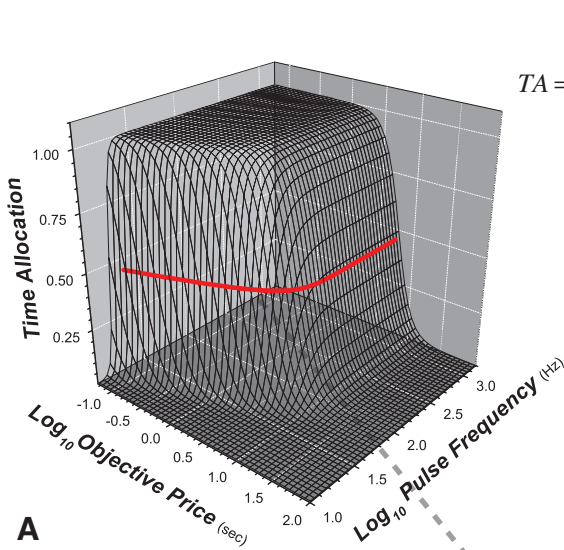
where,

$$SP = e^{k \times OP}$$

(with 4 parameters:  $F_{hm}$ ,  $SP_e$ ,  $g$ ,  $k$ )

The forms of these 4 functions are illustrated in Figure 14 for demonstration, with parameters set to values typically seen in BSR experiments.





$$TA = TA_{\min} + \left[ (TA_{\max} - TA_{\min}) \times \frac{\left( \frac{F^g}{F^g + F_{hm}^g} \right)^a}{\left( \frac{F^g}{F^g + F_{hm}^g} \right)^a + \left( \frac{SP}{SP_e} \right)^a} \right]$$

By substituting for half-maximal time allocation in the *TA* expression above and simplifying, the contour line at half-maximal time allocation is expressed as:

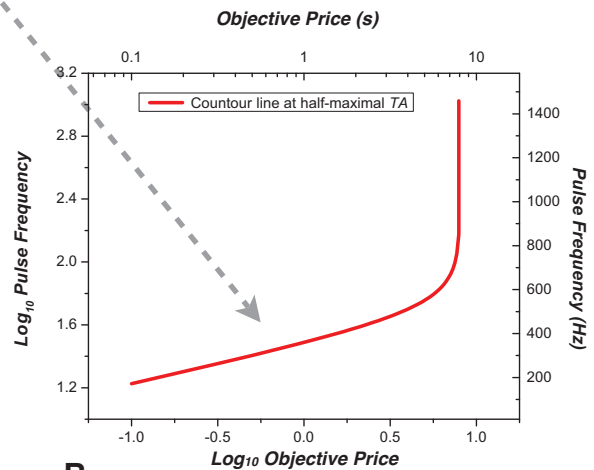
$$\text{Log}F_{M50} = \text{Log}(F_{hm}) + \frac{1}{g} \times \text{Log} \left( \frac{SP}{SP_e - SP} \right)$$

where,

$$SP = OP$$

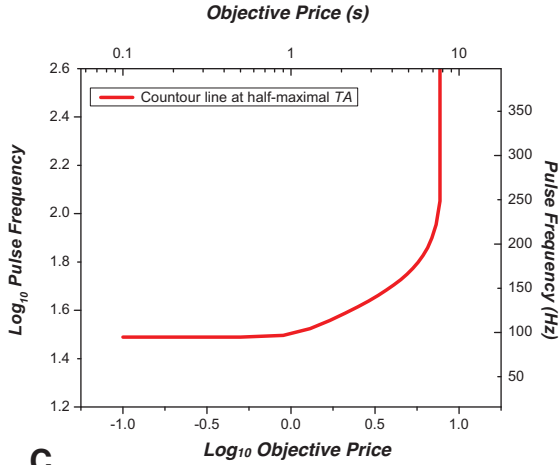
A

The objective price trade-off function



B

Figure 14. A and B. The pulse frequency-objective price trade-off function to be fit to the data points of the (90) psychometric functions is the 3-dimensional reward mountain representation (A) collapsed onto a 2-dimensional plane when time allocation is half-maximal (B). The trade-off function is expressed in terms of independent variables *OP* and *F* and reward mountain parameters *F<sub>hm</sub>*, *g* and *SP<sub>e</sub>* and the associated subjective price parameters. The above 3-dimensional reward mountain and corresponding pulse frequency trade-off function and equations are expressed for the “objective price function” where the transformation of the objective price into subjective price is one to one, *SP = OP*. The same logic applies for deriving the forms of the pulse frequency-objective price trade-off functions for the other subjective price functions to be fit to the data. The derivation for the contour line is described formally in Appendix A.



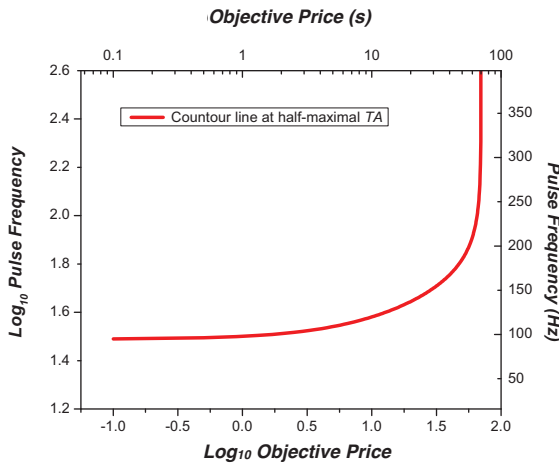
C

**Shizgal's subjective price trade-off function**

$$\text{Log}F_{M50} = \text{Log}(F_{hm}) + \frac{1}{g} \times \text{Log} \left( \frac{SP}{SP_e - SP} \right)$$

where,

$$SP = SP_{\min} + (SP_{bnd} \times \ln(1 + e^{\frac{OP - SP_{\min}}{SP_{bnd}}}))$$



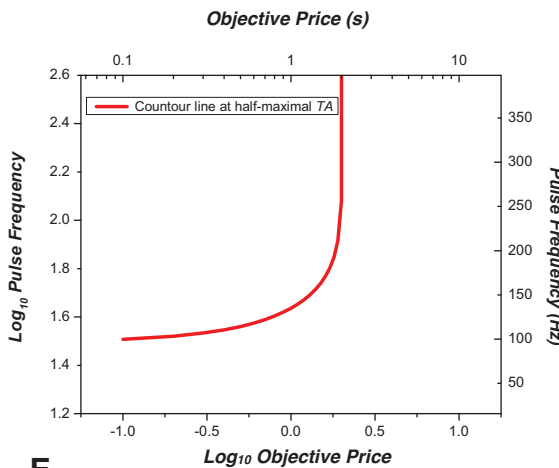
D

**The (inverted) hyperbolic trade-off function:**

$$\text{Log}F_{M50} = \text{Log}(F_{hm}) + \frac{1}{g} \times \text{Log} \left( \frac{SP}{SP_e - SP} \right)$$

where,

$$SP = 1 + (k \times OP)$$



E

**The (inverted) exponential trade-off function:**

$$\text{Log}F_{M50} = \text{Log}(F_{hm}) + \frac{1}{g} \times \text{Log} \left( \frac{SP}{SP_e - SP} \right)$$

where,

$$SP = e^{k \times OP}$$

Figure 14. C-E. The other proposed forms of the pulse frequency-objective price trade-off functions. C. Shizgal's subjective price trade-off function D. The (inverted) hyperbolic trade-off function. E. The (inverted) exponential trade-off function. The parameter values are set to values typically seen in BSR and delay discounting experiments.

### ***2.13. Fitting the pulse frequency-objective price trade-off functions to the data***

The 4 pulse frequency-objective price trade-off functions are fit to the distribution of  $F_{M50}$  and  $OP_{M50}$  values in the  $[\text{Log}_{10}(F) \text{ vs. } \text{Log}_{10}(OP)]$  space (using the Matlab least squares fitting routine in the Statistical Toolbox) and then evaluated to determine the best fitting function.

The resampling method was employed with the goal of obtaining 200 data sets of resampled  $F_{M50}$  and  $OP_{M50}$  distributions and fitting the pulse frequency-objective price trade-off functions to each of the 200 data sets.

To obtain *one* entire set of resampled data, for each sampling matrix, the  $F_{M50}$  values were sampled with replacement for the number of times equivalent to the number of surveys collected (for most rats, 10). For example, considering the distribution of data points for Rat F18 in Table 1, when resampling with replacement for the data obtained in the tests of the pulse frequency sampling vector @ 0.125 s, the resampled data points ( $F_{M50}$  values) could be from survey numbers: 2, 4, 1, 5, 6, 7, 1, 9, 1, 9. For the pulse frequency sampling vector @ 0.25 s, the resampled data points could be from survey numbers: 4, 10, 9, 6, 1, 6, 2, 3, 2, 1. This resampling is employed for all of the sampling matrices. Thus, one entire set of resampled data will look like Table 1 and Figure 13, but with the resampled values. The pulse frequency-objective price trade-off functions are then fit to this resampled data set and parameters (such as  $F_{hm}$ ,  $SP_e$ ,  $g$ ,  $SP_{min}$ ,  $SP_{bnd}$ ,  $k$ ) are derived. The resampling methodology is then repeated, generating a second data set that looks like Table 1 but with resampled values. To this resampled data set, the pulse frequency-objective price trade-off functions are fit, and parameters are again derived. This resampling procedure is repeated 200 times (obtaining 10 tables like Table 1 but

with resampled values). After the 10 iterations, the mean trade-off function parameter values ( $F_{hm}$ ,  $SP_e$ ,  $g$ ,  $SP_{min}$ ,  $SP_{bnd}$ ,  $k$ ) are calculated along with their 95% confidence intervals. The resampling method is employed because it allows for confidence intervals around the parameter means.

For each rat, the mean parameters and corresponding 95% confidence intervals for each function are presented in Tables ApC.1-6 of Appendix C. For each rat, the pulse frequency-objective price trade-off functions are plotted for each subjective price function. Figures 15 and 16 (left panels) illustrates the 4 pulse frequency-objective price trade-off functions for rat F16 and rat F18. The other rats' pulse-frequency-objective price trade-off functions are in Figures ApD.1-4 of Appendix D.

#### ***2.14. Subjective price as a function of objective price***

Once the parameters of each of the objective price-pulse frequency trade-off functions are determined for each rat, each of the 4 proposed subjective price functions can be plotted: subjective price ( $SP$ ) as a function of objective price ( $OP$ ).

Furthermore, for graphical purposes, the data points ( $F_{M50}$  values) of the [ $Log_{10}(F)$  vs.  $Log_{10}(OP)$ ] space can be transformed into subjective price space [ $Log_{10}(SP)$  vs.  $Log_{10}(OP)$ ].

This transformation is completed by the expression below that has been derived from the reward-mountain equation by solving for  $SP$  when time allocation ( $TA$ ) is half-maximal (Appendix E):

$$SP_{MidTA} = RI_{rel} \times SP_e$$

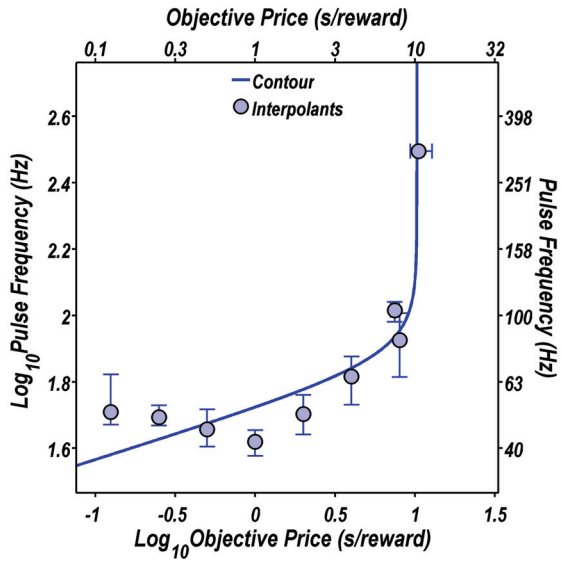
where,

$$RI_{rel} = \frac{F^g}{F^g + F_{hm}^g}$$

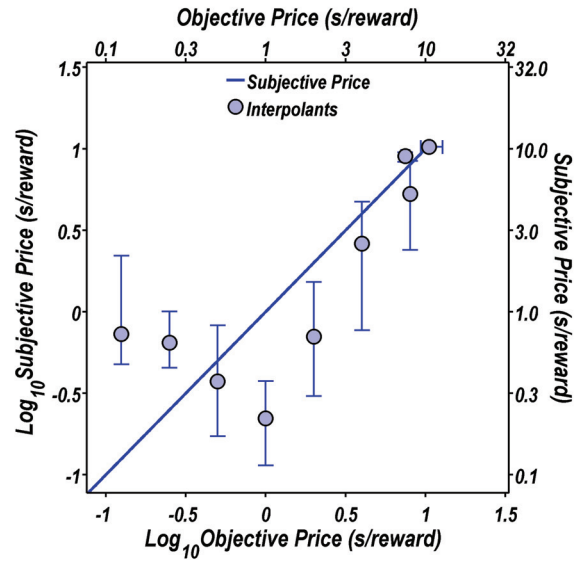
$SP_{MidTA}$  = the subjective price corresponding to half-maximal time allocation

To derive the  $SP_{MidTA}$  values, the parameters,  $g$ ,  $SP_e$ ,  $F_{hm}$  are obtained from the respective pulse-frequency trade-off functions. The subjective price function plots for rats F16 and F18 are illustrated in the right panels of Figure 15 and 16. The plots for the other rats are in Figures ApD.1-4 of Appendix D.

## Objective price function

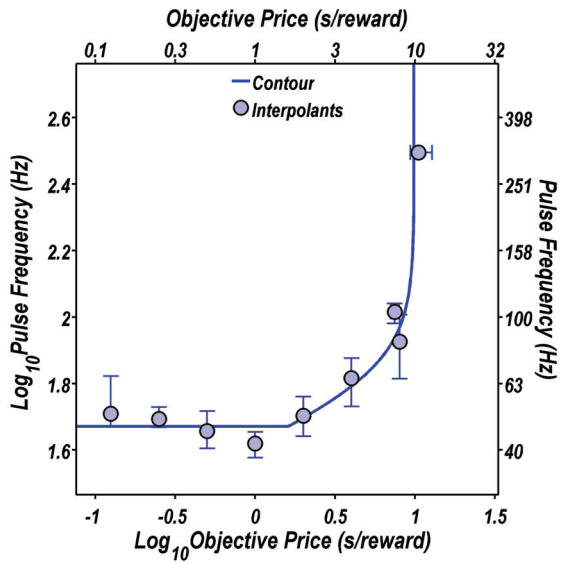


A

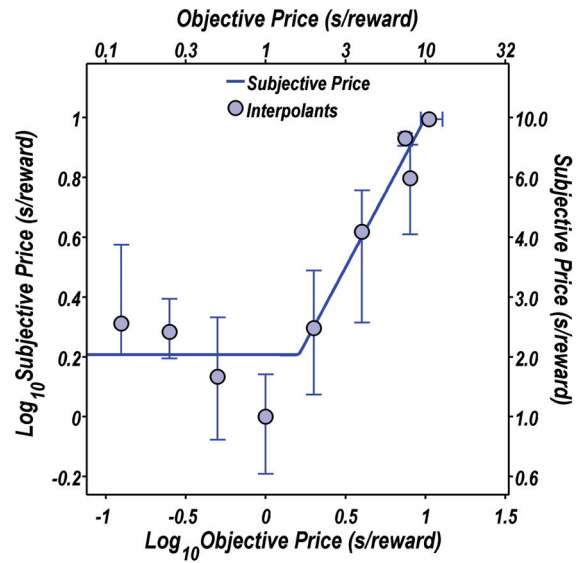


B

## Shizgal's subjective price function

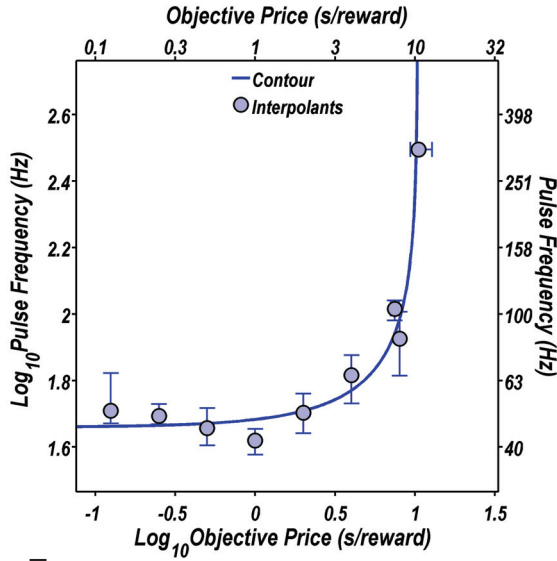


C

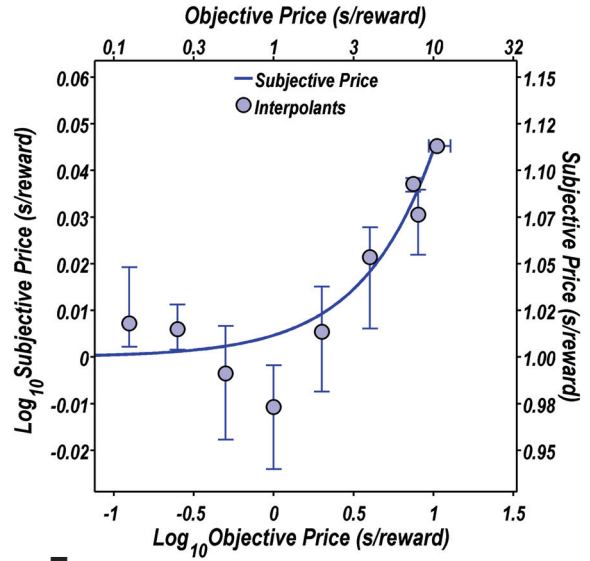


D

### The (inverted) hyperbolic subjective price function

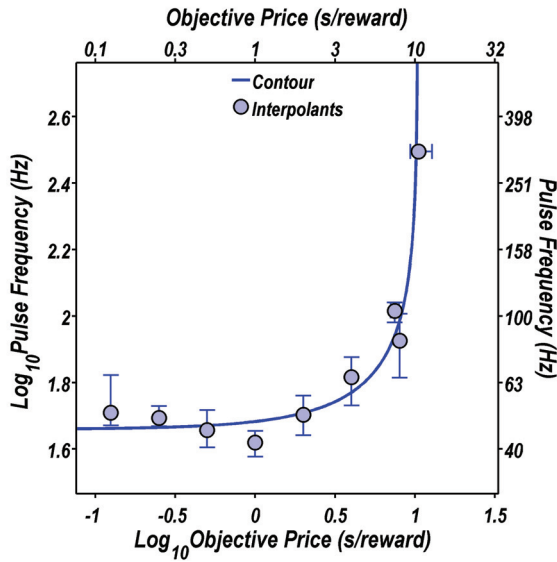


E

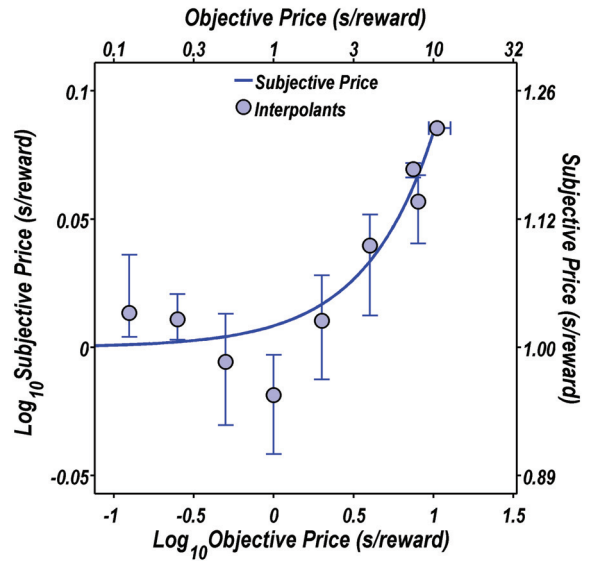


F

### The (inverted) exponential subjective price function



G



H

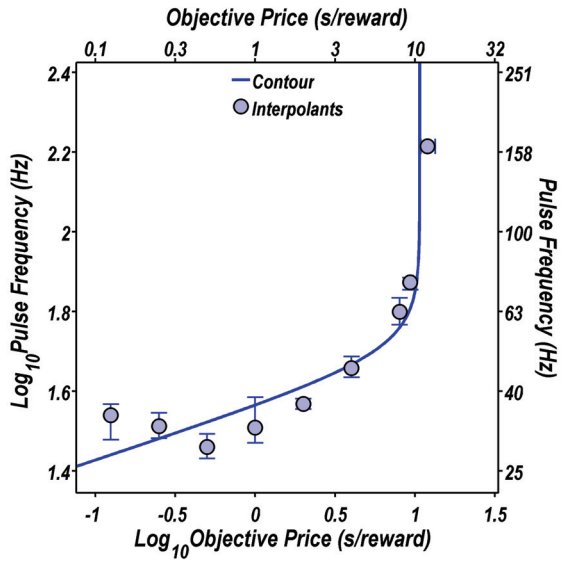


Figure 15. For rat F16, the pulse-frequency trade-off functions (left) and corresponding subjective price functions (right). A & B. The objective price model. C & D. Shizgal's subjective price function. E & F. The inverted hyperbolic delay discounting subjective price function. G & H. The inverted exponential delay discounting subjective price function. "Contour" denotes the function, "interpolants" denotes the data points.

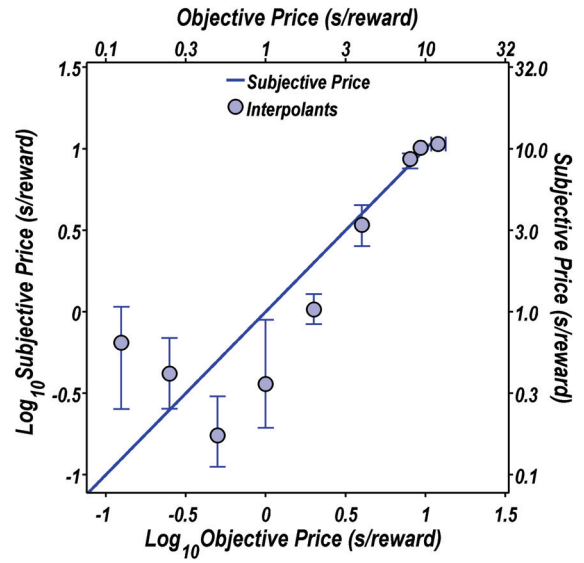
For illustrative purposes, the data points represented on the pulse-frequency trade-off function graphs (left) are the resampled data points means ( $F_{M50}$  and  $OP_{M50}$ ) with associated 95% confidence intervals. Specifically, 1000 resampled distributions were generated: the mean of the  $F_{M50}$  distribution corresponding to each sampling vector was calculated. The data points represent the mean of these 1000 means for each sampling vector. The trade-off function plotted (contour) uses the parameters from the fit to the initial (non-resampled) data.

The corresponding subjective function is plotted to the right of the trade-off function. The parameters derived from the trade-off function are used to plot the subjective price function. The data point means are the transformed pulse frequency values (described in text). Thus, each obtained  $F_{M50}$  value from the spline function corresponds to a subjective price ( $SP$ ) at half-maximal responding. The resampling technique is used as above, 1000 resampled distributions corresponding to each sampling vector was calculated, the mean  $SP$  of each of the 1000 means was calculated.

## Objective price function

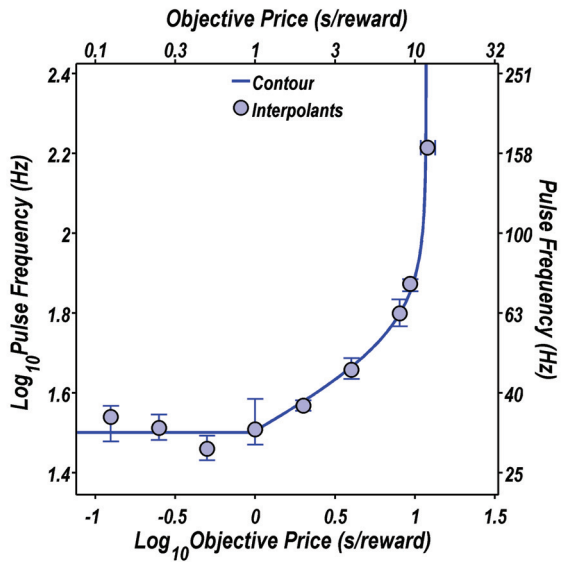


A

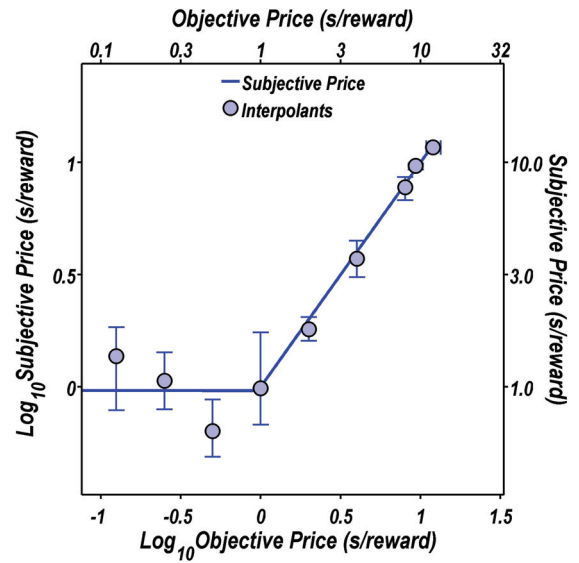


B

## Shizgal's subjective price function

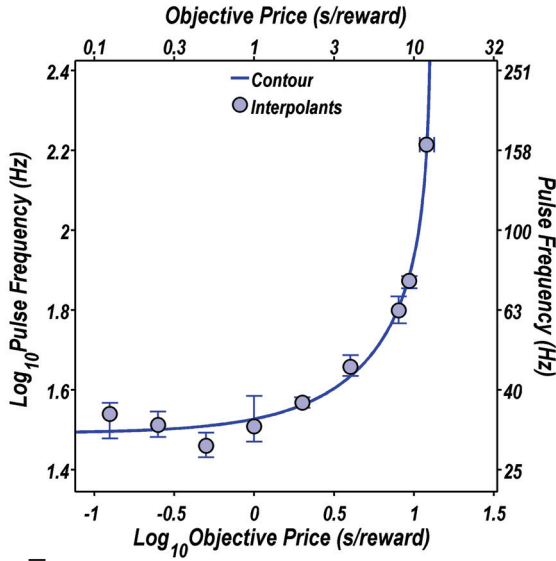


C

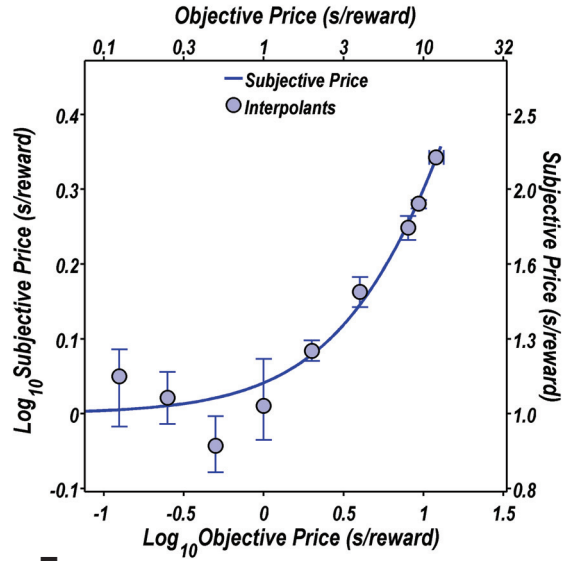


D

The (inverted) hyperbolic subjective price function

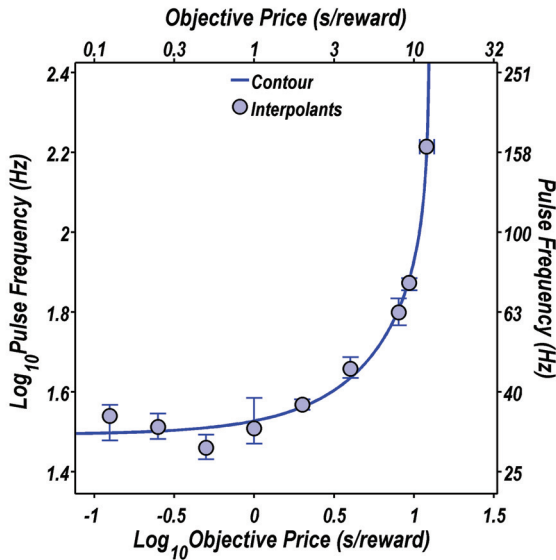


E

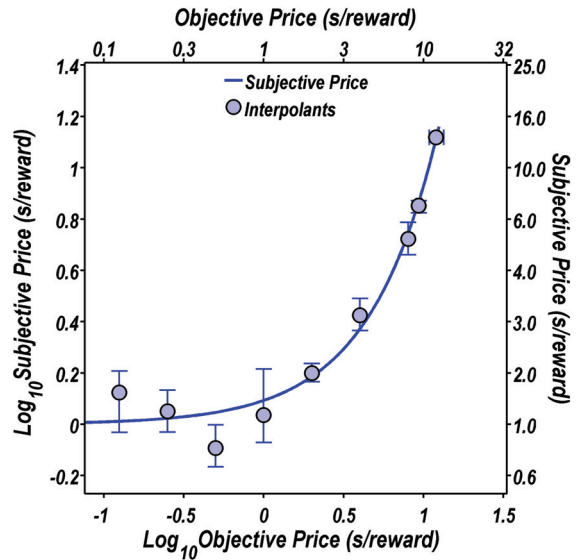


F

The (inverted) exponential subjective price function



G



H

Figure 16. For rat F18, similar caption as for Figure 15.

### **2.15. Model comparisons based on AIC values**

To compare the model fits and determine which model provided the best fit to the data, the Akaike Information Criterion (Akaike, 1974) was employed. Briefly, the AIC statistic provides an estimate of the relative superiority of the tested models by balancing the goodness of fit with the complexity of the model (the number of parameters). For each rat, the AIC statistic was calculated for each model. The *AIC values* and associated statistics for the models for each rat are presented in Tables 2-7. On a relative scale, the more negative the *AIC value*, the better the model. The difference between the AIC for all of the models and the highest ranked model, termed *dAIC*, is determined. The *likelihood* corresponding to each *dAIC* is calculated ( $likelihood = e^{(-dAIC/2)}$ ), which determines the probability that a given model is better than the highest ranked model. The *AIC weight* is the probability that the model is the best model among the whole set of candidate models ( $AIC\ weight = likelihood / \text{sum of } likelihood \text{ of all models}$ ). The *evidence ratio* is the number of times the highest ranked model is more likely to be better than a given model ( $evidence\ ratio = AIC\ weight \text{ of highest ranked model} / AIC\ weight \text{ of given model}$ ). The residual sum of squares (*RSS*) is the sum of the squared errors of prediction, a measure of the difference between the (observed) data and a (predicted) model. In the Tables 2-7, SP refers to Shizgal's subjective price model, HD is the inverted hyperbolic delay discounting model, ED is the inverted exponential discounting model, and OP is the objective price function. The models in the table are presented from best to worst. For the purposes of interpretation and discussion below, it is easiest to discuss the AIC scores with respect to *AIC weights*: the probability that the model is the best model among the whole set of candidate models.

**Table 2. Rat F3**

Model Name	AIC value	dAIC	Likelihood	AIC weight	Evidence Ratio	N	N params	RSS
ED	-461.4674	0	1	0.50001	1	99	4	0.85961
HD	-461.4673	0.0001	0.99995	0.49999	1.00005	99	4	0.85961
SP	-437.0955	24.372	0.00001	0	1.96E+05	99	5	1.07518
OP	-425.9095	35.558	0	0	5.26E+07	99	3	1.25841

**Table 3. Rat F9**

Model Name	AIC value	dAIC	Likelihood	AIC weight	Evidence Ratio	N	N params	RSS
ED	-474.3459	0	1	0.49928	1	90	4	0.42113
HD	-474.3346	0.0114	0.99434	0.49645	1.00569	90	4	0.42118
SP	-464.8207	9.5252	0.00854	0.00427	117.0523	90	5	0.45662
OP	-437.2582	37.088	0	0	1.13E+08	90	3	0.65156

**Table 4. Rat F12**

Model Name	AIC value	dAIC	Likelihood	AIC weight	Evidence Ratio	N	N params	RSS
SP	-471.9116	0	1	0.84334	1	90	5	0.42202
HD	-467.8755	4.036	0.13292	0.1121	7.52343	90	4	0.45252
ED	-465.8793	6.0323	0.04899	0.04132	20.41226	90	4	0.46267
OP	-460.7914	11.12	0.00385	0.00325	259.8491	90	3	0.50165

**Table 5. Rat F16**

Model Name	AIC value	dAIC	Likelihood	AIC weight	Evidence Ratio	N	N params	RSS
ED	-346.9345	0	1	0.37696	1	72	4	0.51619
HD	-346.9114	0.0231	0.98854	0.37264	1.01159	72	4	0.51636
SP	-346.1162	0.8182	0.66425	0.2504	1.50547	72	5	0.5056
OP	-317.5419	29.393	0	0	2.41E+06	72	3	0.80102

**Table 6. Rat F17**

Model Name	AIC value	dAIC	Likelihood	AIC weight	Evidence Ratio	N	N params	RSS
ED	-345.2044	0	1	0.43224	1	81	4	1.02778
HD	-344.9413	0.2631	0.87674	0.37897	1.14059	81	4	1.03113
SP	-343.5363	1.6681	0.4343	0.18772	2.30257	81	5	1.02013
OP	-333.2022	12.002	0.00248	0.00107	403.8734	81	3	1.22497

**Table 7. Rat F18**

Model Name	AIC value	dAIC	Likelihood	AIC weight	Evidence Ratio	N	N params	RSS
SP	-500.5989	0	1	0.49022	1	90	5	0.30683
HD	-499.2927	1.3062	0.52043	0.25512	1.9215	90	4	0.31918
ED	-499.289	1.3099	0.51947	0.25466	1.92503	90	4	0.31919
OP	-458.8359	41.763	0	0	1.17E+09	90	3	0.51266

Tables 2-7. The AIC values and associated statistics for all rats, for all models. The models are presented in order from best to worst. (The more negative the AIC value, the better the model).

For rat F3 (Table 2), the inverted exponential and hyperbolic discounting models are equivalent in terms of AIC statistics score. The AIC weights demonstrate that there is an equal chance that the exponential discounting model or the hyperbolic discounting model is the best model, with a 50% chance corresponding to both models. According to the AIC weights, Shizgal's subjective price function and the objective price function have a 0% chance of being the superior model out of all the models tested.

For rat F9 (Table 3), the AIC values and associated statistics demonstrate the same pattern as F3: a 50% chance that either the exponential delay discounting model or hyperbolic discounting model is the best while Shizgal's subjective price model has a 0.4% chance, followed by the objective price model at 0%.

For rat F12 (Table 4), the AIC demonstrates that Shizgal's subjective price model is the highest ranked model with an 84% probability. The hyperbolic discounting model is second, with an 11% chance of being the best model, followed by the exponential delay discounting model at 4%, and objective price function at 0%.

For rat F16 (Table 5) the exponential delay discounting and hyperbolic discounting models, were of equal probabilities of being the superior model at 37%, followed closely by Shizgal's subjective price function at 25% while the objective price model had a 0% chance of being the superior model.

For rat F17 (Table 6), the exponential discounting function had the highest probability of being the best, at 43% followed by the hyperbolic discounting at 37%, Shizgal's model followed at 18%, the objective price function at 0%.

For rat F18 (Table 7), Shizgal's subjective price function ranked the highest, with a 50% chance of being the best model, followed by the hyperbolic delay discounting and exponential delay model, each at 25%, and objective price function at 0%.

In summary, according to the AIC statistic, for 3 rats (F3, F9, F16), both the hyperbolic and exponential discounting models with the same weight were the superior models, and for 1 rat (F17) the exponential discounting was the highest ranked model. For two rats (F12 and F18), Shizgal's subjective price model was ranked as the superior model. The objective price model was consistently the lowest ranked model.

In addition to the AIC statistic, the best model for each rat was assessed with regards to theoretical validity. First, the assessment was made with respect to the output of the function. Second, the reasonableness of the parameter values of  $g$  and  $F_{hm}$  of the proposed functions were considered based on previous reward mountain experiments;  $k$  was considered based on previous BSR experiments, and  $SP_{min}$  and  $SP_{bnd}$  based on simulations. The theoretical validity of the models for each rat will be evaluated in the next section.



## ***Discussion***

The present experiment tested four psychophysical functions that transform objective opportunity costs into their subjective equivalents to determine the most plausible transformation. These psychophysical functions include: the previously assumed objective price function in which objective prices equal subjective prices along the whole price range; Shizgal's subjective price function in which the function is flat along low objective prices but becomes scalar at a higher range; the inverted hyperbolic discounting function, which describes prices being treated as a delay to the reward, similar to the "hockey-stick" shape of Shizgal's subjective price function but with a smoother transition and a non-scalar relationship; and the inverted exponential delay discounting function, a similar function that describes prices being treated as a traditional delay to a reward. Reward mountain methodology was used such that 9 sampling matrices of the  $[Log_{10}(F) \text{ vs. } Log_{10}(OP)]$  space were tested. The  $F_{M50}$  and  $OP_{M50}$  values were derived from the psychometric curves and plotted in the  $[Log_{10}(F) \text{ vs. } Log_{10}(OP)]$  space. The 4 pulse frequency-objective price trade-off functions with the embedded proposed subjective price functions were fit to the  $F_{M50}$  and  $OP_{M50}$  data points. The parameters of the subjective price functions were derived ( $g, SP_e, F_{hm}, k, SP_{min}, SP_{bnd}$ ) and the subjective price function was plotted in the space  $[Log_{10}(SP) \text{ vs. } Log_{10}(OP)]$ .

The AIC statistic which balances the goodness of fit with the complexity of the model was used to statistically rank the quality of the proposed models. For 3 rats (F3, F9, F16) the hyperbolic and exponential discounting models of the same AIC weight were the superior models, while for 1 rat (F17) the exponential discounting function was the highest ranked. For 2 rats (F12 and F18), Shizgal's subjective price model was

ranked as the superior model. The objective price model was consistently the lowest ranked model.

### ***2.16. Model comparisons based on the theoretical validity of predicted subjective prices***

In addition to statistically ranking the models using the AIC, the plausibility of a model in terms of theoretical validity was assessed. Specifically, the theoretical credibility of each model was examined by considering each given model's predicted subjective prices, as elaborated below.

For F3, for both delay discounting functions, the output, subjective price, spanned a very small range at the low end of the subjective price scale. Examining the range of tested objective prices, 0.125 s to 8 s (~1.8 log unit range), the inverted hyperbolic discounting function predicted subjective prices at a range of only ~ 0.035 common logarithmic units, from about 1 to 1.09 s. For the inverted exponential delay discounting model, the predicted subjective prices over the tested objective price range spanned only ~ 0.08 log units ranging from 1 s to 1.2 s. Visually, the y-axis is magnified to illustrate the shape of the function, but if both the x-axis and y-axis scales matched (as for the plots of Shizgal's subjective price function and the objective price function), the curve would be an almost horizontal line, even shallower than the orange curve representing this function in Figure 3 (a simulation) in the introduction. This implies that the rat interprets an objective 8 s price almost the same as it does a 0.125 s price. According to these models, a 0.125 s and an 8 s would be interpreted as approximately a 1 s value. However, this inference is clearly unsound: by simply considering the 2-dimensional psychometric curves and assessing the  $F_{M50}$  value at 0.125 s, and at 8 s, an increase in the

$F_{M50}$  value of significant magnitude is seen. For rat F3, the increase is about 0.5 common logarithmic units, a 216 % increase from baseline, traditionally considered a very large effect. Thus, although ranked as statistically superior to Shizgal's subjective price function, this function is not logically superior in terms of theoretical validity for this rat. The other models were evaluated for this rat: the subjective price ranges predicted by both Shizgal's subjective price function and the objective price function matched the objective price range more realistically.

For F9, this same pattern of unrealistic subjective price scaling arises for both of the delay discounting functions. Over the tested objective price range (0.125 s to 8 s), the hyperbolic discounting function predicts the subjective prices to span a range of  $\sim 0.045$  log units, from 1 s to 1.1 s. The exponential delay discount function predicts subjective prices that span a range of 0.1 log units, from 1 s to 1.26 s. This extremely small range was not realistic given that the difference in  $F_{M50}$  values was 0.3 log units when testing the pulse frequency matrix at 0.125 and at 8 s, which implies a 100% increase from baseline. The subjective price values for Shizgal's subjective price function and for the objective price function matched the objective price range more closely.

For F12, all of the proposed functions predict subjective prices that are within a realistic range.

For F16, over the tested range (0.125 to 8 s), the hyperbolic discounting function predicted subjective prices of  $\sim 0.03$  log units, from 1 s to 1.07 s; the exponential delay discounting function predicted subjective prices over a range 0.06 log units from 1 s to 1.15 s. Like the other rats, these ranges are too small to be theoretically valid given the large  $F_{M50}$  increase when comparing 0.125 s with 8 s. Both Shizgal's subjective price

function and the objective price function predicted subjective prices over a more plausible range.

For rat F17, over the tested objective price range (0.25 to 8 s), the hyperbolic discounting function predicted subjective prices of  $\sim 0.2$  log units, from 1 s to 1.58 s. The exponential delay discounting function predicted subjective prices over a range of 0.4 log units, from 1 s to 2.51 s. These ranges are considered small, yet theoretically possible; the  $F_{M50}$  increase between 0.125 s and 8 s was 0.2 log units (58%), which is considered a large increase but not as large as for the other rats. Shizgal's subjective price function and the objective price function predicted subjective prices over a plausible range.

For rat F18, over the tested range (0.25 to 8 s), the hyperbolic discounting function predicted subjective prices of  $\sim 0.25$  log units, from 1 s to 1.7 s. This small range is not realistic possible given that the pulse frequency trade-off function demonstrated an effect of 0.3 log units from 0.125 s and 8 s. The exponential delay discounting function, Shizgal's subjective price function and the objective price function predicted subjective prices over a plausible range.

In summary, the interpretation of the transformation of objective prices into subjective ones was considered to be non-realistic if, over the tested objective price range of 0.125 s to 8 s, a model predicted subjective price values of an extremely small range. The occurrences in which the models predicted such a small subjective price range imply that the rat's interpretation of 8 s is almost indistinguishable from 0.125 s. This inference is illogical based on pulse frequency-objective price trade-off magnitudes (large  $F_{M50}$  increases over the tested range). This small subjective price range was seen for the two delay discounting functions. Specifically, the two delay discounting functions predicts

this small subjective price range for rats F3, F9, F16. The range spanned only about 0.1 log units, in which the objective prices were interpreted to be around 1 s. For F17, both delay discounting functions predicted subjective prices in the range of 0.2 to 0.5 log units, a small yet plausible range due to the relatively small pulse frequency-objective price trade-off effect. For rat F18, the hyperbolic delay discounting predicted unrealistic subjective prices, while the exponential function predicted values over a reasonable range.

### ***2.17. Model comparisons based on the credibility of reward mountain and delay discounting parameters***

The tested models were also assessed with regards to the value of the parameters that are typically seen in reward-mountain experiments,  $g$  and  $LogF_{hm}$ . Also, we considered the plausibility of  $LogOP_e$  which is back-solved from parameter  $LogSP_e$  of the present experiment; parameter  $LogOP_e$  is considered because it is this parameter that has been used in previous versions of the reward-mountain model. The plausibility of parameter  $LogOP_e$  is easy to assess: it is the half-maximal value of the time allocation curve of price-sampling matrix condition. It typically spans from 0.85 to 1.08 (7 to 12 s in anti-log units). In previous reward-mountain studies, parameter  $g$  typically spans from 2 to 10; parameter  $LogF_{hm}$  typically spans 1.5 to 2. For F3 and F9, parameters  $g$  and  $LogF_{hm}$  are well out of the typical range for both of the delay discounting functions. For example, for F3 (Table AB1), the inverted hyperbolic discount function reported a  $LogF_{hm}$  value of 0.27 log units (which is 1.86 Hz): this is an implausible pulse frequency at half-maximal reward intensity given that rats will not even respond for a 1.86 Hz train

of electrical stimulation. (Depending on the rat, the typical test range of test pulse frequency values begins at about 20 Hz, an electrical stimulation strength for which the rat spends 10-20% of its time responding.) For F12, the parameters were all within the appropriate range. For F16, for both delay discounting functions,  $LogF_{hm}$  is outside the appropriate range. For F17, for both delay discounting functions, the parameters are within the range. For F18, for the hyperbolic discounting function,  $LogF_{hm}$  is outside of the range. Interestingly, in most cases  $LogOP_e$  is in an appropriate range; the exception is rat F17 for the hyperbolic delay discounting model in which  $LogOP_e$  is outside the plausible range of values.

The  $k$  values that have been previously estimated in terms of hyperbolic delay discounting in BSR studies range from 0.037 to 0.455 (Mazur, Stellar, & Waraczynski, 1987; Fouriezos & Randall, 1997). Using the inverted hyperbolic discounting function, the present study reports  $k$  values: F3: 0.009, F9: 0.027, F12: 1.20, F16: 0.012, F17: 0.13, F18: 0.10. These  $k$  values are within reasonable bounds, although at the lower limit of what has previously been reported in BSR research with the exception of rat F12 which has a high value. The reasonable bounds of the  $k$  values gives strength to the use of the “inverted” form of the function. However, when we assess the delay discounting functions at low  $k$  values by considering the plausibility of the rat’s psychological treatment of objective price, the delay discounting functions do not offer realistic interpretations as explained above (2.15).

The  $k$  values for the exponential function are also within a reasonable range. The  $SP_{min}$  and  $SP_{bnd}$  for Shizgal’s subjective price function were also within realistic bounds based on simulations as demonstrated in Figure 2B and 2C.

### ***2.18. Conclusion: inferring the best model using statistical and theoretical criteria***

Taken together, to deduce the best subjective price function for each rat, the AIC score was considered, as well as the theoretical validity of the subjective price output and reward-mountain parameters. For rats F3 and F9, both of the delay discounting functions are superior with respect to the AIC statistic. However, the subjective price output and the reward-mountain parameters were very far from a reasonable range. This suggests that the next highest ranked model according to the AIC, Shizgal's subjective price function, probably provides the best fit for these rats. Given that the AIC for Shizgal's subjective price function and the objective price function are 0, we calculated the associated AIC statistics without the delay discounting and exponential functions (Tables ApF.1-6 of Appendix F). This calculation demonstrated that Shizgal's subjective price function is the best function out of the two: according to the AIC weight, there is a dramatic 99.6% chance that Shizgal's subjective price function is the best model for this comparison for F3, and a 99.9% chance for F9.

For F12, the AIC statistic demonstrated that Shizgal's subjective price function is the best model with an 85% probability. Furthermore, the subjective prices and reward-mountain parameters were within reasonable bounds for this function. Thus, it was concluded that Shizgal's subjective price function is the superior model for this rat.

For F16, the exponential and hyperbolic discounting function were the best models according to AIC, yet for these models predicted subjective prices and reward mountain parameters outside of reasonable ranges. Thus, it was concluded that the third ranked model, Shizgal's subjective price function, was the best model. Comparing the AIC statistics of Shizgal's subjective price model with the objective price model without

the two delay discounting functions demonstrated that there is a 99.9% chance that Shizgal's subjective price function is the superior model (Table ApF.4).

For F17, the exponential discounting model was the highest ranked according to the AIC, followed by the delay discounting function. The subjective prices were in a smaller range as compared to other rats, yet still fell within a plausible range. Thus, it was concluded that the exponential discounting function was the superior model.

For F18, the AIC statistic predicted Shizgal's subjective price function to be the best model at 50% followed by the hyperbolic discounting function and exponential discounting followed each with a 25% chance. The reward-mountain parameters were out of the reasonable range for the hyperbolic discounting model. It was concluded that for this rat, Shizgal's subjective price was the best model.

Overall, it was concluded, for 5 out of the 6 rats, based on statistical and theoretical criteria, Shizgal's subjective price function was the best model.

### ***2.19. Implications***

Of the 6 rats, the "objective price function" ranked last with regard to the AIC statistic. This consistent ranking confirms the hypothesis that the function previously assumed in reward-mountain experiments is a poor way to describe the transformation. Furthermore, given that the AIC penalizes the addition of a parameter, the objective price function should have an advantage over the other models; it has one fewer parameter than the delay discounting functions and two fewer than Shizgal's subjective price functions. However, even given this advantage, the objective price function is not a convincing way to describe the transformation of objective into subjective prices.



Shizgal's subjective price function was the best model to describe the transformation of objective prices into subjective prices. The parameters of this function  $SP_{min}$  (the minimum subjective price) and  $SP_{bnd}$  (a value controlling the bend of the function) were consistent across rats, which further underscores the strength of this model. The  $SP_{min}$  values are: 0.91 s (F3), 0.94 s (F9), 0.93 s (F12), 1.70 s (F16), 1.00 s (F17), 0.97 s (F18). The  $SP_{bnd}$  parameters are: 0.01 (F3), 0.01 (F9), 0.01 (F12), 0.36 (F16), 0.51 (F17), 0.02 (F18). Across rats, the mean  $SP_{min}$  is 1.08 and mean  $SP_{bnd}$  is 0.15. The function with these parameters resemble the function in the simulation in Figure 3C of the introduction, the green curve. According to the estimated parameters in the experiment, the function begins to be scalar at around 2 s, therefore setting prices above 2 s would be appropriate in operant experiments that use price as a variable. In future reward-mountain models incorporating Shizgal's subjective price function, the  $SP_{min}$  and  $SP_{bnd}$  parameters will be fixed at 1.08 s and 0.15, respectively.

Conventional hyperbolic and exponential models of delay discounting were assessed in their inverted form to determine whether time spent working for the reward (opportunity cost) is treated as a related psychological variable, the delay to reward delivery. In traditional studies of reward delays, there is an experimenter-imposed delay to reward delivery after the subject makes a brief response. In contrast, in studies using opportunity costs, the subject must work for a total required amount of time to reap the reward. However, an inherent delay to reward delivery is set by the imposed price (Figure 2). Thus, it is possible that the same system controls both types of evaluations of the imposed time interval to reward: waiting and working. However, the findings revealed that the delay discounting functions were not able to account for the behaviour

convincingly. Thus, this suggests that two different principles of intertemporal choice are involved in evaluating the time spent working for the reward and time spent waiting for the reward after work requirements have been satisfied. In other words, two different processes are involved in the evaluation of a time interval depending on the context, namely, what the subject is doing during the delay to the reward.

To further strengthen the hypothesis that time spent working and time spent waiting are evaluated by different processes, an analogous experiment can be conducted by directly manipulating delays instead of opportunity costs while the work requirement is set to a constant value. Thus, a pulse frequency-delay trade-off function can be derived in the same fashion: by testing pulse frequency sampling matrices at a range of different delays. Next, the proposed functions, Shizgal's subjective price function, the objective price function, the inverted hyperbolic discounting, and the exponential discounting function can be embedded in the pulse frequency-delay trade-off function and fit to the data. Because the present study suggests that two systems are involved in the evaluation of delays and opportunity costs, it is hypothesized that when delays are imposed, the inverted delay discounting functions would be superior to Shizgal's subjective price function.

The present experiment focused on opportunity costs, one type of cost specified in the reward-mountain model. The other component is that of "effort cost," the amount of exertion per unit time required to obtain the reward. In the present experiment, effort costs have been held constant. However, the effort in an operant experiment can be manipulated by varying the force required to depress the lever via attaching weights to the lever mechanism. The methodology employed to measure the psychophysical

function of opportunity cost can also be used to estimate the psychophysical function of effort costs. Effort costs are of particular interest in operant experiments because it has been suggested that dopamine plays an important role in overcoming this type of costs (Salamone & Correa, 2002; Salamone, Correa, Farrar, & Mingote, 2007; Salamone, Correa, Mingote, & Weber, 2003). The effort (the subjective variable) as a function of force (the objective variable) is predicted to increase as the required force is increased and level off at a certain point, thus, indicating the force at which subjective effort is maximal. Determining the psychophysical function that transforms required force into subjective effort would allow for a refined measurement strategy to evaluate the effects of dopamine manipulations.

## Appendix A: The experimenter-set price versus the computer-generated price

Experimenter-set price (s)	Computer-generated price (s)
0.125	0.137
0.250	0.302
0.500	0.522
1.000	1.016
2.000	2.005
4.000	4.037
8.000	8.047
16.000	16.011

Table ApA.1: Because of the PC clock system constraints, the computer-generated price is not exactly the experimenter-set price, but very close. The error is somewhat larger at the lower prices, but is inconsequential given that the models predict a flat subjective price range along these very low prices. The table demonstrates the differences between the experimenter-set prices and computer-generated prices, from 0.125 s to 16 s.

**Appendix B: The derivation for the expression of the contour line at half-maximal time allocation of the reward mountain model:  $F$  as a function of  $OP$  for each of the 4 proposed subjective price functions.**

First, a general form of the contour line,  $F$  as a function of  $SP$  will be derived.

Second,  $SP$  will be replaced with the 4 proposed subjective price functions.

The time allocation equation can be rearranged and expressed as:

$$\frac{TA - TA_{\min}}{TA_{\max} - TA_{\min}} = \frac{\left(\frac{F^g}{F^g + F_{hm}^g}\right)^a}{\left(\frac{F^g}{F^g + F_{hm}^g}\right)^a + \left(\frac{SP}{SP_e}\right)^a}$$

When time allocation is halfway between  $TA_{\max}$  and  $TA_{\min}$ , the left side of the above equation is equal to 0.5.  $F$  can be expressed as  $F_{M50}$ , the pulse frequency that maintains half-maximal responding. At half-maximal time allocation, the expression above can be simplified:

$$\frac{F_{M50}^g}{F_{M50}^g + F_{hm}^g} = \frac{SP}{SP_e}$$

The next steps are taken to express  $F_{M50}$  as a function of  $SP$ .

$$F_{M50}^g = (F_{M50}^g + F_{hm}^g) \times \left(\frac{SP}{SP_e}\right)$$

$$F_{M50}^g = \left( F_{M50}^g \times \frac{SP}{SP_e} \right) + \left( F_{hm}^g \times \frac{SP}{SP_e} \right)$$

$$F_{M50}^g - \left( F_{M50}^g \times \frac{SP}{SP_e} \right) = F_{hm}^g \times \frac{SP}{SP_e}$$

$$F_{M50}^g \times \left( 1 - \frac{SP}{SP_e} \right) = F_{hm}^g \times \frac{SP}{SP_e}$$

$$F_{M50}^g \times \left( \frac{SP_e - SP}{SP_e} \right) = F_{hm}^g \times \frac{SP}{SP_e}$$

$$F_{M50}^g = F_{hm}^g \times \left[ \frac{SP}{SP_e} \times \left( \frac{SP_e}{SP_e - SP} \right) \right]$$

$$F_{M50}^g = F_{hm}^g \times \left( \frac{SP}{SP_e - SP} \right)$$

$$g \times \text{Log} F_{M50} = g \times \text{Log}(F_{hm}) + \text{Log} \left( \frac{SP}{SP_e - SP} \right)$$

Finally,

$$\text{Log} F_{M50} = \text{Log}(F_{hm}) + \frac{1}{g} \times \text{Log} \left( \frac{SP}{SP_e - SP} \right)$$

The above equation relates  $F_{M50}$  to  $SP$ . Next,  $SP$  is replaced with the 4 proposed subjective price functions such that  $F$  as a function of  $OP$  can be plotted.

**1. The objective price trade-off function:**

The expression below is the contour line at half-maximal time allocation that incorporates the objective price function. It is referred to the *objective price trade-off function*. This function has parameters:  $F_{hm}$ ,  $SP_e$ ,  $g$ .

$$\text{Log}F_{M50} = \text{Log}(F_{hm}) + \frac{1}{g} \times \text{Log} \left( \frac{SP}{SP_e - SP} \right)$$

where,

$$SP = OP$$

**2. Shizgal's subjective price trade-off function**

The expression below is the contour line at half-maximal time allocation that incorporates Shizgal's subjective price function. It is referred to as *Shizgal's subjective price trade-off function*. This function has parameters:  $F_{hm}$ ,  $SP_e$ ,  $g$ ,  $SP_{min}$ ,  $SP_{bnd}$ .

$$\text{Log}F_{M50} = \text{Log}(F_{hm}) + \frac{1}{g} \times \text{Log} \left( \frac{SP}{SP_e - SP} \right)$$

where,

$$SP = SP_{min} + (SP_{bnd} \times \ln(1 + e^{\frac{OP - SP_{min}}{SP_{bnd}}}))$$

### 3. The (inverted) hyperbolic trade-off function:

The expression below is the contour line at half-maximal time allocation that incorporates the (inverted) hyperbolic subjective price function. It is referred to as *the (inverted) hyperbolic trade-off function*. This function has parameters:  $F_{hm}$ ,  $SP_e$ ,  $g$ ,  $k$ .

$$\text{Log}F_{M50} = \text{Log}(F_{hm}) + \frac{1}{g} \times \text{Log} \left( \frac{SP}{SP_e - SP} \right)$$

where,

$$SP = 1 + (k \times OP)$$

### 4. The (inverted) exponential trade-off function:

The expression below is the contour line at half-maximal time allocation that incorporates the (inverted) exponential subjective price function. It is referred to as *the (inverted) exponential trade-off function*. This function has parameters:  $F_{hm}$ ,  $SP_e$ ,  $g$ ,  $k$ .

$$\text{Log}F_{M50} = \text{Log}(F_{hm}) + \frac{1}{g} \times \text{Log} \left( \frac{SP}{SP_e - SP} \right)$$

where,

$$SP = e^{k \times OP}$$

The forms of the 4 functions are illustrated in Figure 14.



## **Appendix C: Reward-mountain parameters for all models, for each rat**

**The objective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	6.18822	5.61318	6.52452	0.91134	0.57504	0.3363	1	25
<b>LogF<sub>hm</sub></b>	1.76617	1.74129	1.79959	0.0583	0.02489	0.03341	0	3
<b>LogSP<sub>e</sub></b>	1.06306	1.00837	1.11793	0.10956	0.05469	0.05487	0	2

**Shizgal's subjective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	6.30369	6.32622	6.32918	0.00296	-0.02253	0.0255	1	25
<b>LogF<sub>hm</sub></b>	1.71939	1.71871	1.72489	0.00618	0.00068	0.0055	0	3
<b>LogSP<sub>e</sub></b>	1.04969	1.04966	1.0509	0.00124	0.00003	0.00121	0	2
<b>LogSP<sub>min</sub></b>	-0.03895	-0.0422	-0.04147	0.00073	0.00325	-0.00252	-1	1
<b>SP<sub>bnd</sub></b>	0.01491	0.01492	0.01492	0	-0.00001	0.00001	0.0001	10

**The inverted hyperbolic delay discounting subjective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	0.67914	0.20405	1.74544	1.54139	0.47509	1.06629	0.1	25
<b>k</b>	0.01363	0.00234	0.07155	0.06921	0.01129	0.05792	0	3
<b>LogF<sub>hm</sub></b>	0.27465	0.00001	1.49954	1.49952	0.27464	1.22488	0	2
<b>LogSP<sub>e</sub></b>	0.0934	0.01496	0.30155	0.28659	0.07844	0.20815	-1	2

**The inverted exponential delay discounting subjective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	0.76168	0.17628	3.36905	3.19277	0.5854	2.60737	0.1	25
<b>k</b>	0.03032	0.00363	0.30269	0.29905	0.02668	0.27237	0	3
<b>LogF<sub>hm</sub></b>	0.84296	0.00439	1.99994	1.99555	0.83857	1.15698	0	2
<b>LogSP<sub>e</sub></b>	0.23314	0.02325	1.64361	1.62036	0.20989	1.41047	-1	2

Table ApC.1. Parameter values for rat F3 for all models. CB Lo indicates the lower 95% confidence band (limit). CB hi indicates the higher 95% confidence band (limit). CB Width indicates the interval from the lower to the higher confidence bounds. Ex Lo indicates the lower 95% confidence error (interval). Ex high indicates the higher 95% confidence error. MB Lo indicates the lower model bound: the lower limit set in the model. MB hi indicates the higher model bound.

**The objective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	7.49174	6.47157	8.08287	1.61131	1.02017	0.59114	1	25
<b>LogF<sub>hm</sub></b>	1.78525	1.75447	1.81559	0.06112	0.03078	0.03034	0	3
<b>LogSP<sub>e</sub></b>	0.98356	0.93123	1.03438	0.10315	0.05233	0.05082	0	2

**Shizgal's subjective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	6.11531	6.11835	6.1188	0.00045	-0.00304	0.00349	1	25
<b>LogF<sub>hm</sub></b>	1.76433	1.76413	1.76662	0.00248	0.0002	0.00229	0	3
<b>LogSP<sub>e</sub></b>	0.97635	0.97609	0.9769	0.00081	0.00026	0.00055	0	2
<b>LogSP<sub>min</sub></b>	-0.02473	-0.02525	-0.02483	0.00041	0.00052	-0.00011	-1	1
<b>SP<sub>bnd</sub></b>	0.01274	0.01276	0.01276	0	-0.00002	0.00002	0.0001	10

**The inverted hyperbolic delay discounting subjective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	1.42049	0.74878	2.7942	2.04542	0.67171	1.37371	0.1	25
<b>k</b>	0.02804	0.00142	0.19031	0.18888	0.02661	0.16227	0	3
<b>LogF<sub>hm</sub></b>	0.80023	0.03139	1.69161	1.66022	0.76883	0.89138	0	2
<b>LogSP<sub>e</sub></b>	0.09892	0.00724	0.48835	0.48111	0.09169	0.38942	-1	2

**The inverted exponential delay discounting subjective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	1.6522	0.73761	4.9183	4.18069	0.91458	3.26611	0.1	25
<b>k</b>	0.06693	0.00271	0.43943	0.43672	0.06423	0.3725	0	3
<b>LogF<sub>hm</sub></b>	1.13966	0.25787	1.99453	1.73666	0.88179	0.85487	0	2
<b>LogSP<sub>e</sub></b>	0.3224	0.01376	1.99992	1.98616	0.30864	1.67752	-1	2

Table ApC.2. Parameter values for rat F9 for all models.

**The objective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	8.03898	7.25637	8.44407	1.1877	0.78261	0.40509	1	25
<b>LogF<sub>hm</sub></b>	1.91636	1.89751	1.93603	0.03851	0.01885	0.01967	0	3
<b>LogSP<sub>e</sub></b>	0.86717	0.84458	0.89138	0.0468	0.02259	0.02421	0	2

**Shizgal's subjective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	5.26347	5.10535	5.48091	0.37557	0.15812	0.21745	1	25
<b>LogF<sub>hm</sub></b>	1.91787	1.89826	1.93852	0.04027	0.01962	0.02065	0	3
<b>LogSP<sub>e</sub></b>	0.88169	0.85575	0.91081	0.05506	0.02594	0.02912	0	2
<b>LogSP<sub>min</sub></b>	-0.03014	-0.04933	0.00956	0.05888	0.01919	0.03969	-1	1
<b>SP<sub>bnd</sub></b>	0.01033	0.00921	0.01104	0.00184	0.00113	0.00071	0.0001	10

**The inverted hyperbolic delay discounting subjective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	5.06409	3.16602	6.85771	3.69168	1.89806	1.79362	0.1	25
<b>k</b>	0.99647	0.23902	2.34897	2.10994	0.75745	1.3525	0	3
<b>LogF<sub>hm</sub></b>	1.88256	1.82874	1.92215	0.09341	0.05382	0.0396	0	2
<b>LogSP<sub>e</sub></b>	0.88392	0.48477	1.26726	0.78249	0.39915	0.38334	-1	2

**The inverted exponential delay discounting subjective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	5.25179	2.70701	7.9359	5.22889	2.54478	2.68411	0.1	25
<b>k</b>	0.41888	0.21502	0.65134	0.43632	0.20387	0.23246	0	3
<b>LogF<sub>hm</sub></b>	1.99984	1.99887	2	0.00113	0.00097	0.00016	0	2
<b>LogSP<sub>e</sub></b>	1.3597	0.80255	1.99982	1.19727	0.55716	0.64012	-1	2

Table ApC.3. Parameter values for rat F12 for all models.

**The objective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	6.62794	6.1007	7.35996	1.25926	0.52724	0.73202	1	25
<b>LogF<sub>hm</sub></b>	1.87018	1.82699	1.90423	0.07723	0.04319	0.03404	0	3
<b>LogSP<sub>e</sub></b>	1.00711	0.94218	1.07943	0.13725	0.06493	0.07232	0	2

**Shizgal's subjective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	4.33162	1.00014	5.52569	4.52555	3.33148	1.19407	1	25
<b>LogF<sub>hm</sub></b>	1.82503	1.6535	1.89739	0.24389	0.17153	0.07236	0	3
<b>LogSP<sub>e</sub></b>	1.00167	0.95378	1.10702	0.15324	0.04789	0.10535	0	2
<b>LogSP<sub>min</sub></b>	0.23003	0.03838	0.79655	0.75817	0.19164	0.56652	-1	1
<b>SP<sub>bnd</sub></b>	0.1245	0.01152	1.86928	1.85776	0.11299	1.74477	0.0001	10

**The inverted hyperbolic delay discounting subjective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	2.38175	1.13749	4.40321	3.26572	1.24425	2.02146	0.1	25
<b>k</b>	0.01236	0.00046	0.14545	0.14499	0.0119	0.13309	0	3
<b>LogF<sub>hm</sub></b>	0.96584	0.16763	1.70322	1.53559	0.79821	0.73738	0	2
<b>LogSP<sub>e</sub></b>	0.03952	0.00208	0.39255	0.39047	0.03744	0.35303	-1	2

**The inverted exponential delay discounting subjective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	2.39535	1.13667	4.54007	3.4034	1.25869	2.14472	0.1	25
<b>k</b>	0.01869	0.00095	0.20769	0.20674	0.01774	0.189	0	3
<b>LogF<sub>hm</sub></b>	1.10578	0.37754	1.91428	1.53674	0.72824	0.8085	0	2
<b>LogSP<sub>e</sub></b>	0.08292	0.00405	0.92297	0.91891	0.07887	0.84005	-1	2

Table ApC.4. Parameter values for rat F16 for all models.

**The objective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	7.64188	6.9337	8.80426	1.87056	0.70818	1.16238	1	25
<b>LogF<sub>hm</sub></b>	1.84302	1.81804	1.87146	0.05342	0.02498	0.02844	0	3
<b>LogSP<sub>e</sub></b>	0.98627	0.95382	1.02036	0.06654	0.03245	0.03409	0	2

**Shizgal's subjective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	6.22397	5.15607	6.77318	1.61711	1.0679	0.54921	1	25
<b>LogF<sub>hm</sub></b>	1.7818	1.64346	1.83417	0.19071	0.13834	0.05237	0	3
<b>LogSP<sub>e</sub></b>	0.96389	0.92697	1.01008	0.08312	0.03692	0.04619	0	2
<b>LogSP<sub>min</sub></b>	0.12072	-0.95625	0.70591	1.66216	1.07697	0.58519	-1	1
<b>SP<sub>bnd</sub></b>	0.57419	0.0103	5.84045	5.83015	0.56388	5.26626	0.0001	10

**The inverted hyperbolic delay discounting subjective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	5.07555	2.45761	6.20239	3.74478	2.61795	1.12683	0.1	25
<b>k</b>	0.19606	0.00289	1.1548	1.15191	0.19317	0.95874	0	3
<b>LogF<sub>hm</sub></b>	1.63262	1.10386	1.80233	0.69848	0.52876	0.16972	0	2
<b>LogSP<sub>e</sub></b>	0.32926	0.01085	1.07566	1.06481	0.31841	0.7464	-1	2

**The inverted exponential delay discounting subjective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	5.03194	2.74122	6.20946	3.46824	2.29072	1.17752	0.1	25
<b>k</b>	0.14927	0.00735	0.44261	0.43526	0.14192	0.29334	0	3
<b>LogF<sub>hm</sub></b>	1.72958	1.35136	1.95304	0.60167	0.37822	0.22345	0	2
<b>LogSP<sub>e</sub></b>	0.5915	0.03119	1.824	1.79282	0.56031	1.23251	-1	2

Table ApC.5. Parameter values for rat F17 for all models.

**The objective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	7.45955	6.40327	8.20457	1.80129	1.05628	0.74501	1	25
<b>LogF<sub>hm</sub></b>	1.69751	1.67349	1.73097	0.05748	0.02402	0.03346	0	3
<b>LogSP<sub>e</sub></b>	1.03196	0.98143	1.08602	0.10459	0.05053	0.05406	0	2

**Shizgal's subjective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	4.62875	4.33616	4.68776	0.35159	0.29259	0.059	1	25
<b>LogF<sub>hm</sub></b>	1.7263	1.71213	1.73575	0.02363	0.01417	0.00945	0	3
<b>LogSP<sub>e</sub></b>	1.0672	1.03828	1.0742	0.03593	0.02892	0.007	0	2
<b>LogSP<sub>min</sub></b>	-0.01378	-0.03409	0.05964	0.09373	0.02032	0.07342	-1	1
<b>SP<sub>bnd</sub></b>	0.02211	0.01484	0.05157	0.03673	0.00727	0.02947	0.0001	10

**The inverted hyperbolic delay discounting subjective price function**

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	2.15797	0.84189	3.53736	2.69547	1.31608	1.37939	0.1	25
<b>k</b>	0.12562	0.00621	0.38895	0.38273	0.11941	0.26333	0	3
<b>LogF<sub>hm</sub></b>	1.39579	0.42695	1.68061	1.25366	0.96884	0.28482	0	2
<b>LogSP<sub>e</sub></b>	0.35687	0.03912	0.73679	0.69767	0.31775	0.37992	-1	2

**The inverted exponential delay discounting subjective price function**

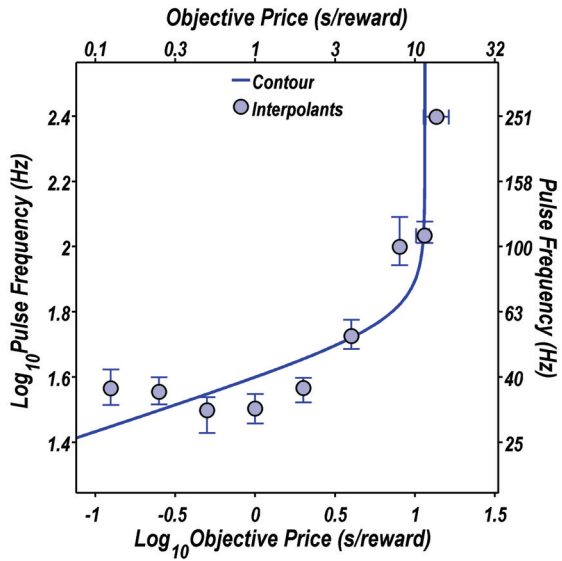
	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
<b>g</b>	3.02836	0.85108	5.20271	4.35163	2.17728	2.17435	0.1	25
<b>k</b>	0.22037	0.01337	0.40956	0.39619	0.20701	0.18918	0	3
<b>LogF<sub>hm</sub></b>	1.74037	0.91616	1.99992	1.08376	0.82422	0.25955	0	2
<b>LogSP<sub>e</sub></b>	1.17576	0.08828	2	1.91172	1.08748	0.82424	-1	2

Table ApC.6. Parameter values for rat F18 for all models.

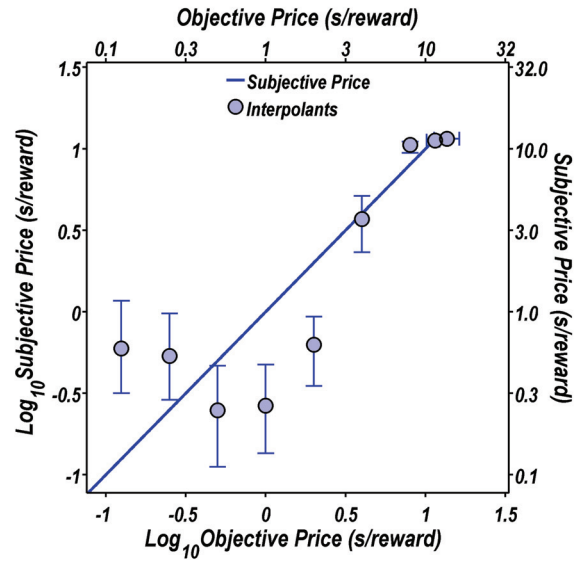
**Appendix D: Pulse frequency trade-off functions and corresponding subjective price functions for all models, for each rat**



## Objective price function

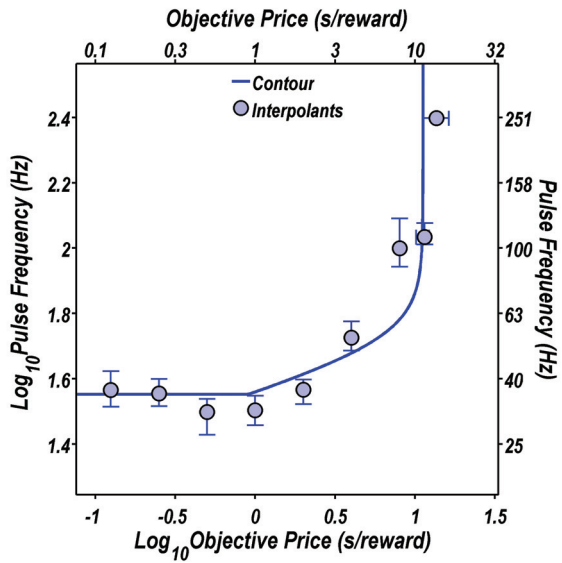


A

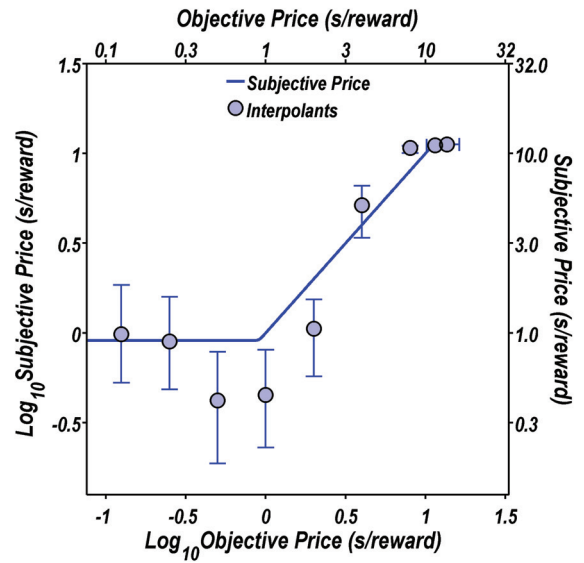


B

## Shizgal's subjective price function

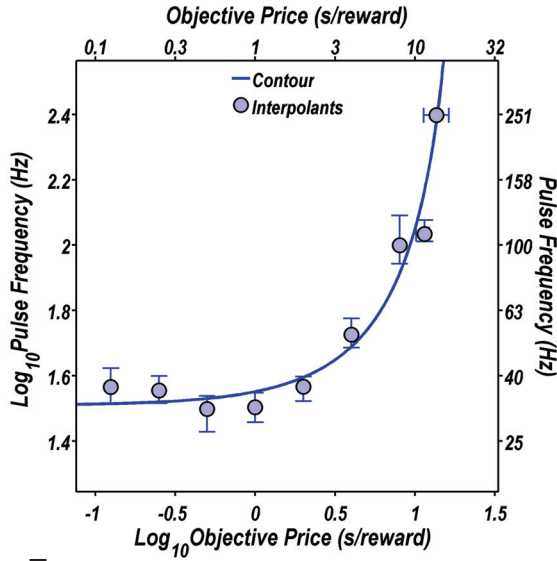


C

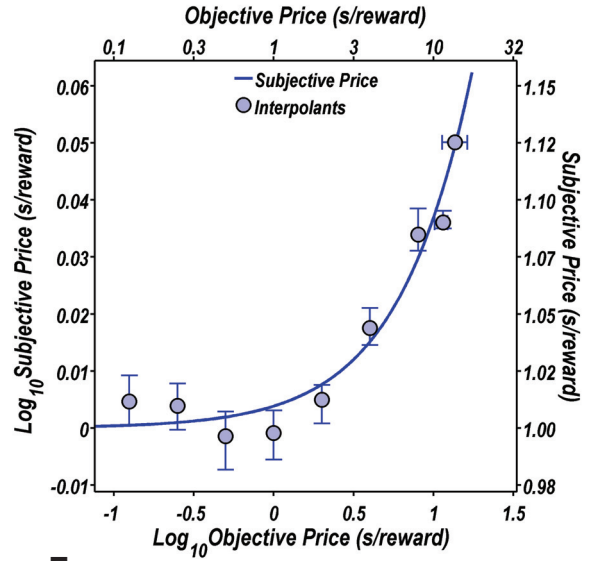


D

### The (inverted) hyperbolic subjective price function

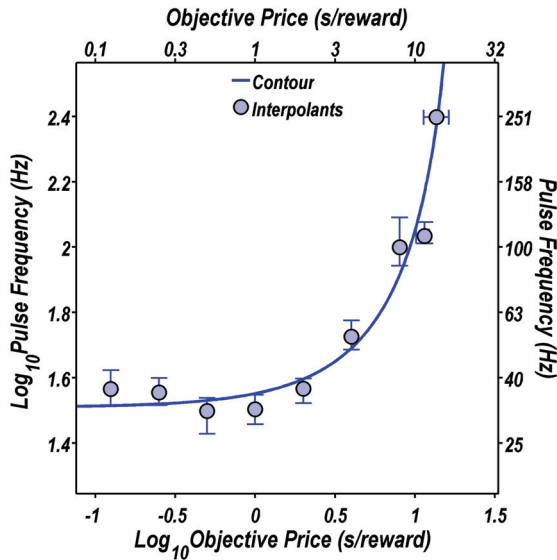


E

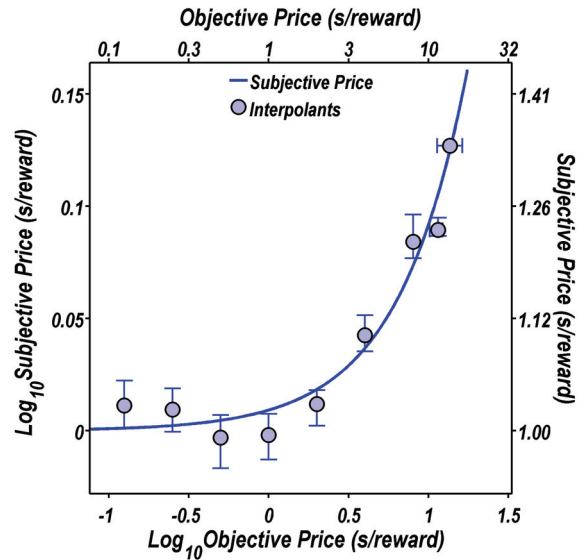


F

### The (inverted) exponential subjective price function



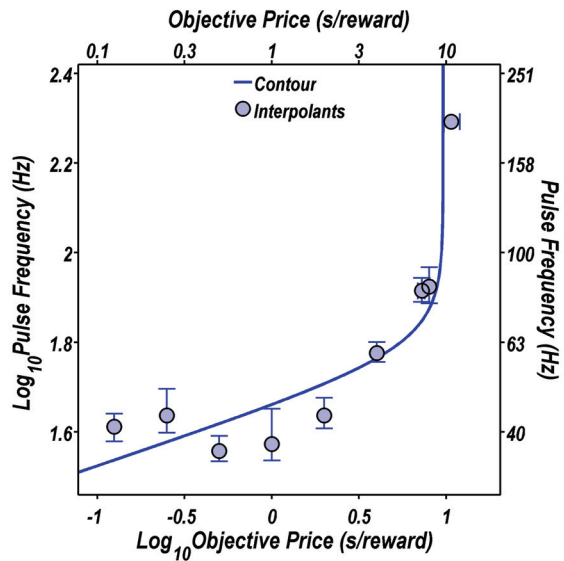
G



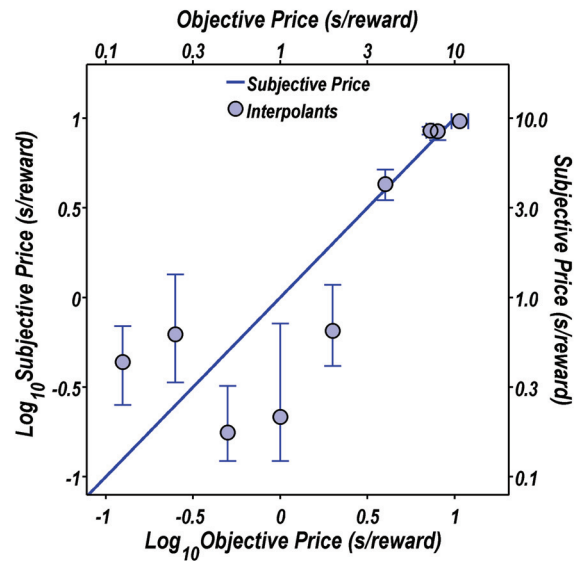
H

Figure ApD.1. Rat F3. Similar caption as for Figure 15.

## Objective price function

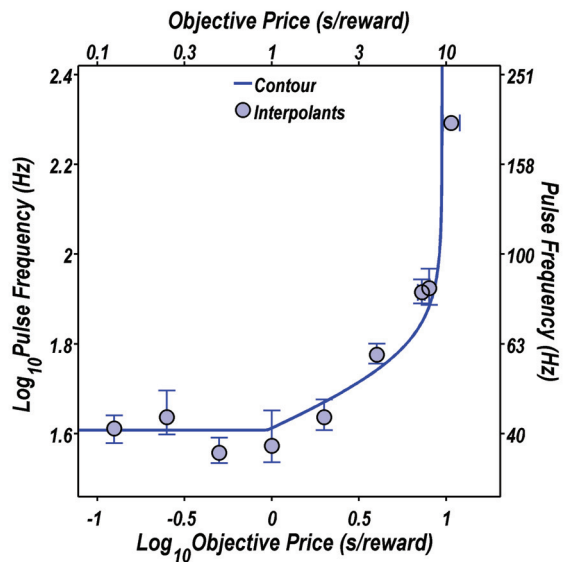


A

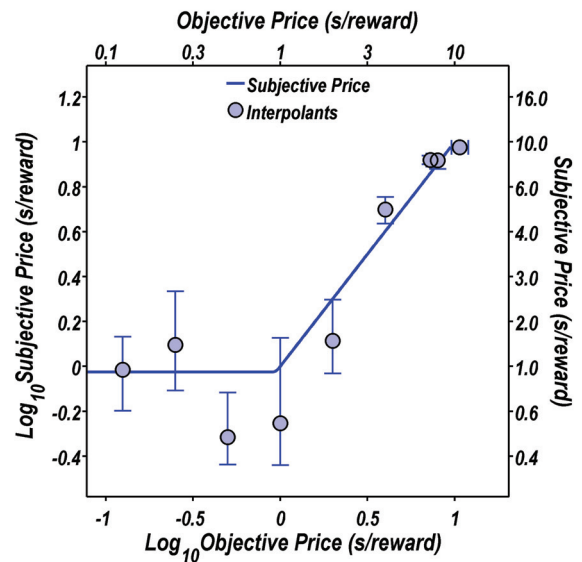


B

## Shizgal's subjective price function

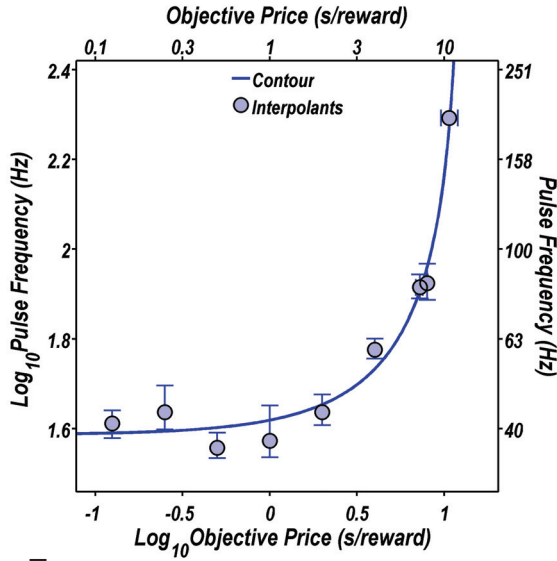


C

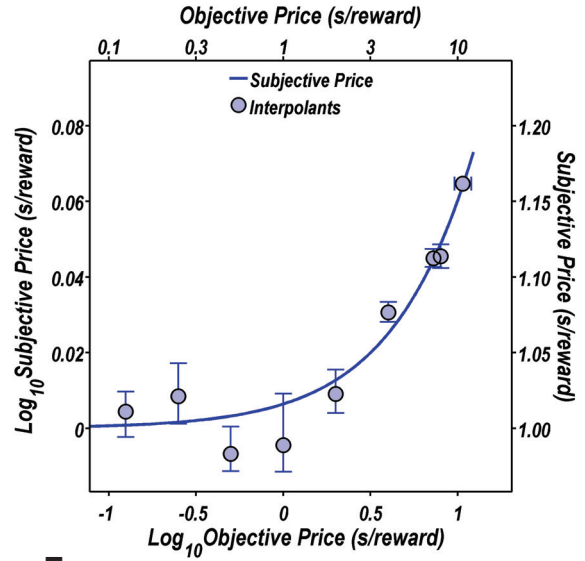


D

### The (inverted) hyperbolic subjective price function

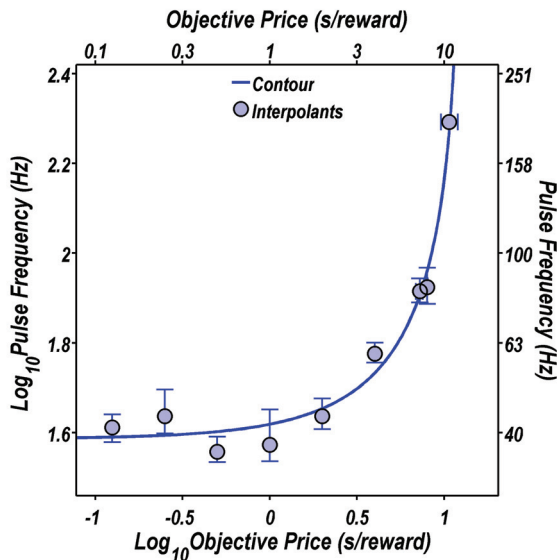


E

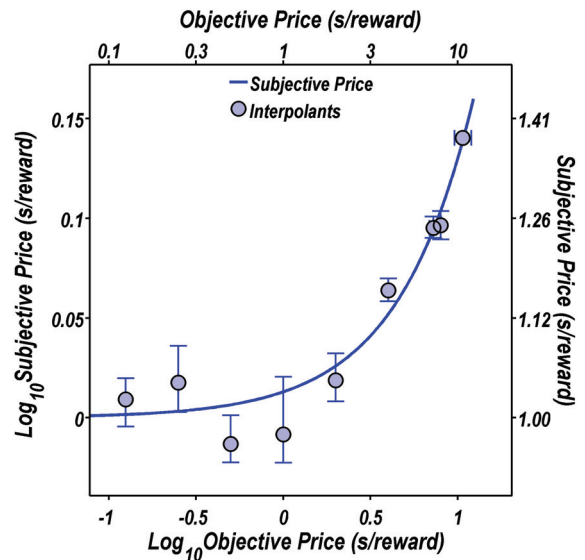


F

### The (inverted) exponential subjective price function



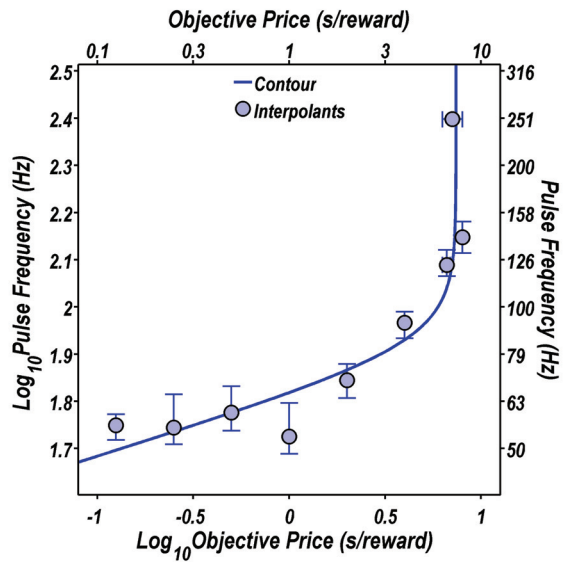
G



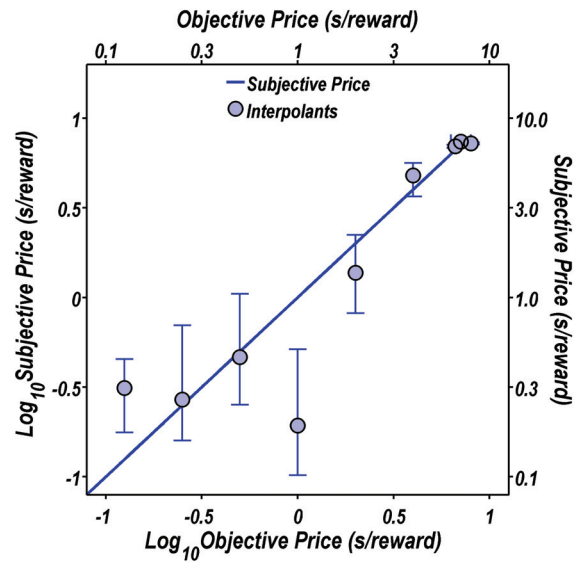
H

Figure ApD.2 Rat F9. Similar caption as for Figure 15.

## Objective price function

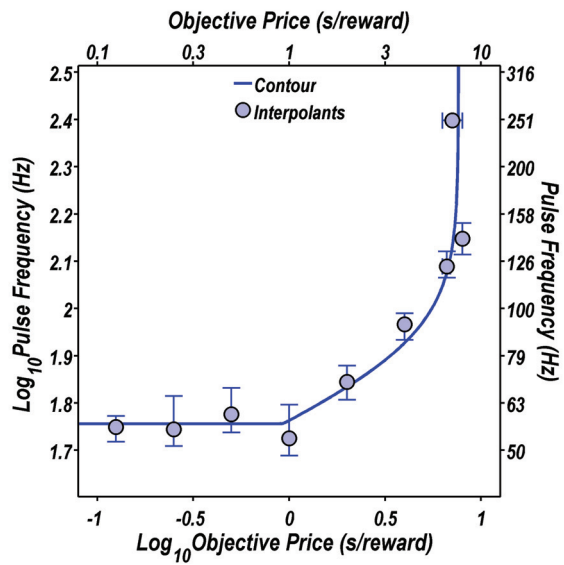


A

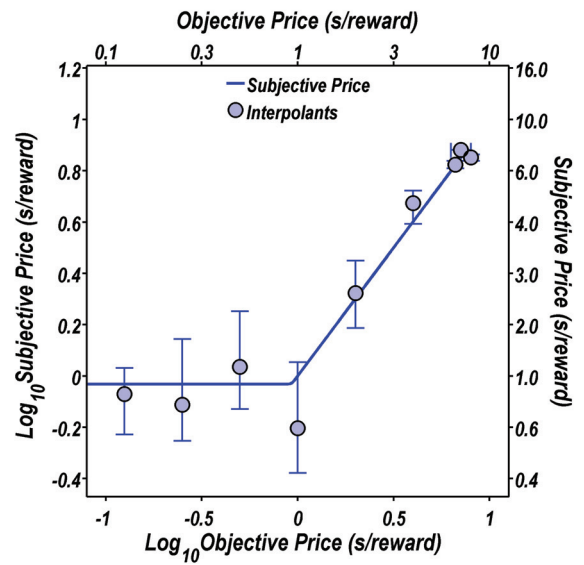


B

## Shizgal's subjective price function

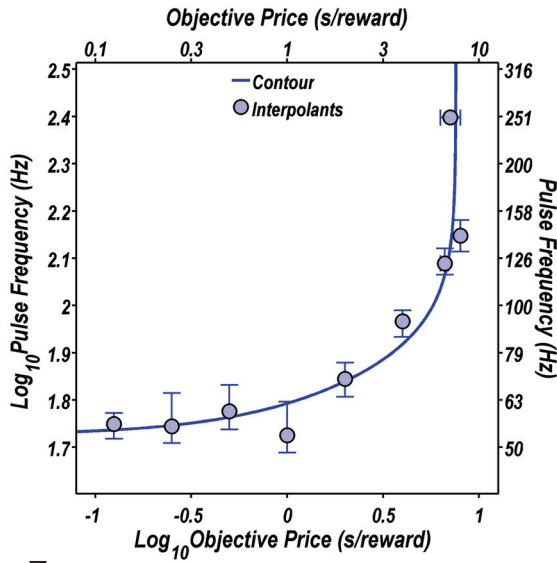


C

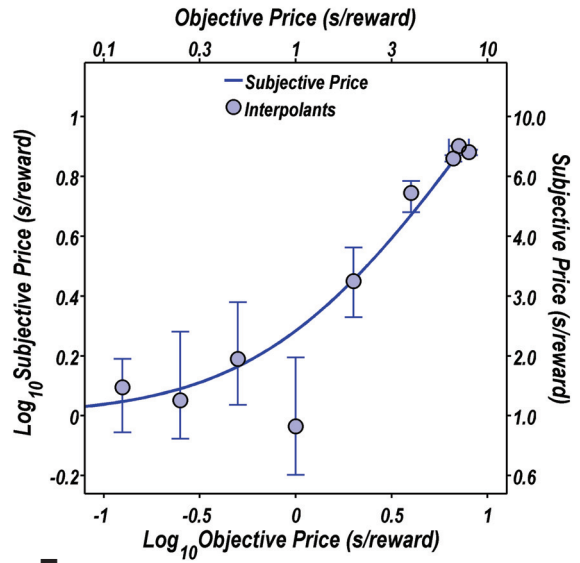


D

### The (inverted) hyperbolic subjective price function

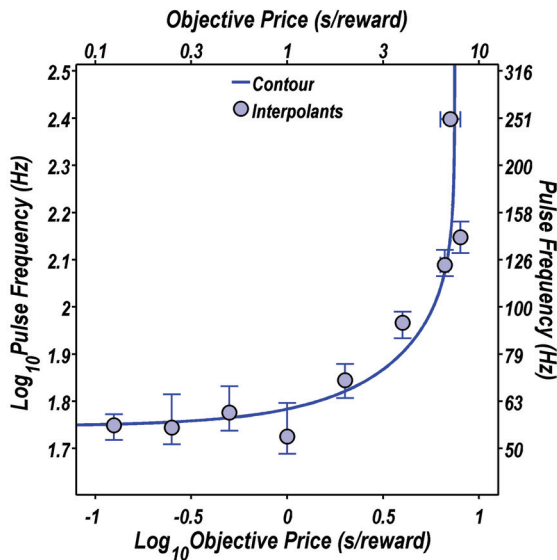


E

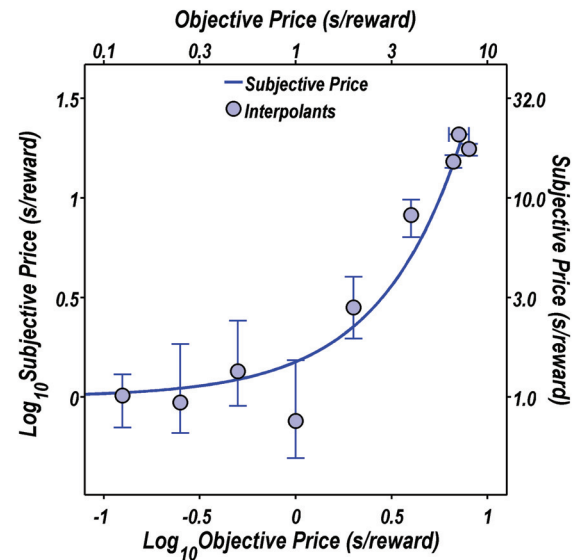


F

### The (inverted) exponential subjective price function



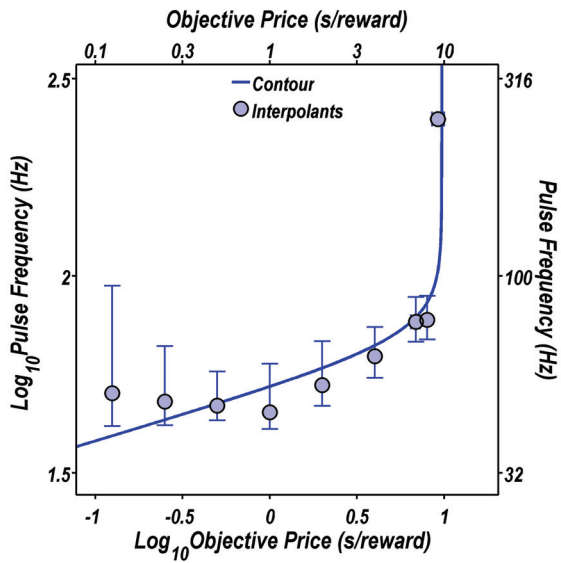
G



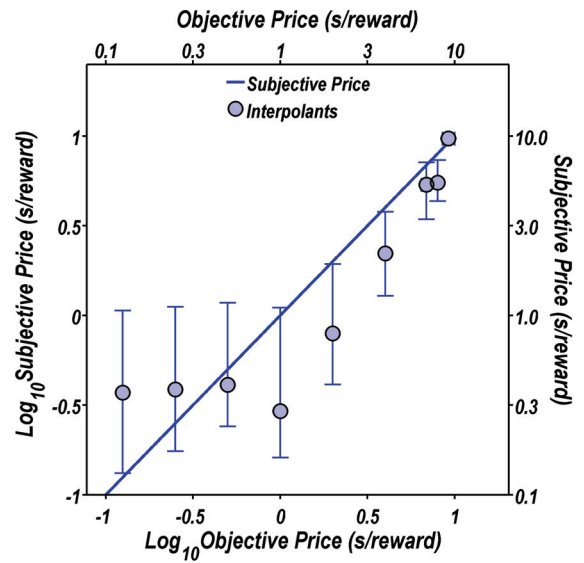
H

Figure ApD.3. Rat F12. Similar caption as for Figure 15.

### Objective price function

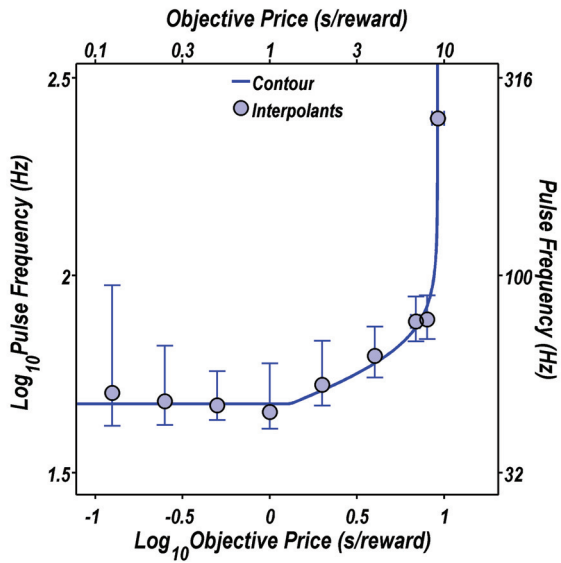


A

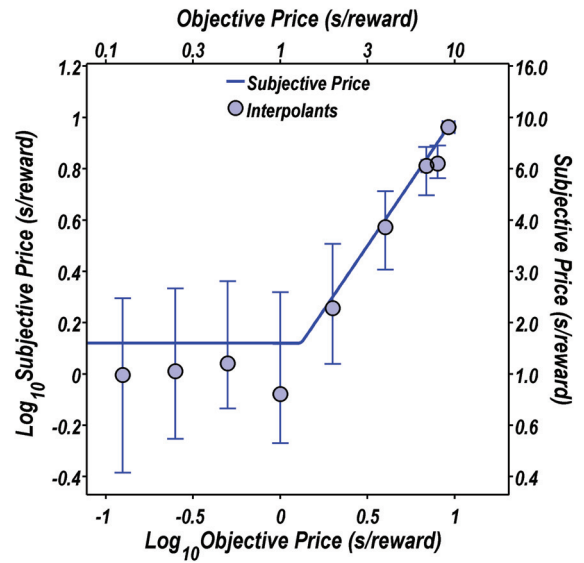


B

### Shizgal's subjective price function

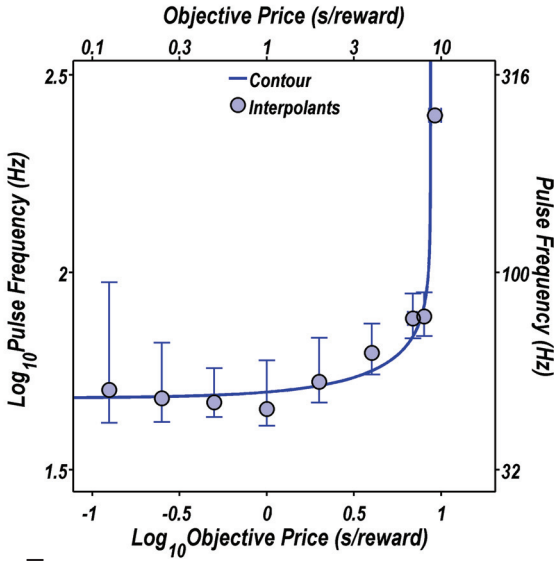


C

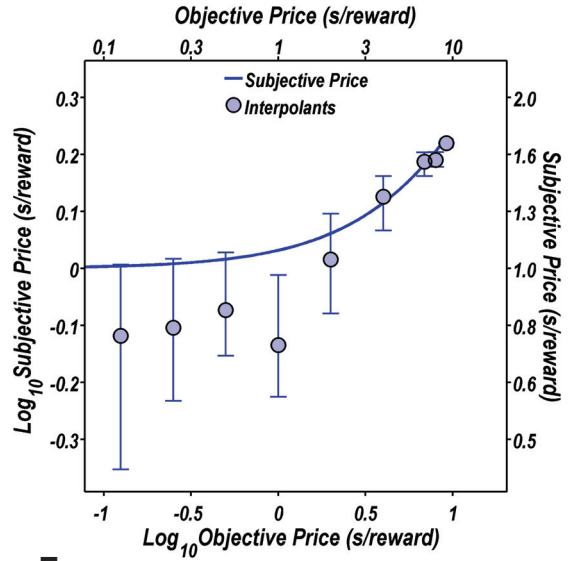


D

The (inverted) hyperbolic subjective price function

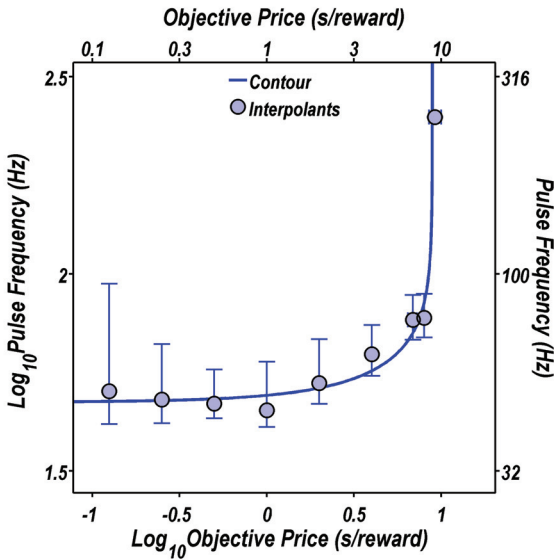


E

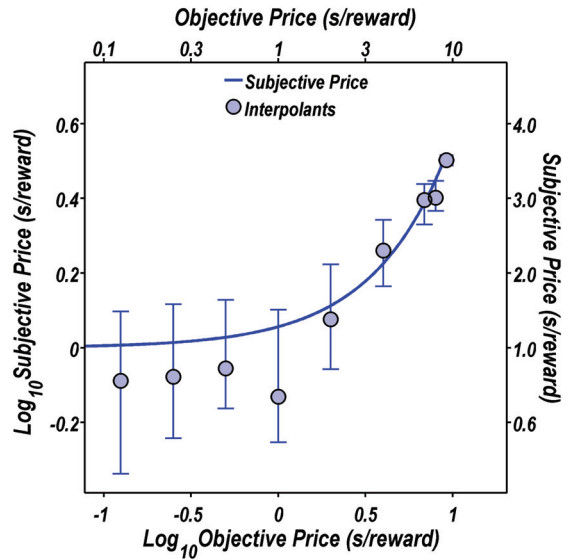


F

The (inverted) exponential subjective price function



G



H

Figure ApD.4. Rat F17. Similar caption as for Figure 15.



**Appendix E: The derivation of the expression for  $SP$  as a function  $F$  at half-maximal time allocation**

First, the general form of the contour line,  $SP_{TAmid}$  as a function of  $RI_{rel}$  will be derived. Next,  $SP_{TAmid}$  will be expressed as a function of  $F$  and parameter  $g$ . This expression is needed to transform the points of the [ $\text{Log}_{10}(F)$  vs.  $\text{Log}_{10}(OP)$ ] space into points of the [ $\text{Log}_{10}(SP)$  vs.  $\text{Log}_{10}(OP)$ ] space (Figure 15 & 16).

First, the goal is to isolate  $SP$  in the general form of the reward mountain equation:

$$TA = TA_{\min} + (TA_{\max} - TA_{\min}) \times \frac{(RI_{rel})^a}{(RI_{rel})^a + \left(\frac{SP}{SP_e}\right)^a}$$

$$\frac{TA - TA_{\min}}{TA_{\max} - TA_{\min}} = \frac{(RI_{rel})^a}{(RI_{rel})^a + \left(\frac{SP}{SP_e}\right)^a}$$

$$\frac{TA - TA_{\min}}{TA_{\max} - TA_{\min}} \times \left( (RI_{rel})^a + \left(\frac{SP}{SP_e}\right)^a \right) = (RI_{rel})^a$$

$$RI_{rel}^a = \frac{TA - TA_{\min}}{TA_{\max} - TA_{\min}} \times (RI_{rel})^a + \frac{TA - TA_{\min}}{TA_{\max} - TA_{\min}} \times \left(\frac{SP}{SP_e}\right)^a$$

$$RI_{rel}^a \times \left( 1 - \frac{TA - TA_{min}}{TA_{max} - TA_{min}} \right) = \frac{TA - TA_{min}}{TA_{max} - TA_{min}} \times \left( \frac{SP}{SP_e} \right)^a$$

$$RI_{rel}^a \times \left( \frac{TA_{max} - TA}{TA_{max} - TA_{min}} \right) = \frac{TA - TA_{min}}{TA_{max} - TA_{min}} \times \left( \frac{SP}{SP_e} \right)^a$$

$$RI_{rel}^a = \frac{\frac{TA - TA_{min}}{TA_{max} - TA_{min}}}{\frac{TA_{max} - TA}{TA_{max} - TA_{min}}} \times \left( \frac{SP}{SP_e} \right)^a$$

$$RI_{rel}^a = \frac{TA - TA_{min}}{TA_{max} - TA} \times \left( \frac{SP}{SP_e} \right)^a$$

$$RI_{rel} = \left( \frac{SP}{SP_e} \right) \times \left( \frac{TA - TA_{min}}{TA_{max} - TA} \right)^{\frac{1}{a}}$$

Solving for  $SP$ :

$$SP = RI_{rel} \times SP_e \times \left( \frac{TA_{max} - TA}{TA - TA_{min}} \right)^{\frac{1}{a}}$$

When  $TA$  falls halfway between  $TA_{max}$  and  $TA_{min}$ :

$$\frac{TA_{\max} - TA}{TA - TA_{\min}} = 1$$

When 1 is raised to any exponent the result is 1,

$$(1)^{\frac{1}{a}} = 1$$

Thus at half-maximal time allocation:

$$SP_{TAmid} = RI_{rel} \times SP_e$$

Expanding,

$$RI_{rel} = \frac{F^g}{F^g + F_{hm}^g}$$

Substituting for  $RI_{rel}$  in  $SP_{TAmid} = RI_{rel} \times SP_e$ :

$$SP_{TAmid} = \frac{F^g}{F^g + F_{hm}^g} \times SP_e$$

Thus, to transform the points plotted on the contour plot ( $F$  as a function of  $OP$ ) the above formula was employed. Parameters  $F_{hm}$ ,  $g$ , and  $SP_e$  are obtained from the respective pulse-frequency trade-off functions.

**Appendix F: AIC tables comparing the subjective price function and objective price function for each rat**

**Table ApF.1. Rat F3**

Model Name	AIC value	dAIC	Likelihood	AIC weight	Evidence Ratio	N	N params	RSS
SP	-437.0955	0	1	0.99629	1	99	5	1.07518
OP	-425.9095	11.186	0.0037237	0.00371	268.5508	99	3	1.25841

**Table ApF.2 Rat F9**

Model Name	AIC value	dAIC	Likelihood	AIC weight	Evidence Ratio	N	N params	RSS
SP	-464.8207	0	1	0.999999	1	90	5	0.45662
OP	-437.2582	27.563	1.03E-06	1.03E-06	966334.2	90	3	0.65156

**Table ApF.3. Rat F12**

Model Name	AIC value	dAIC	Likelihood	AIC weight	Evidence Ratio	N	N params	RSS
SP	-471.9116	0	1	0.996166	1	90	5	0.42202
OP	-460.7914	11.12	0.00385	0.003834	259.8488	90	3	0.50165

**Table ApF.4. Rat F16**

Model Name	AIC value	dAIC	Likelihood	AIC weight	Evidence Ratio	N	N params	RSS
SP	-346.1162	0	1	0.999999	1	72	5	0.5056
OP	-317.5419	28.574	6.24E-07	6.24E-07	1602635	72	3	0.80102

**Table ApF.5. Rat F17**

Model Name	AIC value	dAIC	Likelihood	AIC weight	Evidence Ratio	N	N params	RSS
SP	-343.5363	0	1	0.994331	1	81	5	1.02013
OP	-333.2022	10.334	0.00570122	0.005669	175.401	81	3	1.22497

**Table ApF.6. Rat F18**

Model Name	AIC value	dAIC	Likelihood	AIC weight	Evidence Ratio	N	N params	RSS
SP	-500.5989	0	1	1	1	90	5	0.30683
OP	-458.8359	41.763	8.54E-10	8.54E-10	1.17E+09	90	3	0.51266

Tables ApF.1-6. The AIC values and associated statistics comparing only Shizgal's subjective price model and the objective price model for all rats.

## *Chapter 3*

### *Psychophysical inference of frequency-following fidelity in the neural substrate for brain stimulation reward*

#### **Abstract**

Brain stimulation reward (BSR) has been studied extensively for over 50 years, yet the identity of the neurons directly responsible for this phenomenon is still unknown. Many psychophysical and electrophysiological experiments have been conducted to characterize these neurons. For example the properties of the directly-stimulated (“first-stage”) substrate for self-stimulation of the medial forebrain bundle (MFB) are consistent with those of fine myelinated axons (Bielajew & Shizgal, 1982; Murray & Shizgal, 1996; Shizgal et al. 1980), at least some of which project rostro-caudally (Bielajew & Shizgal, 1986). The present psychophysical experiment estimates an additional property of the first-stage neurons: their maximum firing frequency. When the electrical stimulation was delivered at low pulse frequencies, the first-stage neurons fired once per pulse. As the pulse frequency was increased, the probability of firing in response to each pulse rolled off, and the induced firing frequency reached a plateau of a median of 363 Hz, across 7 rats.

The frequency-following function was derived under the assumption of the “counter model of reward integration,” which stipulates that the rewarding effect of a pulse train of a given duration is determined by the aggregate number of spikes it triggers in the first-stage neurons. This spike count is the product of the number of first-stage

neurons activated (as set by the stimulation current and pulse duration) and the mean rate at which they fire (as set by the pulse frequency) (Gallistel, 1978; Gallistel, Shizgal, & Yeomans, 1981; Simmons & Gallistel, 1994). Within a given time period, spike count can be produced by firing many neurons at a low rate or fewer neurons at a higher rate. Thus, in the computation of the rewarding effect, the current and pulse frequency are reciprocally related and are said to “trade-off.” In self-stimulating rats, we measured current-pulse frequency trade-off functions that were embedded in a 4-dimensional reward model. The stimulation current required to maintain a given level of operant responding at each of a series of pulse frequencies was determined. In accordance with the counter model, as the pulse frequency was increased from low values, the current required to maintain a given level of behavior steadily declined. However, this relationship broke down at higher pulse frequencies, and the required current leveled off. The pulse frequency beyond which further increases in frequencies were ineffective in reducing the required current was inferred as the maximal firing frequency of the first-stage neurons. That the estimated maximum firing frequencies were very high is consistent with the relatively high estimated conduction velocities and rapid onset of recovery from refractoriness in first-stage neurons subserving MFB self-stimulation.



## ***Introduction***

The previous chapter tested an unrealistic assumption that was formerly incorporated in the reward mountain model, the assumption that the rat interpreted the objective-subjective opportunity relationship in a scalar manner. We demonstrated that this scalar relationship did not hold at low opportunity costs. Similarly, the experiment in Chapter 3 investigated another unrealistic assumption that has traditionally been incorporated in BSR experiments; this assumption concerns the firing capabilities of the directly stimulated neurons. Specifically, implicit in models of BSR is perfect frequency following: that each directly stimulated neuron fires once per pulse, regardless of the experimenter-set pulse frequency. However, it is more plausible to assume that the one-to-one mapping breaks down as the pulse frequency becomes high. There must be a limit to the firing rate of any axon due to physiological properties such as synaptic blocking or fatigue as suggested by Gallistel (1978). The goal of the present experiment was to model the progressive failure of frequency following as the pulse frequency is increased and determine the maximal firing frequency for the reward-relevant neurons of the lateral hypothalamus. An extension of the 3-dimensional measurement strategy of the reward mountain model, a 4-dimensional measurement approach, was used.

### ***3.1. The proposed frequency-response function***

The pulse frequency ( $F$ ) of a stimulation train set by the experimenter is the “inducer.” The actual average firing rate of the directly stimulated neurons, termed the firing frequency ( $FF$ ), is the “induced.” Forgie and Shizgal (1993) developed a model describing the relationship between the pulse frequency and firing frequency based on

behavioural trade-off data. The model posits that  $FF$  follows  $F$  perfectly at low  $F$  values (the slope of the function is 1). However, frequency following fidelity breaks down (rolls off) past a certain high  $F$  value (at this  $F$  value, the slope of the function declines, represented by the bend portion of the function in Figure 1). As  $F$  continues to be increased, a maximum  $FF$  is eventually reached and maintained. (The slope approaches 0, represented by the horizontal portion of the function.) In other words, increasing  $F$  above a limit has no effect on firing rate. The frequency following function is expressed as:

$$FF = F_b \left( \ln\left(1 + e^{\frac{F_{nmax}}{F_b}}\right) - \ln\left(1 + e^{\frac{F_{nmax}-F}{F_b}}\right) \right)$$

where,

$FF$  = the average firing frequency (Hz); the induced physiological response

$F$  = the pulse frequency that is experimenter-set (Hz); the inducing stimulus value

$F_{nmax}$  = the pulse frequency at which the firing frequency is *near* the maximal firing frequency that the substrate can attain; the position parameter of the frequency-response function; a demonstration that this value is *near* the maximal firing frequency is described in Appendix B

$F_b$  = the parameter describing the abruptness of the transition between the range of perfect frequency following to the range of frequency roll-off

The frequency-response function was developed by computing the integral of a sigmoidal function that relates the rate at which the induced firing frequency changes (the firing frequency derivative) to pulse frequency. The sigmoidal expression and

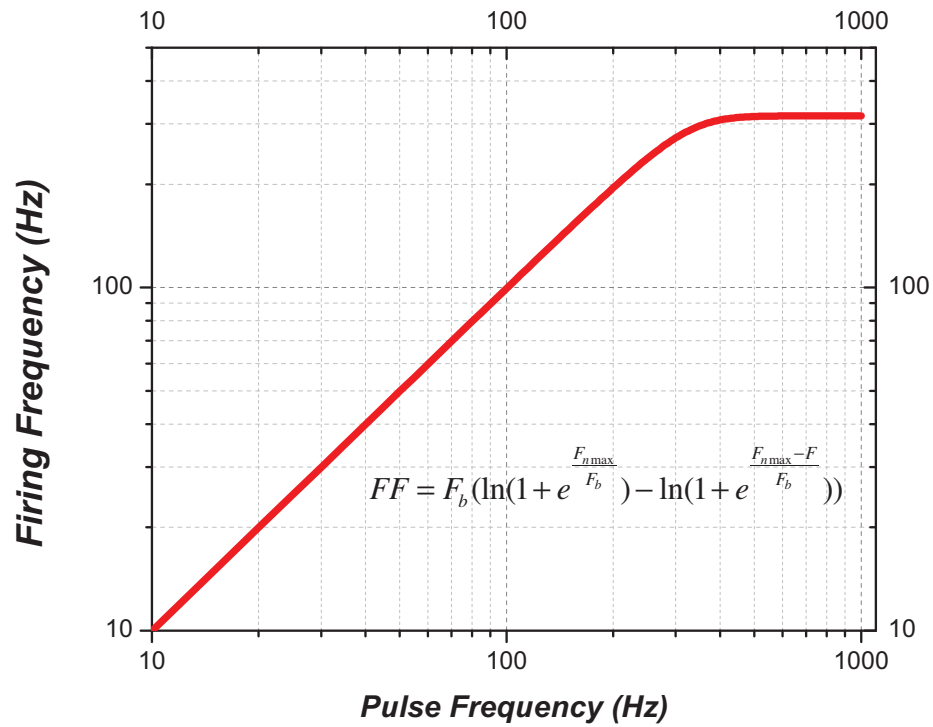


Figure 1. The firing frequency as a function of the (experimenter-set) pulse frequency in common logarithmic space. The firing frequency follows the experimenter-induced pulse frequency at low pulse frequencies but declines smoothly at high frequencies such that a maximum firing frequency is maintained.

corresponding figure are described in Appendix A and B.

### ***3.2. The counter model***

The frequency following experiment described in this chapter is based on the counter model of reward integration. The counter model describes the spatial-temporal integration of the output from the directly stimulated reward-relevant neurons (Gallistel et al., 1981). The post-synaptic substrate, referred to as the integrator, processes the input from the directly stimulated neurons within a given time period. The integrator's output is a function of the overall spike rate and determines the subjective reward intensity that results. Thus,

$$RI = f(S)$$

where,

$RI$  = reward intensity

$S$  = aggregate spike rate

The aggregate spike rate is the product of the rate of firing ( $FF$ ) and the number of neurons ( $Nn$ ) that are fired.

$$S = FF \times Nn$$

For example, according to this model, in a time window of 1 s, it does not matter whether 100 neurons fire 10 times each (aggregate rate = 1000 spikes/1s) or whether 20 neurons fire 50 times each (aggregate rate = 1000 spikes/s). The reward intensity will be equivalent.

The firing frequency ( $FF$ ) is controlled by the experimenter-set pulse frequency ( $F$ ). With pulse duration held constant, the total number of stimulated neurons ( $Nn$ ) is set by the current intensity and is believed to be roughly proportional to the effective current intensity; this relationship is developed in detail in Appendix C.

Note that the number of stimulated neurons is proportional to the *effective* current, which is distinct from the *applied* current. Briefly, the applied current intensity ( $I_P$ ) is proportional to the cross-section area of the stimulated region: the higher the current, the greater the cross-section area of the stimulated region (see Appendix C). However, there is scar tissue that surrounds the electrode. Thus, a given amount of current (a very low intensity) is considered to be “ineffective” (or “waste”) current, denoted as  $I_0$ . Specifically,  $I_0$  is the current that just suffices to fire a hypothetical neuron located at the edge of the scar tissue surrounding the electrode tip. The effective current is the difference between the applied and ineffective current ( $I_P - I_0$ ). It is the effective current that is proportional to the total number of stimulated neurons:

$$Nn = K \times (I_P - I_0)$$

where,

$Nn$  = the total number of directly stimulated nerve fibers

$K$  = density constant; the number of neurons recruited per  $\mu\text{A}$  (neurons/ $\mu\text{A}$ )

$I_p$  = the applied current intensity at a given pulse duration ( $\mu\text{A}$ )

$I_0$  = the “ineffective” or “waste” current; the current that just suffices to fire a hypothetical neuron located at the edge of the scar tissue surrounding the electrode tip ( $\mu\text{A}$ )

Returning to the expression for the aggregate spike rate ( $S$ ):

$$S = FF \times Nn$$

We can substitute the experimenter-set parameters for the “induced” parameters. The effective current ( $I_p - I_0$ ) multiplied by  $K$  is substituted for  $Nn$ , and  $F$  is substituted for  $FF$ :

$$S = F \times K \times (I_p - I_0)$$

Returning to the counter model: reward intensity is a function of the aggregate spike rate arriving at postsynaptic terminals.

$$RI = f(S)$$

Expanding the aggregate spike rate in terms of experimenter-set parameters as described above:

$$RI = f(F \times K \times (I_P - I_0))$$

$K$  and  $I_0$  are constant no matter what pulse frequency ( $F$ ) or applied current intensity ( $I_P$ ) are tested.

For now, we'll consider the effective current  $I_P - I_0$  as roughly equivalent to  $I_P$  because  $I_0$  is very small:

$$(I_P - I_0) \approx I_P$$

So,

$$RI \approx f(K \times F \times I_P)$$

According to the above formulation, pulse frequency ( $F$ ) and current intensity ( $I_P$ ) are related in a reciprocal manner. Doubling the pulse frequency  $F$  (within the firing frequency fidelity range) has the same effect on reward intensity ( $RI$ ) as doubling the current intensity ( $I_P$ ). Similarly, reducing the pulse frequency ( $F$ ) by half (within the firing frequency fidelity range, the range where  $FF$  is equal to  $F$ ) has the same effect on the reward intensity as reducing the current by a half. It is this current-pulse frequency trade-off relationship, demonstrated in numerous studies, that has led to the counter model. The example below describes the conventional approach used to demonstrate the

current-pulse frequency reciprocal relationship and the firing frequency limit.

Furthermore, it provides the basis for the experiment presented in this chapter.

### ***3.3. Illustrative example of the current-pulse frequency trade-off relationship***

Figure 2 is an example of what the counter model predicts as demonstrated by current-pulse frequency trade-off methodology. The dependent measure is time allocated to holding down the lever. Rats are presented with a series of current sweeps, each conducted at a different pulse frequency. A sweep is a series of trials over which the current is decremented systematically from one trial to the next. Figure 3A is a family of such current sweep curves. Time allocation as a function of current is plotted; each current sweep is denoted by a different colour. The current required to sustain half-maximal time allocation (the behavioural criterion) is denoted by the vertical dashed line for every curve. As the pulse frequencies of the current sweeps are doubled, the curves shift leftward. This is consistent with the counter model: at a higher pulse frequency, fewer neurons and therefore a lower current is required to achieve the behavioural criterion. In the range of perfect frequency following, doubling the pulse frequency results in a halving of the required current: the current sweep curves shift leftward by the same magnitude (one half) at each doubling of pulse frequency. It is this finding, demonstrated in numerous studies, that has led to the counter model.

Next, as the pulse frequency approaches the maximal firing frequency of the substrate, each doubling produces a progressively smaller effect on the current required to reach the behavioural criterion. The shifts of the current sweep curves become progressively smaller; at high pulse frequency values, they eventually overlap. For



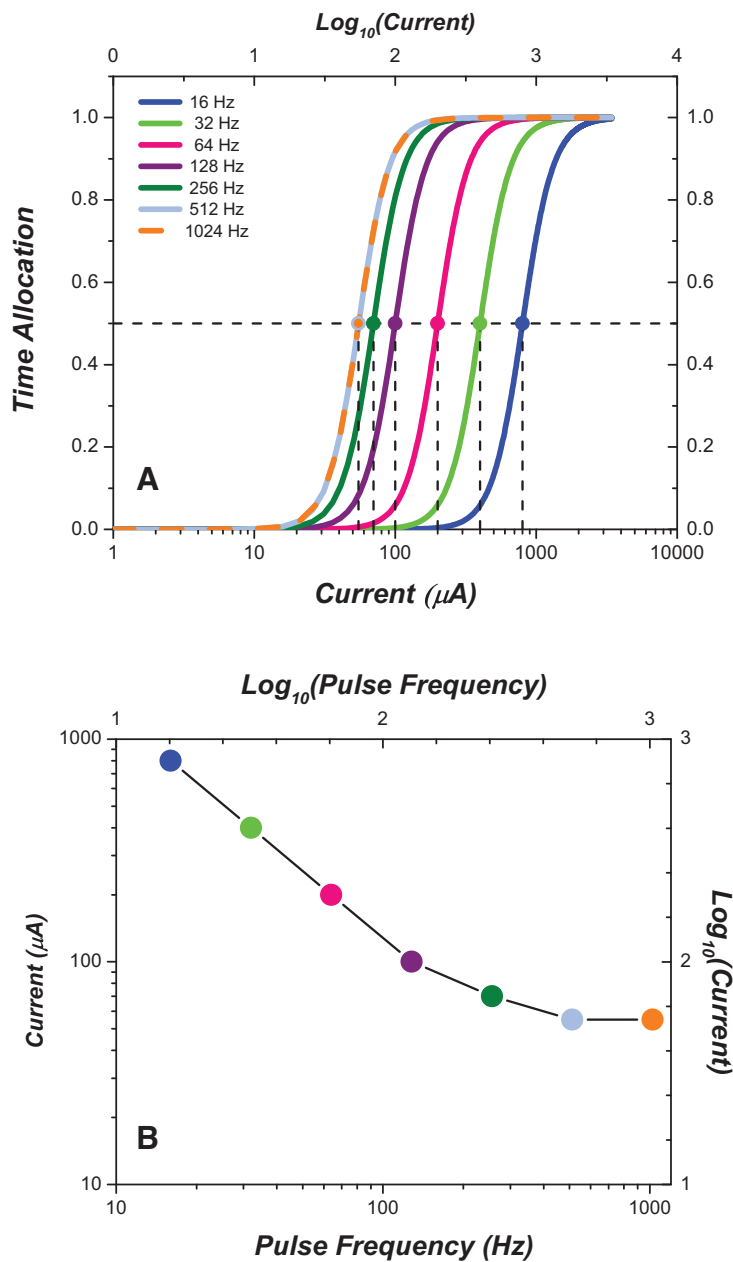


Figure 2. A. Current sweeps (time allocation as a function of current) at various frequencies (spaced by equal multiplicative increments) are plotted in different colours. As the pulse frequency is increased, the curves shift leftward by equal increments at the lower pulse frequency range. At the higher pulse frequency range, this increment becomes smaller and eventually the curves are superimposed. B. Each data point represents the required current at half-maximal time allocation, for each sweep in panel A, denoted by the dotted vertical lines. As the pulse frequency is increased, the current decreases and eventually levels off. This leveling-off reflects the asymptote that is reached at the high pulse frequencies in Figure 1.

instance, at the highest pulse frequencies (500 Hz and 1000 Hz), the current sweep curves overlap, which indicates that the induced firing frequency is the same at these different pulse frequencies. In other words, no reduction in current is required to compensate for an increase in pulse frequency at this very high range because these high pulse frequencies induce equivalent firing frequencies. This breakdown of the current-pulse frequency scalar relationship has been demonstrated in several studies (as will be described below). However, pulse frequencies above 500 Hz have not been tested.

Figure 2B describes the following functional relationship: required current at the behavioural criterion as a function of pulse frequency. As pulse frequency is increased, lower currents are required to meet the behavioural criterion. At high pulse frequencies, the firing frequency response eventually rolls off and saturates as indicated by the leveling off of the trade-off function.

### ***3.4. Previous current-pulse frequency trade-off studies***

The counter model was initially demonstrated by a current-pulse frequency trade-off experiment in which the rat was required to run down an alley to a goal box in order to obtain rewarding electrical stimulation (Gallistel, 1978). For every doubling of pulse frequency, there was (an approximate) halving of required current. A range of pulse frequency-current combinations at pulse frequencies spanning from 12.5 – 50 Hz to 400 Hz spaced by a proportion of 2 was tested; a consistent reduction in required current was demonstrated. Train durations of 0.1 s, 1 s, and 10 s were tested. Pertinent to the present experiment, in this early study, Gallistel (1978) noted that there was a frequency roll-off,

from around 200 Hz to 400 Hz depending on the subject and the length of the train duration.

If in the computation of the rewarding effect, pulse frequency and current intensity are reciprocally related (as the counter model proposes), then on a common logarithmic scale, the plot of the required current as a function of pulse frequency will be a straight line with a slope of -1. Simmons and Gallistel (1994) demonstrated this reciprocal relationship for most rats (4 out of 6) at pulse frequencies ranging from about 10 Hz to 100 Hz - 177 Hz in a lever-pressing task. Furthermore, pulse frequencies above the upper limit to which the regression line was fit (177 Hz to about 350 Hz) were tested. Along this range, the current intensity required to sustain half-maximal responding was roughly equivalent, which implies that the firing frequency reaches a plateau. In a separate dual-operant procedure, the pulse frequency at which the subjective magnitude of the reward attains its maximum possible value was employed as the behavioural criterion. Again, they showed that for most rats (4 out of 6), the slopes were not significantly different than -1. However, the points that diverged from the slope of -1 were at pulse frequencies higher than 250 Hz.

Several other studies are in support of the counter model. For instance, Gallistel, Leon, Waracynski, and Hanau (1991) used current-pulse frequency trade-off methodology: the counter model was supported in the tested frequency range, 13 Hz to 200 Hz. Similarly, in a current-pulse frequency experiment, Forgie and Shizgal (1991) tested pulse frequencies up to 495 Hz; they demonstrated that the counter model held at pulse frequencies up to 250 Hz. Gallistel and Leon (1991) tested pulse frequencies up to 200 Hz in a dual-operant procedure and also demonstrated that the counter model holds:

increasing the current by a given proportion had the same effect on reward intensity as increasing the pulse frequency by that same proportion. This equal effect on reward intensity implies a scalar relationship between current and pulse frequency. Thus, taken together, it is evident that there is substantial support for the counter model.

Frequency-following fidelity can be inferred from the current-pulse frequency relationship. The pulse frequency range at which current-pulse frequency reciprocity holds denotes the range at which frequency following is perfect. The pulse frequency range at which the reciprocal relationship no longer holds implies frequency following roll-off. Although numerous current-pulse frequency trade-off studies have demonstrated both the reciprocal relationship as well as its breakdown at high pulse frequencies, a formal description of frequency following is lacking. Firstly, most studies have not tested a high range of pulse frequencies; typical tests do not exceed  $\sim 350$  Hz. Furthermore, although a study by Forgie and Shizgal (1991) has formally described the shape of the frequency following function (as discussed in section 3.1), the proposed model was used for demonstrative purposes and was not actually fit to the data. Moreover, the electrodes were located in the ventral tegmental area.

The purpose of the experiment in this chapter was to obtain improved estimates of frequency following in the neural substrate of self-stimulation of the lateral hypothalamic level of the medial forebrain bundle. Forgie and Shizgal's (1991) proposed frequency following function was embedded in the 4-dimensional reward model and fit to the data. The 4-dimensional reward model is described below.

### 3.5. The 4-dimensional reward model

The logistic reward growth equation at the core of the mountain model is written as:

$$RI_{rel} = \frac{F^g}{F^g + F_{hm}^g}$$

where,

$F_{hm}$  = the pulse frequency that produces a half-maximal reward intensity (Hz)

$RI_{rel}$  = the ratio of reward intensity ( $RI$ ) to maximum reward intensity ( $RI_{max}$ )

$g$  = the reward growth exponent

In the above expression, current intensity is held constant. Furthermore, this expression is based on the assumption of perfect frequency following whereby each pulse triggers an action potential in every stimulated neuron.

To generalize the above formulation in terms of the *total* induced spikes arriving at the integrator ( $S$ ):

$$RI_{rel} = \frac{S^g}{S^g + S_{hm}^g}$$

where,

$S$  = the total number of spikes per second produced by the electrical stimulation of the directly stimulated neurons

$S_{hm}$  = total number of spikes per second required to produce a reward of half-maximal reward intensity

We can expand  $S$  (total number of spikes per second):

$$S = FF \times K \times (I_p - I_0)$$

where,

$K$  = density constant; the number of neurons recruited per  $\mu\text{A}$  (neurons/ $\mu\text{A}$ )

$FF$  = the firing frequency of the directly stimulated substrate (Hz)

$I_p$  = the applied current at a given pulse duration ( $\mu\text{A}$ )

$I_0$  = the minimum effective current ( $\mu\text{A}$ )

Replacing this expression into the reward intensity formula:

$$RI_{rel} = \frac{(FF \times K \times (I_p - I_0))^g}{(FF \times K \times (I_p - I_0))^g + S_{hm}^g}$$

Dividing through by  $K$ ,

$$RI_{rel} = \frac{(FF \times (I_p - I_0))^g}{(FF \times (I_p - I_0))^g + \left(\frac{S_{hm}}{K}\right)^g}$$

Let,

$$S_{hmin} = \frac{S_{hm}}{K}$$

where,

$S_{hmin}$  = the ratio of the total number of induced spikes per second required to produce a reward with half-maximal reward intensity and the number of neurons recruited per  $\mu\text{A}$ ;  
(spikes  $\times$   $\mu\text{A}$ )/(seconds  $\times$  neurons)

Therefore,

$$RI_{rel} = \frac{(FF \times (I_p - I_0))^g}{(FF \times (I_p - I_0))^g + (S_{hmin})^g}$$

The reward mountain equation developed in the general introduction of this thesis is the ratio of the payoff from the rewarding electrical stimulation to the total sum of the payoffs (from the electrical stimulation and from alternate activities). Because the subjective price function has been developed in Chapter 2, the subjective price ( $SP$ ) can now be substituted for objective price in the reward mountain equation. Thus, time allocation ( $TA$ ) is expressed as:

$$TA = (TA_{\max} - TA_{\min}) \frac{(RI_{rel})^a}{(RI_{rel})^a + \left(\frac{SP}{SP_e}\right)^a} + TA_{\min}$$

where,

$SP$  = subjective price (s)

$SP_e$  = the subjective price at which the rat allocates half of its time to the pursuit of the rewarding electrical stimulation (s)

$RI_{rel}$  = the ratio of reward intensity ( $RI$ ) to maximum reward intensity ( $RI_{max}$ )

$a$  = the price sensitivity exponent

$TA_{min}$  = the minimum time allocation

$TA_{max}$  = the maximum time allocation

Replacing the  $RI_{rel}$  with the expanded version that incorporates firing frequency ( $FF$ ) and current ( $I_p, I_0$ ), the expression is:

$$TA = (TA_{\max} - TA_{\min}) \frac{\left(\frac{(FF \times (I_p - I_0))^g}{(FF \times (I_p - I_0))^g + S_{hmin}}\right)^a}{\left(\frac{(FF \times (I_p - I_0))^g}{(FF \times (I_p - I_0))^g + S_{hmin}}\right)^a + \left(\frac{SP}{SP_e}\right)^a} + TA_{\min}$$

where  $FF$  is the proposed firing frequency function,



$$FF = F_b \left( \ln \left( 1 + e^{\frac{F_{n\max}}{F_b}} \right) - \ln \left( 1 + e^{\frac{F_{n\max} - F}{F_b}} \right) \right)$$

and where  $SP$  is Shizgal's subjective price function,

$$SP = SP_{\min} + (SP_{\text{bnd}} \times \ln \left( 1 + e^{\frac{SP_{\min} - OP}{SP_{\text{bnd}}}} \right))$$

The full 4-dimensional reward model is expressed as:

$$TA = (TA_{\max} - TA_{\min}) \frac{\left( \frac{(F_b (\ln(1 + e^{\frac{F_{n\max}}{F_b}}) - \ln(1 + e^{\frac{F_{n\max} - F}{F_b}})) \times (I_p - I_0))^g}{(F_b (\ln(1 + e^{\frac{F_{n\max}}{F_b}}) - \ln(1 + e^{\frac{F_{n\max} - F}{F_b}})) \times (I_p - I_0))^g + S_{h\min}} \right)^a}{\left( \frac{(F_b (\ln(1 + e^{\frac{F_{n\max}}{F_b}}) - \ln(1 + e^{\frac{F_{n\max} - F}{F_b}})) \times (I_p - I_0))^g}{(F_b (\ln(1 + e^{\frac{F_{n\max}}{F_b}}) - \ln(1 + e^{\frac{F_{n\max} - F}{F_b}})) \times (I_p - I_0))^g + S_{h\min}} \right)^a} + \left( \frac{SP_{\min} - OP}{SP_e} \right)} \right) + TA_{\min}$$

The function described above is a 4-dimensional model because there are 3 independent variables. The first 2 were independent variables in the 3-dimensional version of the model:  $F$  (the pulse frequency in Hz), and  $OP$  (the objective price in seconds). The added independent variable in this version of the reward mountain is  $I_p$  (the stimulation current intensity in  $\mu\text{A}$ ). (In the 3-dimensional reward mountain model described in the introduction, the stimulation current was held constant.)

The parameters of the 4-dimensional reward mountain model are:

$a$ : the price-sensitivity exponent

$F_b$ : the parameter controlling the abruptness of the transition between the range of perfect frequency following to the range of frequency roll-off

$F_{max}$ : the pulse frequency that produces near maximal firing frequency (Hz)

$I_0$ : the ineffective current; the current that just suffices to recruit the neuron nearest to the border of the scar tissue surrounding the electrode tip ( $\mu\text{A}$ )

$SP_e$ : the subjective price at which the rat allocates half of its time to the pursuit of electrical stimulation (s)

$Sh_{min}$ : the ratio of the total number of spikes per second and the number of recruited neurons per  $\mu\text{A}$  required to produce a half-maximal reward ((spikes  $\times$   $\mu\text{A}$ )/(seconds  $\times$  neurons))

$TA_{max}$ : the maximal time allocation

$TA_{min}$ : the minimal time allocation

$SP_{bnd}$ : the parameter controlling the abruptness from the flat portion to the rising (scalar) portion of the subjective price function curve

$SP_{min}$ : the minimum subjective price (s)

One way to visually represent a 4-dimensional model is by two 3-dimensional plots. In the first 3-dimensional plot, time allocation is plotted as a function of pulse frequency and objective price, while current is held constant. This representation includes the pulse frequency and price sampling matrices. Figure 3A plots this representation and Figure 3B plots the corresponding contour plot. In the second representation, time allocation is plotted as a function of pulse frequency and current (while the price is held constant). Figure 4A plots the representation and Figure 4B plots the corresponding contour graph.

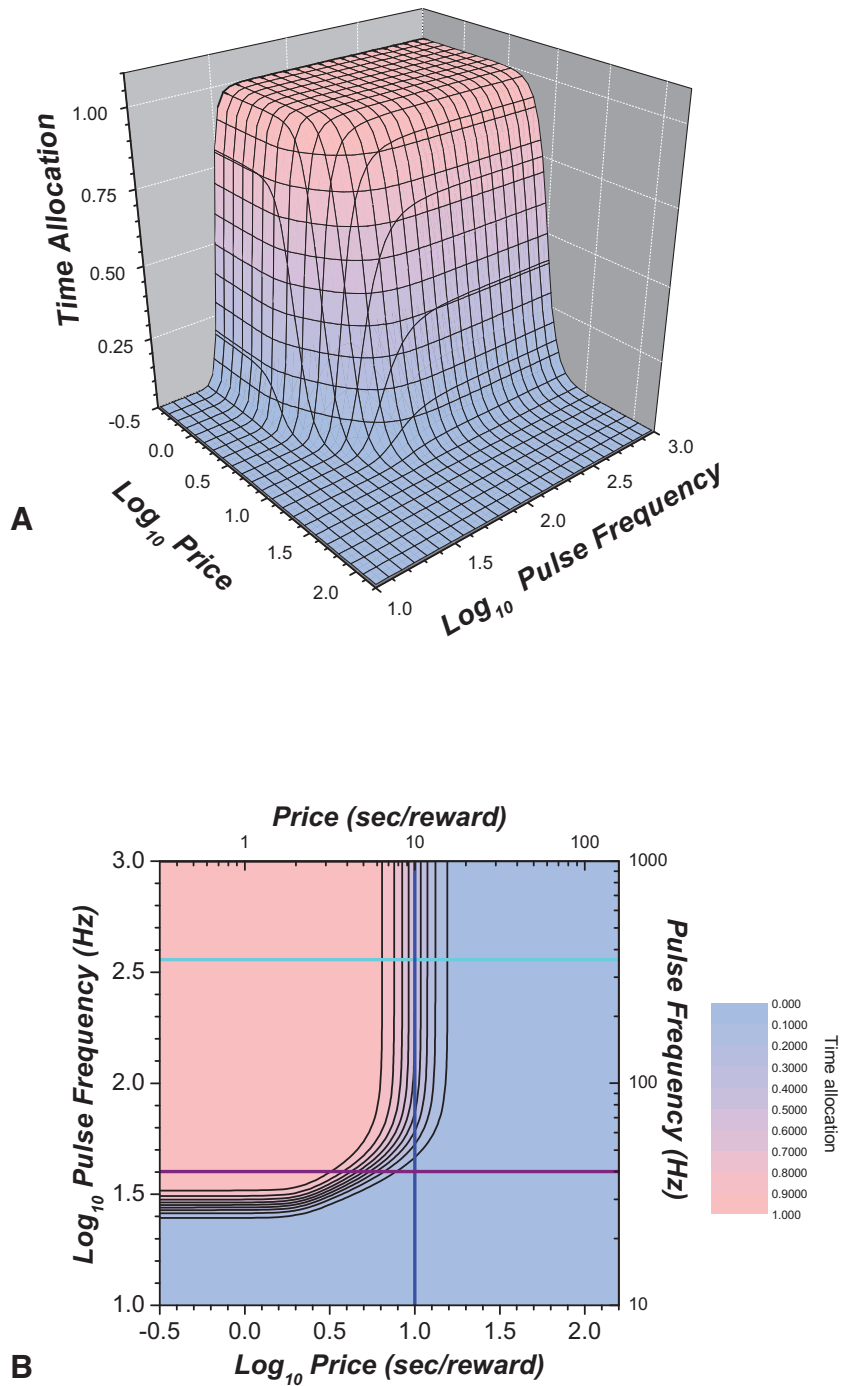


Figure 3. A. The 3-dimensional section (time allocation as a function of price and pulse frequency) of the 4-dimensional reward model. B. The corresponding contour plot. The blue vertical line represents the  $OPE$ ; the purple horizontal line represents  $F_{hm}$ , the cyan line represents  $F_{nmax}$ .

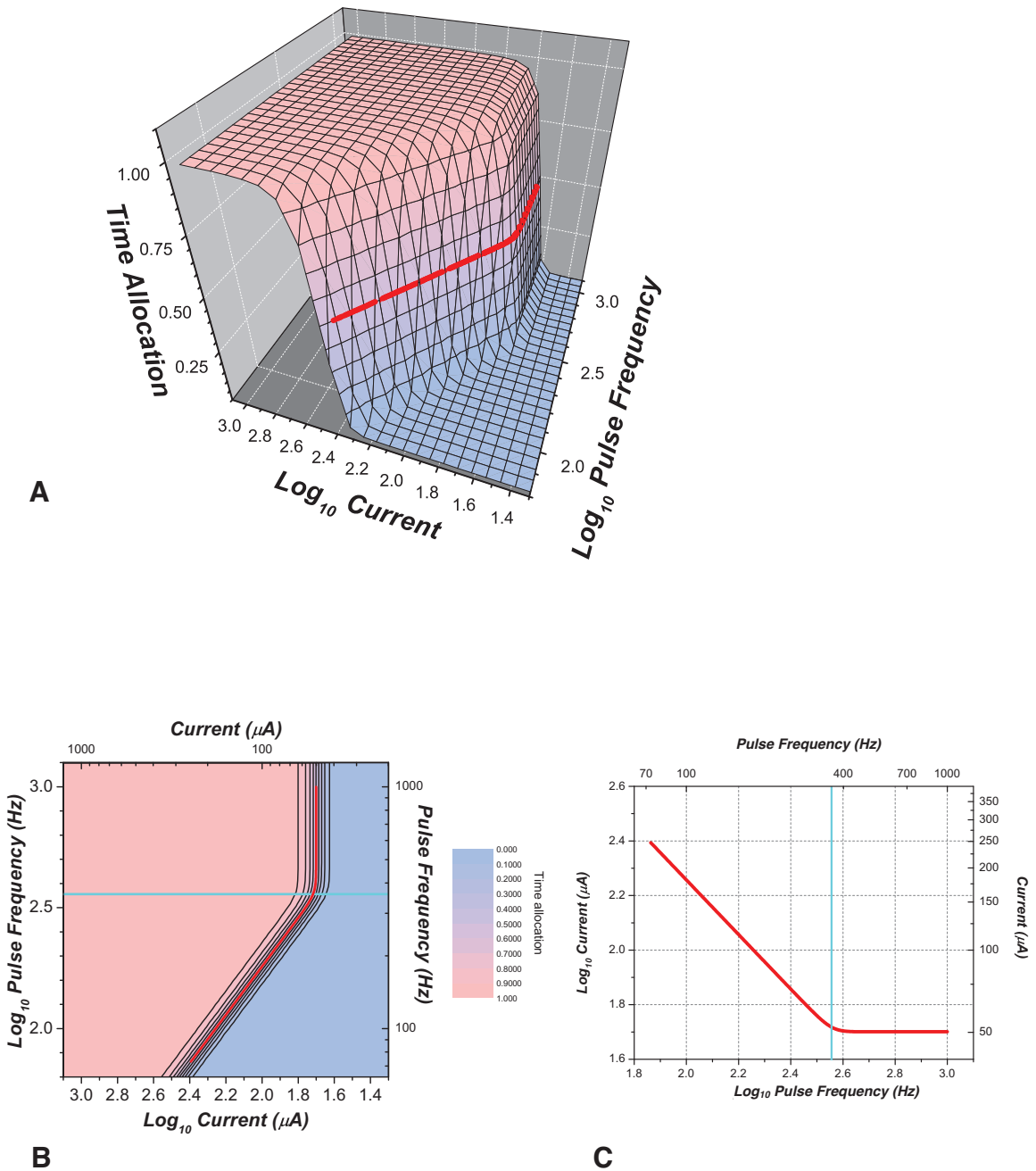


Figure 4. A. The 3-dimensional section (time allocation as a function of current and pulse frequency) of the 4-dimensional reward model. B. The corresponding contour plot. The horizontal cyan line represents  $F_{nmax}$ . C. The 2-dimensional plot of current as a function of pulse frequency at half-maximal time allocation represented by the red curve in panel B.

### 3.6. The current-pulse frequency trade-off function derived from the 4-dimensional model

After the 4-dimensional reward model is fit to the data, required current ( $I_p$ ) as a function of pulse frequency ( $F$ ) can be derived and plotted (at the behavioural criterion of half-maximal time allocation). The form of the current-pulse frequency trade-off function is obtained by solving for  $I_p$ . Below is the function representing required current as a function of pulse frequency at half-maximal time allocation.

$$I_p = \frac{\left(\frac{SP}{SP_e - SP}\right)^{\frac{1}{g}} \times S_{h\min}}{F_b \left(\ln\left(1 + e^{\frac{F_{n\max}}{F_b}}\right) - \ln\left(1 + e^{\frac{F_{n\max} - F}{F_b}}\right)\right) + I_0}$$

This trade-off function is represented by the red curve on the contour plot of Figure 4B. Figure 4C represents this trade-off function rotated 90 degrees and plotted as current as a function of pulse frequency.

The mapping of pulse frequency into firing frequency is reflected in the current-pulse frequency trade-off function. The present experiment uses the 4-dimensional reward method to derive the current-pulse frequency trade-off function for the tested subjects and to estimate parameters  $F_{n\max}$  (the pulse frequency that produces near maximal firing) and  $F_b$  (the parameter controlling the abruptness of the transition between the range of perfect frequency following to the range of frequency roll-off).

## ***Method***

### ***3.7. Subjects***

In total, 7 male Long Evans rats served as subjects: the 6 rats that were used in the subjective price experiment from Chapter 2 (rats F3, F9, F12, F16, F17, F18), and 1 additional subject (rat LesRO1).

### ***3.8. Apparatus and Materials & Surgical Procedure***

See Chapter 2.

### ***3.9. Training: current sweeps***

After the subjects had completed the subjective price experiment, they were trained in a current-sweep condition. A current sampling matrix consisted of 14 elements (points) in which 14 current intensity values were spaced by equal logarithmic units while the pulse frequency and price remained constant. The points were adjusted such that there were several along the upper and lower asymptotes, and along rising portion of the psychometric curves relating time allocation to current intensity. Eight separate current sampling matrices were tested at 8 different pulse frequencies separated by 0.15 logarithmic units: 1000 Hz, 708 Hz, 502 Hz, 354 Hz, 252 Hz, 178 Hz, 126 Hz, 90 Hz. The price was set to 4 s. A single sampling matrix was tested one to two times to determine the range of currents that drove performance from its lower to its upper asymptotes.

### ***3.10. 4-dimensional testing paradigm***

The parameter space for this condition was [ $\text{Log}_{10}(OP)$  vs.  $\text{Log}_{10}(F)$  vs.  $\text{Log}_{10}(I)$ ] where  $OP$  is the price,  $F$  is the pulse frequency and  $I$  is the current intensity. This condition consisted of 11 sampling matrices, including the pulse-frequency sampling matrix @ 4 s, price-sampling matrix, and radial-sampling matrix. (These were roughly the same matrices that were used in the previous subjective price experiment). In addition, there were 8 current sampling matrices at the different pulse frequencies listed in the current sweep training condition described above (in which the price was held constant at 4 s).

For one survey (one complete test of all points of the 11 sampling matrices), the elements of the sampling matrices were chosen at random, without replacement. Bracketing trials were employed in the same fashion as in the subjective price experiment. One survey was comprised of 2 daily sessions; each daily session was about 6 to 7 hours in duration. Per survey, there were 462 trials (154 experimental trials, and 308 bracketing trials).

After collecting 5 surveys, the data were analyzed (described below). If the psychometric curves (time allocation as a function of price, current, or pulse frequency) did not have well-defined upper and lower asymptotes, the sampling matrix was adjusted, and 5 additional surveys were collected and the psychometric functions were re-assessed. In total, 10 surveys were collected.

### ***3.11. Data analysis***

The data were analyzed using the bootstrapping technique described in Appendix D (4-dimensional model). Briefly, 1000 resampled data sets were generated. The 4-dimensional model was fit to each of the 1000 resampled data sets; the mean parameters and corresponding 95% confidence interval were estimated.



## ***Results***

Figure 5A shows the 3-dimensional section (time allocation as a function of pulse frequency and objective price) of the 4-dimensional reward model, for rat F12. Figure 5B illustrates the corresponding contour plot. The vertical blue line represents the location parameter  $OP_e$ . The horizontal red line represents the location parameter  $F_{hm}$ . The horizontal cyan line represents  $F_{nmax}$ . The dashed lines represent the 95% confidence intervals around the location parameters. Figure 6A is the 3-dimensional section (time allocation as a function of pulse frequency and current) of the 4-dimensional reward model, for rat F12. Figure 6B is the corresponding contour plot, pulse frequency as a function of current intensity. The red line denoted in both of the 3-dimensional and contour plot is the current-pulse frequency trade-off function. Figure 6C illustrates the trade-off function; the  $F_{nmax}$  parameter is denoted by the vertical cyan line.

The 3-dimensional sections of the 4-dimensional reward model for the other rats are illustrated in Figures ApE.1-12 of Appendix E. The trade-off functions for all of the subjects are illustrated in Figure 7. The fitted  $F_{nmax}$  parameters of the subjects range from 236 Hz to 380 Hz: rat F3 (380 Hz), rat F9 (338 Hz), rat F12 (363 Hz), rat F16 (380 Hz), rat F17 (380 Hz), rat F18 (331 Hz), rat LesRO1 (236 Hz). The  $F_b$  ranged from 3.5 to 50 across rats: F3 (10.2), F9 (22.3), F12 (20.6), F16 (6.3), F17 (3.5), F18 (31.4), LesRO1 (50.3). The full set of parameters for each rat is presented in Tables ApF.1-7 of Appendix F. The median and interquartile range for  $F_{nmax}$  is 363, +/- 45.5 Hz. The median and interquartile range for  $F_b$  is 21, +/- 25.

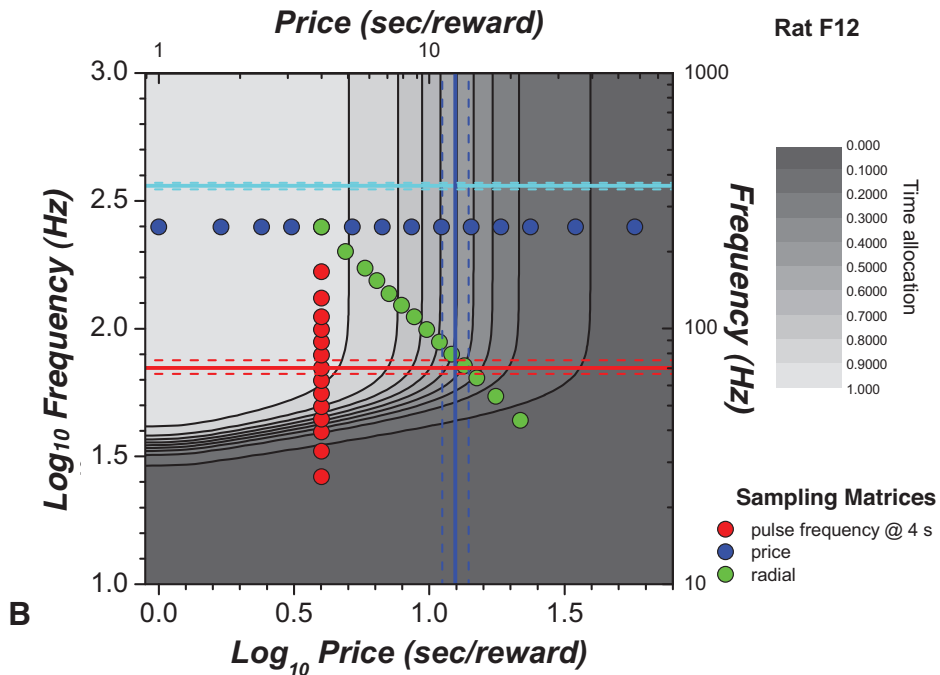
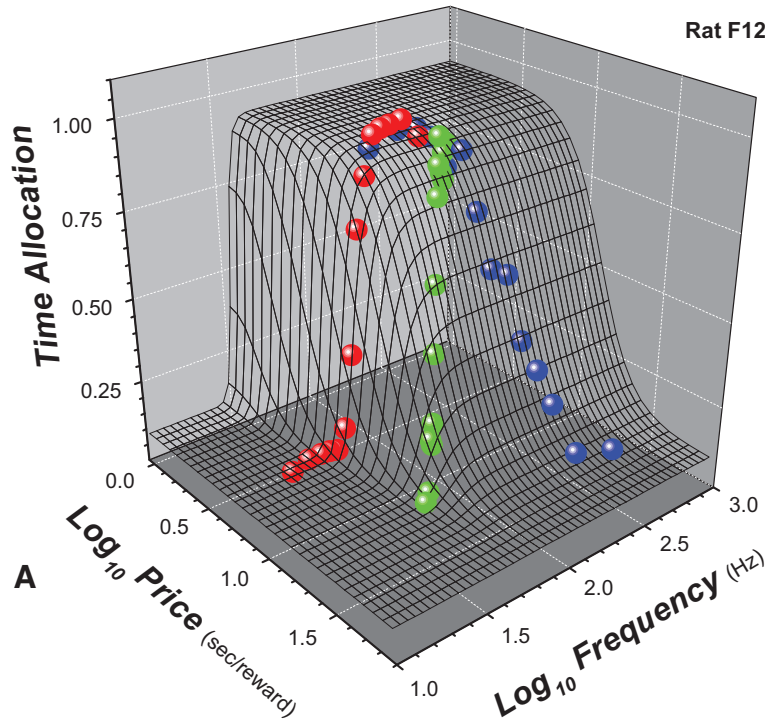


Figure 5. For rat F12, a 3-dimensional section (time allocation as a function of price and pulse frequency) of the 4-dimensional reward model. B. The corresponding contour plot. The vertical blue line represents  $OP_e$ , the horizontal cyan line represents  $F_{nmax}$ , the horizontal red line represents  $F_{hm}$ . The dashed lines represent the 95% confidence intervals. (**Price** refers to *objective price*; **Frequency** refers to *pulse frequency*.)

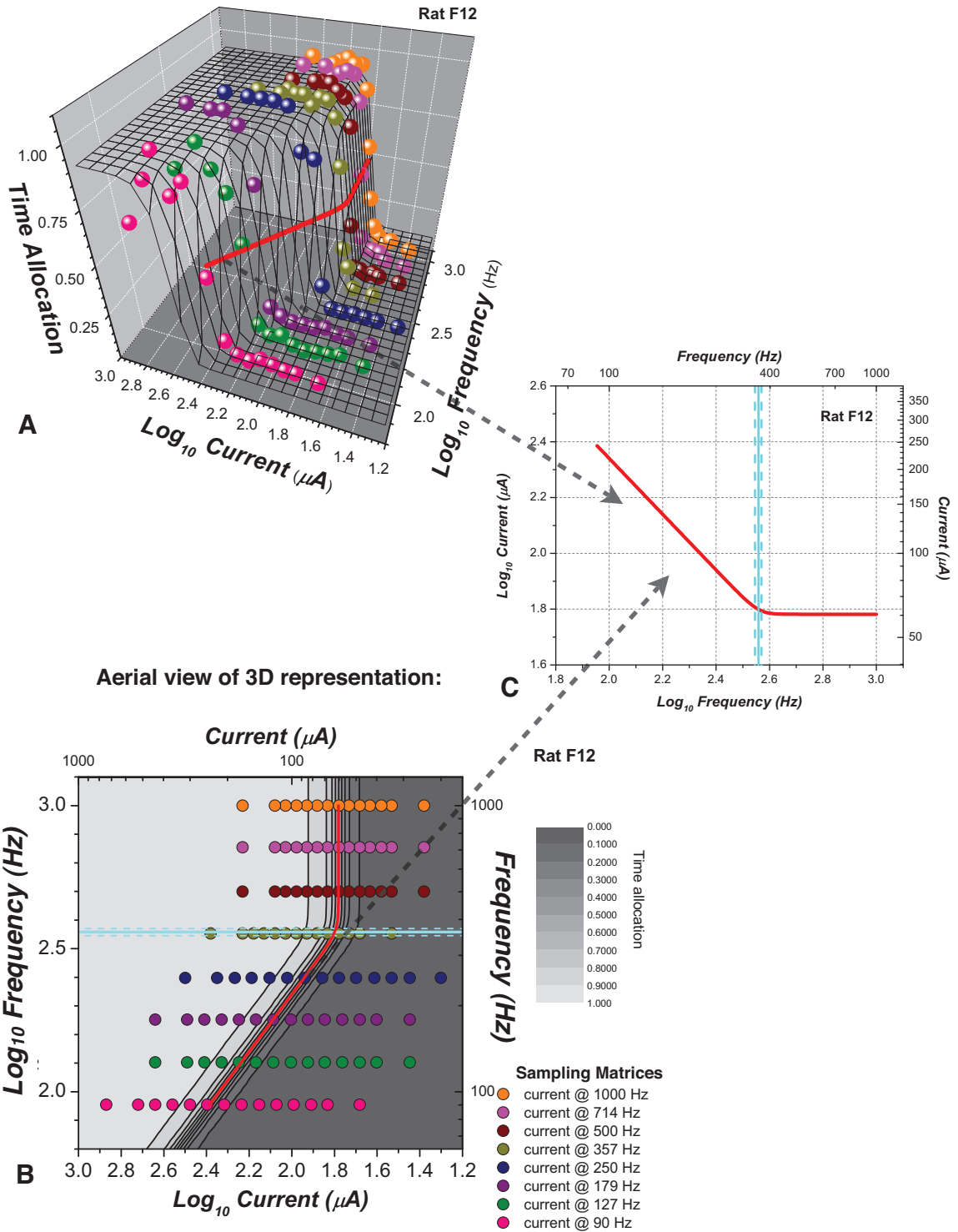


Figure 6. For rat F12, a 3-dimensional section (time allocation as a function of current and pulse frequency) of the 4-dimensional reward mountain model. B. The corresponding contour plot. The horizontal cyan line represents  $F_{nmax}$ . The dashed lines represent the 95% confidence intervals. C. The current-pulse frequency trade-off function at half-maximal time allocation. This function corresponds to the red curve plotted in A and B.

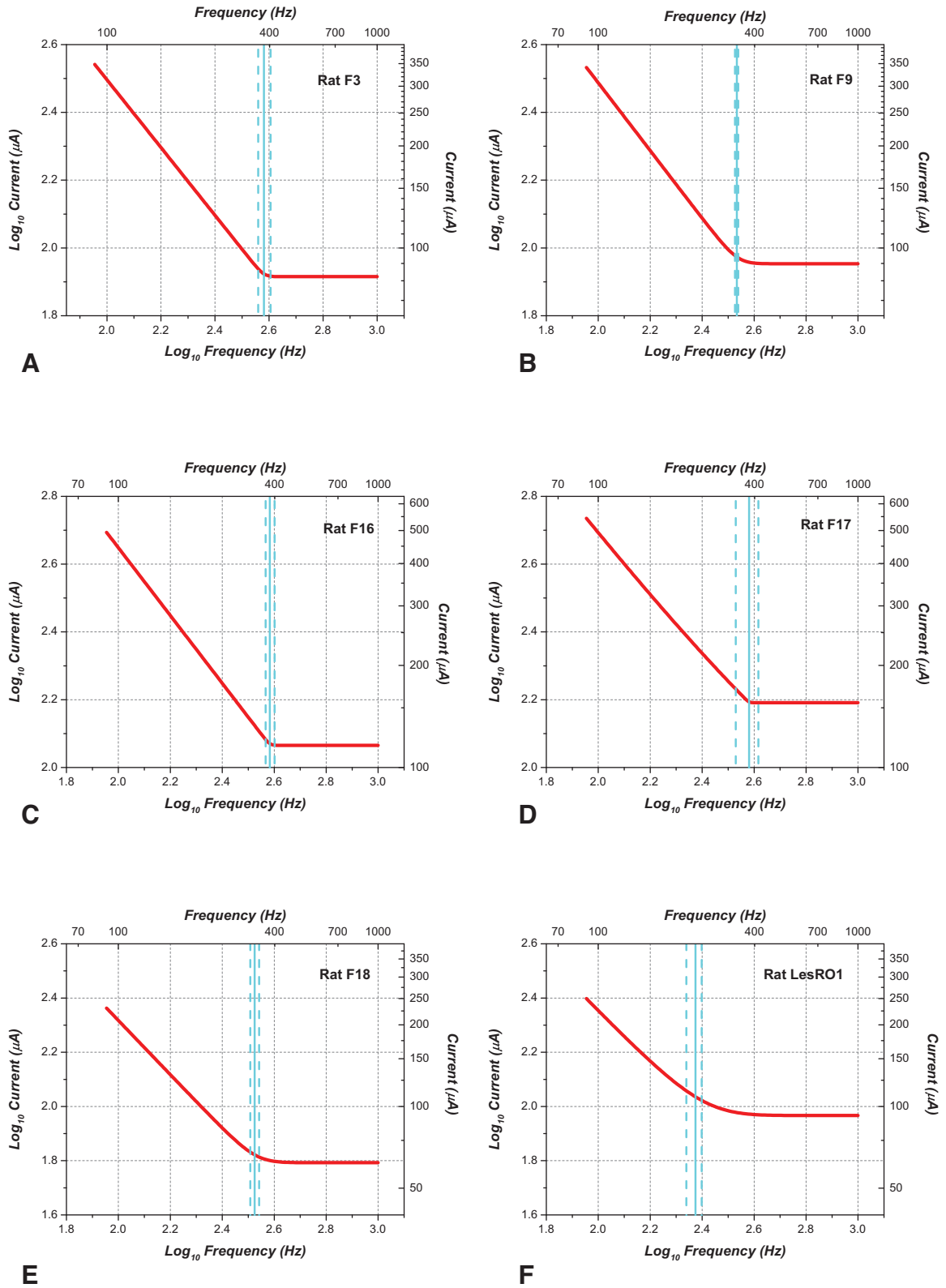


Figure 7A-F. The current-pulse frequency trade-off functions for the other subjects. The solid cyan line represents  $F_{nmax}$  and the dashed lines represent the 95% bootstrapped confidence intervals.

In common logarithmic space, the slope of the diagonal line of the current-pulse frequency trade-off function (prior to  $F_{max}$ ) for rats F3, F9, F12, F16, F18, and Les RO1 was -1 and close to -1 for rat F17 (-0.89). A slope of -1 indicates a reciprocal relationship between current and pulse frequency. This reciprocal relationship implies that the counter model holds for a given range of pulse frequencies (before firing frequency roll-off occurs).

## ***Discussion***

### ***3.12. The counter model was supported***

The present experiment tested the unlikely assumption that the directly stimulated neurons fire once per pulse, regardless of the pulse frequency. Rats performed in an operant task in which price, pulse frequency, and current intensity varied. A 4-dimensional reward model that incorporated the frequency-response function was fit to the data. From this fit, the current-pulse frequency trade-off function was derived: the contour function at half-maximal time allocation in the space [ $\text{Log}_{10}(I)$  vs.  $\text{Log}_{10}(F)$ ]. This trade-off function reflects the firing capabilities of the behaviourally reward-relevant neurons.

The present experiment supports the counter model of reward integration. The counter model states that reward intensity is a function of aggregate spike rate (the product of the number of neurons simulated and the rate at which they fire). By extension, reward intensity reflects the conjoint contributions of pulse frequency and current intensity. In other words, to maintain a given level of reward intensity, increasing the pulse frequency by a given proportion is compensated by a reduction in current intensity by that same proportion. This relationship implies that the two variables are reciprocally related: in common logarithmic space, this relationship predicts that the slope of the current versus pulse-frequency trade-off function will be -1. For most rats in this experiment, the trade-off function had a slope of -1 or close to -1, prior to the bend of the function (Figure 7). Thus, the present experiment adds to the numerous studies that support the counter model.

### 3.13. Frequency-following fidelity broke down at high pulse frequencies

As predicted, over the highest range of pulse frequencies, no reduction of current was needed to maintain the behavioural criterion (half-maximal time allocation). The firing frequency eventually reaches a physiological limit, therefore a reduction in the number of stimulated neurons (set by the current intensity) is not required to maintain the behavioural criterion. The firing frequency roll-off is reflected in the bend portion of the current-frequency trade-off function, its slope denoted by parameter  $F_b$ . The median near maximal firing frequency of the substrate ( $F_{nmax}$ ) across rats was demonstrated to be a median and interquartile range of 363 +/- 45.5 Hz. The median and interquartile range for  $F_b$  is 21, +/- 25. (The median was used because the value of 236 Hz for rat LesRO1 was considered to be an outlier). This range is in accordance with other studies that have reported that the first-stage neurons have the ability to follow high pulse frequencies but reach an asymptote past a certain pulse frequency.

Note that what distinguishes rat LesRO1 from the other rats is that the bend is gradual rather than sharp, not that the curve flattens out at a drastically lower pulse frequency. Parameter  $F_{nmax}$  (near maximal firing frequency) indicates the location of the middle of the bend of the derivative of the frequency-response function. When the bend is gradual,  $F_{nmax}$  is lower than the maximal firing frequency ( $F_{nmax}$  is to the left of where the curve flattens). For the other rats, the bend is sharper and thus  $F_{nmax}$  is close to the maximal firing frequency (indicated by the start of the flat portion of the curve). The plot for Rat LesRO1 demonstrates that raising the pulse frequency from ~ 250 Hz (2.49 log units) to ~ 350 Hz (2.54 log units) actually does decrease the required current as it does for the other rats. Thus, the discrepancy between the data from rat LesRO1 and the other

rats is not as pronounced when considering the bend and its effect on  $F_{max}$  and relation to the maximal firing frequency.

### ***3.14. Why does frequency following fail at high pulse frequencies?***

To account for this roll-off, it makes sense that if the pulse period of the stimulation train (the time between pulses) is less than the absolute refractory period, then frequency following will fail. According to this hypothesis, given that the directly stimulated reward relevant neurons have an absolute refractory period of 0.4 ms to 1.2 ms (Yeomans, 1979), frequency following should start failing at a pulse frequency of at least 800 Hz at the train duration tested. However, the present experiment showed that frequency following began to fail from at most 380 Hz. Thus, another mechanism must contribute to the frequency roll-off. Accordingly, Gallistel (1978) suggested that pulse frequency roll-off may be due to synaptic blocking or fatigue. In Gallistel's study, current-pulse frequency trade-off functions were estimated using a range of train durations. Frequency following broke down sooner (at lower pulse frequencies) at the longer train duration (10 s) than at the shorter train durations. The earlier breakdown in the case of the longer train durations suggests that synaptic fatigue and negative feedback may play a role in frequency-following failure.

### ***3.15. Reward intensity saturation is not caused by frequency following failure***

The present study demonstrated that reward intensity saturates before frequency following fails at the current used for frequency, price, and radial sampling matrices. This is revealed in the pulse frequency versus price contour plots (Figure 5B, Appendix E:



Figures ApE.1B, 3B, 5B, 7B, 9B, 11B): the near maximal frequency ( $F_{max}$ , represented by the cyan line) occurs at a significantly higher pulse frequency than the pulse frequency at which the reward intensity saturates (represented by the shoulder of the curves). Thus, reward intensity saturation is probably due to a mechanism other than frequency following roll-off. This idea is supported in a previous study that measured the reward intensity function at different currents using a dual-operant procedure (Simmons & Gallistel, 1994). For the same subject, when the reward intensity function was measured at a high current intensity, the reward intensity leveled-off at a lower pulse frequency than when the function was measured at a lower current intensity. If the reward intensity function saturated due to the frequency roll-off, then both functions (measured at low and high current intensities) would saturate at the same pulse frequency. As was clearly not the case, reward intensity must saturate due to some other mechanism. The present study further demonstrated that reward intensity saturation is not due to frequency firing roll-off.

### ***3.16. Implications: properties of the reward-relevant neurons***

The exact identity of the directly stimulated neurons is presently unknown; due to the heterogeneity of the medial forebrain bundle, electrical stimulation of the substrate activates several neural systems. Consequently, many of the systems that are activated are probably not responsible for the rewarding effect. A purely electrophysiological approach will not give much insight into the properties of the neurons responsible for the rewarding effect. By employing behavioural trade-off experiments, previous studies have revealed the quantitative neurophysiological and anatomical properties of the neurons responsible

for reward-seeking behaviour (described below). The results of the present experiment provide further insight into the nature of the stimulated substrate.

The absolute refractory periods of the directly stimulated neurons of the medial forebrain bundle have been shown to range from 0.4 to 1.2 ms as revealed through behavioural C-T experiments (Yeomans, 1979). Conduction velocities are in the range of 1 to 8 m/s which suggest myelinated, fine fibers that are 0.5 to 2  $\mu$ m in diameter (Bielajew & Shizgal, 1982). As well, the reward-relevant direction of conduction in at least some of these fibers is rostro-caudal (Bielajew & Shizgal, 1986). The results of the present experiment add to the criteria for identifying the first stage neurons: the ability of neurons to follow high pulse frequencies (median and interquartile range of 363 +/-45.5 Hz).

The known electrophysiological properties of the first-stage neurons allow experimenters to rule out candidate neurons that do not match these properties. In particular, the quantitative properties identified by behavioural trade-off experiments do not support the “catecholamine hypothesis,” the early hypotheses concerning the nature of the first-stage neurons (Wise, 1978). The correlation between the brain regions supporting self-stimulation and the location of the catecholamine pathways initially contributed to this proposal. Also, numerous pharmacological studies had demonstrated an involvement of dopamine (DA) in BSR. Thus, it was evident that DA neurons play an important role in BSR. However, the known physiological properties of DA cannot account for the behaviourally derived trade-off data: DA neurons are unmyelinated, have slower conduction velocities (0.3 – 1.5 m/sec), longer refractory period periods (1.8-20 msec). As well, the direction of conduction of DA neurons along the MFB is caudal-

rostral whereas at least a subset of the reward-relevant neural projections have been demonstrated to be rostral-caudal (Bielajew & Shizgal, 1986). The findings of the present study contribute to the evidence that suggests that the directly stimulated neurons are non-dopaminergic: it is unlikely that DA neurons could follow frequencies up to such a high rate due to their long refractory periods and slow conduction time.

### ***3.17 A direct application of the frequency following function: behavioural-neural activity correspondence***

The behavioural trade-off functions in previous and present studies have identified physiological properties that do not support the hypothesis that DA neurons constitute the directly stimulated substrate. However, results from the BSR curve shift and neurochemical experiments imply that DA neurons are implicated in reward-seeking behaviour. One proposal is that the DA neurons are in series with the neurons directly activated by the electrical stimulation. The DA neurons are suggested to receive input from the directly activated neurons (Wise, 1980). Because at least some neurons of the MFB send descending projections to the VTA (Shizgal & Bielajew, 1986), it is possible that DA neurons of the VTA are transynaptically activated. The ventral tegmental area-substantia nigra complex of the midbrain sends DA projections throughout the brain. One particular projection from the posteromedial VTA to the nucleus accumbens (NAC) shell is of particular interest to studies on BSR because it is implicated in goal-directed behaviours (Ikemoto, 2010). Thus, much of the focus on DA and ICSS has been on the VTA-NAC shell projection. It is this projection and thus DA release in the NAC shell that can provide insight into the nature of VTA DA neurons in BSR. The behavioural

trade-off function can be used as a tool in combination with neurochemical and electrophysiological methods to test the hypothesis that DA neurons of the VTA are in series with the directly activated neurons.

In this procedure, the behavioural and neural activity trade-off functions are measured in succession. The neural activity trade-off function is analogous to the behavioural trade-off function except that the dependent measure is DA activity, such as DA neuronal firing rate (measured via electrophysiology) or DA transient magnitude (measured by fast-scan cyclic voltammetry FSCV, in which peak oxidation current reflects DA magnitude). If the required current levels off for DA activity over the same pulse frequency range as it does for behaviour, then this would suggest that the VTA DA neurons are in series with the directly activated neurons. The rationale for the argument that this correspondence can be used as an indicator that a given neuronal population is a stage of the reward-relevant circuitry was originally discussed by Gallistel and colleagues (1981) and is explained in the following paragraph.

The function that relates behavioural output (time allocation to pursuit of electrical stimulation) to stimulus input (the electrical stimulation parameters) arises from the concatenation of “buried” or “hidden” functions, each describing a stage of reward processing. Among these hidden stages is the early stage during which the aggregate spike rate is transformed into reward intensity and the later stage during which the penultimate output is transformed into behaviour. An important property of the measured behavioural allocation function (time allocation versus stimulation parameters) is monotonicity. In a monotonically increasing function, each output corresponds to one input value (along the ascending portion of the curve), or a range of consecutive input

values (illustrated by a leveling-off of the function). In reward evaluation, we assume a serial linking of stages (concatenation), in which the output of a stage of processing is the input to the next stage. Assuming the stages are concatenated, then for this monotonic behavioural allocation function to occur, it is necessary for all of the preceding hidden functions to be monotonic as well (Gallistel et al., 1981). Thus, if the experimenter holds the behavioural output constant, then this implies that all of the inputs of the preceding stages are held constant as well. In this view, there is a one-to-one correspondence between behaviour and any given input. Therefore, determining whether activity of a given neural substrate corresponds to the behavioural output would indicate the possibility that the substrate in question is an input in the series of concatenated stages. (This method is correlative, therefore a behavioural-neural activity correspondence does not indicate definitively, but rather the possibility that an input is implicated in reward-seeking behaviour.)

Using this rationale, Moisan and Rompré (1998) employed behavioural and electrophysiological techniques to test whether midbrain DA neurons comprise a stage of the reward circuitry that is activated by electrical stimulation of the posterior mesencephalon (PM). The PM axons project to the ventral midbrain where DA neurons are one of the constituents. The goal was to determine whether the DA neuron cell bodies of the ventral midbrain track the current-frequency combinations. Behavioural measurements entailed two pulse-frequency sweep curves (obtained at a low and high current) that ranged from 10 Hz to 110 Hz. Four current-pulse frequencies combinations were chosen based on the behavioural data. They demonstrated that the magnitude of DA activation was not correlated with the stimulation frequency or current intensity of

the stimulation train alone, but with its rewarding intensity. These findings support the hypothesis that DA neurons are transynaptically activated by the directly activated reward-relevant neurons. In accordance, shifts of the reward-mountain structure along the price axis following a DA manipulation suggests that DA may be activated at stages of the reward circuitry further downstream (Hernandez et al., 2010, Hernandez et al., 2012; Trujillo-Pisanty et al., 2013).

Cossette (2011) used the same rationale as Moisan and Rompré (1998) but employed neurochemical detection of DA in the NAC shell. Current sweeps were obtained at 4 fixed pulse frequencies (60 Hz, 130 Hz, 250 Hz, 1000 Hz). The required current intensity at the behavioural criterion (half-maximal time allocation) as a function of pulse frequency was plotted in logarithmic space. The behavioural results were similar to the findings in this chapter. Specifically, the function declined monotonically over the tested range and leveled off over the final decrement, from 250 Hz to 1000 Hz. Cossette (2011) then measured DA transients in the NAC using voltammetry at each of the 4 current-pulse frequency combinations. The DA trade-off function declined as the pulse frequency was increased from 60 Hz to 130 Hz, which was consistent with the behavioural trade-off function. However, further increases in pulse frequency resulted in a leveling off or even a rise, which was inconsistent with the behavioural data. This divergence of the neurochemical trade-off function from the behavioural trade-off function is problematic for the hypothesis that DA neurons are in series with the directly activated reward relevant neurons.

The discrepancy between Cossette's (2011) and Moisan and Rompré's (1998) findings may be due to methodological differences. Because detection of DA transients

by means of FSCV is restricted to the location of the carbon-fiber detection electrode, only a small proportion of DA terminals in the NAC can be described. It is possible that the set of local DA terminal responses detected by FSCV could jointly comprise a stage of the reward evaluation. For instance, one set of DA neurons could track the aggregate firing rate along the low range of pulse frequencies, while another set could track the aggregate firing rate along the intermediate range, and yet another along the high range. The total aggregate firing rate across all pulse frequencies may be reconstructed by integrating the DA response subsets. However, the problem with this proposal is that none of the DA transients in Cossette's (2011) study tracked the decline in the behavioural trade-off function past 130 Hz. Thus, there was no evidence of DA tracking the middle and high pulse frequency range. A further limitation to Cossette's study is that the FSCV data were obtained in anesthetized subjects, which may affect the results. The behavioural current-pulse frequency experiment and complementary FSCV experiment will be repeated using FSCV in behaving subjects. In previous studies (Cossette, 2011; Moisan & Rompré, 1998), the trade-off functions were not formally derived and only four points in the sampling space were tested. Formally estimating the function and tracking the DA response along a broader range of points would provide greater resolution with which to assess the neural activity-behavioural correspondence.

### ***3.18. Further implications: first-stage neurons and glutamate***

The findings from the present 4-dimensional frequency following experiment can also be used in conjunction with other techniques to help identify the first-stage neurons. One hypothesis is that glutamate forebrain neurons projecting to the ventral tegmental

area (VTA) of the midbrain constitute at least a subset of the first-stage neurons. Several lines of evidence support this proposal. For example, Geisler & Zahm (2005) showed that numerous forebrain regions send glutamatergic projections to the VTA through the MFB. In addition, You, Chen, and Wise (2001) demonstrated glutamate and DA release in the VTA during self-stimulation of the MFB. Kempadoo and colleagues (2013) showed that LH to VTA optical stimulation promotes peptide neurotensin and glutamate release in the VTA. In addition, Jennings and colleagues (2013) demonstrated that glutamatergic and GABAergic projections from the bed nucleus of the stria terminalis produced functionally opposing rewarding and aversive states. Thus, a plausible mechanism of MFB stimulation is that the electrical stimulation directly activates glutamate neurons in the MFB, which in turn activate the VTA DA neurons. The first test of this hypothesis will employ optogenetic methods to determine whether rats will self-stimulate for optical stimulation of glutamate terminals of the VTA. Optogenetic techniques allow for specificity and temporal control, manipulations not previously achievable with the use of drugs. The opsin will be injected into a specific candidate forebrain area and transported to the VTA (provided that there is a projection of the target area to the VTA). The candidate forebrain areas will be chosen based on results from lesion studies. Whether the opsin is transported to the VTA will be apparent if the rat self-stimulates for light activation of the VTA, and can also be visualized in tissue slices post mortem.

If the rats self-stimulate for light activation of glutamate terminals in the VTA, the next step would be to use the collision method to determine whether the glutamate terminals (stimulated by the optical probe in the VTA) and the MFB fibers (stimulated by an electrode that supports electrical self-stimulation) constitute a common pathway.



Collision studies can reveal a common pathway linking two implants. If the fibers contributing to the reward signals are common (and if the experimenter-set interval between these two signals is sufficiently short), then orthodromic action potentials triggered by the electrode site will collide with antidromic action potentials triggered by the optical probe. Collision of action potentials causes conduction failure, which in turn causes a reduction in the subjective reward intensity (reflected in the behaviour). The same optical probe can potentially be used for recordings or coupled with a recording microelectrode. The physiological properties of the stimulated glutamatergic neurons can be compared with the characteristic properties of the first-stage neurons such as the frequency following limit, the absolute refractory period, and conduction velocity. If the properties match, this would further strengthen the hypothesis that the first-stage neurons are glutamatergic neurons projecting from the forebrain to the VTA.

### ***3.19. Implications for reward-mountain methodology***

The frequency following roll-off experiment findings are important in the context of the reward-mountain methodology. To drive performance to its upper bounds, the experimenter manipulates the range of tested pulse frequencies and currents. However, it is necessary to know at what pulse frequency the roll-off occurs. Consider a case in which a lesion causes a disruption of first-stage neurons: an increase in pulse frequencies (due to a shift of the mountain along the pulse frequency axis) is required to restore baseline behavioural levels. If the experimenter-set pulse frequency range includes values that exceed  $F_{max}$ , the findings and conclusions could be inaccurate and misleading. In this case, the reward intensity function could saturate due to frequency

firing failure, before the output of the reward integrator can reach its maximal level ( $RI_{max}$ ). Thus, if the experienced maximal reward is actually due to frequency following failure, the experienced maximal reward intensity will be less than the maximal reward intensity ( $RI_{max}$ ) that the substrate can attain. For instance, we choose a pulse frequency well-above  $F_{nmax}$  and call it  $F_{hi}$ , such that  $F_{hi} \gg F_{nmax}$ . If the experimenter-set pulse frequencies are below  $F_{nmax}$ , then the maximal reward intensity,  $RI_{max}$ , can be reached and will equal the reward intensity produced by  $F_{hi}$  (denoted as  $RI_{hi}$ ): thus,  $RI_{max} = RI_{hi}$  (Figure 8A). However, in another scenario, if several of the tested pulse frequencies are above  $F_{nmax}$ , it is possible that the maximal reward intensity of the integrator ( $RI_{max}$ ) is not reached. Consequently, it is likely that the experienced reward intensity (the measured asymptotic reward intensity, red curve in Figure 8B) is due to frequency following failure rather than reward saturation. In this case,  $RI_{max} > RI_{hif}$  (Figure 8B).

In the scenario in which  $RI_{max}$  is not reached, the experienced maximal reward intensity ( $RI_{hi}$  which is less than  $RI_{max}$ ) has a major influence on the detection of changes in parameter  $SP_e$ . Recall that  $SP_e$  (the subjective price of “everything else”) is calculated as

$$SP_e = \frac{RI_{max}}{U_e}$$

When the output of the integrator cannot reach its maximum value, the subject experiences a value less than  $RI_{max}$ , referred to as  $RI_{hi}$ . Therefore,

$$SP_e = \frac{RI_{hi}}{U_e}$$

where,

$$RI_{hi} < RI_{max}$$

Therefore, in the second case,  $SP_e$  is lower than what it would have been if the reward integrator had been able to reach the maximal reward intensity of the substrate. This is problematic in studies measuring the effects of manipulations on the location parameters. For instance, if a manipulation such as a lesion to a particular brain area caused  $F_{hm}$  to increase, the experimenter might increase the range of pulse frequencies to very high levels in order to capture the sigmoidal shape of the time allocation curve. However, some of the experimentally chosen pulse frequency values could be above  $F_{nmax}$ ; as a result, the  $SP_e$  will be reported to have decreased when in fact the manipulation may have had no true effect on  $SP_e$ . The transformation of pulse frequencies ( $F$ ) into firing frequencies ( $FF$ ) described in this chapter circumvents this potential erroneous inference. Future studies will incorporate the frequency roll-off expression in the 3-dimensional testing paradigm with  $F_{nmax}$  set to the first quartile value measured in this study, 316 Hz (rounded down from 317.5 Hz) and median  $F_b$  of 21.

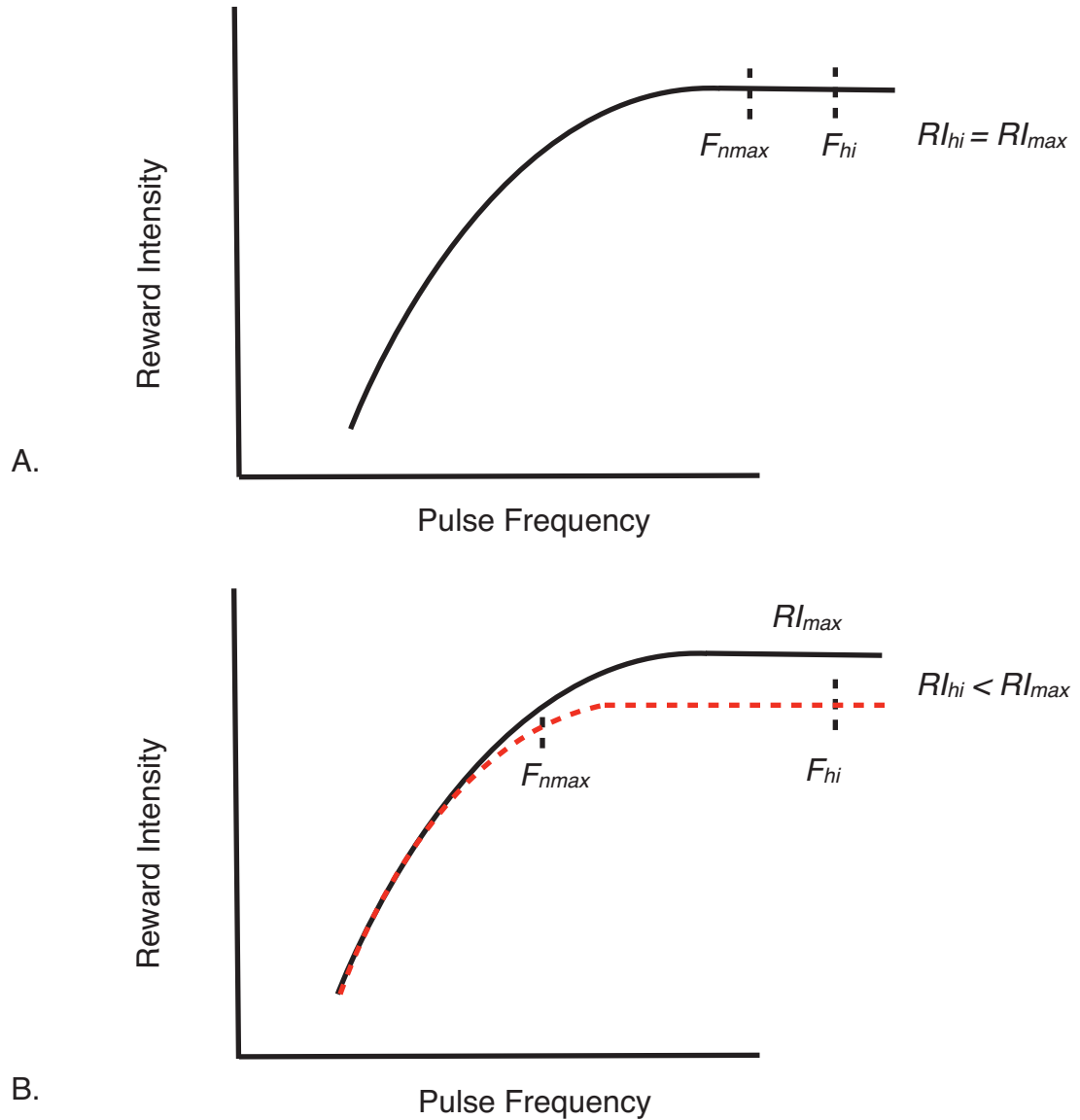


Figure 8. A. The reward intensity function. As pulse frequency increases, reward intensity increases and then levels off. If the tested pulse frequencies are below  $F_{nmax}$  (denoted by the dotted vertical line), then the maximal reward intensity  $RI_{max}$  is reached. If a pulse frequency higher than  $F_{nmax}$  is chosen, referred to as  $F_{hi}$ , the reward intensity corresponding to these two pulse frequencies would be equivalent. B. In another scenario, if several of the tested pulse frequencies are above  $F_{nmax}$ , the reward intensity function will saturate due to  $F_{nmax}$ , and will not reach  $RI_{max}$ . The dotted red line represents the reward intensity function that saturates due to  $F_{nmax}$ . The black solid line represents the output of the reward intensity function if  $F_{nmax}$  had not been reached. The reward intensity corresponding to the high frequency pulse train ( $RI_{hi}$ ) is less than the maximal output that could have been attained by the reward intensity function ( $RI_{max}$ ) if  $F_{nmax}$  had not been reached.

## ***Appendix A: Derivation of the frequency-response function***

The frequency-response function was developed by computing the integral of a sigmoidal expression that describes the rate at which the induced firing frequency changes (the frequency-response derivative) as a function of pulse frequency. The shape of this derivative function is sigmoidal. Figure A1 plots this functional derivative in linear space (upper panel) and in common logarithmic space (lower panel) with hypothetical parameter values. The firing frequency follows low pulse frequencies perfectly, which is represented by the flat line with a derivate of 1. However, at a certain point, the derivative of the function (slope) decreases and reaches an asymptotic value. (The derivative never reaches 0, but is very close.) The derivative of the firing frequency,  $FF'$  is expressed as:

$$FF' = \frac{e^{-\left(\frac{F-F_{nmax}}{F_b}\right)}}{1 + e^{-\left(\frac{F-F_{nmax}}{F_b}\right)}}$$

where,

$FF$  = the average firing frequency in Hz

$F$  = the pulse frequency that is experimentally-set in Hz

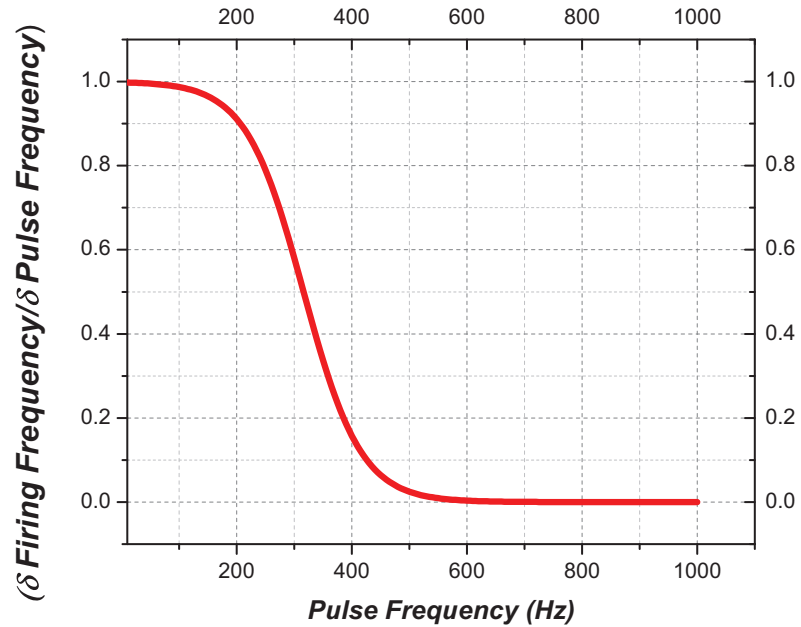
$F_{nmax}$  = the position parameter of the sigmoid, the pulse frequency at which the slope of the function is half-maximal; the firing frequency at this pulse frequency is actually near maximal and thus the reason why “*nmax*” is in the subscript (demonstrated in Appendix B)

$F_b$  = parameter governing the slope of the sigmoidal function

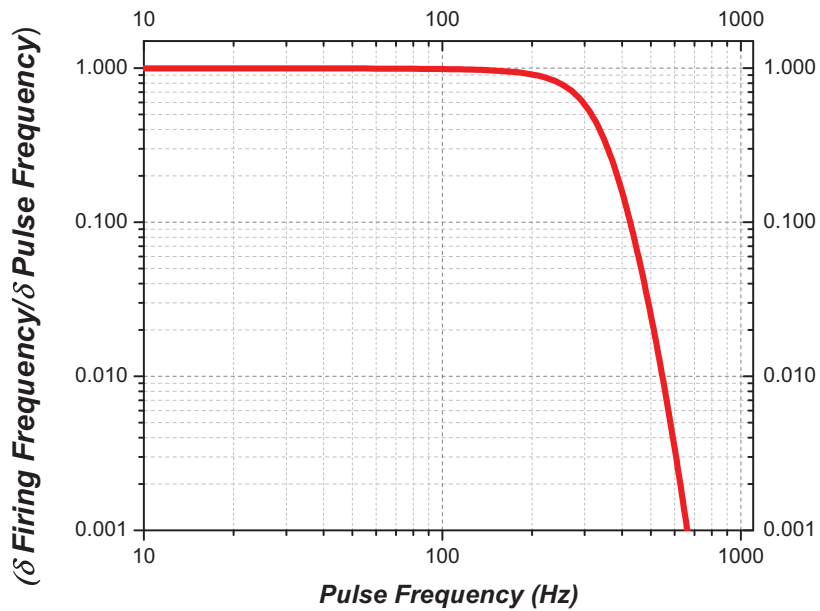
The integral of the above function (as illustrated in Figure 1) is the actual function that predicts frequency firing and is expressed as:

$$FF = F_b \left( \ln \left( 1 + e^{\frac{F_{nmax}}{F_b}} \right) - \ln \left( 1 + e^{\frac{F_{nmax} - F}{F_b}} \right) \right)$$

As  $F$  is increased,  $FF$  increases and reaches an asymptote at around  $F_{nmax}$ .



**A**



**B**

Figure ApA1. A. The derivative of the firing frequency (the induced) as a function of pulse frequency (the inducer). B. The same function plotted in common logarithmic space. The firing frequency follows perfectly at low pulse frequencies, which is represented by the flat line at a derivate of 1. However, at a certain point, the derivative decreases which reflects the gradual bending and reaching of an asymptotic value of the function relating firing frequency to pulse frequency. The derivative never reaches 0, but is very close.

***Appendix B: A numerical example demonstrating that the near maximal firing frequency parameter ( $F_{nmax}$ ) of the frequency roll-off comes very close to the maximal value of frequency firing  $FF$ .***

The example below shows that the near maximal firing frequency parameter ( $F_{nmax}$ ) of the frequency roll-off comes very close to the maximal value of frequency firing  $FF$ .

Given the background literature, the maximum firing frequency is achieved at a pulse frequency well below 1000 Hz. Thus, we can set  $F$  to 1000 Hz to drive the neurons to their maximal firing frequency. The numerical example below shows that the maximal firing frequency  $FF$  will be very close to parameter  $F_{nmax}$  when driving the neurons to their maximum firing frequency.

Let,

$$F_{nmax} = 316 \text{ Hz}$$

$$F_b = 50$$

$$F = 1000 \text{ Hz}$$

$$FF = F_b \left( \ln \left( 1 + e^{\frac{F_{nmax}}{F_b}} \right) - \ln \left( 1 + e^{\frac{F_{nmax} - F}{F_b}} \right) \right)$$

$$FF = 50 \times \left( \ln \left( 1 + e^{\frac{316}{50}} \right) - \ln \left( 1 + e^{\frac{316 - 1000}{50}} \right) \right)$$



$$FF = 316.089916175$$

Thus, the maximal firing frequency  $FF$  is very close to  $F_{nmax}$ .

Also, increasing  $F$  from 1000 Hz to 10 000 Hz further shows that the maximal firing frequency  $FF$  stays nearly the same as  $F_{nmax}$ :

$$FF = 50 \times (\ln(1 + e^{\frac{316}{50}}) - \ln(1 + e^{\frac{316-10000}{50}}))$$

$$FF = 316.089916277$$

Thus, this example further underscores the premise that “near maximal” firing represented by  $F_{nmax}$  is very close to the maximal firing frequency.

It is important to note that the discrepancy between  $F_{nmax}$  and the maximal firing frequency depends on the bend parameter,  $F_b$ . The higher the value of  $F_b$ , the broader the bend, and the further away  $F_{nmax}$  is from the maximal firing. For example,

Let,

$$F_{nmax} = 316 \text{ Hz}$$

$$F_b = 100$$

$$F = 1000 \text{ Hz}$$

$$FF = 100 \times (\ln(1 + e^{\frac{316}{100}}) - \ln(1 + e^{\frac{316-1000}{100}}))$$

$$FF = 320.1507 \text{ Hz}$$

Here, the maximal firing frequency ( $FF$ ) at parameters  $F_{nmax}$  (316 Hz) and  $F_b$  (100) is 320 Hz. Although the maximal value computed here is higher than the “near maximal” ( $F_{nmax}$ ) value, the difference is not very large (a difference of  $\sim 4$  Hz). In the previous example,  $F_b$  was set to a lower value (50 Hz); the difference between the calculated maximal firing frequency and  $F_{nmax}$  was negligible.

***Appendix C: The relationship between current and the number of stimulated neurons***

***The relationship between the current intensity and stimulated surface area***

The current density threshold ( $I_{dt}$ ) is the current density ( $\mu\text{A}/\text{mm}^2$ ) needed to activate a neuron 1 mm away from the electrode tip (Tehovnick, 1996) and varies depending on the properties of the neuron (size, myelination, etc).

Thus, knowing the current threshold constant  $I_{dt}$  of a particular substrate, the threshold current required to stimulate a neuron  $Re$  from the electrode tip will be:

$$I_p = I_{dt} \times K_{geom1} \times Re^2$$

where,

$I_p$  = the applied current intensity at a given pulse duration in units  $\mu\text{A}$

$I_{dt}$  = the current density threshold in units  $\mu\text{A}/\text{mm}^2$

$Re$  = the radius of excitation; the mean distance from the center of the electrode to the edge of the region in which neurons are directly stimulated, in units mm

$K_{geom1}$  = a scalar that translates  $Re^2$  into a surface area; for example, if the field were spherical  $K_{geom1}$  would be  $4\pi$

Alternatively, if the experimenter wants to determine the radius of excitation  $Re$  when applying a particular current:

$$Re = \sqrt{\frac{I_p}{I_{dt} \times Kgeom1}}$$

The product of current intensity and the scalar that translates the radius ( $Re$ ) into surface area,  $I_{dt} \times Kgeom1$ , is traditionally simplified and expressed as:  $Kidist$ , the current distance constant in units  $\mu A/mm^2$ .

Therefore,

$$Kidist = I_{dt} \times Kgeom1$$

Simplifying the above expression:  $I_p = I_{dt} \times Kgeom1 \times Re^2$

$$I_p = Kidist \times Re^2$$

Current distance constants have been described ranging from 100 to 4000  $\mu A/mm^2$  (Tehovnick, 1996). Fouriezos and Wise (1984) demonstrated that the  $Kidist$  mediating MFB self-stimulation is 1300  $\mu A/mm^2$  (at a pulse duration of 0.1 ms). Also, Yeomans, Maidment, and Bunney (1988) demonstrated  $Kidist$  values ranging from 1800 to 4000  $\mu A/mm^2$  (also corresponding to a pulse duration of 0.1 ms) in the MFB which has been suggested to correspond to high threshold dopamine neurons.

Next, to relate the current to the number of stimulated neurons, we need to take into account that, a glial scar surrounds the electrode. There is a minimum current that will be “ineffective” near the electrode tip. This “waste” current is expressed as:

$$I_0 = Kidist \times Rs^2$$

where,

$I_0$  = the “ineffective” or “waste” current; the current that just suffices to fire a hypothetical neuron located at the edge of the scar tissue surrounding the electrode tip

$Rs$  = the average distance from the electrode tip to the edge of the scar tissue

The difference between  $I_P$ , the applied current and  $I_0$ , the ineffective current is the “effective” current.

$$I_P - I_0 = Kidist \times Re^2 - Kidist \times Rs^2$$

$$I_P - I_0 = Kidist(Re^2 - Rs^2)$$

Rearranging the expression (to be used in the next section):

$$(Re^2 - Rs^2) = \frac{I_P - I_0}{Kidist}$$

*The relationship between the current intensity and the number of neurons stimulated*

The formula that relates the number of neurons to the surface area is simply geometric principles:

$$Nn = Kpack \times Kgeom2(Re^2 - Rs^2)$$

where.

$Nn$  = the total number of directly stimulated nerve fibers

$Kpack$  = a scalar that determines spatial density (packing) of the stimulated neurons in units neurons/mm<sup>2</sup>; for simplicity, we assume a homogenous spatial density around the electrode tip

$Kgeom2$  = a scalar that translates  $Re^2$  into cross sectional area; if the stimulated cross-section were circular, it would be  $\pi$

$Re^2 - Rs^2$  = the difference between the radius of excitation and radius of scar tissue

Substituting  $\frac{I_p - I_0}{Kidist}$  for  $(Re^2 - Rs^2)$ ,

$$Nn = Kpack \times Kgeom2 \left( \frac{I_p - I_0}{Kidist} \right)$$

Next, let,

$$K = \frac{K_{pack} \times K_{geom2}}{K_{dist}}$$

where  $K$  has units neurons/ $\mu$ A

Therefore,

$$Nn = K \times (I_p - I_0)$$

**Thus, assuming a homogenous distribution of equally excitable neurons at the electrode tip, the number of stimulated neurons  $Nn$  grows in a scalar fashion as a function of effective current,  $I_p - I_0$ .**

## ***Appendix D: The resampling technique***

### ***Why the resampling technique is used***

It is conventionally assumed that the measure of dispersion around the parameters of a data set is symmetrical about the mean. However, in the present data sets, this dispersion may be uneven. Specifically, as time allocation is constrained from 0 to 1, there could be floor and ceiling effects. Past a certain point, the independent variable can no longer affect the dependent variable: a rat cannot respond more than 100% or less than 0% of the time. These constraints on the dependent variable can lead to skewness in the parameter estimates' distributions (e.g,  $a$ ,  $g$ ,  $F_{hm}$ ,  $OP_e$ ,  $TA_{max}$ ,  $TA_{min}$ ). To free us from the assumption of normality, we use a resampling technique with replacement (bootstrapping) that allows for empirically derived parameter distributions and 95% (asymmetrical) confidence intervals around the parameter means (Efron & Tibshirani, 1986).

Furthermore, the empirically derived distributions permit us to avoid making unrealistic assumptions concerning the lack of correlation between parameters. By estimating the parameter distributions, correlations between the parameters can be determined.

### ***How the raw data are obtained***

For the 3-dimensional mountain model, one survey consists of a test of a total of 3 sampling matrices: pulse-frequency sampling matrix @ 4 s, price-sampling matrix, radial-sampling matrix. One daily session is required to collect data from 1 survey. In



total, for most rats, 10 surveys make up a complete experiment. The conventional statistical analysis would be to fit the reward mountain model to these data by a non-linear least squares regression technique. Instead, a resampling with replacement procedure was used for the reasons explained above. The resampling approach involves generating 1000 resampled data sets and obtaining a distribution for each of the parameters ( $a$ ,  $g$ ,  $F_{hm}$ ,  $OP_e$ ,  $TA_{max}$ ,  $TA_{min}$ ).

### ***Fitting the 3-dimensional model***

The 3-dimensional mountain model was fit to the data from the early training condition, described in Chapter 2. The pool from which we resample consists of the data from the 10 surveys of the complete experiment. A *single* resampled data set consists of data from 10 surveys that are resampled with replacement from the pool. For example, the list of surveys comprising a single resampled data set could entail survey numbers: 2,4,5,2,6,6,9,1,9,3. In another resampled data set, the survey numbers could be: 1,5,6,2,10,10,9,3,6,7. In total, 1000 resampled data sets resembling the previous examples are generated. Figure D1 illustrates this example. The reward mountain model is individually fit to each of the 1000 data sets using the non-linear least-squares routine in the MATLAB Optimization Toolbox (The Mathworks, Natick, MA). The mean parameters ( $g$ ,  $OP_e$ ,  $F_{hm}$ ,  $TA_{max}$ ,  $TA_{min}$ ,  $a$ ) are determined and their respective 95% confidence intervals are calculated by excluding the lowest 2.5% and highest 2.5% of the estimates.

Two methods are used to determine the parameters of a resampled data set. The first method is employed to minimize the bias that would be introduced to the slope parameters ( $a, g$ ) if these were estimated by computing a conventional average of the data. Such an average would introduce bias because the data includes daily variation due to hidden factors that occur during the daily sessions. This problem and its solution are illustrated in a simplified 2-dimensional example (time allocation as a function of pulse frequency) in Figure D2. In this example, there is substantial drift in the lateral displacement of the curves (a drift in  $F_{hm}$ ). Fitting the function to the data from *each* individual survey and subsequently *averaging the parameters* yields a curve (denoted by a thick grey line) with a slope representative of the individual curves. However, fitting the function to the averaged time allocation values produces a curve with a slope, shallower than the slopes of the individual curves (black dotted curve). The first method termed the “location-specific” approach, captures across-survey drift of the location parameters while minimizing the number of free parameters. The model is fit separately to the data from each survey such that the location parameters  $F_{hm}$  and  $OP_e$  are free to vary across surveys while the remaining parameters ( $g, TA_{max}, TA_{min}, a$ ) are kept common. Thus, by this method, a 24-parameter model is fit (2 location parameters that are free to vary  $\times$  10 surveys + 4 common parameters = 24 parameters). The number of free parameters is minimized with this approach: if all 6 parameters had been free to vary, there would be 60 estimated parameters (6 location parameters that are free to vary  $\times$  10 surveys = 60 parameters).

The second method, termed the “all parameters common” method, entails fitting the mountain model to the pooled data from the resampled surveys of an individual data set. With this method, only 6 parameters are estimated ( $g, OP_e, F_{hm}, TA_{max}, TA_{min}, a$ ). If across-survey drift is small, this method may be more appropriate than the first method due to the reduced number of parameters.

To determine which method (the location-specific vs. all parameters common approach) provided the best fit to the data, the AIC statistic (Akaike information criterion) was calculated. The AIC value balances the goodness of fit with the number of parameters whereby each additional parameter penalizes the AIC value. In all rats tested in the experiments presented in this thesis, according to the AIC value, the location-specific approach was best.

#### ***4-dimensional model***

The 4-dimensional model fit to the data described in Chapter 3 is expressed as:

$$TA = (TA_{\max} - TA_{\min}) \frac{\left( \frac{(F_b (\ln(1 + e^{\frac{F_{u\max}}{F_b}}) - \ln(1 + e^{\frac{F_{u\max}-F}{F_b}}))) \times (I_P - I_0))^g}{(F_b (\ln(1 + e^{\frac{F_{u\max}}{F_b}}) - \ln(1 + e^{\frac{F_{u\max}-F}{F_b}}))) \times (I_P - I_0))^g + S_{hmin}} \right)^a}{\left( \frac{(F_b (\ln(1 + e^{\frac{F_{u\max}}{F_b}}) - \ln(1 + e^{\frac{F_{u\max}-F}{F_b}}))) \times (I_P - I_0))^g}{(F_b (\ln(1 + e^{\frac{F_{u\max}}{F_b}}) - \ln(1 + e^{\frac{F_{u\max}-F}{F_b}}))) \times (I_P - I_0))^g + S_{hmin}} \right)^a + \left( \frac{SP_{\min} - OP}{SP_e} + (SP_{\min} + (SP_{bnd} \times \ln(1 + e^{\frac{SP_{\min} - OP}{SP_{bnd}}})) \right)} \right)} + TA_{\min}$$

where,

$a$ : the price-sensitivity exponent

$F_b$ : the parameter controlling the abruptness of the transition between the range of perfect frequency following to the range of frequency roll-off

$F_{max}$ : the pulse frequency that produces near maximal firing frequency (Hz)

$I_0$ : the ineffective current; the current that just suffices to recruit the neuron nearest to the border of the scar tissue surrounding the electrode tip ( $\mu\text{A}$ )

$Sh_{min}$ : the ratio of the total number of spikes per second and the number of recruited neurons per  $\mu\text{A}$  required to produce a half-maximal reward ((spikes  $\times$   $\mu\text{A}$ )/(seconds  $\times$  neurons))

$TA_{max}$ : the maximal time allocation

$TA_{min}$ : the minimal time allocation

$SP_{bnd}$ : the parameter controlling the abruptness of the change between the flat portion and the rising (scalar) portion of the subjective price function curve

$SP_{min}$ : the minimum subjective price (s)

$SP_e$ : the subjective price at which the rat allocates half of its time to the pursuit of electrical stimulation (s)

The fitting approach was carried out by generating 1000 resampled data sets, in the same manner as the fit of the 3-dimensional model (described above). However, only the “all parameters common” method was employed. In total, there were 8 parameters ( $a$ ,  $F_b$ ,  $F_{max}$ ,  $I_0$ ,  $Sh_{min}$ ,  $TA_{max}$ ,  $TA_{min}$ ,  $SP_e$ ) that were estimated. Parameters  $SP_{min}$  and  $SP_{bnd}$  were set to fixed values that were estimated in previous subjective price analyses. (That is, each rat, had unique values of  $SP_{min}$  and  $SP_{bnd}$ .)

### An example of the resampling procedure

A single survey consists of a (daily) test of 3 sampling matrices (pulse-frequency sampling matrix @ 4 s, price-sampling matrix, radial-sampling matrix). The time allocation corresponding to each element of a sampling matrix is measured. A complete experiment consists of 10 surveys. The data from each of the 10 surveys make-up the pool of data from which we resample.

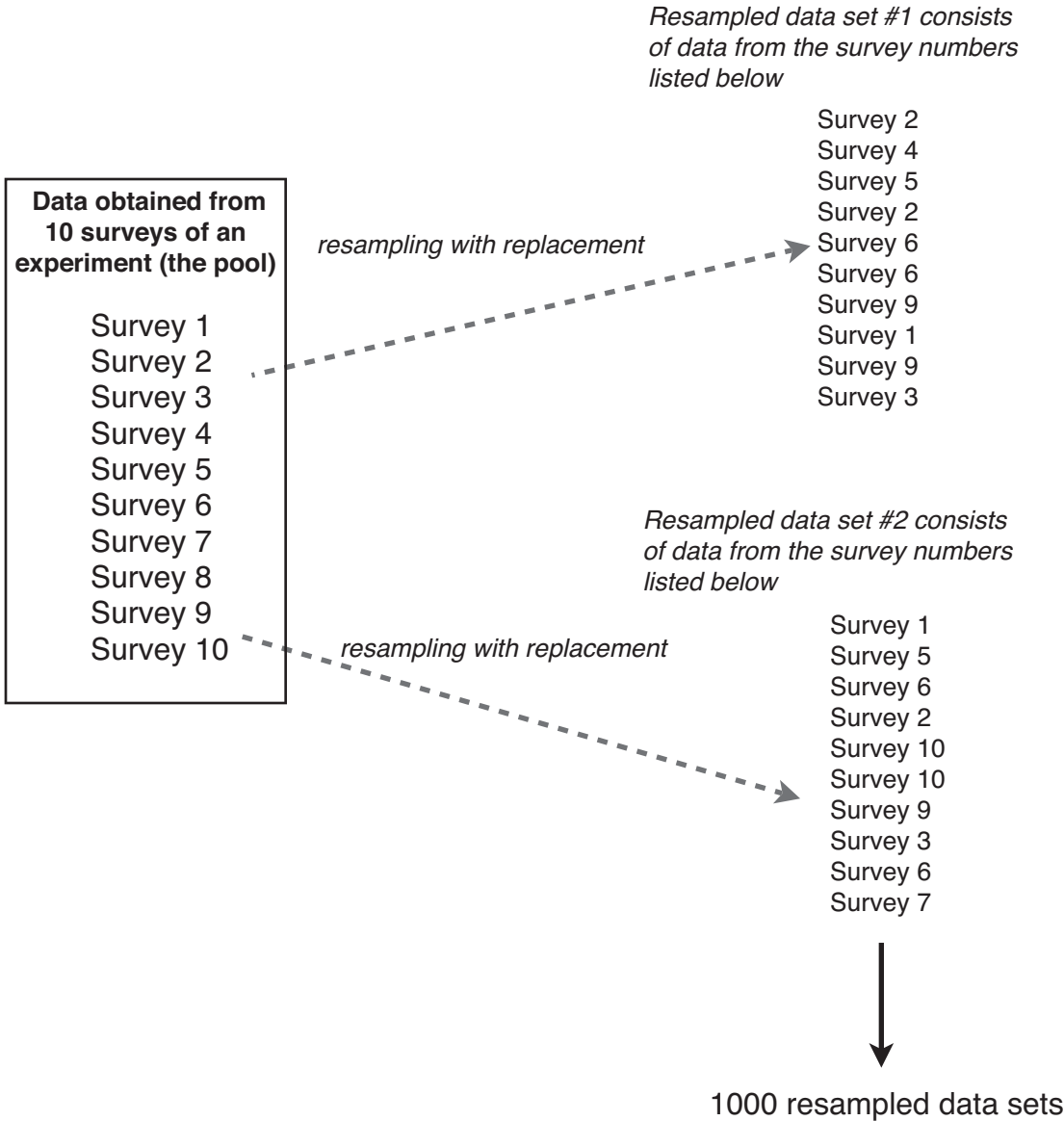
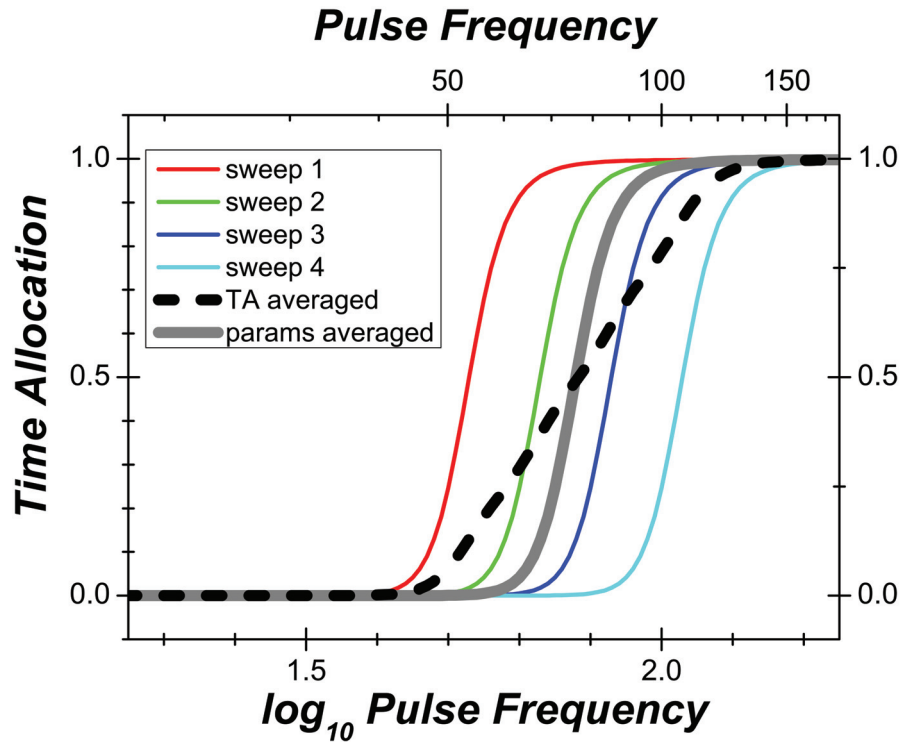


Figure ApD.1. An example of the resampling approach. A single resampled data set consists of data from 10 resampled surveys from the pool. One thousand resampled data sets are generated.



*From Hernandez et al. 2010*

Figure ApD.2. Fitting the function to the data from each individual survey and subsequently averaging the parameters yields a curve (denoted by a thick grey line) with a slope representative of the individual curves. Fitting the function to the averaged time allocation values produces a curve with a slope (black dotted curve), not representative of the slopes of the individual curves.

***Appendix E:***

***The 3-dimensional sections of the 4-dimensional model for all rats***



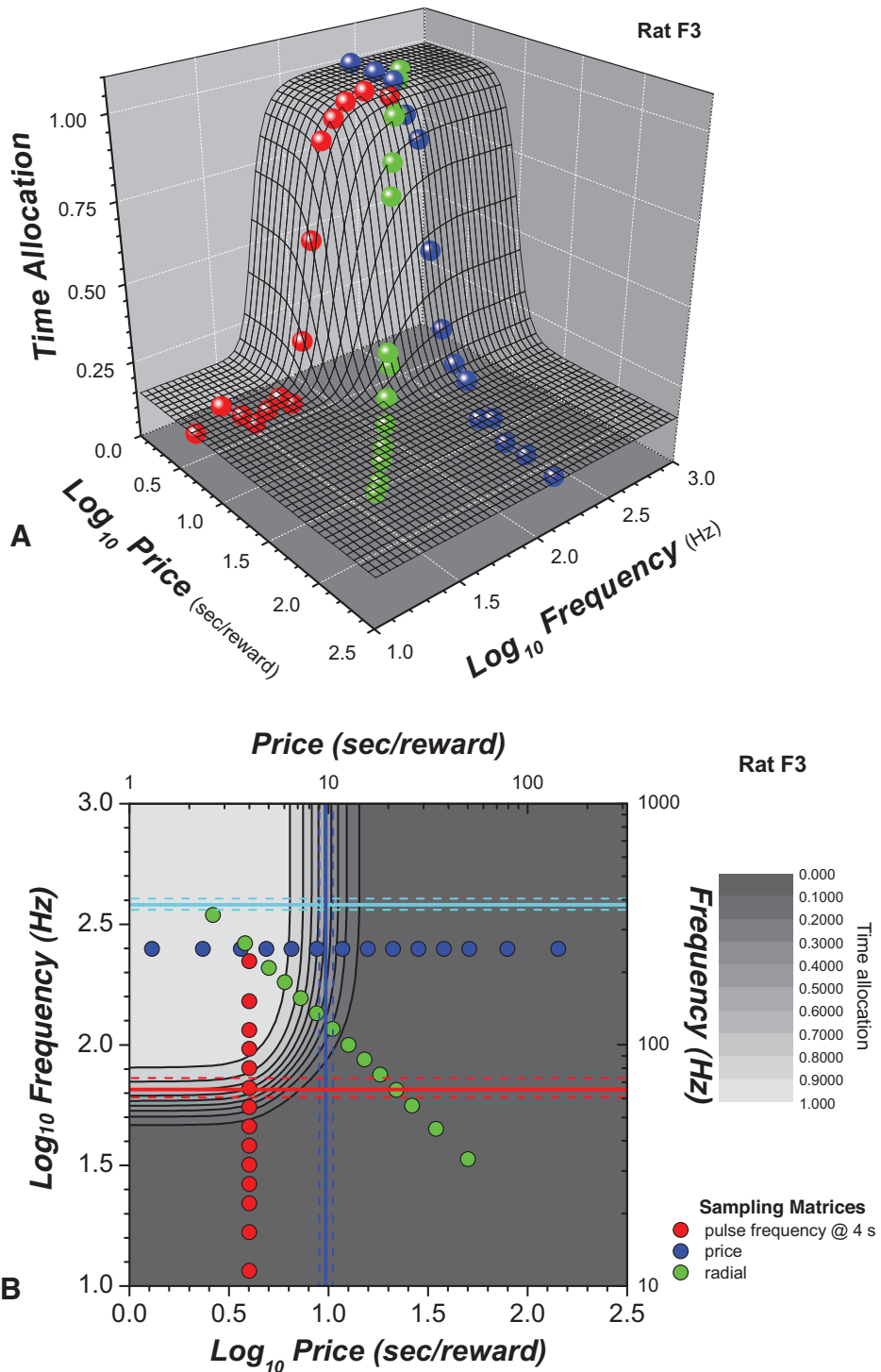


Figure ApE1. For rat F3, a 3-dimensional section (time allocation as a function of price and pulse frequency) of the 4-dimensional reward model. B. The corresponding contour plot. The vertical blue line represents  $OP_e$ , the horizontal cyan line represents  $F_{nmax}$ , the horizontal red line represents  $F_{hm}$ . The dashed lines represent the 95% confidence intervals.

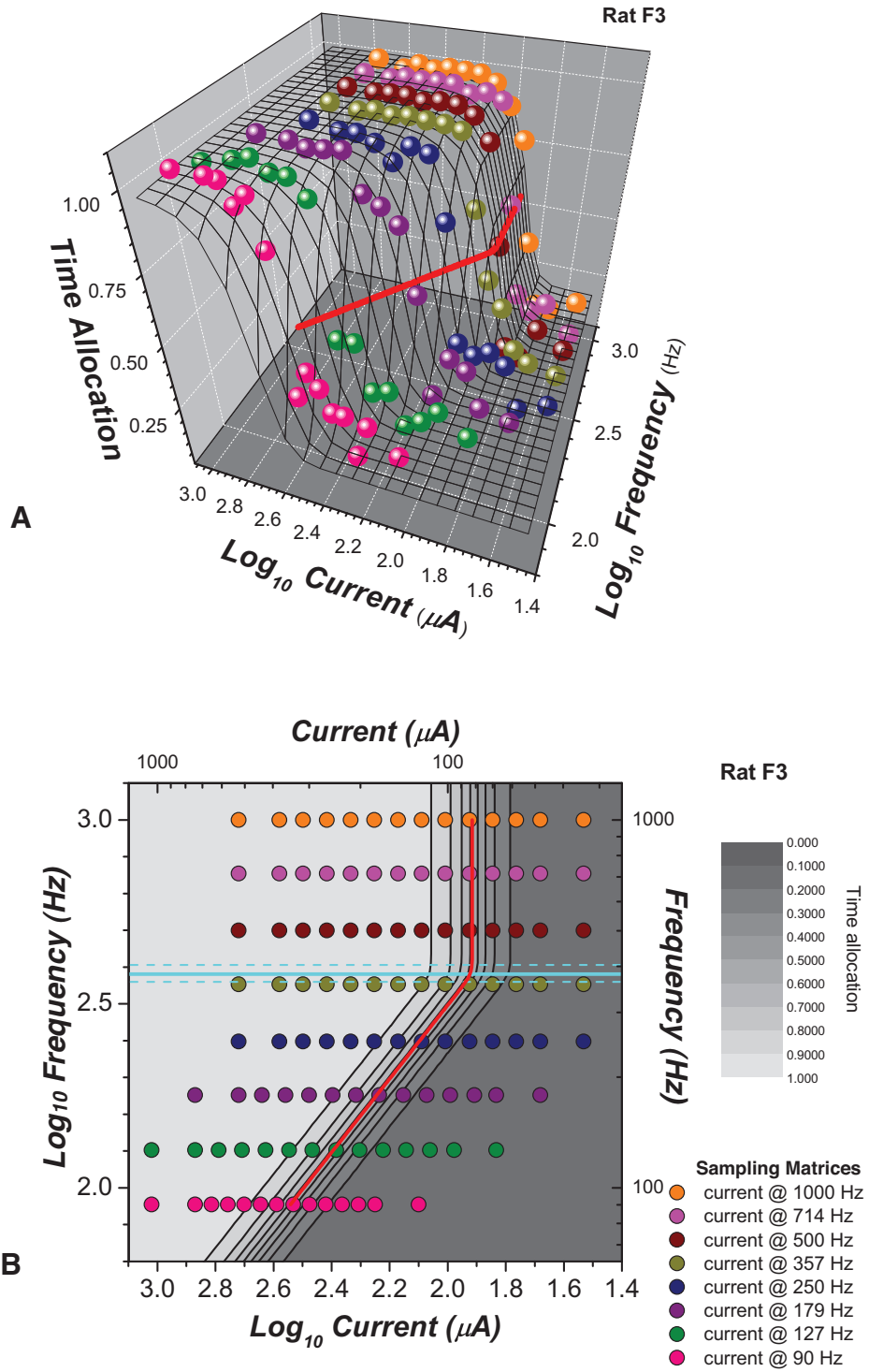


Figure ApE.2. For rat F3, a 3-dimensional section (time allocation as a function of current and pulse frequency) of the 4-dimensional reward mountain model. B. The corresponding contour plot. The horizontal cyan line represents  $F_{nmax}$ . The dashed lines represent the 95% confidence intervals.

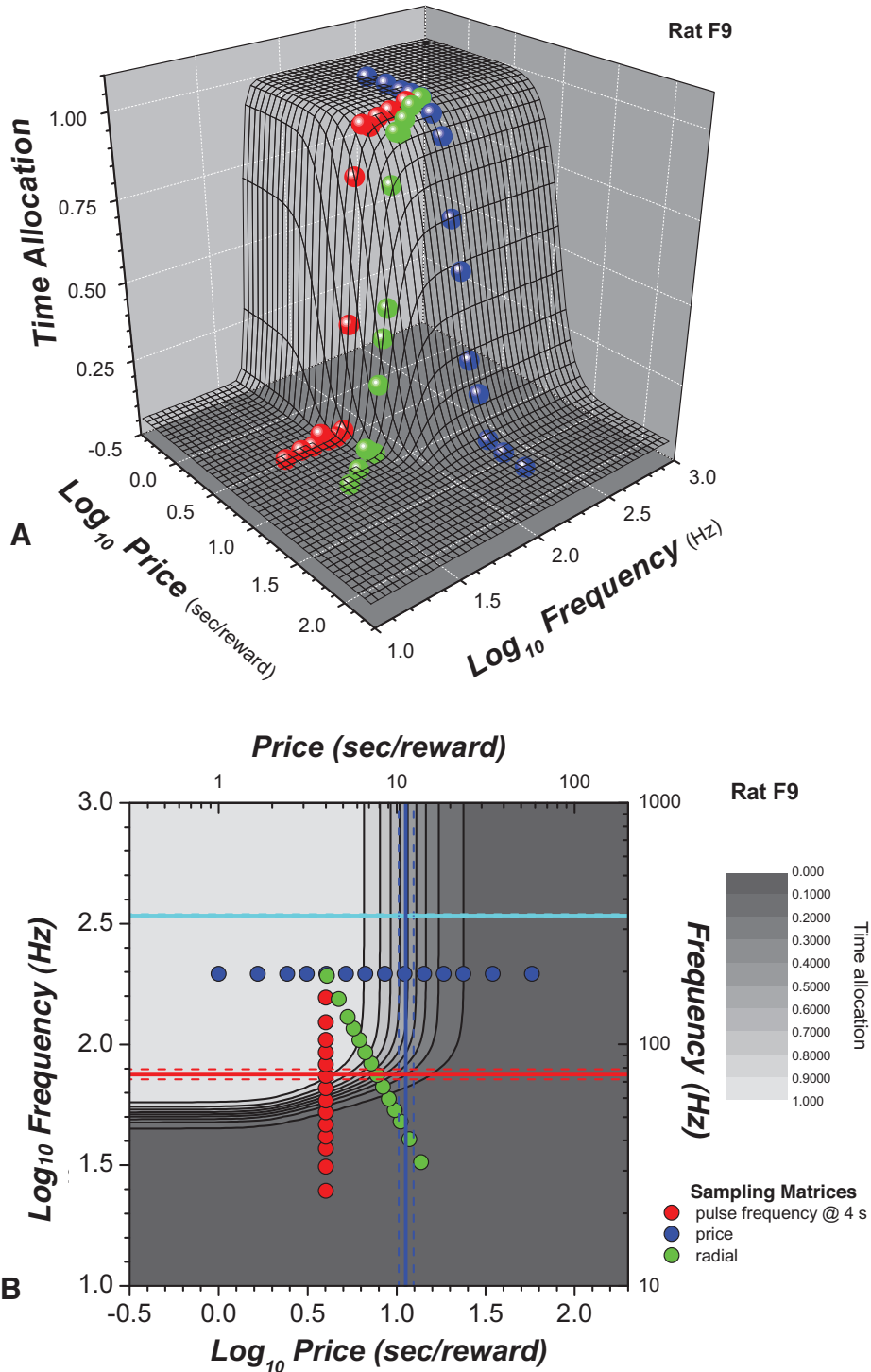


Figure ApE.3. For rat F9, a 3-dimensional section (time allocation as a function of price and pulse frequency) of the 4-dimensional reward model. B. The corresponding contour plot. The vertical blue line represents  $OP_e$ , the horizontal cyan line represents  $F_{nmax}$ , the horizontal red line represents  $F_{hm}$ . The dashed lines represent the 95% confidence intervals.

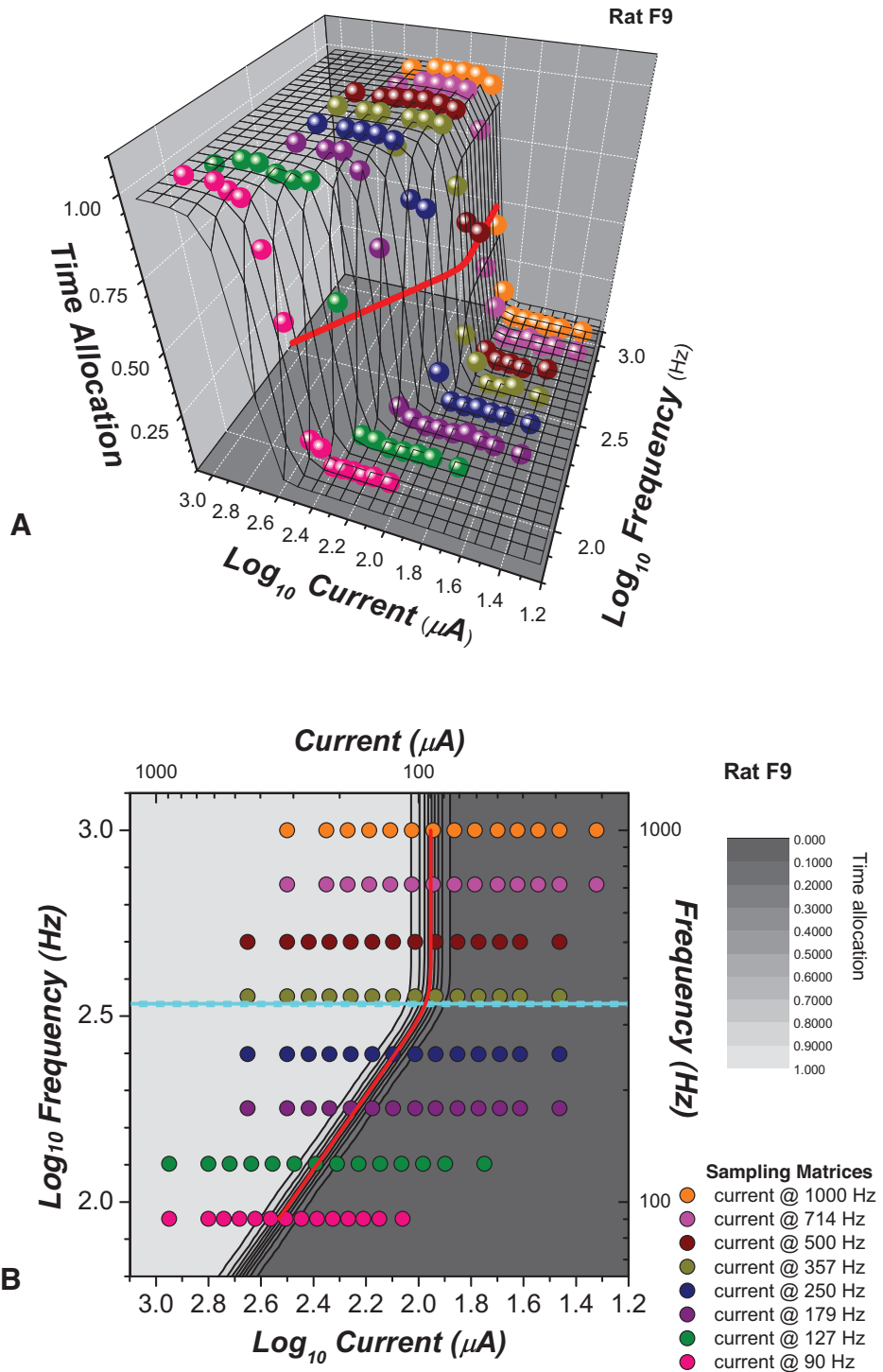


Figure ApE.4. For rat F9, a 3-dimensional section (time allocation as a function of current and pulse frequency) of the 4-dimensional reward mountain model. B. The corresponding contour plot. The horizontal cyan line represents  $F_{nmax}$ . The dashed lines represent the 95% confidence intervals.

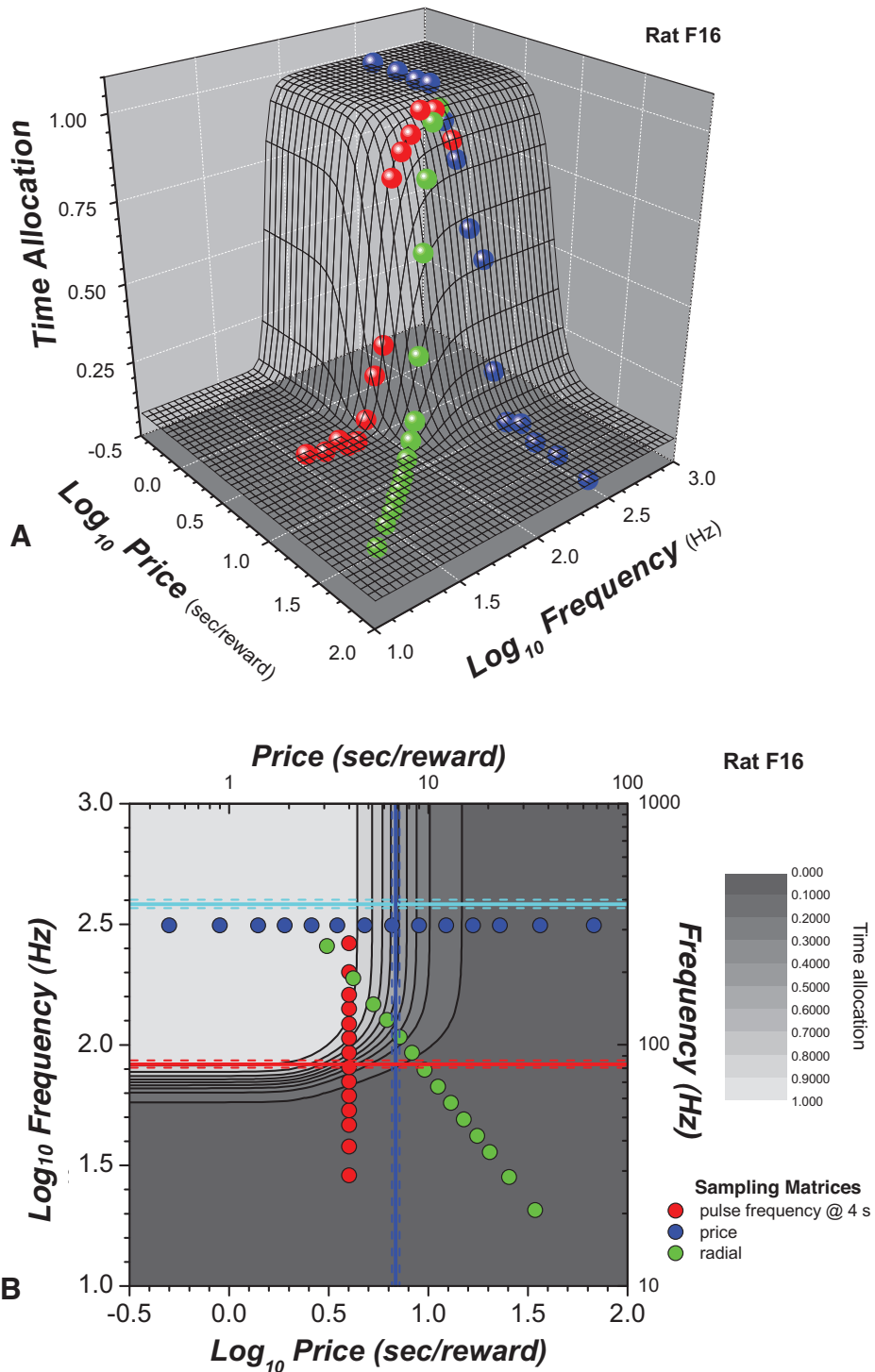


Figure ApE.5. For rat F16, a 3-dimensional section (time allocation as a function of price and pulse frequency) of the 4-dimensional reward model. B. The corresponding contour plot. The vertical blue line represents  $OP_e$ , the horizontal cyan line represents  $F_{nmax}$ , the horizontal red line represents  $F_{hm}$ . The dashed lines represent the 95% confidence intervals.

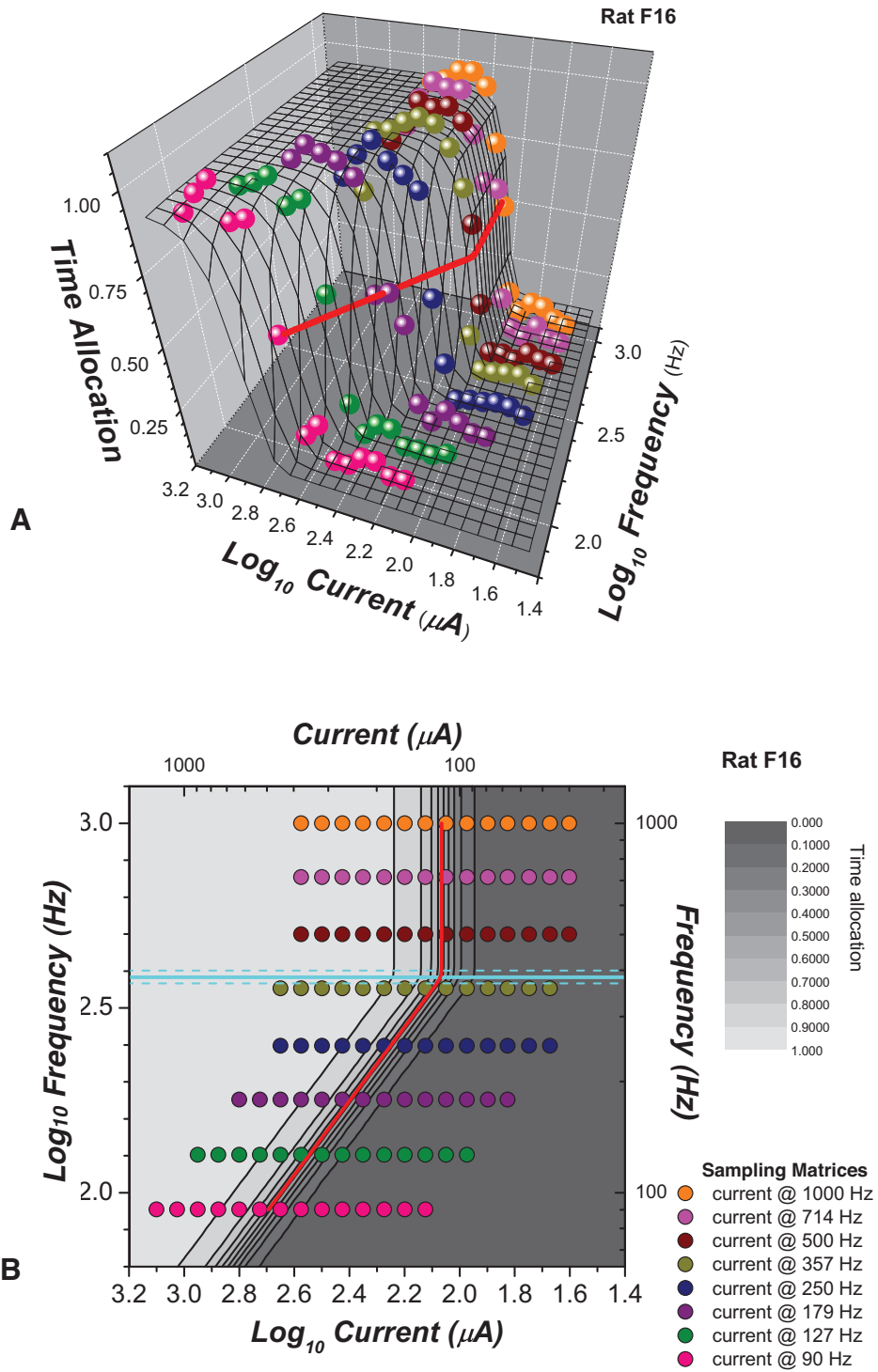


Figure ApE.6. For rat F16, a 3-dimensional section (time allocation as a function of current and pulse frequency) of the 4-dimensional reward mountain model. B. The corresponding contour plot. The horizontal cyan line represents  $F_{nmax}$ . The dashed lines represent the 95% confidence intervals.

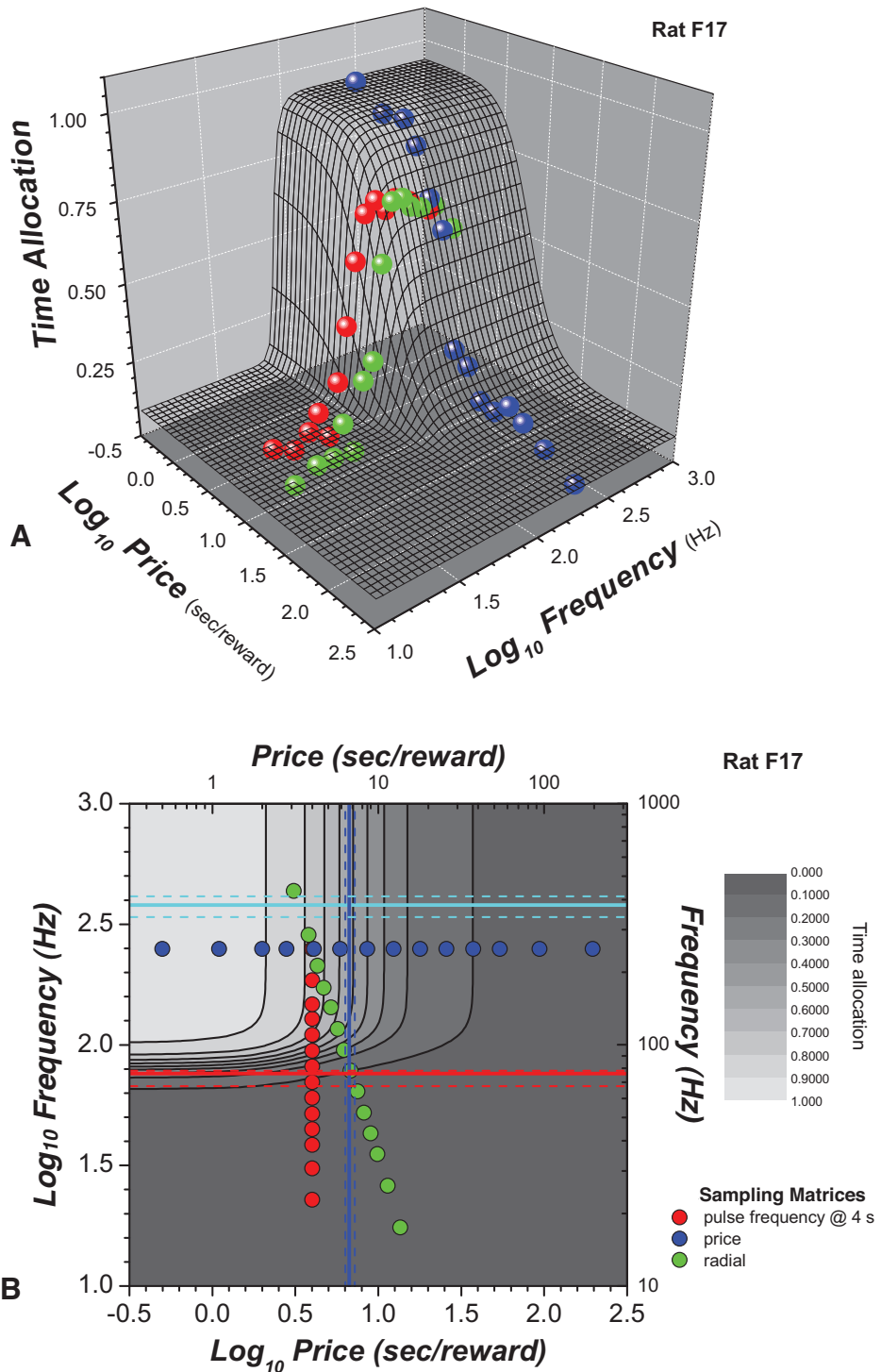


Figure ApE.7. For rat F17, a 3-dimensional section (time allocation as a function of price and pulse frequency) of the 4-dimensional reward model. B. The corresponding contour plot. The vertical blue line represents  $OP_e$ , the horizontal cyan line represents  $F_{nmax}$ , the horizontal red line represents  $F_{hm}$ . The dashed lines represent the 95% confidence intervals.

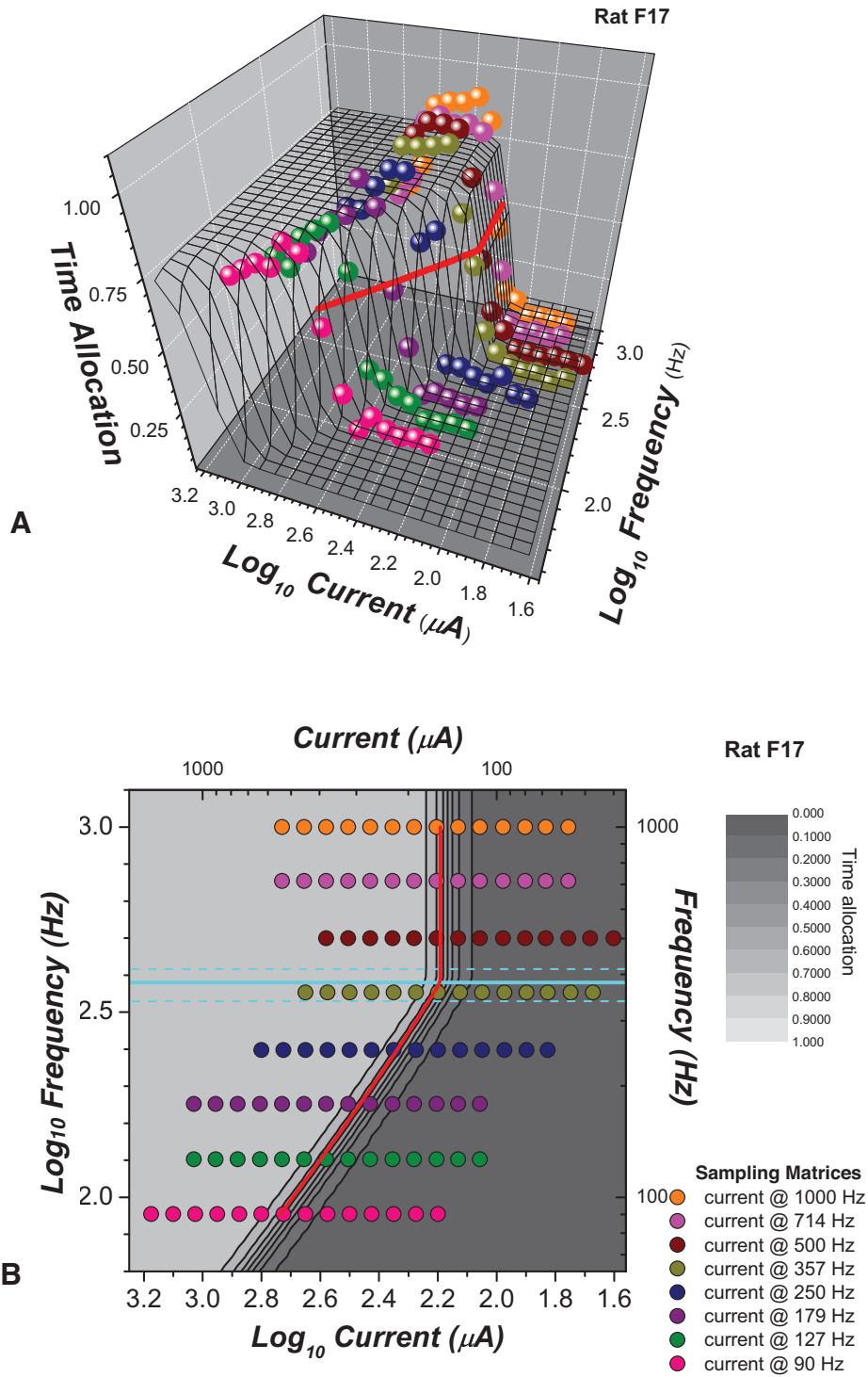


Figure ApE.8. For rat F17, a 3-dimensional section (time allocation as a function of current and pulse frequency) of the 4-dimensional reward mountain model. B. The corresponding contour plot. The horizontal cyan line represents  $F_{nmax}$ . The dashed lines represent the 95% confidence intervals.



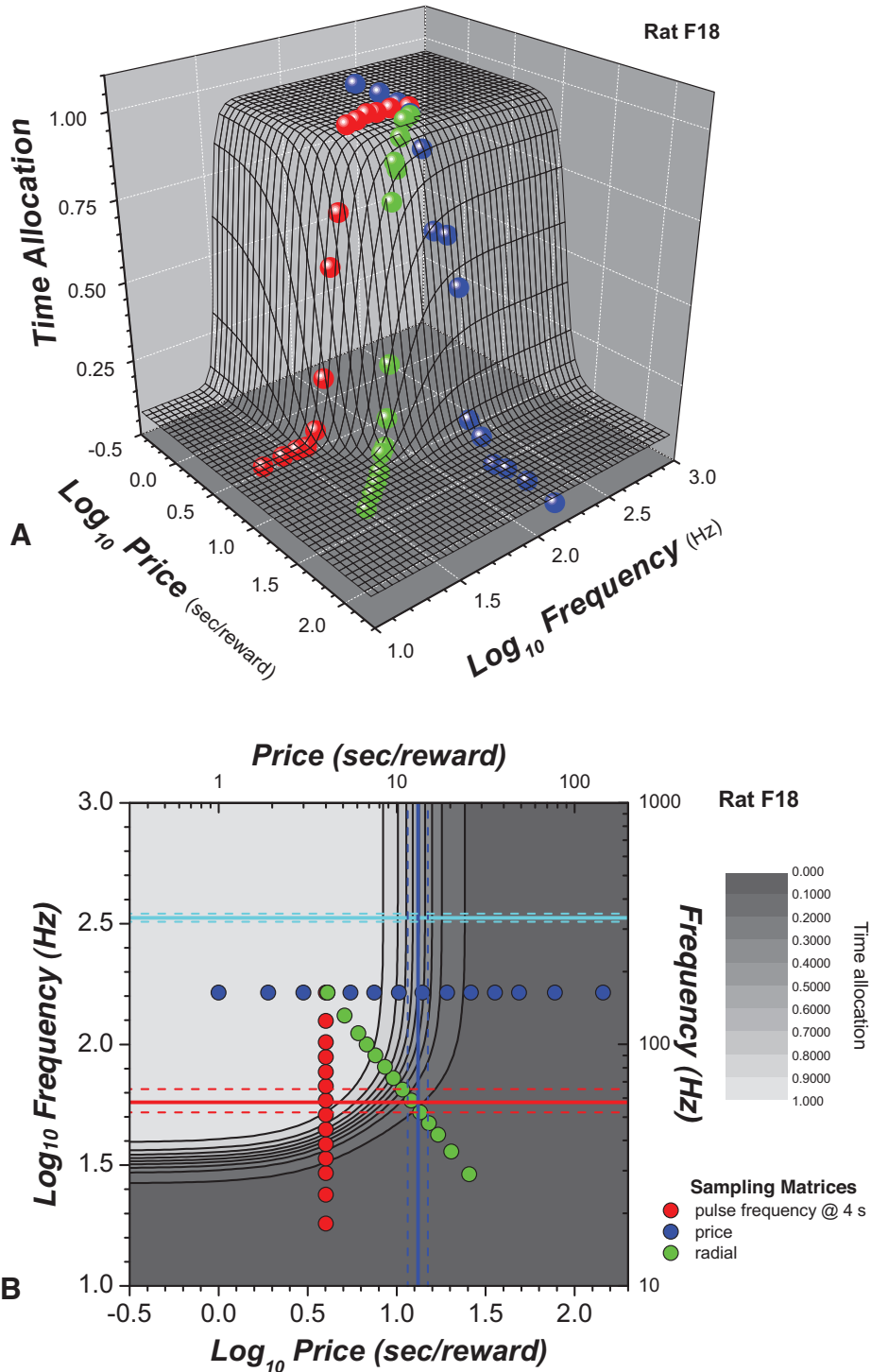


Figure ApE.9. For rat F18, a 3-dimensional section (time allocation as a function of price and pulse frequency) of the 4-dimensional reward model. B. The corresponding contour plot. The vertical blue line represents  $OP_e$ , the horizontal cyan line represents  $F_{nm}$ , the horizontal red line represents  $F_{hm}$ . The dashed lines represent the 95% confidence intervals.

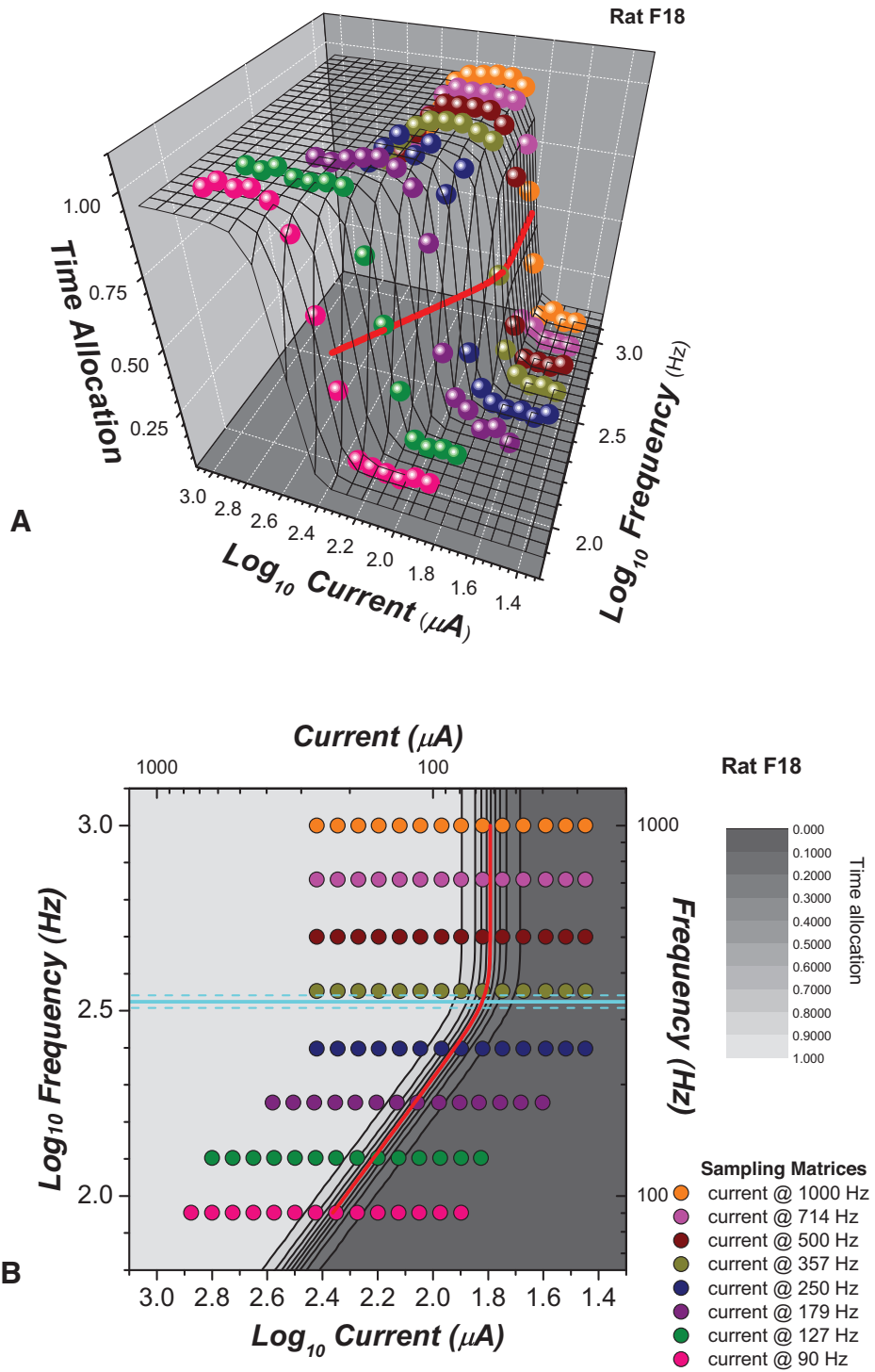


Figure ApE.10. For rat F18, a 3-dimensional section (time allocation as a function of current and pulse frequency) of the 4-dimensional reward mountain model. B. The corresponding contour plot. The horizontal cyan line represents  $F_{nmax}$ . The dashed lines represent the 95% confidence intervals.

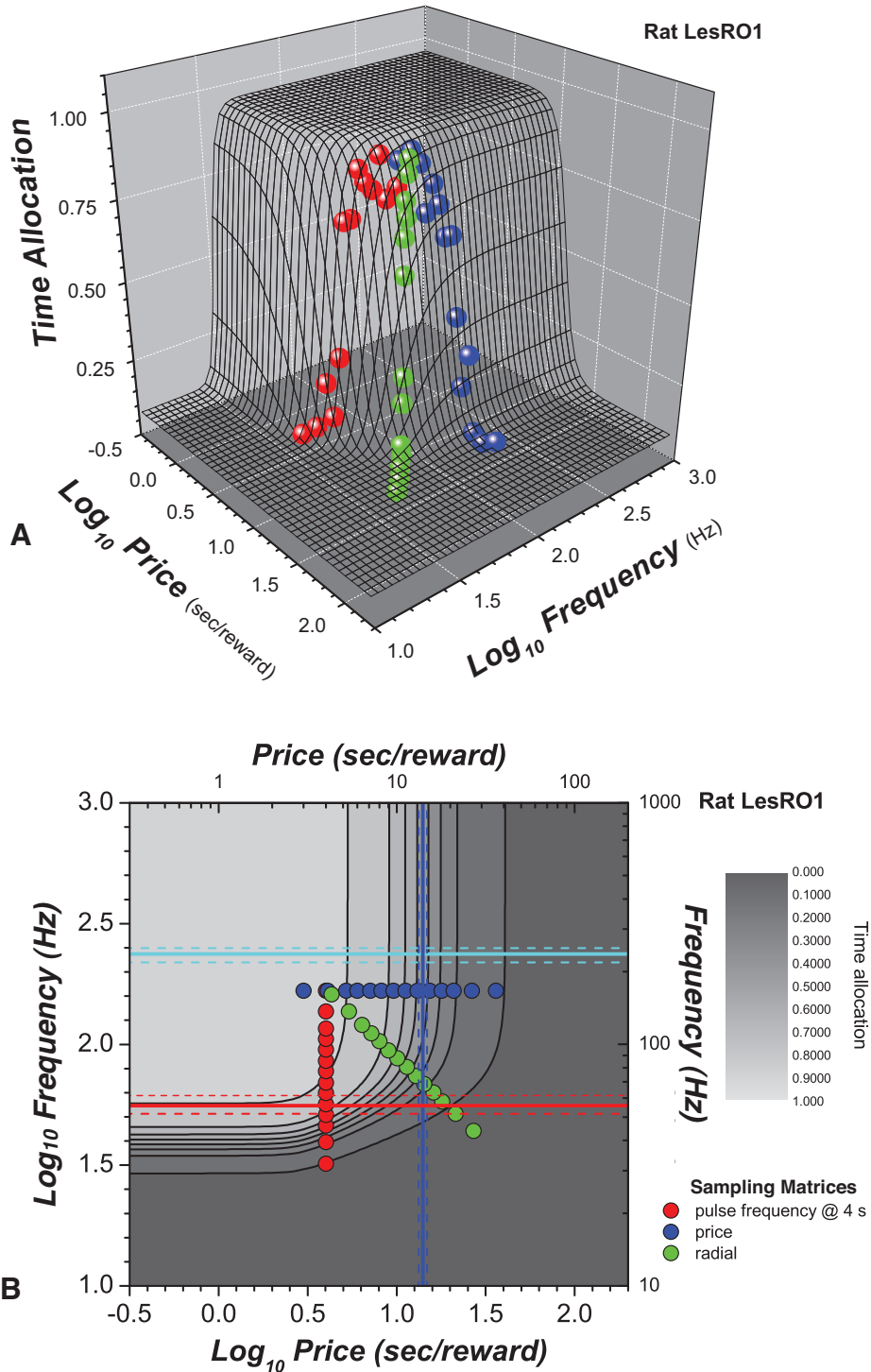


Figure ApE.11. For rat LesRO1, a 3-dimensional section (time allocation as a function of price and pulse frequency) of the 4-dimensional reward model. B. The corresponding contour plot. The vertical blue line represents  $OP_e$ , the horizontal cyan line represents  $F_{nmax}$ , the horizontal red line represents  $F_{hm}$ . The dashed lines represent the 95% confidence intervals.

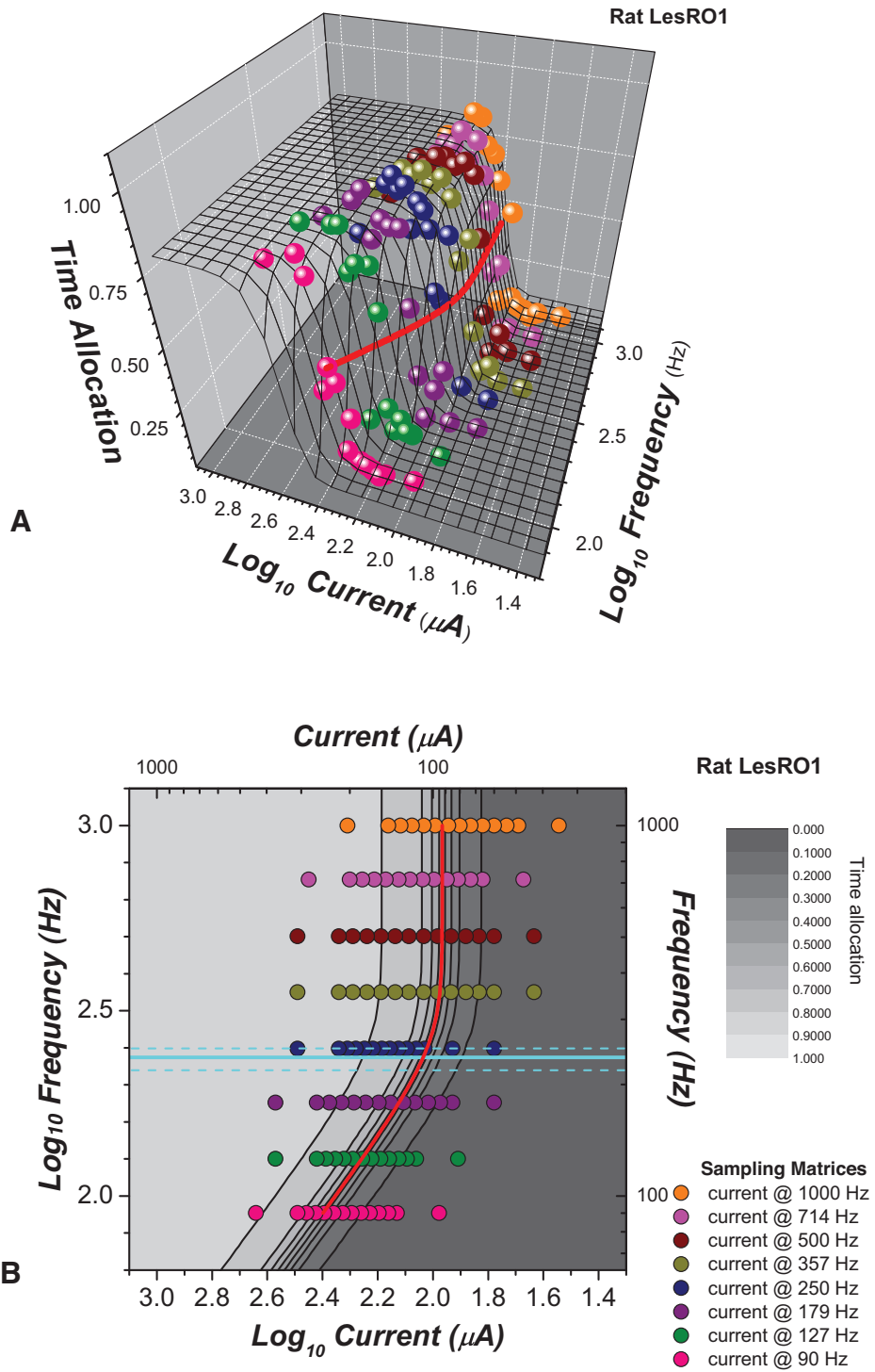


Figure ApE.12. For rat LesRO1, a 3-dimensional section (time allocation as a function of current and pulse frequency) of the 4-dimensional reward mountain model. B. The corresponding contour plot. The horizontal cyan line represents  $F_{nmax}$ . The dashed lines represent the 95% confidence intervals.

***Appendix F:***

***The parameter values of the 4-dimensional model for all rats***

**Table ApF.1**

Rat F3								
	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
A	5.55035	5.15229	5.99767	0.84538	0.39806	0.44732	1	20
F <sub>b</sub>	10.18933	1.00001	53.99854	52.99853	9.18932	43.80921	1	1000
LogF <sub>nmax</sub>	2.58098	2.55988	2.60602	0.04614	0.0211	0.02503	1	3
G	2.66615	2.30137	3.25642	0.95506	0.36478	0.59027	1	20
LogI <sub>0</sub>	-4.99884	-3.84154	-3.84154	0	-1.1573	1.1573	-5	3
LogSP <sub>e</sub>	0.98583	0.95403	1.02279	0.06875	0.03179	0.03696	0	3
LogS <sub>hm</sub>	4.51443	4.48171	4.56285	0.08113	0.03271	0.04842	1	6
TA <sub>max</sub>	1	0.99999	1	0.00001	0.00001	0	0.5	1
TA <sub>min</sub>	0.14752	0.1391	0.15611	0.017	0.00842	0.00859	0	0.5

**Table ApF.2**

Rat F9								
	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
A	3.97283	3.34181	5.19815	1.85634	0.63102	1.22532	1	20
F <sub>b</sub>	22.32744	1.001	34.63579	33.63479	21.32644	12.30835	1	1000
LogF <sub>nmax</sub>	2.53373	2.52651	2.53899	0.01248	0.00723	0.00526	1	3
G	6.21196	5.10514	11.35535	6.25021	1.10682	5.14339	1	20
LogI <sub>0</sub>	-4.99462	0.02498	0.30536	0.28038	-5.01959	5.29997	-5	3
LogSP <sub>e</sub>	1.05026	1.01134	1.0947	0.08335	0.03892	0.04444	0	3
LogS <sub>hm</sub>	4.52845	4.50718	4.54861	0.04143	0.02127	0.02015	1	6
TA <sub>max</sub>	0.99999	0.99591	0.99701	0.0011	0.00408	-0.00297	0.5	1
TA <sub>min</sub>	0.05373	0.04979	0.05827	0.00848	0.00394	0.00454	0	0.5

**Table ApF.3**

Rat F12								
	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
A	3.39885	3.02449	3.82298	0.79848	0.37436	0.42412	1	20
F <sub>b</sub>	20.55646	1.00009	32.244	31.24391	19.55637	11.68754	1	1000
LogF <sub>nmax</sub>	2.55859	2.5453	2.56999	0.02469	0.01329	0.0114	1	3
G	6.5581	5.45054	8.70698	3.25644	1.10756	2.14888	1	20
LogI <sub>0</sub>	-4.96056	0.16669	0.26713	0.10045	-5.12724	5.22769	-5	3
LogSP <sub>e</sub>	1.09564	1.0479	1.14531	0.09741	0.04774	0.04967	0	3
LogS <sub>hm</sub>	4.38972	4.36546	4.41872	0.05326	0.02426	0.029	1	6
TA <sub>max</sub>	0.93775	0.92313	0.95027	0.02713	0.01461	0.01252	0.5	1
TA <sub>min</sub>	0.08312	0.07575	0.09248	0.01673	0.00737	0.00936	0	0.5

**Table ApF.4**

Rat F16								
	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
A	4.77438	4.21574	5.29505	1.07931	0.55864	0.52067	1	20
F <sub>b</sub>	6.28139	1.77759	63.61433	61.83674	4.5038	57.33295	1	1000
LogF <sub>nmax</sub>	2.58313	2.56692	2.60151	0.03459	0.01621	0.01838	1	3
G	4.90426	4.51351	5.33733	0.82382	0.39075	0.43307	1	20
LogI <sub>0</sub>	-4.99998	-4.99906	-4.99499	0.00407	-0.00092	0.00499	-5	3
LogSP <sub>e</sub>	0.83638	0.8181	0.85521	0.03711	0.01828	0.01883	0	3
LogS <sub>hm</sub>	4.61859	4.6041	4.63794	0.03384	0.01449	0.01935	1	6
TA <sub>max</sub>	0.99967	0.97811	1	0.02189	0.02156	0.00032	0.5	1
TA <sub>min</sub>	0.0764	0.07137	0.0826	0.01123	0.00504	0.0062	0	0.5

**Table ApF.5**

Rat F17

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
A	2.36171	2.02514	3.65886	1.63372	0.33657	1.29715	1	20
F <sub>b</sub>	3.49399	2.19379	96.63067	94.43688	1.3002	93.13668	1	1000
LogF <sub>nmax</sub>	2.58033	2.5294	2.61611	0.0867	0.05093	0.03578	1	3
G	8.87163	7.69477	20	12.30523	1.17686	11.12837	1	20
LogI <sub>o</sub>	1.54249	-1.08297	1.63059	2.71355	2.62546	0.08809	-5	3
LogSP <sub>e</sub>	0.82532	0.80062	0.85759	0.05696	0.0247	0.03227	0	3
LogS <sub>hm</sub>	4.6337	4.60745	4.66201	0.05456	0.02625	0.02831	1	6
TA <sub>max</sub>	0.97186	0.82077	1	0.17923	0.15109	0.02814	0.5	1
TA <sub>min</sub>	0.08477	0.07643	0.09302	0.01658	0.00834	0.00824	0	0.5

**Table ApF.6**

Rat F18

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
A	5.88592	4.4587	7.26925	2.81055	1.42722	1.38333	1	20
F <sub>b</sub>	31.38301	1.42266	50.46804	49.04537	29.96035	19.08502	1	1000
LogF <sub>nmax</sub>	2.5245	2.50755	2.54172	0.03417	0.01695	0.01722	1	3
G	3.29755	2.68018	5.00318	2.32299	0.61737	1.70563	1	20
LogI <sub>o</sub>	-4.9999	-4.99993	-4.97038	0.02955	0.00003	0.02953	-5	3
LogSP <sub>e</sub>	1.11983	1.06174	1.17604	0.11429	0.05809	0.0562	0	3
LogS <sub>hm</sub>	4.41408	4.37195	4.46977	0.09782	0.04213	0.05569	1	6
TA <sub>max</sub>	0.95952	0.94723	0.97394	0.02671	0.01229	0.01442	0.5	1
TA <sub>min</sub>	0.07545	0.0705	0.08165	0.01115	0.00494	0.0062	0	0.5

**Table ApF.7**

Rat LesRO1

	Estimate	CB Lo	CB Hi	CB Width	Ex Lo	Ex Hi	MB Lo	MB Hi
A	3.8212	3.37212	4.38458	1.01246	0.44907	0.56338	1	20
F <sub>b</sub>	50.27374	39.49126	62.17827	22.68701	10.78248	11.90454	1	1000
LogF <sub>nmax</sub>	2.37412	2.33913	2.3984	0.05927	0.035	0.02428	1	3
G	4.08223	3.59446	4.73072	1.13626	0.48777	0.64849	1	20
LogI <sub>o</sub>	-4.99718	-4.99975	-4.52519	0.47456	0.00256	0.47199	-5	3
LogSP <sub>e</sub>	1.14701	1.12276	1.17083	0.04807	0.02424	0.02382	0	3
LogS <sub>hm</sub>	4.43918	4.41626	4.45566	0.0394	0.02292	0.01648	1	6
TA <sub>max</sub>	0.81762	0.79937	0.8336	0.03423	0.01826	0.01597	0.5	1
TA <sub>min</sub>	0.0877	0.07915	0.0979	0.01875	0.00855	0.0102	0	0.5

Table ApF.1-7. Parameter values for all rats. CB Lo indicates the lower 95% confidence band (limit). CB hi indicates the higher 95% confidence band (limit). CB Width indicates the interval from the lower to the higher confidence bounds. Ex Lo indicates the lower 95% confidence error (interval). Ex high indicates the higher 95% confidence error. MB Lo indicates the lower model bound: the lower limit set in the model. MB hi indicates the higher model bound.

## *Chapter 4*

### *Three-dimensional analysis of the effects of a lesion challenge to the rewarding effect of electrical brain stimulation*

#### **Abstract**

Rats will work to obtain electrical stimulation of various brain sites. The effect that leads the rat to seek out the stimulation is called “brain stimulation reward” (BSR). Extensively studied BSR sites are arrayed along the medial forebrain bundle (MFB) which contains dozens of different fiber populations. Although many characteristics of the directly stimulated (“first-stage”) fibers underlying BSR have been described, it is as yet unknown exactly which MFB fibers are responsible for the rewarding effect. Lesions are among the many methods that have been marshaled to meet this challenge. We report a 3-dimensional, psychophysical measurement method that can link lesion-induced changes in intracranial self-stimulation to particular stages of neural processing and can thus help identify first-stage fibers. In this method, operant performance for BSR is measured as a function of both the strength (pulse frequency) and opportunity cost (work time required to earn a reward) of the rewarding electrical stimulation. The dependent variable is time allocated to holding down the lever. A 3-dimensional surface is obtained by fitting the reward-mountain model to these data. One location parameter of this model,  $F_{hm}$ , sets the position of the 3-dimensional surface along the pulse-frequency axis whereas the other,  $OP_e$ , sets the position along the opportunity cost axis. Changes in  $F_{hm}$  result from perturbations of components that precede the output of the reward-intensity



function that translates the induced volley of action potentials into the intensity of BSR. These early-stage components include the first-stage neurons. Changes in  $OP_e$  result from perturbations of components beyond the output of the reward-intensity function (after the computation of the rewarding effect has occurred). These components entail processes such as the evaluation of subjective effort, opportunity costs, and competing rewards.

The model predicts that a lesion that damages first-stage fibers will increase  $F_{hm}$ . To test this hypothesis, we made small electrolytic lesions through the stimulation electrode. In 6 out of 7 subjects, the value of  $F_{hm}$  increased, as predicted. These increases were often accompanied by decreases in  $OP_e$  that may reflect the consequences of multiple integrators, and/or damage to local dopaminergic fibers excited transynaptically by the stimulation. The results from this proof-of-principle experiment demonstrate that the 3-dimensional reward model can play a useful role in the quest to identify the first-stage neurons subserving BSR. Namely, the 3-dimensional reward model and testing paradigm can be used in conjunction with more specific lesion techniques as well as with optogenetic silencing methods.

## ***Introduction***

A long-term goal of BSR research is the identification of neurons mediating lateral hypothalamic self-stimulation. One model of reward circuitry is the “descending path hypothesis” which proposes that the directly stimulated reward fibers of the medial forebrain bundle (MFB) originate in the basal forebrain and descend through the MFB. Several lines of research have led to this hypothesis. For example, anatomical studies show that the MFB extends from the forebrain to the VTA and traverses the lateral hypothalamus (LH) (Nieuwenhuys, Geeraedts & Veening, 1982). Bielajew and Shizgal (1986) demonstrated through the use of a psychophysical adaptation of the hyperpolarization-block technique that at least some of the reward-relevant neurons project in the rostral-caudal direction linking the lateral hypothalamus (LH) with the VTA. Together, Nieuwenhuys’ and Bielajew and Shizgal’s results suggest the possibility that the neurons in question arise from the forebrain structures. Accordingly, single unit recording studies have shown that several forebrain nuclei have electrophysiological properties that are consistent with the properties of the first-stage neurons (Bielajew, Shizgal & 1982; Murray & Shizgal, 1996a; Murray & Shizgal, 1996b; Rompre & Shizgal, 1986; Shizgal, Shindler, Rompre, 1989). Taken together, these data suggest that there is a strong possibility that the first-stage neurons originate in the forebrain areas.

Lesions are among the many techniques that have been employed to investigate the identity of the first-stage neurons. The destruction of a “candidate” brain area and then subsequent assessment of the behaviour under study (e.g. bar pressing for a reward) tests the necessity of the destroyed brain area for that particular behaviour. Although lesions to several brain areas have been shown to reduce the reward efficacy of the

electrical stimulation, the effects have been variable. Early lesion studies in which the rostral sources of MFB fibers were destroyed often failed to significantly and consistently alter reward efficacy of more caudal MFB sites (e.g., Boyd & Gardner, 1967; Huston, Orenstein, Lehner, 1982; Olds & Olds, 1969; Lorens, 1966). However, these early studies compared self-stimulation rates collected at a single combination of stimulation parameters, thus rendering the results difficult to interpret due to the limitations of the 1-dimensional approach (discussed in Chapter 1). More recent studies using the rate-frequency curve shift method developed by Gallistel (1978) have largely confirmed the conclusions of early work. For instance, Colle and Wise (1987) destroyed the forebrain via suction ablations and noted only small effects or no effects at all. Furthermore, in experiments in which the lesion was more specific (reviewed below), the effects of lesions to the forebrain areas have been either variable within and across studies, not long-lasting or unsubstantial, casting doubt on the validity of the descending path hypothesis.

Selected findings of the BSR-lesion literature will be reviewed as the basis for conducting the present proof-of-principle experiment that validates the 3-dimensional reward mountain methodology for detecting lesion-induced effects. The 3-dimensional approach has the potential to reconcile several controversies noted in past lesion studies. Importantly, it also provides a means to link a candidate brain area with a stage of the circuitry implicated in reward seeking, an inference not previously feasible with the rate-frequency curve shift method.

#### ***4.1. Measuring and interpreting lesion effects***

The rate-frequency curve is plotted as the behavioural response (such as the number of bar presses) as a function of pulse frequency (while the current is held constant). Under the assumption of the counter model, the magnitude of lateral shifts is proportional to the lesion-induced reduction in reward-relevant neurons. The pulse frequency required to maintain half-maximal responding is the index of measurement for reward efficacy; it is referred to as the “pulse-frequency threshold” and denoted as  $F_{M50}$ . Changes in reward efficacy are reported as changes in pulse-frequency thresholds from baseline. Typically, these pulse-frequency threshold changes from baseline are reported in common logarithmic units. Increases less than 0.1 log units are considered negligible, attributed to day to day variability, increases from 0.1 to 0.2 log units are considered small and of interest, whereas increases greater than 0.2 log units are considered moderate to large. Further discussion of the meaning of pulse frequency thresholds on a logarithmic scale can be found in Appendix A.

#### ***4.2. The effect of lesions to forebrain and hypothalamic regions on ICSS***

In an extensive study, Waraczynski (1988) made unilateral coronal knife cuts to transect many MFB projection systems and measured the pulse frequency thresholds of rewarding electrical stimulation of the lateral hypothalamus (LH) during approximately 2 weeks following the cut. Overall, the effects on the pulse-frequency threshold were variable and not pronounced for the various areas that were transected. For rats in which knife cuts transected the diagonal band of Broca and septal region, the pulse frequency threshold increases were on average small to moderate, only around 0.1 log units to 0.2

log units; most of these increases were transient. For subjects with cuts in the lateral preoptic area (LPO), 4 rats showed small increases in pulse-frequency thresholds (0.1 to 0.2 log units) lasting over a week, 2 rats showed smaller increases (around 0.1 log units), while 4 other subjects with similarly placed cuts showed only transient and variable changes in pulse frequency thresholds. Subjects that had knife cuts in the medial preoptic area (MPO) showed no effects on pulse frequency thresholds or a slight decrease in pulse-frequency thresholds.

The most surprising results of Waraczynski's study (1988) were the rats that received cuts just anterior to the stimulating electrode of the lateral hypothalamus (ALH). If the rewarding effect arises from the forebrain and is carried by the MFB neurons, then transecting the ALH should presumably interrupt most of the transmission from the forebrain and thus substantial changes in the pulse-frequency threshold should be seen. However, only half of these rats showed increases in pulse-frequency threshold; most of these shifts, which were approximately 0.2 log units on average, did not persist beyond one week. The other half of subjects showed no changes in pulse-frequency threshold. The damage specific to the rats with the behaviourally effective lesions was unclear. Further studies (Murray & Shizgal, 1991, 1996c; Gallistel, Leon, Lim, Sim & Waraczynski, 1996) confirmed these mixed effects of ALH lesions. Thus, in Waraczynski's knife cut study, transections of many MFB projection systems did not have a pronounced or consistent effect on the pulse frequency threshold as one might expect if the reward-relevant cell bodies resided in the forebrain areas.

Another surprising result is the relative ineffectiveness of lesions to the nucleus accumbens and ventral pallidum. These areas are probably not the first-stage neurons but have traditionally been implicated in reward modulation given that the nucleus accumbens receives dopaminergic projections from the ventral tegmental area and interacts with the ventral pallidum (Ikemoto, 2010). Johnson and Stellar (1994) damaged the nucleus accumbens (NAC) and ventral pallidum (VP) and the juncture between these two areas using excitotoxic lesions (neurotoxin N-methyl-D-aspartic acid, NMDA) to assess their effects on LH self-stimulation. This type of lesion selectively destroys cell bodies, leaving fibers of passage largely intact. Of the 5 subjects with lesions to the nucleus accumbens, only one displayed a marginal increase (0.13 log units) in pulse-frequency threshold. Of the 5 subjects with lesions to the ventral pallidum, only 2 showed moderate effects (reductions in pulse frequency of 0.2 log units). Of the subjects with lesions to the juncture between the nucleus accumbens and ventral pallidum, there were both moderate increases and decreases in pulse-frequency threshold. Waraczynski and Demco (2006) later confirmed the ineffectiveness of ventral pallidum inactivation on self-stimulation of the LH. Thus, the effects of lesions to these areas are mixed, and not pronounced.

To follow-up on studies by Waraczynski (1988) as well as Janas and Stellar (1987) that detected small to moderate effects of lesions to the LPO, Arvanitogiannis, Waraczynski, and Shizgal (1996) targeted the LPO using excitotoxic lesions (NMDA). Electrolytic lesions and knife cuts do not have the specificity of excitotoxic lesion, thus it was not clear whether the effects of damage to the LPO seen previously were due to fibers of passage passing through the LPO or neurons originating in the LPO. In

Arvanitogiannis and colleagues' (1996) study, the lesions destroyed at least some of the LPO, as well surrounding regions such as the anterior lateral hypothalamus, and subnucleus reticulatus (SLEA) in all rats. Five of the 15 subjects displayed large and long-lasting increases in pulse frequency threshold spanning from 0.2 to more than 0.4 log units above baseline. Seven of the subjects displayed a moderate or transient increase: pulse-frequency threshold increases spanned from 0.15 to 0.4 log units but stabilized to an increase of about 0.1 log units. However, for 2 subjects in which the damage was similar to the damage in the rats who showed an effect, no change in pulse-frequency threshold was detected. Thus, excitotoxic damage rostral to the LH at the LPO and surrounding structures has a relatively stable effect on the pulse frequency threshold as compared to other brain areas.

Arvanitogiannis and colleagues (1996) showed that damage often included the subnucleus reticulatus (SLEA). Expanding on these findings, Waraczynski (2003) further assessed the involvement of the SLEA in MFB self-stimulation. The medial and lateral components of the SLEA and neighbouring areas were temporarily inactivated (via lidocaine). The magnitude of threshold increases ranged from 0.2 to 0.4 log units; on average damage to central structures result in larger threshold increases.

#### ***4.3. The effect of lesions to the mid and hindbrain on ICSS***

The descending path hypothesis led to numerous studies of the effects forebrain lesions on LH and VTA electrical self-stimulation. The studies reviewed above focused on lesions rostral to the stimulation electrode. Given that at least some of the reward-relevant neurons extend from the LH to the VTA, disruption of the neurons *posterior* to

the LH should disrupt the reward processing. Gallistel and colleagues (1996) demonstrated that cuts just outside the anterior border to the VTA or dorsal and somewhat caudal to it tended to have more of an effect than cuts actually transecting the VTA (Waraczynski, 2006). The lack of substantial lesion-induced effects of the VTA was also shown by Janas and Stellar (1987). These findings suggest that although MFB neurons send and receive projections to and from the VTA, there may be a subset of reward-relevant neurons that turn dorsally to continue caudally to mid- and/or hindbrain structures; conversely, the reward-relevant projections could arise from mid- and/or hindbrain structures (Waraczynski, 2006).

Waraczynski's (2006) proposal that a subset of reward-relevant MFB neurons project to or arise from the mid or hindbrain structures is consistent with the findings that many midbrain and hindbrain cell bodies send efferents to and/or receive afferents from the MFB (Nieuwenhuys, Geeraerds, & Veening, 1982; Veening, Swanson, Cowan, Nieuwenhuys, & Geeraerds, 1982). Also in support of this proposal is the fact that electrical stimulation of some of these mid- and hindbrain areas is rewarding. In accordance, Gallistel and colleagues (1996) proposed that the neurons carrying the reward signal could be located in the midbrain or hindbrain rather than in the forebrain. They suggested that these neurons are bipolar: both their nuclei and terminals are located in the midbrain with projections to the MFB. Therefore, subsequent lesion studies targeted structures of the midbrain and hindbrain to assess the extent of reward-relevant neuronal damage.

Several studies of inactivation or lesions to the mid and hindbrain structures showed little or no effect on lateral hypothalamic self-stimulation. Targeted structures



included the parabrachial nucleus (Waraczynski & Shizgal, 1995), peduncolopontine tegmentum and neighbouring lateral dorsal tegmentum (Waraczynski & Perkins, 1998), mesencephalic grey and dorsal raphe (Waraczynski, Carlton, and Perkins, 1998), superior cerebellar peduncle, and the median raphe and interpeduncular nucleus (Waraczynski, Perkins & Acheson, 1999). However, lesions to the lateral habenula (Morissette & Boye, 2008) caused long-lasting pulse frequency threshold increases that ranged from 0.12 to 0.28 log units; in one rat the pulse frequency threshold increased by 0.54 log units.

#### ***4.4. Summary of lesion-induced effects***

Taken together, lesions to the basal forebrain, midbrain, and hindbrain tend not to have substantial effects on pulse frequency threshold. The largest effects were seen in 2 subjects (Murray & Shizgal, 1996c): immediate increases of around 0.5 to 0.7 log units after lesions of the ALH. Another large increase of about 0.54 log unit was seen in one rat with a lateral habenula lesion. The more consistent lesion effects were observed after LPO and SLEA damage; these threshold increases ranged from 0.2 to 0.4 log units. Nonetheless, a 0.4 log unit increase in pulse-frequency threshold implies only a 60% reduction in reward relevant fibers according to the logic advanced by the proponents of the curve-shift method. The majority of the reported effects were even lower than this increase.

#### ***4.5. Validity of the rate-frequency curve method to assess lesion-induced effects***

Given that lesions of many brain regions do not have substantial effects on pulse-frequency thresholds that are measured using the rate-frequency curve method, Gallistel and colleagues (1996) assessed the validity of this method by conducting a proof-of-principle study. Electrolytic lesioning through the stimulation electrode was the chosen manipulation. By targeting the electrical stimulation site, the authors were assured that the lesions removed reward-relevant fibers. Substantial increases in pulse-frequency threshold were seen: progressively bigger lesions through the stimulating electrode had a progressively increasing effect on pulse frequency thresholds. After the most extensive lesion, the threshold increased dramatically by 1.2-1.3 log units (a 15 to 19 fold increase or about a 1400% to 1800% & increase).

It is unlikely that such large increases in thresholds will occur when destroying neurons that are not directly in the field of stimulation. Nonetheless, the importance of Gallistel's (1996) study is that it demonstrates the ability rate-frequency curve shift to detect potentially large lesion effects.

#### ***4.6. Validity of the reward-mountain model to assess lesion-induced effects***

Although the rate-frequency curve shift method is a useful method for evaluating the effects of lesions on ICSS, it is limited in terms of what it can reveal regarding the role of brain areas in reward processing. Changes in lateral displacements of the rate-frequency curve do not distinguish between the effects of manipulations acting before or after the output of the reward integrator. This method does not reveal whether a given manipulation perturbs the first-stage neurons or the neurons that are implicated in a later

stage of reward processing, such as the perceived cost of the reward or evaluation of competing activities. Reward-mountain methodology can distinguish between the two: displacements along the pulse frequency axis (reflected in changes in parameter  $F_{hm}$ ) indicate disruptions of the directly activated first-stage neurons. Displacements along the objective price axis (changes in parameter  $OP_e$ ) indicate disruptions of the circuitry that occur after reward integration. The reward-mountain measurement strategy is elaborated on in Chapter 1.

The 3-dimensional measurement strategy could also reveal effects that are traditionally hidden by the use of traditional 2-dimensional rate-frequency curves. Specifically, after destruction of a given brain area, rate-frequency methodology may not detect an effect on parameter  $F_{M50}$ ; this leads to the conclusion that brain area under investigation is not implicated in reward pursuit. However, for this same region, it is possible that reward mountain methodology can detect a significant effect on parameter  $OP_e$ . Accordingly, the use of the reward-mountain measurement strategy to assess the effects of a cannabinoid blocker on self-stimulation of the LH demonstrated that while there was no significant effect on pulse-frequency threshold values ( $F_{M50}$ ), a significant effect on parameter  $OP_e$  was detected (and not  $F_{hm}$ ) (Trujillo-Pisanty, Hernandez, Moreau-Debord, Cossette, Conover, Cheer, Shizgal, 2011). Therefore, the reward-mountain measurement strategy could potentially reveal lesion-induced effects of brain regions that conventionally showed no effects using past methodologies.

The present experiment is analogous to Gallistel's (1996) proof-of-principle experiment. While Gallistel's study validated the rate-frequency curve methodology, the present study is aimed at validating the reward-mountain measurement strategy to detect

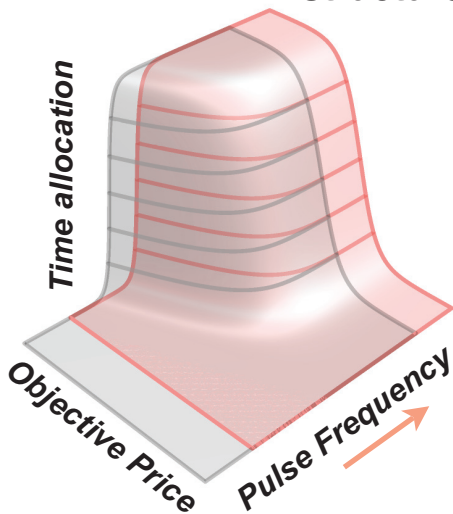
a lesion-induced effect on lateral hypothalamic self-stimulation. The lesion is electrolytic and delivered through the stimulation electrode. The prediction is that there will be a rightward shift of the reward mountain structure along the pulse frequency axis (an increase in location parameter  $F_{hm}$ ). That is, eliminating the directly activated first-stage neurons would affect the stage of processing preceding reward integration by disrupting the actual input to the integrator, thus causing a rightward shift along the pulse frequency axis. Specifically, the surviving substrate would require more action potentials to drive behaviour back to baseline levels. The secondary possibility is that there may be displacement of the structure along the objective price axis (reflected in changes in parameter  $OP_e$ ) for several possible reasons. Among them are the consequences of multiple integrators, elaborated on in the discussion section. The more obvious possible reason for a change in  $OP_e$  is that due to (1) the intermixing of different populations of ascending and descending fibers traversing through the lesion site and (2) the non-selectivity of the electrolytic damage, neurons implicated in *post*-reward integration stages could have also been disrupted. Parameter  $OP_e$  could either increase or decrease depending on the role of the damaged neurons.

In the present experiment, electrolytic damage of the LH through the stimulation electrode is predicted to cause an increase in location parameter  $F_{hm}$ . These changes in location parameters would confirm that the reward-mountain measurement strategy could be used to assess lesion-induced effects. This methodology and prediction is summarized in Figure 1. Furthermore, a group of sham rats not receiving the lesion is tested to assess the variability of pulse-frequency threshold changes over time and to guide us in our interpretation of “meaningful” effects.

### Baseline:



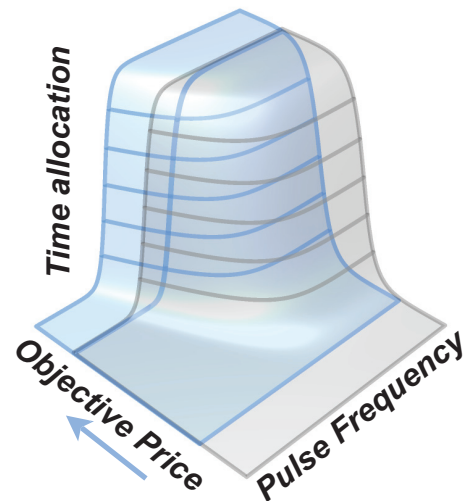
### Possible displacements of the mountain structure after a manipulation:



Displacement of the mountain along the **pulse frequency axis** is reflected in a change in parameter

$F_{hm}$

*This implies disruption of the circuitry before the output of reward integration*



Displacement of the mountain along the **objective price axis** is reflected in a change in parameter

$OP_e$

*This implies disruption of the circuitry after the output of reward integration*

**A lesion through the stimulating electrode is predicted to cause an *increase* in parameter  $F_{hm}$  and a possible *increase or decrease* in parameter  $OP_e$ .**

Figure 1. Summary of the possible displacements of the mountain structure and the prediction of the displacement of the structure after a lesion through the stimulating electrode.

## ***Materials and Methods***

### ***4.7. Subjects***

The same 7 male Long Evans rats that served as subjects in the frequency-following fidelity experiment in Chapter 3 served as subjects in the lesion group of the present study. Furthermore, a control group consisting of 3 additional rats with electrodes implanted in the LH was added.

### ***4.8. Baseline condition***

After completion of the frequency-following experiment (Chapter 3), the rats were trained in the 3-sampling matrix mountain condition. The sampling matrices were: pulse-frequency sampling matrix @ 4 s, price-sampling matrix, radial-sampling matrix. If the radial-sampling matrix did not pass through or near the point defined by the fitted position parameters of the mountain [ $\text{Log}_{10}(OP_e)$ ,  $\text{Log}_{10}(F_{hm})$ ], it was adjusted. Furthermore, if the psychometric curves did not have well defined upper and lower asymptotes, the corresponding sampling matrices were adjusted. In total 5 to 10 sessions were collected.

The control rats were first trained in pulse frequency and price-sweep conditions as described in Chapter 2 (section 2.94 and 2.95). Next, they were trained in the 3-sampling matrix mountain condition. The radial-sampling matrix and psychometric curves were assessed and adjusted as described above. After stable responding was evaluated (confidence interval widths of 0.1 log units or less), 10 sessions were collected.

#### ***4.9. Electrolytic lesion and post-lesion condition***

Rats were lesioned through the stimulating electrode at  $100 \mu\text{A}$  for 15 s. They were not tested for 24 hours after the lesion. Subsequently, the subjects underwent the post-lesion condition, which was a 3-sampling matrix mountain condition (described above). If no effect (no change in parameter  $F_{hm}$  or  $OP_e$ ) was seen after 2 sessions, then a second lesion was given at  $200 \mu\text{A}$  for 15 s and the subject was tested again. After every 1 or 2 sessions, the mountain model was fit to the data. Adjustments were made to the experimental parameters (pulse frequency and objective price) for subsequent sessions so that the sigmoidal form of the psychometric curves could be captured. This usually involved raising the pulse frequencies of the pulse-frequency sampling matrix and raising the pulse frequency of the price-sampling matrix. (If the pulse frequency of the price-sampling matrix is close to the  $F_{hm}$  value, then the mountain model cannot be fit successfully to the data because of mathematical constraints.) Once the parameters were adjusted and the mountain model could be fit to the data, about 15 sessions were collected.

The control rats were connected to the lesion-maker for 15 s, but the stimulation was not turned on. After a 24 hour rest period, the control rats underwent the 3-sampling matrix mountain condition for 20 sessions.

#### ***4.10. Histology***

After completion of the experiment, the rats were anesthetized with a lethal dose of sodium pentobarbital. For 15 s, through the stimulation electrode, a 1 mA anodal current was delivered to deposit iron ions at the site of the electrode tip. The animals



were perfused intracardially with 0.9% sodium chloride, followed by a formalin-Prussian Blue solution (10% formalin, 3% potassium ferricyanide, 3% potassium ferrocyanide, and 0.5% trichloroacetic acid) that forms a blue reaction with the iron deposited at the tip of the electrode. The brains were removed and were fixed with a 10 % formalin solution for at least two weeks. Coronal sections, 30 to 40  $\mu\text{m}$  thick, were cut with a cryostat and tip locations were determined under low magnification with reference to the stereotaxic atlas of Paxinos and Watson (2007).

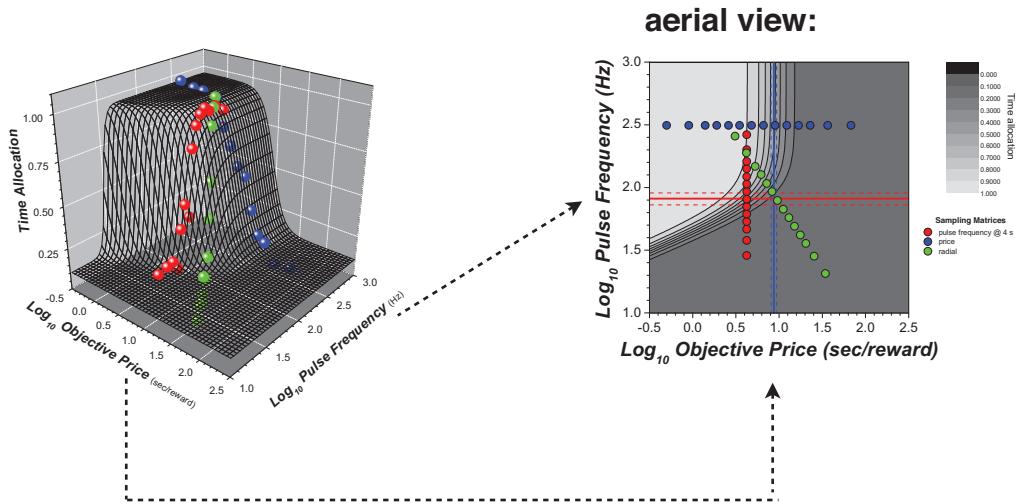
## ***Statistics and Results***

The rats that received lesions once were F16, and F18, LESRO1. The rats that received lesions twice were F3, F9, F17.

The resampling procedure described in Chapter 3 was used to fit the mountain model to the data and estimate the model parameter values ( $F_{hm}$ ,  $OP_e$ ,  $g$ ,  $TA_{max}$ ,  $TA_{min}$ ,  $a$ ). The goal of this study was to determine whether there were differences in location parameters ( $F_{hm}$  and  $OP_e$ ) between the baseline and post-lesion conditions. This was achieved by obtaining 1000 bootstrapped *difference values* of the  $F_{hm}$  and  $OP_e$  parameters from baseline. To obtain these difference values, first 1000 bootstrapped estimates of  $F_{hm}$  and  $OP_e$  are estimated for the baseline condition. Second, 1000 bootstrapped estimates of  $F_{hm}$  and  $OP_e$  for the post-lesion condition are estimated. Third, a difference value (post-lesion parameter value minus baseline parameter value) was calculated for the 1000 estimates. The mean difference value of the 1000 estimates and associated 95% confidence interval was determined. An example of this strategy is described in Appendix B. If the 95 % confidence interval of the difference value included 0, the effect was considered non-significant.

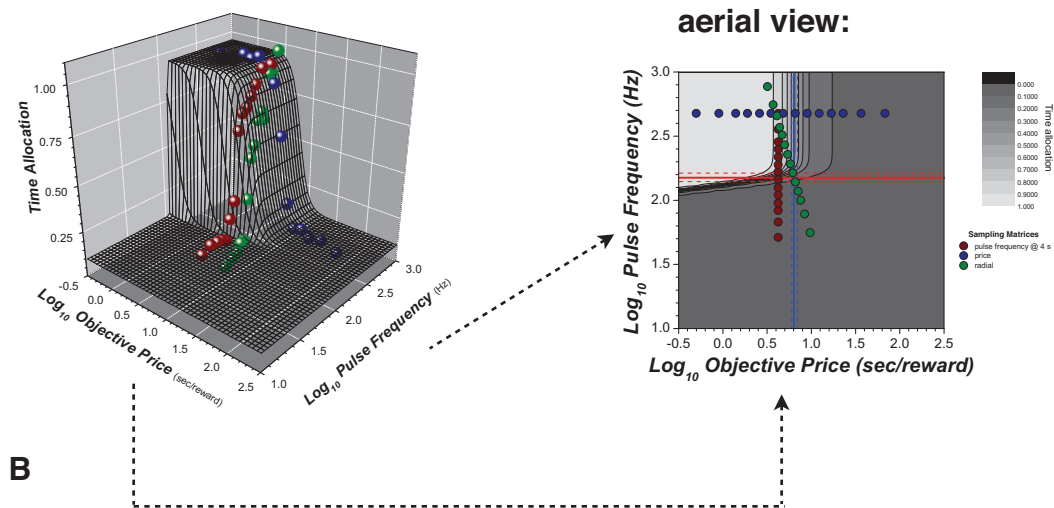
The 3-dimensional and corresponding contour plots of data from the averaged sessions for the baseline condition and post-lesion condition are illustrated in Figure 2, for Rat F16. An illustration of the comparison of parameters  $F_{hm}$  and  $OP_e$  is shown in Figure 3. The contour plots corresponding to the baseline condition are displayed twice for clarity: the top left-hand panel and bottom right-hand panel. The location parameters of  $F_{hm}$  (represented by the red line) and  $OP_e$  (represented by the blue line) can be easily compared with the lesion condition in the bottom left-hand panel. For parameter  $F_{hm}$ ,

### Baseline condition:



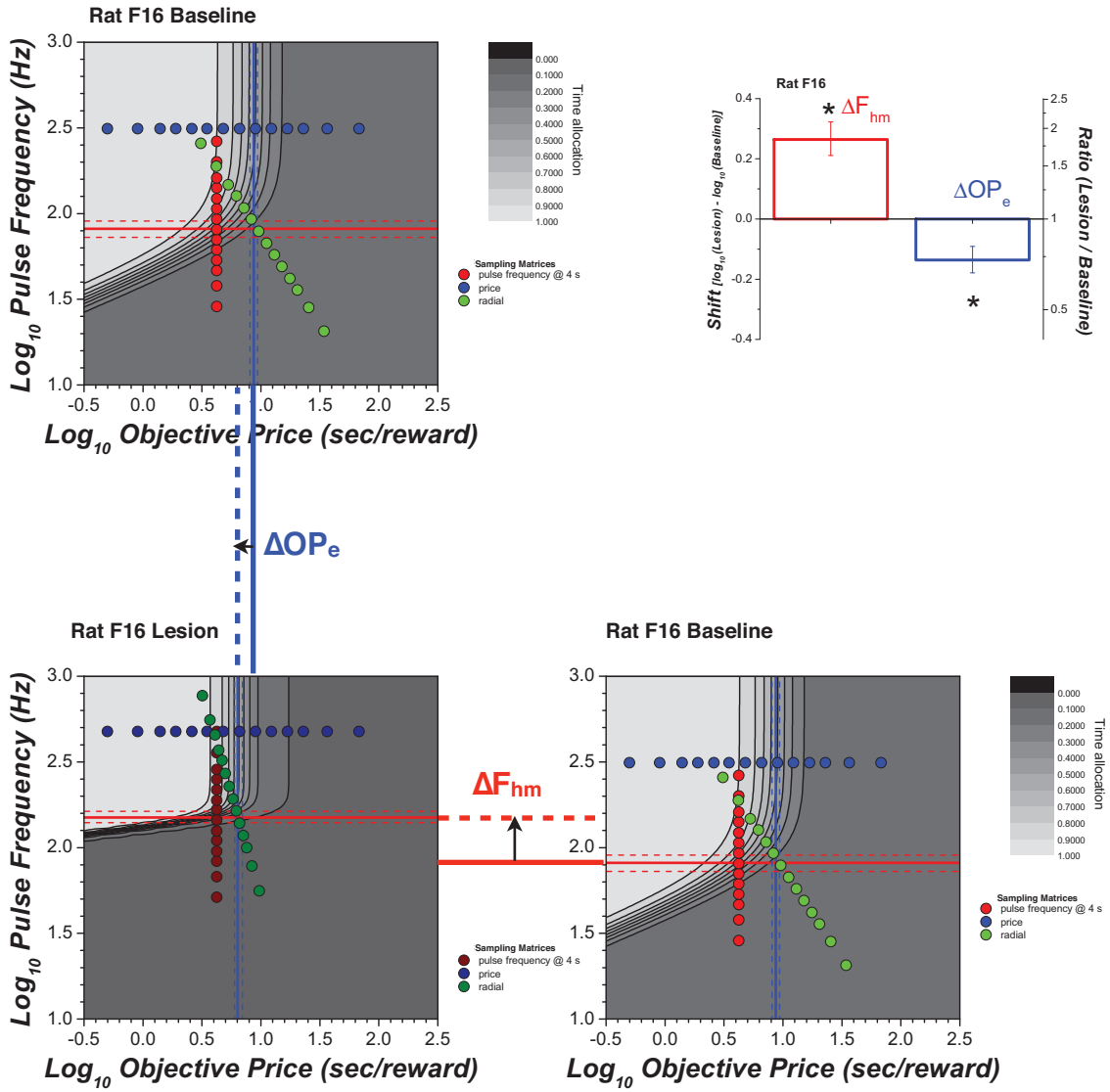
**A**

### Post-lesion condition:



**B**

Figure 2. For rat F16, The 3-dimensional plot and corresponding contour plot (an aerial view) for the average of all tested days of the baseline condition (A) and the post-lesion condition (B).



A

Figure 3A. For rat F16, the contour plots of the baseline condition and of the post-lesion condition. The contour plots corresponding to the baseline condition are displayed twice, the top left-hand panel and bottom right-hand panel such that the location parameters of  $F_{hm}$  (represented by the blue line) and  $OP_e$  (represented by the red line) can be compared with the lesion condition in the bottom left-hand corner. For parameter  $F_{hm}$ , there is an over 0.2 log unit increase and for parameter  $OP_e$ , an over 0.1 log unit decrease from baseline. These increases and decreases in the parameters from baseline are represented in the bar graph.

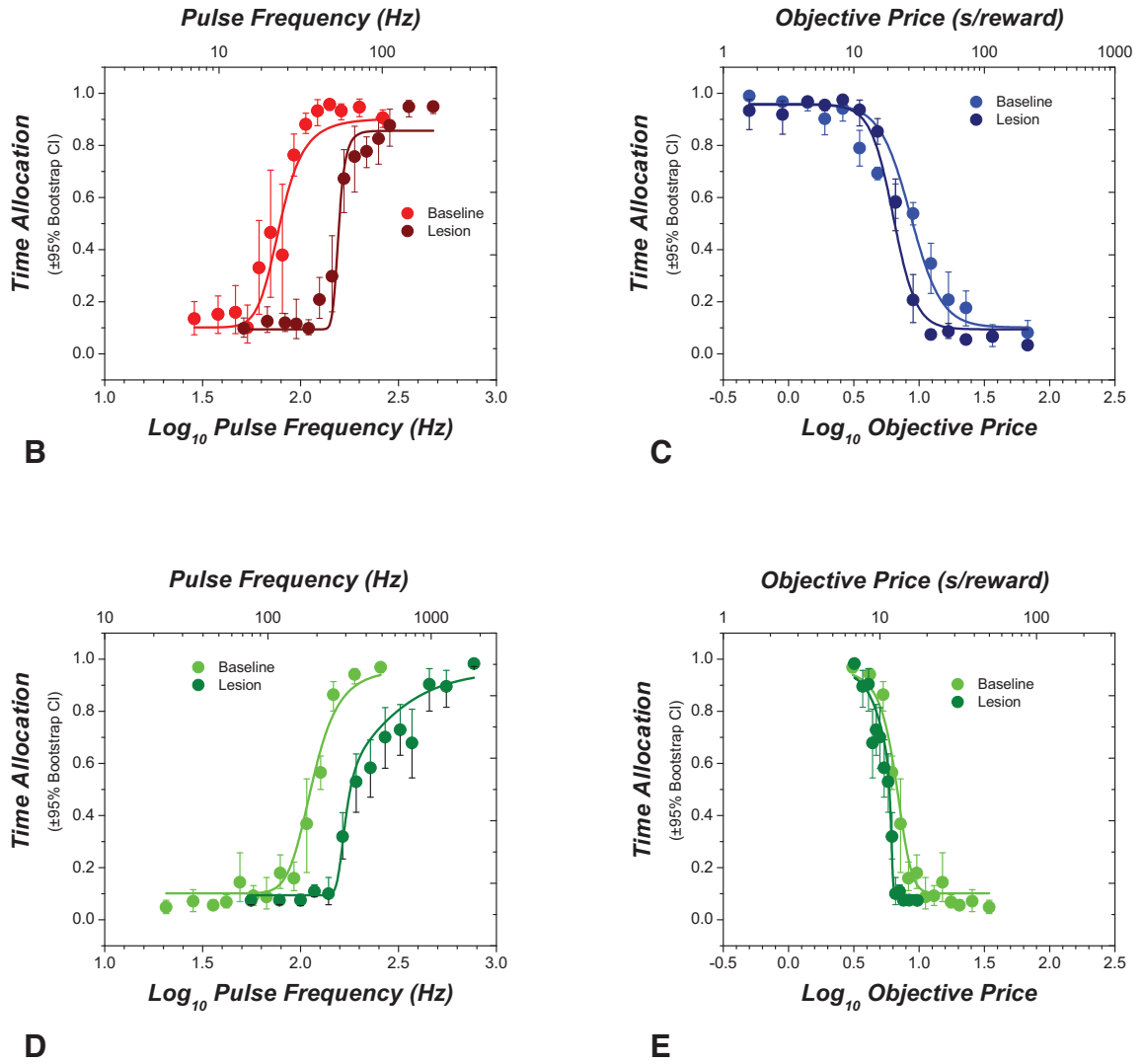


Figure 3B-E. A 3 dimensional representation can be represented by several 2 - dimensional plots whereby one independent variable varies while the other stays constant. Panels B-E are the corresponding 2-dimensional plots of Figure 3. B. Time allocation as a function of pulse frequency (corresponding to the *pulse frequency sampling vector @ 4 s*). C. Time allocation as a function of objective price (corresponding to the *price sampling vector*). D & E. The tests of the *radial sampling vector* can be represented by time allocation as a function of pulse frequency, and time allocation as a function of objective price.

there is an *increase* of over 0.2 log units and for parameter  $OP_e$ , a *decrease* of over 0.1 log units from baseline. These increases and decreases in the parameters from baseline are represented in the bar graph in the upper right-hand panel. Figures 3B-D are the corresponding 2-dimensional psychometric curves for each tested sampling matrix: the pulse-frequency sampling matrix @ 4 s (red curve), the price-sampling matrix (blue curve), the radial-sampling matrix (green curve: 2 plots). The shifts in parameters viewed in Figure 3A are also seen from the 2-dimensional perspective: time allocation as a function of pulse frequency or objective price.

Figure 4 displays the time course of post-lesions changes in parameters  $F_{hm}$  (red) and  $OP_e$  (blue) for Rat F16. Testing started on “post-lesion day 1”. Each data point represents the averaged (bootstrapped) difference from baseline for the parameters collected over 2-4 days. For rat F16, the parameter differences were sustained throughout the post-lesion test days without evidence of recovery.

The other rats showed changes of similar magnitude in the location-parameter estimates (Figure 5). In 6 out of 7 rats, significant increases of parameter  $F_{hm}$  were seen. In these rats, the magnitude of the increase was around a 0.2 log unit increase from baseline (58% increase). Rat F18 was the only rat to have an  $F_{hm}$  difference less than 0.05 log units (< 12%). In 5 out of 7 rats, decreases in parameter  $OP_e$  were seen ranging from 0.1 to 0.3 log units (20-50% reduction). In one rat (LesRO1) there was an increase in  $OP_e$  of about 0.1 log units, but this was not significant ( $p < 0.05$ ).

Figure 6 displays the time course of the changes in the parameter values over the tested post-lesion days. In most rats, recovery did not occur, except for Rat F18. For rat F12 and F17, the parameter magnitude changes peaked somewhat over the initial post-



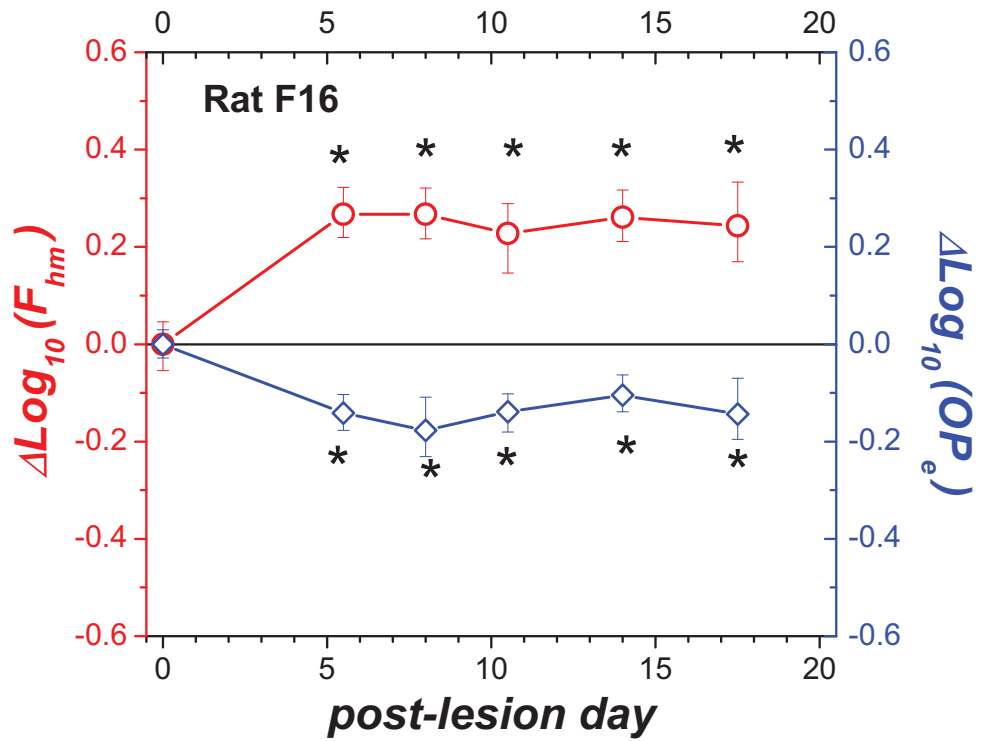


Figure 4. For rat F16, the time course of post-lesions changes in parameters  $F_{hm}$  and  $OP_e$ . Day 0 is the day after the lesion in which the rat rests for 24 hours. Testing starts on post-lesion day 1. Each data point represents the averaged (bootstrapped) difference of the parameters from baseline around the represented post-lesion day.

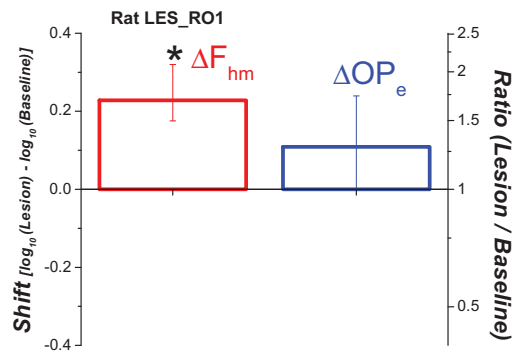
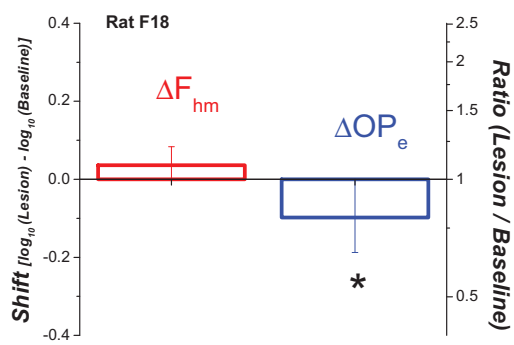
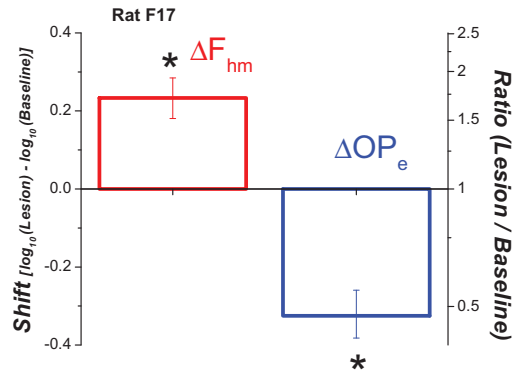
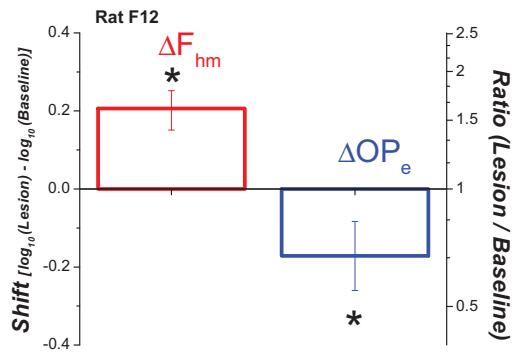
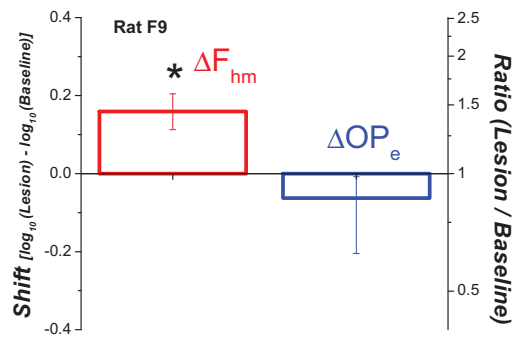
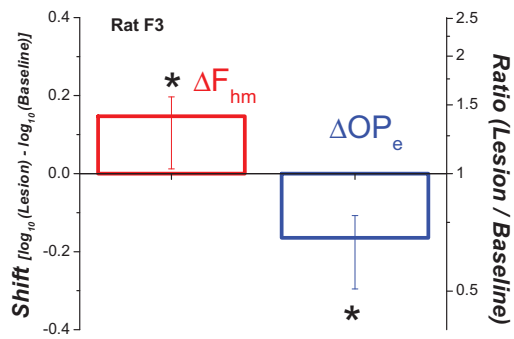
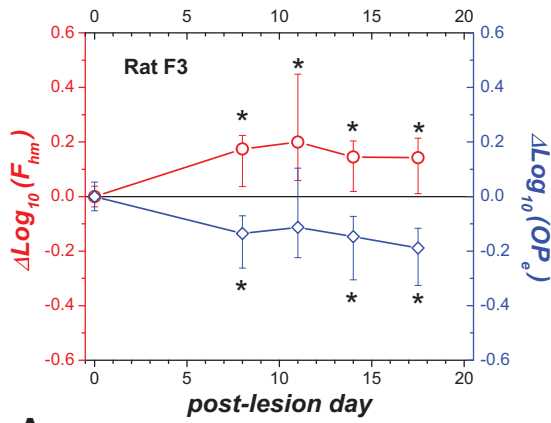
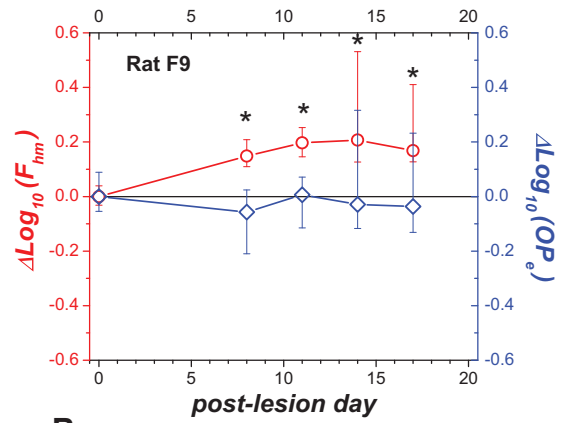


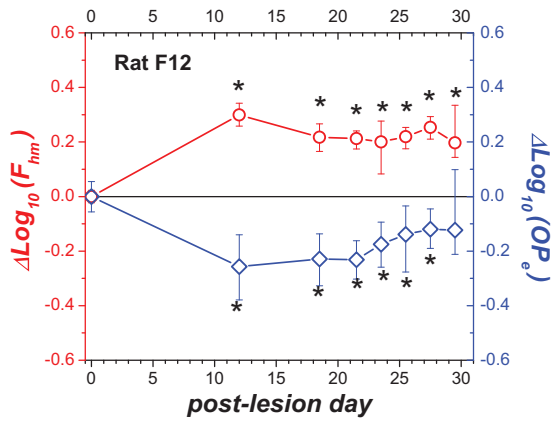
Figure 5. Shifts along the frequency axis ( $\Delta F_{hm}$ ) are shown on the left and shifts along the objective price axis ( $\Delta OP_e$ ) on the right for each rat. In all subjects, the lesion shifted the 3 dimensional mountain rightward along the pulse frequency axis (positive direction). Leftward shifts (negative direction) along the price axis were reliable in 5 out of the 7 subjects. (\* denotes cases in which the 95% confidence interval around the shift fails to include 0).



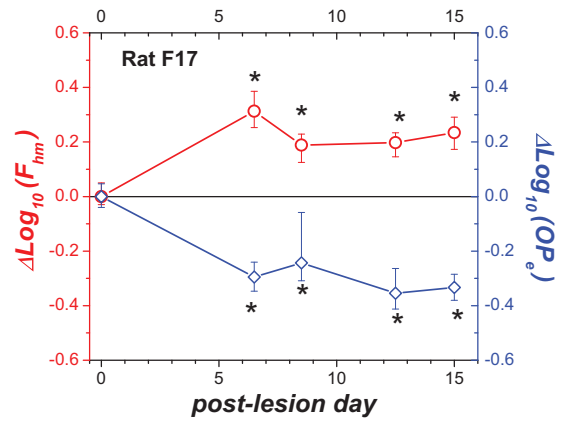
**A**



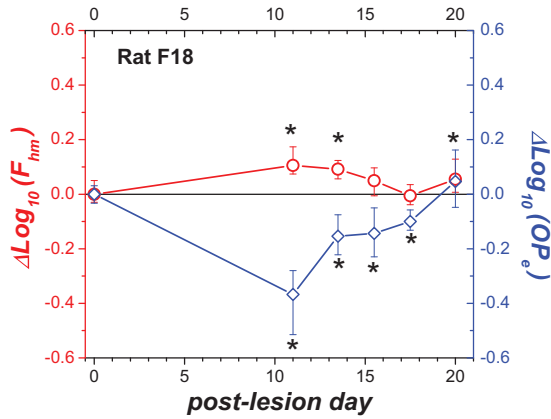
**B**



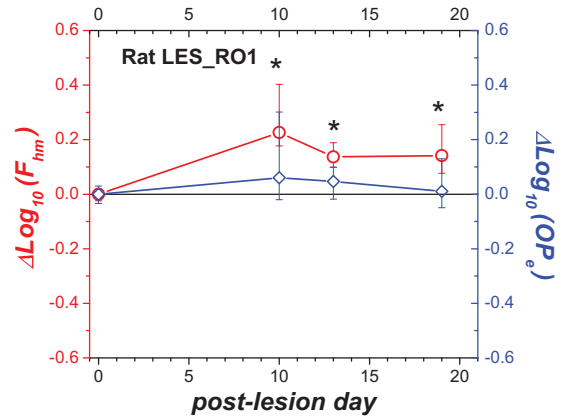
**C**



**D**



**E**



**F**

Figure 6. For all rats, the time course of post-lesions changes in parameters  $F_{hm}$  and  $OP_e$ . Day 0 is the day after the lesion in which the rat rests for 24 hours. Testing starts on post-lesion day 1. Each data point represents the averaged (bootstrapped) difference of the parameters from baseline around the represented post-lesion day.

lesion sessions and progressed closer to baseline over the course of the post-lesion testing days.

Figure 7 shows changes in the location-parameter estimates for the 3 control rats. (These rats were connected to the lesion-maker, but not lesioned). Rat G3 and Rat G4 showed very small parameter differences from baseline (less than 0.05 log units). For Rat G5, there was an initial 0.13 log unit increase from baseline for parameter  $OP_e$  (not shown). However, after eliminating the first ten sessions (from the baseline condition) and comparing the 10 post-lesion sessions with an additional 10 sessions (collected after the post-lesion condition), the difference in the parameter value  $OP_e$  was no longer seen (Figure 7C) which indicates that the rat required more sessions for the behaviour to reach stability. Figure 8 illustrates the electrode locations on a coronal plane for each rat, all located within the lateral hypothalamus.

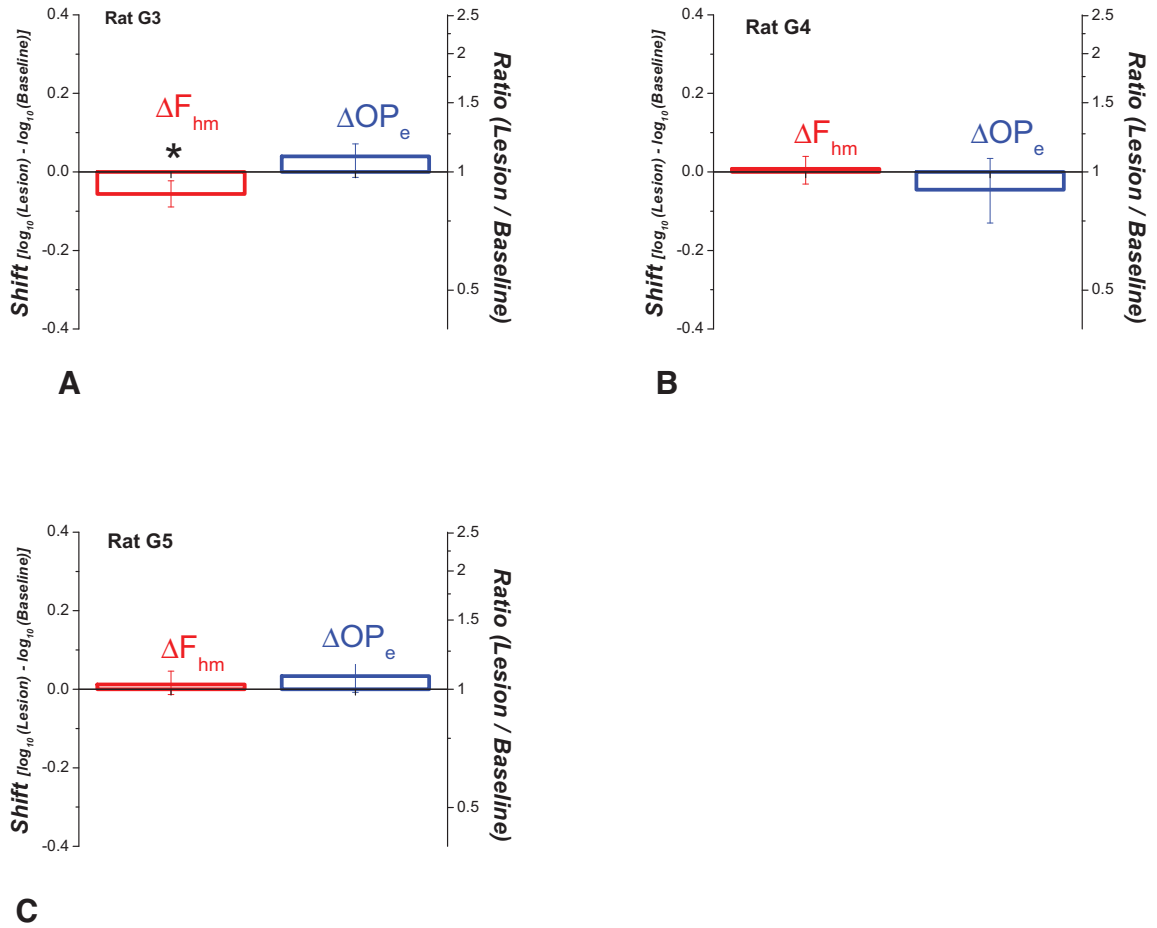


Figure 7. Control rats: (A-C) Shifts along the frequency axis ( $\Delta F_{hm}$ ) are shown on the left and shifts along the objective price axis ( $\Delta OP_e$ ) on the right for each rat. The magnitudes of the shifts in parameters are small in comparison to the experimental subjects .

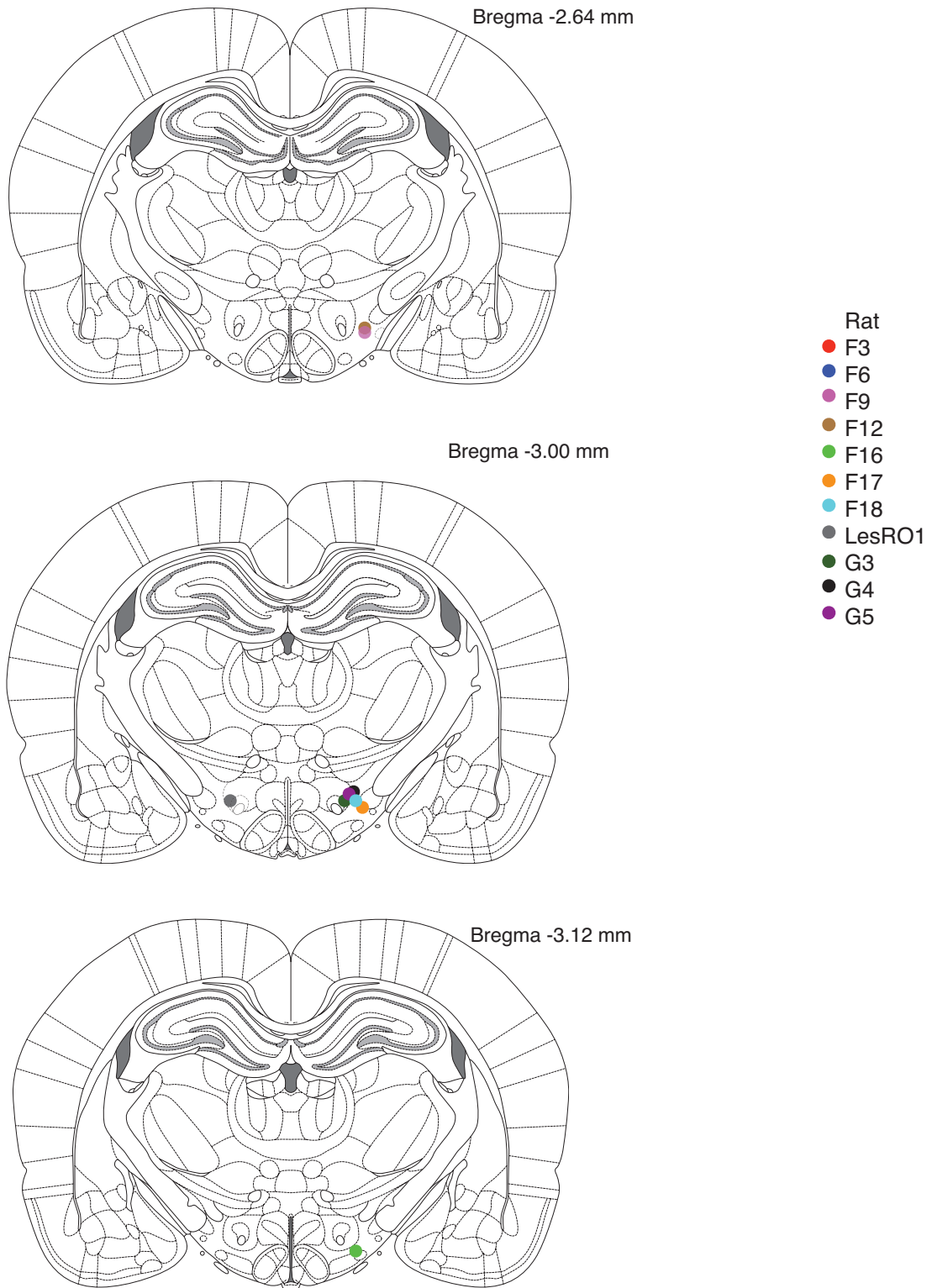


Figure 8. Location of the electrode tips for all of the rats in the experiments.



## ***Discussion***

As predicted by the model, a lesion that damaged the directly stimulated substrate shifted the reward mountain structure rightward along the frequency axis (reflected in an increase in  $F_{hm}$ ). Proof-of-principle was demonstrated: to compensate for the damage, the surviving neurons had to fire at higher pulse frequencies to restore baseline levels of behaviour. These changes in  $F_{hm}$  were often accompanied by changes in  $OP_e$ . In 5 out of 7 rats, the lesion of the directly stimulated substrate shifted the 3-dimensional structure leftward along the objective price axis (reflected in a reduction in  $OP_e$ ) implying that some of the fibers destroyed act beyond the output of the integrator (section 4.14). Alternatively, the shift on the objective price axis may be a consequence of multiple integrators (section 4.13).

The magnitude of parameter changes from baseline in this study were substantially smaller (0.1 – 0.3 log units) than the magnitudes reported by Gallistel and colleagues (1996), in which  $F_{M50}$  values were shown to increase by 1.2 -1.3 log units after lesions were made through the stimulating electrode. However, Gallistel and colleagues made progressively larger lesions at 300  $\mu A$ , at 10 s, then 20 s, then 40 s, (the currents of the tested rats ranged from 500 to 700  $\mu A$ ) while the present study entailed smaller lesions of 100  $\mu A$  to 200  $\mu A$  for 15 s (the tested currents ranged from 350 to 600  $\mu A$ ). If the lesions used in this study were made progressively larger, it is predicted that the magnitudes of parameter changes from baseline would increase to the level seen in Gallistel and colleagues' (1996) study. Also, the magnitudes of parameter changes are in the range detected in previous studies using the reward mountain methodology (Arvanitogiannis et al., 2008; Hernandez et al., 2010; Trujillo-Pisanty et al. 2011;

Hernandez et al., 2012; Breton et al., 2013; Trujillo-Pisanty et al. 2013), which further validates this methodology for lesion-induced effects.

There was no relationship between the current intensity used during testing and whether the rat required a second lesion to be behaviourally affected. For example, rat F3, tested at a current of  $500 \mu\text{A}$  required a second lesion for an effect to be detected, while rat F16 also tested at  $500 \mu\text{A}$ , required only one lesion. This suggests that for rat F3, the reward-relevant neurons were situated further from the electrode tip as compared to rat F16.

#### **4.11. Time course analyses**

For most rats, no evidence of recovery to baseline was seen for both parameters ( $F_{hm}$  and  $OP_e$ ), (Figure 6). The exception was rat F18: both parameters recovered around day 20 of the post-lesion sessions. Furthermore, for F12 and F17, the magnitude of the parameter differences from baseline peaked somewhat during the initial post-lesion sessions. In general, the magnitude of parameter changes from baseline were consistent across subjects.

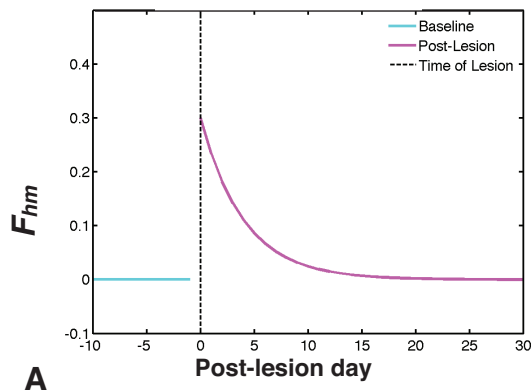
For the time course plots (Figure 6), each data point represents the averaged bootstrapped difference from baseline (as described in Appendix B) for parameters  $F_{hm}$  and  $OP_e$  collected over 2-4 sessions. The limitation to the bootstrapping approach employed in the study is that it does not allow for the analysis of parameter changes from baseline for a *single* session, which is the reason why each data point represents an average of several sessions rather than a single session. Future statistical approaches may allow improved assessment of bootstrapped changes from baseline over time such that

the mountain model can be fit to a *single* session and the corresponding 95% confidence interval around the parameters can be obtained. This involves resampling at an even “finer grain” than what is employed here. Here the time allocation values are resampled across sessions, whereas the ensuing statistical analysis will resample the distribution of “hold times” of the lever, within a trial of a single session.

Furthermore, the reward mountain model could usually not be fit to the data collected during the first and second post-lesion sessions. These data were typically noisy and the parameters needed to be adjusted to capture the sigmoidal form of the psychometric curves. Appropriate changes to the experimental parameters were usually only in effect on the third session. Thus, it is possible that the magnitude of  $F_{hm}$  or  $OP_e$  changes from baseline are considerably higher on the first and second days post-lesion. That is, a large transient effect, which has been reported in previous studies, may have occurred but was undetected with the present approach. The resampling approach described above may provide a way to model the individual noisy sessions with a bootstrapping approach, across time. The procedure described above fixes parameters  $g$ ,  $a$ ,  $TA_{min}$ ,  $TA_{max}$  across the sessions but the location parameters  $F_{hm}$  and  $OP_e$  are free to vary. Thus, the later sessions in which the data is less noisy provides a good estimate of parameters  $g$ ,  $a$ ,  $TA_{min}$ ,  $TA_{max}$  such that they can guide the estimate of the location parameters of the initial individual noisy sessions. With this approach, it is predicted that for rat F18 (in which no significant change in  $F_{hm}$  was detected), the magnitude of the increase in parameter  $F_{hm}$  from baseline could be large at the initial sessions of the post-lesion condition. Furthermore, it is possible that this statistical approach could reveal large transient effects during the initial post-lesion days for the other rats.

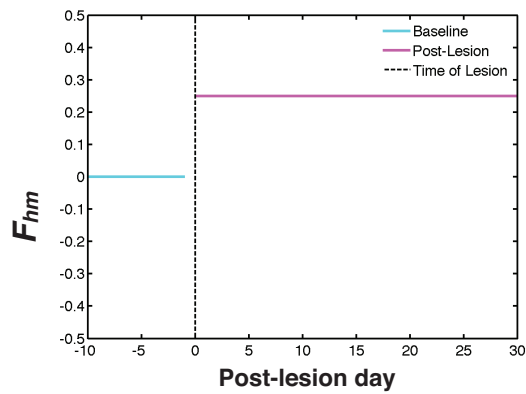
As well, in the present time course graphs, changes in the parameter values  $F_{hm}$  and  $OP_e$  across time are plotted. Future statistical analyses will be aimed at modeling the post-lesion time course of lesion-induced effects such that persistent and transient effects can be identified (e.g., Trujillo-Pisanty, Martel, Conover, Arvanitogiannis, & Shizgal, 2013). Two examples of potential models are shown in Figure 9.

In addition, the reward-mountain analysis used for this experiment does not take into account the transformation of objective prices into subjective prices or pulse frequencies into firing frequencies. Subsequent analyses will incorporate these transformations. Because the objective price values and pulse frequency values are within or close to the identity range (the range in which objective prices equal subjective prices and the range in which pulse frequencies equals firing frequencies) the results obtained with the present analysis are expected to be close to the results of the future analysis that will incorporate the two transformations.



Exponential, transient effect

**A**



Linear, persistent effect

**B**

Figure 9. Examples of two potential models that can be used to fit the time course of post-lesion effects. A. The lesion effect is transient such that the change in parameter  $F_{hm}$  increases abruptly and decreases rapidly over time. B. The lesion effect causes an increase of the  $F_{hm}$  parameter change and persists.

*Figure adapted from Kent Conover*

#### ***4.12. The control group and stability***

Two rats in the control group initially did not show substantial changes in parameter values from baseline. However, rat G5 initially showed a change in  $OP_e$  of about 0.13 log units. Yet, after testing rat G5 for an additional ten sessions, and eliminating the first 10 sessions, the  $OP_e$  difference from baseline was no longer seen. This finding implies that rat G5 required more training in order for stable behaviour to be achieved. Thus, some rats may require a longer training phase as compared to others. Future experiments should be guided at ensuring that changes in parameter values are not detected before commencing the experimental condition. To confirm stability, the conventional criterion is to make certain that the confidence intervals around the parameter estimates do not exceed 0.1 log units. Nonetheless, as demonstrated by rat G5, a further test may be required: for example, the first test of 10 sessions with a succeeding test of 10 sessions can be compared before the experimental phase is commenced. If the rat is persistently variable between the averaged sessions (a greater than 0.1 log unit difference) this individual variability could be later statistically taken into account when comparing the baseline condition with the experimental condition.

#### ***4.13. The dual-integrator model and implications for lesion-induced effects***

A dual-integrator model of reward integration may have implications for the variability detected by the rate-frequency curve shift method that is traditionally observed between rats with lesions to the same areas. Furthermore, in the context of reward mountain methodology, the dual-integrator model may also explain the (leftward) shifts along the objective prices axis (decreases in  $OP_e$ ). A description of this model and its

relationship to lesion effects is provided below.

The most parsimonious models of BSR and the reward mountain model assume a homogenous reward substrate and thus a single integrator. However, it is possible and even likely that the rewarding effect of stimulating particular brain areas such as the lateral hypothalamus is due to the direct activation of heterogeneous neural populations. That is, different brain areas could potentially serve different initial reward-related functions and provide input to multiple integrators. Their axons may be intermixed in the MFB; the electrical stimulation could be activating this intermingling of functionally different populations of axons. In the dual-integrator model, the outputs of two integrators (the reward intensity values from two populations) are summed to obtain a total value of reward intensity. In contrast, the single integrator model conventionally assumed (discussed in Chapter 1) entails one population of neurons undergoing reward integration, and thus one integrator. The two models make different predictions in the ability of curve-shift methodology to detect lesion-induced effects: the dual-integrator model could explain why lesion-induced effects have traditionally been difficult to discern.

The output of the reward integrator is described as a logistic function that grows and levels off (termed the reward-intensity function). The rate of growth (steepness of the function) is controlled by parameter  $g$ . According to the single integrator model, destroying half of the reward-relevant neurons would require a doubling of pulse frequency to restore behaviour to its baseline level (time allocation as a function of pulse frequency), regardless of the form or steepness of the reward intensity function (controlled by parameter  $g$ ). Thus, the detection of effects should prove to be relatively

simple if a lesion to an area is considerably large.

According to the dual-integrator model, damaging half of the substrate within the stimulation field would not require a doubling of the pulse frequency to restore behavior to baseline. Consider a simple scenario in which half of the fibers coursing through the stimulation site arise from a given brain area and provide input to one integrator. The other half of the fibers within the stimulation site arises from another brain area that provides input to a second integrator. If the neurons of the first population are totally destroyed, then half of the fiber population within the stimulation site has been damaged, and there is no more input to the first integrator. To restore behaviour how much should the pulse frequency be proportionally increased to fire the intact population to achieve baseline behavioural levels? If the outputs of the two integrators are combined by simple addition so as to achieve a value of total reward intensity of both neural populations, and  $g$  is large, then only a very small proportional increase in pulse frequency is needed to compensate for the damaged population. These very small required increases in pulse frequency may go undetected using the rate-frequency curve method because of the small signal to noise ratio. Therefore, the variability traditionally seen in lesion-induced effects could be due to activation of a heterogeneous substrate and across site differences in reward growth (controlled by parameter  $g$ ) (Arvanitogiannis et al., 1996).

The dual-integrator model also has implications for reward mountain methodology and lesion-induced effects, namely, the detection of a change in parameter  $OP_e$  after the destruction of first-stage neurons. If there are two populations of neurons, and thus two integrators, each integrator can have differing properties (Breton et al., 2013). For instance, the rate at which the reward intensity grows (controlled by parameter  $g$ ) may



differ across integrators. In addition, the maximal reward intensity that a given substrate can achieve can differ across integrators. It is proposed that the total maximum reward intensity is the sum of the maximum reward intensities from both integrators. Assuming the simplest case in which the integrators are weighted equally:

$$RI_{max1} + RI_{max2} = RI_{maxTotal}$$

where,

$RI_{max1}$  = the maximal reward intensity from integrator 1

$RI_{max2}$  = the maximal reward intensity from integrator 2

$RI_{maxTotal}$  = the total maximal reward intensity of both integrators

For example, if a lesion destroys the substrate that provides the input to integrator 1, then the total maximal reward intensity could only be provided by integrator 2. However, firing the intact population more (by increasing the pulse frequency) to compensate for the lesion will *not* bring the maximal firing rate back to the (pre-lesion) baseline total maximal reward intensity level. This is because integrator 2 will have reached its maximum, which by definition is less than the total maximal reward intensity.

Thus, if only the input to integrator 2 is intact, then the maximum output of integrator 2 (denoted  $RI_{max2}$ ) will never be able to reach the previous baseline total reward intensity value, ( $RI_{maxTotal}$ ) no matter how much the neurons are fired:

$$RI_{max2} < RI_{maxTotal}$$

This reduction in the total maximal reward intensity has implications for the effects on parameter  $OP_e$ .

Recall that parameter  $SP_e$  is the subjective price at which working for the *maximal* reward intensity of the stimulation ( $RI_{max}$ ) occurs at half-maximal time allocation. At half-maximal time allocation, the payoff obtained from the stimulation can be denoted  $U_{midBSR}$ .

Thus,

$$U_{midBSR} = \frac{RI_{max}}{SP_e}$$

Rearranging,

$$SP_e = \frac{RI_{max}}{U_{midBSR}}$$

However, in the example above, after the lesion to the population of neurons providing input to integrator 1, the total maximum reward intensity achievable now only comes from integrator 2, and thus  $RI_{max2}$ .

$$SP_e = \frac{RI_{max2}}{U_{midBSR}}$$

But remember:  $RI_{max2} < RI_{maxTotal}$

When a reduction in total reward intensity has occurred, the rat spends half-maximal time allocation to the lever at a *reduced* value of  $SP_e$ . A reduced value of  $SP_e$  implies a reduced value of  $OP_e$  (when testing prices within the scalar range of the subjective price function). This reduction in  $OP_e$  implies a shift of the mountain structure leftward along the objective price function.

Therefore, in the present lesion study, if the lesion caused a shift of the structure along the pulse frequency axis due to damage to the first-stage neurons and if dual integration occurred, the maximal attainable reward intensity experienced could be reduced (compared to the baseline condition). Consequently, the objective price at which the rat spends half of its time working for the reward of maximal intensity ( $OP_e$ ) would be reduced, post-lesion. This model may explain why most of the rats in the present experiment showed a reduction in parameter  $OP_e$  from baseline.

#### ***4.14. Dopamine and $OP_e$ changes from baseline***

The above hypothesis regarding  $OP_e$  changes from baseline assume that only the first-stage neurons have been disrupted, consequently, the associated  $OP_e$  changes could potentially be due to a lowering of maximal reward intensity values for various reasons. However, the decreases seen in parameter  $OP_e$  could also be due to the disruption of neurons that are transynaptically activated by the stimulation, neurons located *downstream* from the first-stage neurons. The non-specificity of the electrolytic lesion

which damages cell bodies and fibers of passage that ascend and descend through the lateral hypothalamus strongly suggests the possibility that later-stage neurons could be damaged. In particular, the decreases in  $OP_e$  may reflect damage to local dopaminergic fibers that have been previously proposed to be excited transynaptically by the stimulation. This hypothesis is consistent with the finding that drugs that alter dopamine tone, such as cocaine, GBR, and pimozide which change the value of  $OP_e$  (Hernandez et al., 2010; Hernandez et al. 2012; Trujillo-Pisanty, 2013). The hypothesis that the reductions in  $OP_e$  were due to damage to local dopaminergic fibers can be tested by measuring phasic and tonic dopamine release in rats working for BSR after a lesion is made through the stimulating electrode. If a reduction in phasic or tonic dopamine is found, this would suggest that a lesion to the directly stimulated substrate damages local dopaminergic fibers excited transynaptically by the stimulation.

#### ***4.15. Further implications***

These results show that the reward mountain model can play a useful role in the search to identify the first stage neurons: damage to the first stage fibers shifts the mountain rightward along the pulse frequency axis. The model can be used in conjunction with more specific lesion techniques such as excitotoxic lesions of different neural populations of interest. One such population is the source of glutamatergic inputs to the VTA. These inputs are of interest because glutamatergic transmission in the ventral tegmental area is necessary for MFB stimulation to drive DA transients in the nucleus accumbens (Sombers et al., 2009). Also, MFB stimulation releases glutamate in the VTA (You et al., 2001). Lesions producing increases in  $F_{hm}$  will be interpreted to have

damaged fibers involved in early stages of the BSR substrate, prior to the output of the integrator. These damaged neurons showing this effect are candidates for the directly stimulated stage of the circuit. Candidate areas include the lateral preoptic area (LPO) and sublenticular extended amygdala (SLEA) that have produced relatively stable shifts using the rate-frequency curve shift methodology. Whether these neurons are the first-stage neurons or transynaptically activated can be elucidated with the reward mountain methodology. While not impossible, it may be difficult to obtain shifts exclusively along the pulse frequency axis without the associated  $OP_e$  shifts for the various reasons outlined above. However, because in past studies it has been possible to achieve shifts exclusively in parameter  $OP_e$ , consistent and sizeable changes in parameter  $F_{hm}$  with possible associated shifts in  $OP_e$  can be interpreted to have affected the first-stage neurons.

Furthermore, this measurement strategy can be used in conjunction with optogenetic methods (Yizhar et al., 2011) for temporally controlled silencing and stimulation of activity in neurochemically and anatomically defined neural populations. The combination of the new neuroanatomical techniques that provide considerable specificity with the refined behavioural measurement strategy validated here could offer novel insights into the nature of the reward circuitry not previously accessible with past approaches.

## ***Appendix A: The meaning of common logarithmic unit changes using rate-frequency curve shift methodology***

The interpretation of logarithmic changes in the rate-frequency curve that has traditionally been assumed is presented below.

A 0.1 log unit increase in pulse frequency threshold ( $F_{M50}$ ) is a 26% increase from baseline ( $10^{0.1}=1.26$ ,  $1.26-1=0.26$ ), for example, an increase from 63 Hz to 80 Hz.

According to adherents of the curve shift-method, an increase in reward threshold implies a reduction in reward-relevant neurons\*: reductions in reward-relevant neurons are of the same magnitude but in the opposite direction on a logarithmic scale. Therefore, a 0.1 log unit increase in pulse frequency threshold implies a 0.1 log unit reduction in reward relevant neurons, or a 20 % reduction in reward-relevant neurons ( $10^{-0.1}=0.80$ ,  $1-0.8=0.2$ ). A 0.2 log unit increase implies a 58% increase from baseline ( $10^{0.2}=1.58$ ,  $1.58-1=0.58$ ), for example, an increase from 63 Hz to 100 Hz. This threshold increase refers to a 37% reduction in reward relevant neurons ( $10^{-0.2}=0.63$ ,  $1-0.63=0.37$ ). A 0.3 log unit increase is a 100% increase from baseline ( $10^{0.3}=2$ ,  $2-1=1$ ), for instance an increase from 63 Hz to 126 Hz which refers to a 50% reduction in reward relevant neurons ( $10^{-0.3}=0.5$ ,  $1-0.5=0.5$ ).

It is best to work in common logarithmic units when describing changes from baseline because equal proportional changes in magnitude (from baseline) in the positive and negative direction are not equal in magnitude when presented on a scale of percentage values (examples provided above). These equivalent proportional changes are equal on a logarithmic scale.

*\*The reward-mountain model shows that increases and decreases in  $F_{M50}$  do not necessarily imply that the directly stimulated neurons are affected by a physiological manipulation.*

## Appendix B

Example of how the 1000 bootstrapped difference value estimates of a parameter are obtained.

<b>LogF<sub>hm</sub> Baseline</b>	<b>LogF<sub>hm</sub> Post- lesion</b>	<b>LogF<sub>hm</sub> Difference value</b>
1.9	2.1	0.2
1.89	2	0.11
1.91	2.1	0.19
1.85	2.2	0.35
1.88	2	0.12
1.99	2.1	0.11



1000 estimates

1000 bootstrapped estimates of  $F_{hm}$  for the baseline and post-lesion are estimated as well as the corresponding “Difference value”. An example is illustrated above.

The mean of the 1000  $F_{hm}$  Difference values and corresponding 95% confidence interval is reported.

## *Chapter 5*

### *General discussion and conclusion*



The marked and impressive systematic behaviour that the rats demonstrated from the onset of the first sessions underscores the nature of reward-cost integration: a process of fast encoding and combining of variables and consequent storing of the result in memory. The experimental tasks were complex in the sense that they were characterized by several frequently changing variables. The methodology presented in this thesis, a variant of the reward-mountain paradigm which conventionally consists of 3, and sometimes 4 sampling matrices (Hernandez et al., 2010; Hernandez et al., 2012; Trujillo-Pisanty et al., 2013; Breton et al., 2013), tests additional sampling matrices within a session. In the first experiment in Chapter 2, the rat was presented with 9 sampling matrices, which in total included 126 unique pulse frequency-objective price combinations, randomized throughout a daily session. After completion of this experiment, the same rats were then presented with 11 different sampling matrices, which in total included 154 unique pulse frequency-objective price-current intensity combinations, also presented in a randomized fashion (Chapter 3). The rats were methodical, performing consistently in their decision to attend to the lever or engage in other activities. At low prices as well as at high reward intensities, the rats allocated a high proportion of their time to the lever. As the price was progressively increased or reward intensity was decreased, the rats adjusted their behaviour such that their time allocation to the lever was steadily reduced. This systematic behaviour was displayed even though the numerous points of the sampling matrices were randomized within a session. Notably, even when a complete test of the sampling matrices was divided over 2 days as was the case for the condition in which 11 sampling matrices were tested (Chapter 3), consistent and logical behaviour was demonstrated.

The presentation of trials in a triad manner allowed the rats to (indirectly) compare the pulse frequency of an experimental trial to two extreme values of reward intensity: a constant very high intensity on the leading trial and a very low intensity on the trailing trial. However, the rat's methodical performance suggests that it is not simply comparing the pulse frequency of the experimental trial with that of the leading and trailing trials and making a binary decision as to whether it is of higher or lower value. Rather, the rat's choices imply a processing of benefits and costs, a computation of the overall value or "payoff" that is encoded in memory.

Survival depends on an animal's ability to rapidly and accurately integrate reward and costs in the face of changing variables, thus it is not surprising that the rat is exceptionally efficient at this complicated experimental task. Quantitative models have been successful at describing goal-directed behaviours. For instance, matching studies have demonstrated that rewards and costs are combined in a scalar (multiplicative) manner.

An important initial stage of reward processing, often overlooked in neuroscientific studies of decision-making, occurs before the integration of reward intensity and cost. Namely, the translation of each objective, external variable into the corresponding subjective, internal value. The formal relationship between the external and internal variables, termed psychophysical functions, are often non-linear: what is subjectively experienced is not necessarily a direct reflection of the external world. An analysis of the transformations of the reward-related variables is important for the full understanding of cost-benefit decision-making. Traditionally, psychophysical functions have described sensory experience such as the translation of the physical intensity of light

into perceived brightness, or sound waves into perceived loudness. However, studies of reward pursuit have extended psychophysical scaling to the realm of valuation, motivation and decision-making.

### ***5.1. Psychophysical functions of reward-seeking variables***

#### ***5.1.1. The psychophysical function of reward intensity: previous studies***

In BSR experiments, several studies have shown the transformation of pulse frequency into reward intensity is a logistic (sigmoidal) growth function, which can be approximated by a power function with an exponent from 2 to 10, over a portion of its domain (Hamilton, Stellar, & Hart, 1985; Gallistel & Leon, 1991; Leon & Gallistel, 1992; Mark & Gallistel, 1993; Simmons & Gallistel, 1994). As well, the time allocation function of the 3-dimensional reward mountain model incorporates and validates the form of this reward intensity function (Arvanitogiannis & Shizgal, 2008; Hernandez et al., 2010; Hernandez et al., 2010; Trujillo-Pisanty et al., 2011; Trujillo-Pisanty et al., 2013; Breton et al., 2013). The reward intensity function is expressed as:

$$RI_{rel}(D, F) = \frac{F^g}{F^g + [F_{hm}(D)]^g}$$

where,

$RI_{rel}$  = the relative reward intensity: the reward intensity as a proportion of maximal reward intensity

$F_{hm}$  = the pulse frequency that produces half-maximal reward intensity (Hz)

$g$  = the intensity growth exponent; controls the steepness of the reward-intensity function

$D$  = duration of the stimulation train in seconds (held constant at 0.5 s in these experiments)

$F$  = pulse frequency (Hz)

### ***5.1.2. The psychophysical function of opportunity costs: Chapter 2 summary***

When confronted with an option, the subject not only evaluates the reward intensity, but the cost of the reward as well. The cost of the reward in the behavioural paradigm employed in this thesis has two components: the effort cost, defined as the amount of exertion per unit time to obtain the reward, and the opportunity cost, the experimenter-set time required to work or engage in the task to secure the reward. The opportunity cost is incurred at the expense of the foregone opportunities or benefits arising from the next best option. In the present experiments, effort cost is held constant while the opportunity cost varies. The function that describes the translation of the experimenter-set opportunity cost (termed the objective price) into the equivalent subjective domain (subjective price) was measured in Chapter 2:

$$SP = SP_{\min} + (SP_{bnd} \times \ln(1 + e^{\frac{OP - SP_{\min}}{SP_{bnd}}}))$$

where,

$SP_{\min}$  = minimum subjective price (s)

$SP_{bnd}$  = controls the abruptness of the transition from the “blade” to “handle”

$OP$  = objective price (s)

$SP$  = subjective price (s)

The shape of the function resembles a hockey stick. It has a flat horizontal “blade” that extends over the range of initial, low costs. The flat part of the function along the low range of objective costs indicates that very low opportunity costs are subjectively equal to each other. For example, comparing an opportunity cost of 0.5 s to 0.25 s, the foregone benefits arising from the alternate activities that could be performed within these two intervals are equivalent. Further up on the scale, at around 2 s, the flat region is joined by a steep upward-curving portion of the curve that represents the range at which animals begin to discriminate between opportunity costs. Further, a straight “handle” extends over the range of higher costs in which the objective and subjective costs mirror each other. Parameters  $SP_{min}$  (the minimum subjective price that the rat estimates) and  $SP_{bnd}$  (the parameter that controls the bend from the blade to the handle) are within a narrow range across rats which strengthens the validity of the form of this function. The mean  $SP_{min}$  is 1.08 s, the mean  $SP_{bnd}$ , 0.15. The form of the function is described in Figure 1A.

### ***5.1.3. The psychophysical function of the frequency-following response: Chapter 3 summary***

The subjective components that comprise the rat’s computation of payoff have been described above: the subjective reward intensity and opportunity cost. However, the psychophysical translation of pulse frequency into reward intensity has an embedded stage that deserved testing and a formal description: the translation of the experimenter-

set pulse frequency ( $F$ , the inducing stimulus) into the actual firing frequency of the neuron ( $FF$ , the induced physiological response). The translation of pulse frequency into reward intensity as described earlier assumes that the firing frequency ( $FF$ ) perfectly follows the experimenter induced pulse frequency ( $F$ ). However, the experiment in Chapter 3 recognizes that this following must break down: as the pulse frequency becomes high, a physiological limit of maximal firing frequency that the substrate can attain is reached. The frequency-following response function was estimated in Chapter 3 using the 4-dimensional reward mountain model and testing paradigm (the additional independent variable: current intensity). The form of the function measured in Chapter 3 is expressed below:

$$FF = F_b \left( \ln\left(1 + e^{\frac{F_{nmax}}{F_b}}\right) - \ln\left(1 + e^{\frac{F_{nmax}-F}{F_b}}\right) \right)$$

where,

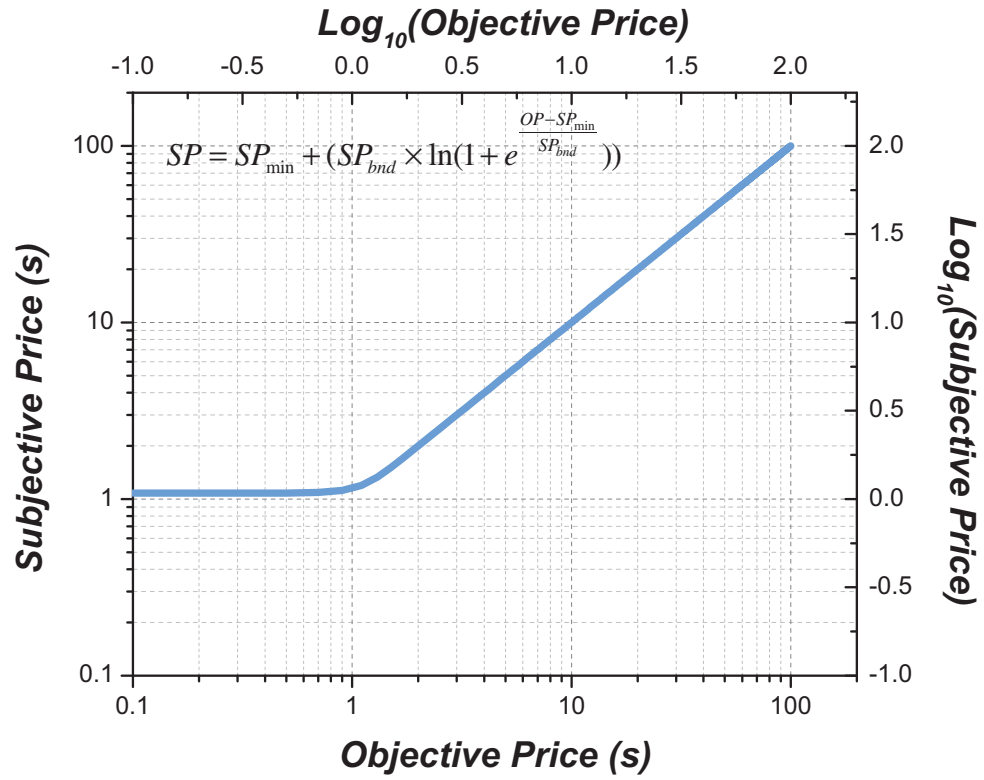
$FF$  = the average firing frequency (Hz); the induced physiological response

$F$  = the pulse frequency that is experimenter-set (Hz); the inducing stimulus value

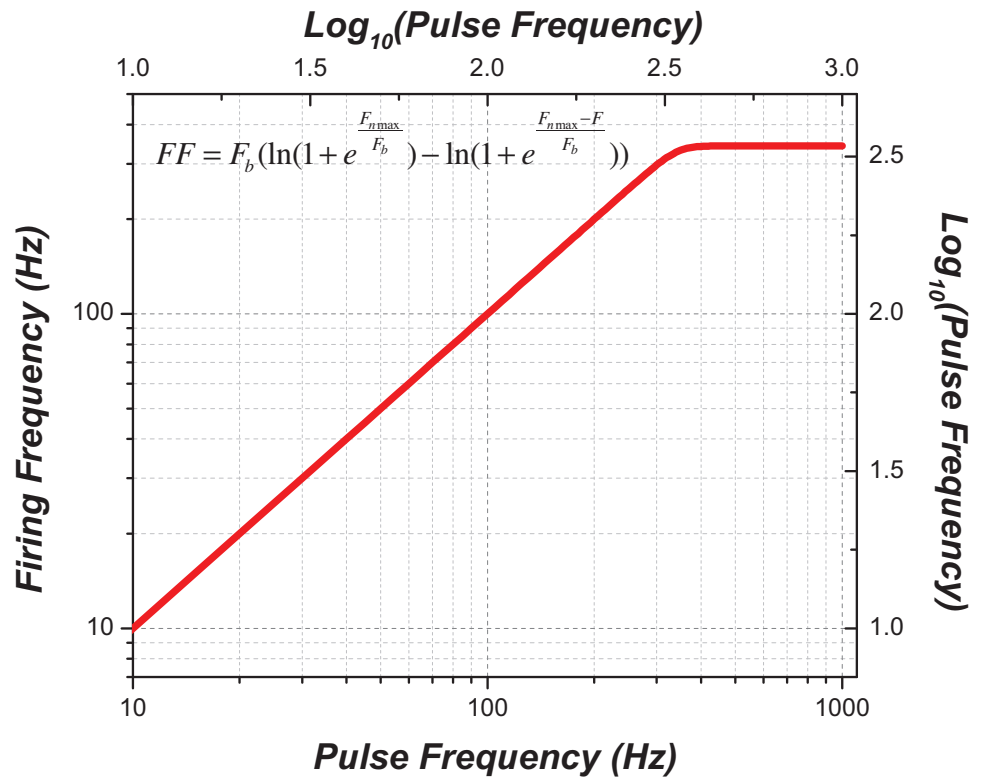
$F_{nmax}$  = the pulse frequency at which the firing frequency is *near* the maximal firing frequency that the substrate can attain; the position parameter of the frequency-response function

$F_b$  = the parameter describing the abruptness of the transition between the range of perfect frequency following to the range of frequency roll-off

Parameter  $F_{nmax}$  and  $F_b$  were both within a narrow range across rats. The median  $F_{nmax}$  is 363 Hz and  $F_b$  is 21. The form of the function is described in Figure 2B. To be conservative, the first quartile value of  $F_{nmax}$  (rounded to 316 Hz) will be used in future reward-mountain analyses that incorporate the psychophysical function of the frequency-following response.



**A**



**B**



Figure 1. A. The subjective price function estimated in Chapter 2. The mean values of the parameters across parameters are used here:  $SP_{\min}$ , 1.08 s and  $SP_{\text{bnd}}$ , 0.15. B. The frequency-following function estimated in Chapter 3. The median values of the parameters are  $F_{\text{max}}$ , 361 Hz and  $F_b$ , 21.

## ***5.2. Implications: psychophysical functions and neural signals***

Psychophysical functions can be used to identify the neural correlates of a given process by assessing whether the predicted encoding matches a neural signal. The frequency-following response function derived in this thesis has applications for this endeavor. Specifically, whether dopamine tracks the behaviourally derived frequency-following response function can reveal whether these neurons are transynaptically activated by the directly activated substrate. Several studies have used this rationale to investigate the role of dopamine in reward pursuit. Moisan and Rompré (1998) tested four behaviourally derived trade-off current-pulse frequencies combinations and showed (via electrophysiology) that dopamine neurons of the ventral tegmental induced by posterior mesencephalon electrical stimulation tracked the combinations. Cossette (2011) tested four behaviourally derived current-pulse frequency trade-off combinations and measured the electrochemical dopamine response in the nucleus accumbens (via cyclic voltammetry) that was induced by ventral tegmental area stimulation. Unlike Moisan and Rompré's results (1998), dopamine did not track the rewarding effect along the whole range of tested pulse frequency-current combinations. However, in both Moisan and Rompré and Cossette's experiments, the trade-off functions were not formally derived and only four points in the sampling space were tested. Formally estimating the function and tracking the dopamine response along a broader range of points provides greater resolution with which to assess the neural activity-psychophysical correspondence.

### 5.3. Implications: the psychophysical functions bolster the reward-mountain model

The psychophysical functions strengthen the computational framework of reward seeking employed in this thesis: the reward-mountain model. Through a series of stages, the reward-mountain model ties the electrical activation of the substrate that supports self-stimulation to the subject's operant behaviour. Hitherto, the behaviour (time allocated to holding down the lever) is expressed as:

$$TA(D, F, P) = TA_{\min} + \left[ (TA_{\max} - TA_{\min}) \times \frac{\left( \frac{F^g}{F^g + [F_{hm}(D)]^g} \right)^a}{\left( \frac{F^g}{F^g + [F_{hm}(D)]^g} \right)^a + \left( \frac{OP}{OP_e} \right)^a} \right]$$

where,

$F_{hm}$  = pulse frequency that produces half-maximal reward intensity (Hz)

$g$  = the intensity growth exponent; controls the steepness of the reward-intensity function

$D$  = duration of the stimulation train in seconds (held constant at 0.5 s in these experiments)

$a$  = the payoff-sensitivity exponent; represents how sensitive the rat is to the price of the reward; it accounts for over or under-matching

$TA_{\max}$  = the maximum time allocation

$TA_{\min}$  = the minimum time allocation

$OP$  = the objective price: the opportunity cost in seconds set by the experimenter (s)

$F$  = pulse frequency (Hz)

The two experimenter-set independent variables used to measure time allocation ( $TA$ ), pulse frequency ( $F$ ) and objective price ( $OP$ ) have now been expanded into the subjective realm: the firing frequency ( $FF$ ) and subjective price ( $SP$ ), respectively. Time allocation is now expressed as:

$$TA(D, F, P) = TA_{\min} + \left[ (TA_{\max} - TA_{\min}) \times \frac{\left( \frac{FF^g}{FF^g + [FF_{hm}(D)]^g} \right)^a}{\left( \frac{FF^g}{FF^g + [FF_{hm}(D)]^g} \right)^a + \left( \frac{SP}{SP_e} \right)^a} \right]$$

where,

$$FF = F_b \left( \ln \left( 1 + e^{\frac{F_{n\max}}{F_b}} \right) - \ln \left( 1 + e^{\frac{F_{n\max} - F}{F_b}} \right) \right)$$

and,

$$SP = SP_{\min} + (SP_{bnd} \times \ln \left( 1 + e^{\frac{OP - SP_{\min}}{SP_{bnd}}} \right))$$

The principal goal of the reward-mountain model and the associated measurement strategy is to determine the stage of neural circuitry that has been disrupted by a manipulation such as a pharmacological or physiological challenge. The 2-dimensional measurement strategies, wherein performance is measured as a function of one independent variable (either pulse frequency or cost), have the ability to detect an effect on self-stimulation but cannot distinguish at what stage of processing the effect has occurred. Specifically, a given manipulation could have impacted the directly activated

neurons or alternatively, a subsequent processing stage such as the one that computes the perceived cost of the reward. The 3-dimensional measurement strategy has the ability to reduce this ambiguity by measuring performance as a function of both the reward intensity and opportunity cost. The direction of the displacements of the 3-dimensional structure can reveal whether a given manipulation affects the directly activated neurons (as indicated by displacements along the pulse frequency axis), or a stage further downstream, subsequent to the integration of the rewarding effect (as indicated by displacements along the cost axis). The stages of reward processing that now incorporate the forms of the subjective price function and frequency-following response function are illustrated in Figure 2.

Thus far, the reward-mountain model has been validated (Arvanitogiannis et al. 2008; Breton et al. 2013). Experiments using this testing paradigm have challenged a long-standing viewpoint regarding the role of dopamine in reward processing (Hernandez et al., 2010; Hernandez et al., 2012, Trujillo-Pisanty et al., 2013). In addition, the ability of the 3-dimensional paradigm to detect a lesion challenge as demonstrated by the proof of principle study in Chapter 4 implies that the model is well suited for the detection and measurement of lesion-induced effects of targeted brain regions. Aptly, new optogenetic tools allow for the activation and silencing of specific neural populations. When matched with the sophisticated computational approach of the reward-mountain model, the potential to deconstruct the complexity of the underlying circuitry is unprecedented.

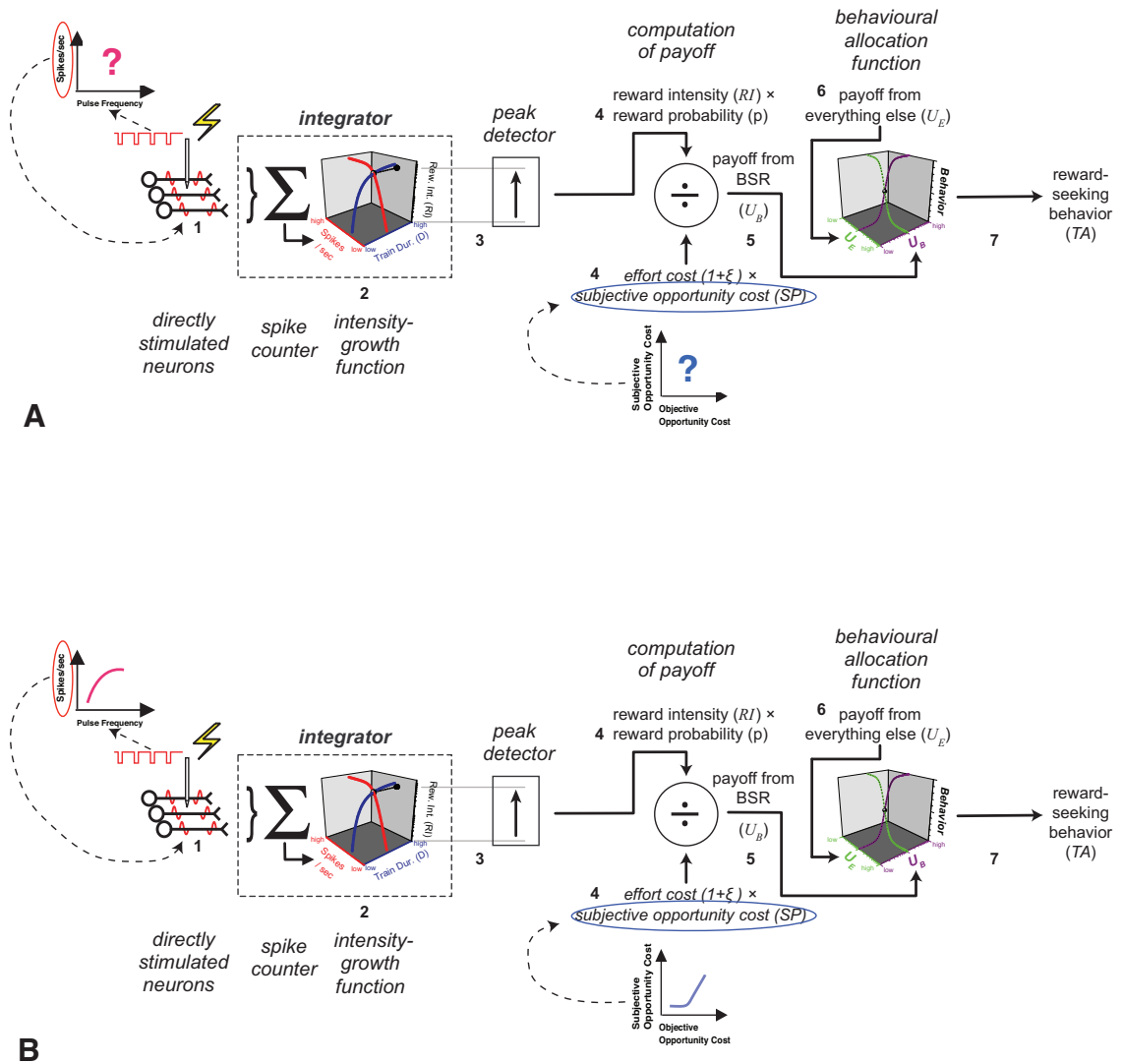


Figure 2. Graphical representation of the reward-mountain model. A. The psychophysical functions estimated in this thesis are denoted by the question marks: (i) the frequency-following response function represented in pink and (ii) the subjective price function represented in blue. B. The question marks are replaced by the psychophysical functions.

#### ***5.4. The necessity for specifying psychophysical functions in the reward-mountain model***

If the assumption is that the objective and subjective variables are identical along the whole range, the detection of the pharmacological or lesion manipulations could potentially be obfuscated. Consider a case in which, a lesion causes a disruption of first-stage neurons: an increase in pulse frequencies (due to a shift along the pulse frequency axis) to restore baseline behavioural levels would be required. However, if the tested pulse frequencies exceed the firing frequency limit ( $F_{max}$ ), a premature reduction in the maximal achievable reward intensity could occur. This reduction in maximal reward intensity could lead to a displacement of the structure along the cost axis, thus resulting in flawed inferences regarding the affected stages. Knowing the limits within which to set the pulse frequencies (as estimated in Chapter 3) circumvents this potential problem.

A test of the subjective-opportunity cost function was needed: if the cost is used to scale the form of the reward-intensity function, it is necessary for the scale to embody the subjective interpretation of the cost. Making incorrect assumptions regarding the form of the cost scale could lead to an incorrect estimation of the reward-intensity function, which in turn could result in imprecise measures of the displacements of the reward-mountain structure. The experiment presented in Chapter 2 demonstrated the objective-subjective price relationship. The two are scalar over a range, and it is the scalar range that should be employed in the testing paradigm.

### ***5.5. Additional psychophysical functions***

In addition to reward intensity and opportunity cost, the reward-mountain model includes other variables: probability, effort cost, and delay (all held constant within the present and previous experiments). A full generalization of the reward-mountain model would entail extending these variables into the subjective domain. The same logic used in this thesis can be used to estimate the psychophysical functions of these variables.

### ***5.6. Summary***

Reward-seeking behaviour is complex. The multifaceted nature of the goal and context, such as the intensity of the reward, associated costs, selection, execution, state of the subject, and competing options make for a demanding computational task. Remarkably, the brain performs these multistage computations seemingly at ease. How such computations are carried out and the nature of the encoding has broad implications, from animal foraging to how humans make choices. Furthermore, working out the circuitry has implications for disorders such as drug dependence and obesity, impairments understood within the context of dysfunctions in reward processing and decision-making.

Accordingly, behavioural ecologists have long considered animal behaviour from the perspective of benefits and costs (Stephens and Krebs, 1986). In recent years, numerous neuroscientific studies have been directed at the analysis of cost-benefit based decision-making and behaviour. A non-exhaustive list of the subjects of such studies includes rodents (Salamone, Cousins, Bucher, 1994; Denk, Waltong, Jennings, Sharp, & Rushworth, 2005), humans (Croxxson, Walton, O'Reilly, Behrens, & Rushworth, 2010;



Basten, Biele, Heekeren, & Fiebach, 2010), nonhuman primates (Hosokawa, Kennerley, Sloan, Wallis, 2013; Pasquereau & Turner, 2013) and lower organisms (Hirayama, Catanho, Brown, & Gilette, 2012). The ultimate purpose of these studies is to link variables of reward seeking to the underlying neural mechanisms.

Computational models of information processing and sophisticated testing paradigms such as the 3-dimensional model of brain stimulation reward in the rodent afford a way to identify and understand the neural representation of the reward and associated variables, motivated behaviour, and the decision-making process.

## REFERENCES

- Aberman, J. E., & Salamone, J. D. (1999). Nucleus accumbens dopamine depletions make rats more sensitive to high ratio requirements but do not impair primary food reinforcement. *Neuroscience*, *92*(2), 545-552.
- Andersson, B. (1953). The effect of injections of hypertonic NaCl-solutions into Different Parts of the hypothalamus of goats. *Acta Physiologica Scandinavica*. *28*(2-3), 188-201.
- Arvanitogiannis, A., Waraczynski, M., & Shizgal, P. (1996). Effects of excitotoxic lesions of the basal forebrain on MFB self-stimulation. *Physiology & Behavior*, *59*(4), 795-806.
- Arvanitogiannis, A., Flores, C., & Shizgal, P. (1997). Fos-like immunoreactivity in the caudal diencephalon and brainstem following lateral hypothalamic self-stimulation. *Behavioural Brain Research*, *88*(2), 275-279.
- Arvanitogiannis, A., & Shizgal, P. (2008). The reinforcement mountain: allocation of behavior as a function of the rate and intensity of rewarding brain stimulation. *Behavioral Neuroscience*, *122*(5), 1126-1138.
- Balagura, S., & Hoebel, B. G. (1967). Self-stimulation of the lateral hypothalamus modified by insulin and glucagon. *Physiology & Behavior*, *2*(4), 337-340.
- Basten, U., Biele, G., Heekeren, H. R., & Fiebach, C. J. (2010). How the brain integrates costs and benefits during decision making. *Proceedings of the National Academy of Sciences*, *107*(50), 21767-21772.
- Baum, W. M., & Rachlin, H. C. (1969). Choice as time allocation. *Journal of the Experimental Analysis of Behavior*, *12*(6), 861-874.
- Beyene, M., Carelli, R. M., & Wightman, R. M. (2010). Cue-evoked dopamine release in the nucleus accumbens shell tracks reinforcer magnitude during intracranial self-stimulation. *Neuroscience*, *169*(4), 1682-1688.
- Bielajew, C., & Shizgal, P. (1982). Behaviorally derived measures of conduction velocity in the substrate for rewarding medial forebrain bundle stimulation. *Brain Research*, *237*(1), 107-119.
- Bielajew, C., & Shizgal, P. (1986). Evidence implicating descending fibers in self-stimulation of the medial forebrain bundle. *The Journal of Neuroscience*, *6*(4), 919-929.

- Bielajew, C., Thrasher, A., & Fouriez, G. (1987). Self-stimulation sites in the lateral hypothalamic and lateral preoptic areas are functionally connected. *Canadian Psychology*, 28, 36.
- Bielajew, C., Bushnik, T., Konkle, A., & Schindler, D. (2000). The substrate for brain-stimulation reward in the lateral preoptic area: II. Connections to the ventral tegmental area. *Brain Research*, 881(2), 112-120.
- Bielajew, C., Konkle, A. T. M., Fouriez, G., Boucher-Thrasher, A., & Schindler, D. (2001). The substrate for brain-stimulation reward in the lateral preoptic area: III. Connections to the lateral hypothalamic area. *Behavioral Neuroscience*, 115(4), 900-909.
- Breton, Y. A., Mullett, A., Conover, K., & Shizgal, P. (2013). Validation and extension of the reward-mountain model. *Frontiers in Behavioral Neuroscience*, 7.
- Breton, Y. A. (2013b). *Molar and Molecular Models of Performance for Rewarding Brain Stimulation*. Unpublished doctoral dissertation, Concordia University, Montreal, Canada.
- Boye, S. M., & Rompré, P. P. (1996). Mesencephalic substrate of reward: axonal connections. *The Journal of Neuroscience*, 16(10), 3511-3520.
- Boyd, E. S., & Gardner, L. C. (1967). Effect of some brain lesions on intracranial self-stimulation in the rat. *American Journal of Physiology--Legacy Content*, 213(4), 1044-1052.
- Breton, Y. A., Marcus, J. C., & Shizgal, P. (2009). *Rattus Psychologicus*: construction of preferences by self-stimulating rats. *Behavioural Brain Research*, 202(1), 77-91.
- Breton, Y. A., Mullett, A., Conover, K., & Shizgal, P. (2013). Validation and extension of the reward-mountain model. *Frontiers in Behavioral Neuroscience*, 7.
- Carter, D. A., & Phillips, A. G. (1975). Intracranial self-stimulation at sites in the dorsal medulla oblongata. *Brain Research*, 94(1), 155-160.
- Cheer, J. F., Aragona, B. J., Heien, M. L., Seipel, A. T., Carelli, R. M., & Wightman, R. M. (2007). Coordinated accumbal dopamine release and neural activity drive goal-directed behavior. *Neuron*, 54(2), 237-244.
- Chergui, K., Charlety, P. J., Akaoka, H., Saunier, C. F., Brunet, J. L., Buda, M., ... & Chouvet, G. (1993). Tonic activation of NMDA receptors causes spontaneous burst discharge of rat midbrain dopamine neurons in vivo. *European Journal of Neuroscience*, 5(2), 137-144.

- Christoph, G. R., Leonzio, R. J., & Wilcox, K. S. (1986). Stimulation of the lateral habenula inhibits dopamine-containing neurons in the substantia nigra and ventral tegmental area of the rat. *The Journal of Neuroscience*, *6*(3), 613-619.
- Colle, L. M., & Wise, R. A. (1987). Opposite effects of unilateral forebrain ablations on ipsilateral and contralateral hypothalamic self-stimulation. *Brain Research*, *407*(2), 285-293.
- Conover, K. L., & Shizgal, P. (1994a). Competition and summation between rewarding effects of sucrose and lateral hypothalamic stimulation in the rat. *Behavioral Neuroscience*, *108*(3), 537-548.
- Conover, K. L., & Shizgal, P. (1994b). Differential effects of postingestive feedback on the reward value of sucrose and lateral hypothalamic stimulation in rats. *Behavioral Neuroscience*, *108*(3), 559-572.
- Conover, K. L., Woodside, B., & Shizgal, P. (1994). Effects of sodium depletion on competition and summation between rewarding effects of salt and lateral hypothalamic stimulation in the rat. *Behavioral Neuroscience*, *108*(3), 549-558.
- Conover, K. L., & Shizgal, P. (2005). Employing labor-supply theory to measure the reward value of electrical brain stimulation. *Games and Economic Behavior*, *52*(2), 283-304.
- Corbett, D., & Wise, R. A. (1980). Intracranial self-stimulation in relation to the ascending dopaminergic systems of the midbrain: a moveable electrode mapping study. *Brain Research*, *185*(1), 1-15.
- Corbett, D., Fox, E., & Milner, P. M. (1982). Fiber pathways associated with cerebellar self-stimulation in the rat: a retrograde and anterograde tracing study. *Behavioural Brain Research*, *6*(2), 167-184.
- Cossette, M.P. (2011). *Transient release of dopamine in response to rewarding stimulation of the medial forebrain bundle*. Unpublished master's dissertation, Concordia University, Montreal, Canada.
- Cousins, M. S., Wei, W., & Salamone, J. D. (1994). Pharmacological characterization of performance on a concurrent lever pressing/feeding choice procedure: effects of dopamine antagonist, cholinomimetic, sedative and stimulant drugs. *Psychopharmacology*, *116*(4), 529-537.
- Crosson, P. L., Walton, M. E., O'Reilly, J. X., Behrens, T. E., & Rushworth, M. F. (2009). Effort-based cost-benefit valuation and the human brain. *The Journal of Neuroscience*, *29*(14), 4531-4541.

- Denk, F., Walton, M. E., Jennings, K. A., Sharp, T., Rushworth, M. F. S., & Bannerman, D. M. (2005). Differential involvement of serotonin and dopamine systems in cost-benefit decisions about delay or effort. *Psychopharmacology*, *179*(3), 587-596.
- Deutsch, J. A. (1964). Behavioral measurement of the neural refractory period and its application to intracranial self-stimulation. *Journal of Comparative and Physiological Psychology*, *58*(1), 1.
- Edmonds, D. E., & Gallistel, C. R. (1974). Parametric analysis of brain stimulation reward in the rat: III. Effect of performance variables on the reward summation function. *Journal of Comparative and Physiological Psychology*, *87*(5), 876.
- Efron, B., & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, 54-75.
- Forgie, M. L., & Shizgal, P. (1993). Mapping the substrate for brain stimulation reward by means of current-number trade-off functions. *Behavioral Neuroscience*, *107*(3), 506-524.
- Fouriez, G., & Randall, D. (1997). The cost of delaying rewarding brain stimulation. *Behavioural Brain Research*, *87*(1), 111-113.
- Fouriez, G., Bielajew, C., & Pagotto, W. (1990). Task difficulty increases thresholds of rewarding brain stimulation. *Behavioural Brain Research*, *37*(1), 1-7.
- Frank, R. A., & Williams, H. P. (1985). Both response effort and current intensity affect self-stimulation train duration thresholds. *Pharmacology Biochemistry and Behavior*, *22*(4), 527-530.
- Franklin, K. B. J. (1978). Catecholamines and self-stimulation: Reward and performance effects dissociated. *Pharmacology Biochemistry and Behavior*, *9*(6), 813-820.
- Fulton, S., Woodside, B., & Shizgal, P. (2000). Modulation of brain reward circuitry by leptin. *Science*, *287*(5450), 125-128.
- Gallistel, C. R., Rolls, E., & Greene, D. (1969). Neuron function inferred from behavioral and electrophysiological estimates of refractory period. *Science*, *166*(3908), 1028-1030.
- Gallistel, C. R. (1974). Motivation as central organizing process: The psychophysical approach to its functional and neurophysiological analysis. In *Nebraska symposium on motivation*. University of Nebraska Press.
- Gallistel, C. R., Stellar, J. R., & Bubis, E. (1974). Parametric analysis of brain stimulation reward in the rat: I. The transient process and the memory-containing process. *Journal of Comparative and Physiological Psychology*, *87*(5), 876-883.

- Gallistel, C. R. (1978). Self-stimulation in the rat: quantitative characteristics of the reward pathway. *Journal of Comparative and Physiological Psychology*, 92(6), 977-998.
- Gallistel, C. R., Shizgal, P., & Yeomans, J. S. (1981). A portrait of the substrate for self-stimulation. *Psychological Review*, 88(3), 228-273.
- Gallistel, C. R., Boytim, M., Gomita, Y., & Klebanoff, L. (1982). Does pimozide block the reinforcing effect of brain stimulation? *Pharmacology Biochemistry and Behavior*, 17(4), 769-781.
- Gallistel, C. R., & Leon, M. (1991). Measuring the subjective magnitude of brain stimulation reward by titration with rate of reward. *Behavioral Neuroscience*, 105(6), 913.
- Gallistel, C. R., Leon, M., Waraczynski, M., & Hanau, M. S. (1991). Effect of current on the maximum possible reward. *Behavioral Neuroscience*, 105(6), 901-912.
- Gallistel, C. R., Leon, M., Lim, B. T., Sim, J. C., & Waraczynski, N. (1996). Destruction of the medial forebrain bundle caudal to the site of stimulation reduces rewarding efficacy but destruction rostrally does not. *Behavioral Neuroscience*, 110(4), 766-790.
- Garris, P. A., Kilpatrick, M., Bunin, M. A., Michael, D., Walker, Q. D., & Wightman, R. M. (1999). Dissociation of dopamine release in the nucleus accumbens from intracranial self-stimulation. *Nature*, 398(6722), 67-69.
- Geisler, S., Derst, C., Veh, R. W., & Zahm, D. S. (2007). Glutamatergic afferents of the ventral tegmental area in the rat. *The Journal of Neuroscience*, 27(21), 5730-5743.
- German, D. C., & Bowden, D. M. (1974). Catecholamine systems as the neural substrate for intracranial self-stimulation: a hypothesis. *Brain Research*, 73(3), 381-419.
- Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, 84(3), 279.
- Glickman, S. E., & Schiff, B. B. (1967). A biological theory of reinforcement. *Psychological Review*, 74(2), 81-109.
- Hamilton, A. L., Stellar, J. R., & Hart, E. B. (1985). Reward, performance, and the response strength method in self-stimulating rats: Validation and neuroleptics. *Physiology & Behavior*, 35(6), 897-904.
- Herberg, L. J. (1963). Seminal ejaculation following positively reinforcing electrical stimulation of the rat hypothalamus. *Journal of Comparative and Physiological Psychology*, 56(4), 679-685.

- Hernandez, L., & Hoebel, B. G. (1988). Food reward and cocaine increase extracellular dopamine in the nucleus accumbens as measured by microdialysis. *Life Sciences*, *42*(18), 1705-1712.
- Hernandez, G., Hamdani, S., Rajabi, H., Conover, K., Stewart, J., Arvanitogiannis, A., & Shizgal, P. (2006). Prolonged rewarding stimulation of the rat medial forebrain bundle: neurochemical and behavioral consequences. *Behavioral Neuroscience*, *120*(4), 888-904.
- Hernandez, G., Haines, E., Rajabi, H., Stewart, J., Arvanitogiannis, A., & Shizgal, P. (2007). Predictable and unpredictable rewards produce similar changes in dopamine tone. *Behavioral Neuroscience*, *121*(5), 887-895.
- Hernandez, G., Breton, Y. A., Conover, K., & Shizgal, P. (2010). At what stage of neural processing does cocaine act to boost pursuit of rewards? *PLoS One*, *5*(11), e15081.
- Hernandez, G., Trujillo-Pisanty, I., Cossette, M. P., Conover, K., & Shizgal, P. (2012). Role of dopamine tone in the pursuit of brain stimulation reward. *The Journal of Neuroscience*, *32*(32), 11032-11041.
- Herrnstein, R. J. (1970). On the Law of Effect. *Journal of the Experimental Analysis of Behavior*, *13*(2), 243-266.
- Herrnstein, R. J. (1974). Formal properties of the matching law. *Journal of the Experimental Analysis of Behavior*, *21*(1), 159-164.
- Heyman G (1988). How drugs affect cells and reinforcement affects behavior: Formal analogies. In: Commons ML, Church RM, eds. *Biological determinants of reinforcement Quantitative analyses of behavior*, Vol 7. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc. pp 157–182.
- Hirayama, K., Catanho, M., Brown, J. W., & Gillette, R. (2012). A core circuit module for cost/benefit decision. *Frontiers in Neuroscience*, *6*.
- Hoebel, B. G., & Teitelbaum, P. (1962). Hypothalamic control of feeding and self-stimulation. *Science*, *135*(3501), 375-377.
- Hoebel, B. G. (1965). Hypothalamic lesions by electrocauterization: Disinhibition of feeding and self-stimulation. *Science*, *149*, 452-453.
- Hoebel, B. G. (1968). Inhibition and disinhibition of self-stimulation and feeding: hypothalamic control and postingestional factors. *Journal of Comparative and Physiological Psychology*, *66*(1), 89-100.
- Hoebel, B. G. (1969). Feeding and self-stimulation. *Annals of the New York Academy of Sciences*, *157*(2), 758-778.

Hoebel, B. G., & Thompson, R. D. (1969). Aversion to lateral hypothalamic stimulation caused by intragastric feeding or obesity. *Journal of Comparative and Physiological Psychology*, 68(4), 536-543.

Hoebel, B. G. (1976). Satiety: Hypothalamic stimulation, anorectic drugs, and neurochemical substrates. *Hunger; Basic Mechanisms and Clinical Implications*, D. Novin, W. Wyrwicka & GA Bray, eds, 1976.

Hosokawa, T., Kennerley, S. W., Sloan, J., & Wallis, J. D. (2013). Single-neuron mechanisms underlying cost-benefit analysis in frontal cortex. *The Journal of Neuroscience*, 33(44), 17385-17397.

Huston, J. P., Ornstein, K., & Lehner, R. (1982). The diencephalic peninsula: self-stimulation after unilateral precollicular transection and removal of the telencephalon. *Brain Research*, 245(1), 187-191.

Ikemoto, S. (2010). Brain reward circuitry beyond the mesolimbic dopamine system: a neurobiological theory. *Neuroscience & Biobehavioral Reviews*, 35(2), 129-150.

Janas, J. D., & Stellar, J. R. (1987). Effects of knife-cut lesions of the medial forebrain bundle in self-stimulating rats. *Behavioral Neuroscience*, 101(6), 832-845.

Jennings, J.H. et al., 2013. Distinct extended amygdala circuits for divergent motivational states. *Nature*, 496(7444), 224–228.

Johnson, S. W., Seutin, V., & North, R. A. (1992). Burst firing in dopamine neurons induced by N-methyl-D-aspartate: role of electrogenic sodium pump. *Science*, 258(5082), 665-667.

Johnson, P. I., & Stellar, J. R. (1994). N-Methyl-D-aspartic acid-induced lesions of the nucleus accumbens and/or ventral pallidum fail to attenuate lateral hypothalamic self-stimulation reward. *Brain Research*, 646(1), 73-84.

Kalivas, P. W., Duffy, P., & Barrow, J. (1989). Regulation of the mesocorticolimbic dopamine system by glutamic acid receptor subtypes. *Journal of Pharmacology and Experimental Therapeutics*, 251(1), 378-387.

Keesey, R. E., & Powley, T. L. (1973). Self-stimulation and body weight in rats with lateral hypothalamic lesions. *The American Journal of Physiology*, 224(4), 970-978.

Kempadoo, K. A., Tourino, C., Cho, S. L., Magnani, F., Leininger, G. M., Stuber, G. D., ... & Bonci, A. (2013). Hypothalamic neurotensin projections promote reward by enhancing glutamate transmission in the VTA. *The Journal of Neuroscience*, 33(18), 7618-7626.



- Killeen, P. (1972). The matching law. *Journal of the Experimental Analysis of Behavior*, 17(3), 489-495.
- Leon, M., & Gallistel, C. R. (1992). The function relating the subjective magnitude of brain stimulation reward to stimulation strength varies with site of stimulation. *Behavioural Brain Research*, 52(2), 183-193.
- Leon, M. I., & Gallistel, C. R. (1998). Self-stimulating rats combine subjective reward magnitude and subjective reward rate multiplicatively. *Journal of Experimental Psychology: Animal Behavior Processes*, 24(3), 265-277.
- Lorens, S. A. (1966). Effect of lesions in the central nervous system on lateral hypothalamic self-stimulation in the rat. *Journal of Comparative and Physiological Psychology*, 62(2), 256-262.
- Margules, D. L., & Olds, J. (1962). Identical "feeding" and "rewarding" systems in the lateral hypothalamus of rats. *Science*, 135(3501), 374-375.
- Mark, T. A., & Gallistel, C. R. (1993). Subjective reward magnitude of medial forebrain stimulation as a function of train duration and pulse frequency. *Behavioral Neuroscience*, 107(2), 389.
- Mazur J.E (1987). An adjusting procedure for studying delayed reinforcement. In: Commons M.L, Mazur J.E, Nevin J.A, Rachlin H, editors. Quantitative analyses of behavior: Vol. 5. The effect of delay and of intervening events on reinforcement value. Hillsdale, NJ: Erlbaum; 1987. pp. 55–73.
- Mazur, J. E., Stellar, J. R., & Waraczynski, M. (1987). Self-control choice with electrical stimulation of the brain as a reinforcer. *Behavioural Processes*, 15(2), 143-153.
- McDowell, J. J. (2005). On the classic and modern theories of matching. *Journal of the Experimental Analysis of Behavior*, 84(1), 111-127.
- Miliaressis, E., Rompre, P. P., Laviolette, P., Philippe, L., & Coulombe, D. (1986). The curve-shift paradigm in self-stimulation. *Physiology & Behavior*, 37(1), 85-91.
- Miller, N. E. (1957). Experiments on motivation: Studies combining psychological, physiological, and pharmacological techniques. *Science*, 126(3286), 1271-8.
- Miller, H. L. (1976). Matching-based hedonic scaling in the pigeon. *Journal of the Experimental Analysis of Behavior*, 26(3), 335-347.
- Mogenson, G. J., & Morgan, C. W. (1967). Effects of induced drinking on self-stimulation of the lateral hypothalamus. *Experimental Brain Research*, 3(2), 111-116.

- Moisan, J., & Rompré, P. P. (1998). Electrophysiological evidence that a subset of midbrain dopamine neurons integrate the reward signal induced by electrical stimulation of the posterior mesencephalon. *Brain Research*, *786*(1), 143-152.
- Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of Neuroscience*, *16*(5), 1936-1947.
- Morissette, M. C., & Boye, S. M. (2008). Electrolytic lesions of the habenula attenuate brain stimulation reward. *Behavioural Brain Research*, *187*(1), 17-26.
- Murray, B., & Shizgal, P. (1991). Anterolateral lesions of the medial forebrain bundle increase the frequency threshold for self-stimulation of the lateral hypothalamus and ventral tegmental area in the rat. *Psychobiology*, *19*(2), 135-146.
- Murray, B., & Shizgal, P. (1994). Evidence implicating both slow-and fast-conducting fibers in the rewarding effect of medial forebrain bundle stimulation. *Behavioural Brain Research*, *63*(1), 47-60.
- Murray, B., & Shizgal, P. (1996a). Physiological measures of conduction velocity and refractory period for putative reward-relevant MFB axons arising in the rostral MFB. *Physiology & Behavior*, *59*(3), 427-437.
- Murray, B., & Shizgal, P. (1996b). Behavioral measures of conduction velocity and refractory period for reward-relevant axons in the anterior LH and VTA. *Physiology & Behavior*, *59*(4), 643-652.
- Murray, B., & Shizgal, P. (1996c). Attenuation of medial forebrain bundle reward by anterior lateral hypothalamic lesions. *Behavioural Brain Research*, *75*(1), 33-47.
- Nakahara, D., Ozaki, N., Miura, Y., Miura, H., & Nagatsu, T. (1989). Increased dopamine and serotonin metabolism in rat nucleus accumbens produced by intracranial self-stimulation of medial forebrain bundle as measured by in vivo microdialysis. *Brain Research*, *495*(1), 178-181.
- Neuringer, A. J. (1967). Effects of reinforcement magnitude on choice and rate of responding. *Journal of the Experimental Analysis of Behavior*, *10*(5), 417-424.
- Nicola, S. M. (2010). The flexible approach hypothesis: unification of effort and cue-responding hypotheses for the role of nucleus accumbens dopamine in the activation of reward-seeking behavior. *The Journal of Neuroscience*, *30*(49), 16585-16600.
- Nieuwenhuys, R., Geeraedts, L. M., & Veening, J. G. (1982). The medial forebrain bundle of the rat. I. General introduction. *Journal of Comparative Neurology*, *206*(1), 49-81.

- Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507-520.
- Olds, J., & Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology*, *47*(6), 419.
- Olds, J. (1958). Self-stimulation of the brain its use to study local effects of hunger, sex, and drugs. *Science*, *127*(3294), 315-324.
- Olds, J., Yuwiler, A., Olds, M. E., & Yun, C. (1964). Neurohumors in hypothalamic substrates of reward. *American Journal of Physiology--Legacy Content*, *207*(1), 242-254.
- Olds, M. E., & Olds, J. (1969). Effects of lesions in medial forebrain bundle on self-stimulation behavior. *American Journal of Physiology--Legacy Content*, *217*(5), 1253-1264.
- Omelchenko, N., Bell, R., & Sesack, S. R. (2009). Lateral habenula projections to dopamine and GABA neurons in the rat ventral tegmental area. *European Journal of Neuroscience*, *30*(7), 1239-1250.
- Paxinos, G., & Watson, C. (2007). *The rat brain in stereotaxic coordinates: hard cover edition*. Academic press.
- Pasquereau, B., & Turner, R. S. (2013). Limited Encoding of Effort by Dopamine Neurons in a Cost-Benefit Trade-off Task. *The Journal of Neuroscience*, *33*(19), 8288-8300.
- Rachlin, H. (1971). On the tautology of the matching law. *Journal of the Experimental Analysis of Behavior*, *15*(2), 249-251.
- Rada, P. V., Mark, G. P., Yeomans, J. J., & Hoebel, B. G. (2000). Acetylcholine release in ventral tegmental area by hypothalamic self-stimulation, eating, and drinking. *Pharmacology Biochemistry and Behavior*, *65*(3), 375-379.
- Redgrave, P., & Horrell, R. I. (1976). Potentiation of central reward by localised perfusion of acetylcholine and 5-hydroxytryptamine. *Nature*, *262*(5566), 305-7.
- Rolls, E. T. (1971). Involvement of brainstem units in medial forebrain bundle self-stimulation. *Physiology & Behavior*, *7*(3), 297-310.
- Rompré, P. P., & Shizgal, P. (1986). Electrophysiological characteristics of neurons in forebrain regions implicated in self-stimulation of the medial forebrain bundle in the rat. *Brain Research*, *364*(2), 338-349.

- Rompré, P. P., & Bauco, P. (1990). GBR 12909 reverses the SCH 23390 inhibition of rewarding effects of brain stimulation. *European Journal of Pharmacology*, *182*(1), 181-184.
- Routtenberg, A., & Lindy, J. (1965). Effects of the availability of rewarding septal and hypothalamic stimulation on bar pressing for food under conditions of deprivation. *Journal of Comparative and Physiological Psychology*, *60*(2), 158.
- Routtenberg, A., & Malsbury, C. (1969). Brainstem pathways of reward. *Journal of Comparative and Physiological Psychology*, *68*(1), 22-30.
- Salamone, J. D., Cousins, M. S., & Bucher, S. (1994). Anhedonia or anergia? Effects of haloperidol and nucleus accumbens dopamine depletion on instrumental response selection in a T-maze cost/benefit procedure. *Behavioural Brain Research*, *65*(2), 221-229.
- Salamone, J. D., & Correa, M. (2002). Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behavioural Brain Research*, *137*(1), 3-25.
- Salamone, J. D., Correa, M., Mingote, S., & Weber, S. M. (2003). Nucleus accumbens dopamine and the regulation of effort in food-seeking behavior: implications for studies of natural motivation, psychiatry, and drug abuse. *Journal of Pharmacology and Experimental Therapeutics*, *305*(1), 1-8.
- Salamone, J. D., Correa, M., Farrar, A., & Mingote, S. M. (2007). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology*, *191*(3), 461-482.
- Samuelson, P. A. (1937). A note on measurement of utility. *The Review of Economic Studies*, *4*(2), 155-161.
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593-1599.
- Shizgal, P., & Matthews, G. (1977). Electrical stimulation of the rat diencephalon: Differential effects of interrupted stimulation on on-and off-responding. *Brain Research*, *129*(2), 319-333.
- Shizgal, P., Bielajew, C., Corbett, D., Skelton, S. R., & Yeomans, J. (1980). Behavioral methods for inferring anatomical linkage between rewarding brain stimulation sites. *Journal of Comparative and Physiological Psychology*, *94*(2), 227.
- Shizgal, P. (1989). Toward a cellular analysis of intracranial self-stimulation: contributions of collision studies. *Neuroscience & Biobehavioral Reviews*, *13*(2), 81-90.

- Shizgal, P., Schindler, D., & Rompré, P. P. (1989). Forebrain neurons driven by rewarding stimulation of the medial forebrain bundle in the rat: comparison of psychophysical and electrophysiological estimates of refractory periods. *Brain Research*, 499(2), 234-248.
- Shizgal, P. (1997). Neural basis of utility estimation. *Current Opinion in Neurobiology*, 7(2), 198-208.
- Shizgal, P. (2012). Scarce means with alternative uses: robbins' definition of economics and its extension to the behavioral and neurobiological study of animal decision making. *Frontiers in Neuroscience*, 6.
- Simmons, J. M., & Gallistel, C. R. (1994). Saturation of subjective reward magnitude as a function of current and pulse frequency. *Behavioral Neuroscience*, 108(1), 151.
- Sonnenschein, B., Conover, K., & Shizgal, P. (2003). Growth of brain stimulation reward as a function of duration and stimulation strength. *Behavioral Neuroscience*, 117(5), 978-994.
- Stein, L. (1980). The chemistry of reward. *Biology of Reinforcement: Facets of Brain Stimulation Reward*, Academic Press, New York, 109-130.
- Stephens, D. W., and Krebs, J. R. (1986). Foraging Theory. Princeton, NJ: Princeton University Press.
- Tehovnik, E. J. (1996). Electrical stimulation of neural tissue to evoke behavioral responses. *Journal of Neuroscience Methods*, 65(1), 1-17.
- Trujillo-Pisanty, I., Hernandez, G., Moreau-Debord, I., Cossette, M. P., Conover, K., Cheer, J. F., & Shizgal, P. (2011). Cannabinoid receptor blockade reduces the opportunity cost at which rats maintain operant performance for rewarding brain stimulation. *The Journal of Neuroscience*, 31(14), 5426-5435.
- Trujillo-Pisanty, I., Conover, K., & Shizgal, P. (2013). A new view of the effect of dopamine receptor antagonism on operant performance for rewarding brain stimulation in the rat. *Psychopharmacology*, 1-14.
- Trujillo-Pisanty, I., Martel, A., Arvanitogiannis, A., Shizgal, P. (2013). *On the role of basal-forebrain MFB efferents in self-stimulation of the rat medial forebrain bundle*. Poster session presented at the annual meeting of the Society for Neuroscience. San Diego, CA.
- Ungerstedt, U. (1970). Stereotaxic mapping of the monoamine pathways in the rat brain. *Acta physiologica Scandinavica. Supplementum*, 367, 1-48.

- Valenstein, E. S. (1964). Problems of measurement and interpretation with reinforcing brain stimulation. *Psychological Review*, 71(6), 415.
- Valenstein, E. S., Cox, V. C., & Kakolewski, J. W. (1970). Reexamination of the role of the hypothalamus in motivation. *Psychological Review*, 77(1), 16.
- Veening, J. G., Swanson, L. W., Cowan, W. M., Nieuwenhuys, R., & Geeraedts, L. M. (1982). The medial forebrain bundle of the rat. II. An autoradiographic study of the topography of the major descending and ascending components. *Journal of Comparative Neurology*, 206(1), 82-108.
- Waraczynski, M. A. (1988). Basal forebrain knife cuts and medial forebrain bundle self-stimulation. *Brain Research*, 438(1), 8-22.
- Waraczynski, M., & Shizgal, P. (1995). Self-stimulation of the MFB following parabrachial lesions. *Physiology & Behavior*, 58(3), 559-566.
- Waraczynski, M., Carlton, E., & Perkins, M. (1998). Midbrain periaqueductal lesions do not degrade medial forebrain bundle stimulation reward. *Behavioural Brain Research*, 95(2), 167-177.
- Waraczynski, M., & Perkins, M. (1998). Lesions of pontomesencephalic cholinergic nuclei do not substantially disrupt the reward value of medial forebrain bundle stimulation. *Brain Research*, 800(1), 154-169.
- Waraczynski, M., Perkins, M., & Acheson, A. (1999). Lesions of midline midbrain structures leave medial forebrain bundle self-stimulation intact. *Behavioural Brain Research*, 103(2), 175-184.
- Waraczynski, M. (2003). Lidocaine inactivation demonstrates a stronger role for central versus medial extended amygdala in medial forebrain bundle self-stimulation. *Brain Research*, 962(1), 180-198.
- Waraczynski, M. A. (2006). The central extended amygdala network as a proposed circuit underlying reward valuation. *Neuroscience & Biobehavioral Reviews*, 30(4), 472-496.
- Waraczynski, M., & Demco, C. (2006). Lidocaine inactivation of the ventral pallidum affects responding for brain stimulation reward more than it affects the stimulation's reward value. *Behavioural Brain Research*, 173(2), 288-298.
- Waraczynski, M. (2008). GABA receptor agonism in the sublenticular central extended amygdala impairs medial forebrain bundle self-stimulation but GABA blockade does not enhance it. *Behavioural Brain Research*, 187(2), 396-404.

- Waraczynski, M., Salemme, J., & Farral, B. (2010). Brain stimulation reward is affected by D2 dopamine receptor manipulations in the extended amygdala but not the nucleus accumbens. *Behavioural Brain Research*, 208(2), 626-635.
- Waraczynski, M., Zwifelhofer, W., & Kuehn, L. (2012). Brain stimulation reward is altered by affecting dopamine–glutamate interactions in the central extended amygdala. *Neuroscience*, 224, 1-14.
- Wise, R. A. (1978). Neuroleptic attenuation of intracranial self-stimulation: reward or performance deficits? *Life Sciences*, 22(7), 535-542.
- Wise, R. A. (1980). Action of drugs of abuse on brain reward systems. *Pharmacology Biochemistry and Behavior*, 13, 213-223.
- Wise, R. A. (1996). Neurobiology of addiction. *Current opinion in Neurobiology*, 6(2), 243-251.
- Wise, R. A. (1998). Drug-activation of brain reward pathways. *Drug and Alcohol Dependence*, 51(1), 13-22.
- Witten, I. B., Steinberg, E. E., Lee, S. Y., Davidson, T. J., Zalocusky, K. A., Brodsky, M., Yizhar, O., Gong, S., Ramakrishnan, C., Stuber, G.D., Tye, K.M., Janak, P.H., & Deisseroth, K. (2011). Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron*, 72(5), 721-733.
- Woolf, N. J. (1991). Cholinergic systems in mammalian brain and spinal cord. *Progress in Neurobiology*, 37(6), 475-524.
- Woolf, N. J., & Butcher, L. L. (1986). Cholinergic systems in the rat brain: III. Projections from the pontomesencephalic tegmentum to the thalamus, tectum, basal ganglia, and basal forebrain. *Brain Research Bulletin*, 16(5), 603-637.
- Yeomans, J. S. (1975). Quantitative measurement of neural post-stimulation excitability with behavioral methods. *Physiology & Behavior*, 15(5), 593-602.
- Yeomans, J. S. (1979). The absolute refractory periods of self-stimulation neurons. *Physiology & Behavior*, 22(5), 911-919.
- Yeomans, J. S., Kofman, O., & McFarlane, V. (1985). Cholinergic involvement in lateral hypothalamic rewarding brain stimulation. *Brain Research*, 329(1), 19-26.
- Yeomans, J. S., Mathur, A., & Tampakeras, M. (1993). Rewarding brain stimulation: Role of tegmental cholinergic neurons that activate dopamine neurons. *Behavioral Neuroscience*, 107(6), 1077.

Yeomans, J. S., Takeuchi, J., Baptista, M., Flynn, D. D., Lepik, K., Nobrega, J., ... & Ralph, M. R. (2000). Brain-stimulation reward thresholds raised by an antisense oligonucleotide for the M5 muscarinic receptor infused near dopamine cells. *The Journal of Neuroscience*, *20*(23), 8861-8867.

Yeomans, J., Forster, G., & Blaha, C. (2001). M5 muscarinic receptors are needed for slow activation of dopamine neurons and for rewarding brain stimulation. *Life Sciences*, *68*(22), 2449-2456.

Yizhar, O., Fenno, L. E., Davidson, T. J., Mogri, M., & Deisseroth, K. (2011). Optogenetics in neural systems. *Neuron*, *71*(1), 9-34.

You, Z. B., Chen, Y. Q., & Wise, R. A. (2001). Dopamine and glutamate release in the nucleus accumbens and ventral tegmental area of rat following lateral hypothalamic self-stimulation. *Neuroscience*, *107*(4), 629-639.