Further Validation of a Model for Inferring the Value of Rewarding Brain Stimulation

Yannick Breton

A Thesis in

The Department of Psychology

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Arts (Psychology) at

Concordia University

Montreal, Quebec, Canada

August 2006

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-20686-7
Our file Notre référence
ISBN: 978-0-494-20686-7

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

Further Validation of a Model for Inferring the Value of Rewarding Brain Stimulation

Yannick Breton

Animals will work vigorously for electrical stimulation delivered to the medial forebrain bundle. The activity thus elicited is believed to mimic the effects of naturally rewarding stimuli, such as food and sex. This thesis tests a model of performance for this rewarding electrical stimulation. The proportion of time allocated to self-stimulation activities is evaluated with respect to the subjective intensity and opportunity cost of rewards. The 3D surface that comes from assessing self-stimulation performance based on the pulse frequency and stimulation price of a rewarding train of brain stimulation can then be used to determine the stage at which a given manipulation contributes its effects to reward processing. The pulse frequency determines the intensity of the rewarding effect of the stimulation; the amount of time an animal must invest in harvesting each reward controls the price. Previous attempts to validate a model of performance for rewarding brain stimulation relied on the assumption that an animal's preference for self-stimulation activities is revealed by the testing situation. The present thesis provides evidence that rat preferences are constructed by the testing situation much like human beings' preferences. Presenting animals with randomized test trials of varying price-frequency pairs provides a solution that minimizes the systematic biases to which the testing situation contributes. This improved methodology was used to validate that the model can correctly detect the effect of manipulations that act prior to the output of the spatio-temporal integrator that summates the activity elicited in primary reward neurons by the electrical stimulation.

AKNOWLEDGEMENTS

No man is an island, and rarely can a scientific endeavour be sought out and brought to completion by a single person, alone. I would therefore like to acknowledge the positive role of all those who contributed in no trivial way to the present thesis. First, I would like to thank my academic supervisor, Dr. Peter Shizgal, for being an essential source of inspiration and wisdom. I can truly state—without hyperbole—that I can only hope to one day have an iota of this gentleman and scholar's insight and knowledge. I would also like to thank Dr. Kent Conover for his help and problem-solving know-how. I have learned most of my technical, statistical and programming expertise from this brilliant man. Kudos go to Steve Cabilio, network manager, and Dave Munro, the centre's technician, for their patience and hard work in designing the software and hardware necessary to run these experiments.

Intelligent people become exponentially smarter when surrounded by other brilliant minds. It goes without saying that I thank everyone at the CSBN for providing a most stimulating and passion-instilling environment. In particular, I would like to thank all members of the Shizgal lab, past and present, for their continued help and support: Renu Chitra, Stephanie Fulton, Selma Hamdani, Nafissa Ismail, Ada Mullett, Jennifer Smith, Catherine Cowan, Giovanni Hernandez, Rebecca Solomon, Jamie Marcus (who did a commendable amount of work for the first experiment) and others that I've surely forgotten to mention here. By the same token, I would like to thank Dr. Andreas Arvanitogiannis and Dr. Uri Shalev for providing feedback and comments on this work. It is this kind of scientific discussion and scholarly debate that make the pursuit of knowledge so worthwhile. I should also acknowledge my debt to our furry quadruped

subjects: I suspect I have learned substantially more from them about running the experiments than the rats have learned from me throughout the experiments themselves.

Finally, I thank the most important people of all, the ones who have struggled to follow me as I painstakingly explained my craft and the object of my devotion of the past two years. I thank my mother, Miguelina, and my father, Jean-Daniel, for having been a constant source of encouragement and a boundless source of generosity and support. I thank my brothers, Sacha and Michael, for their continued interest and relentless curiosity. I am certain they are all as proud of this thesis as I am of having written it.

TABLE OF CONTENTS

Experiment 1

Experiment 2

LIST OF FIGURES

LIST OF TABLES AND EQUATIONS

Further Validation of a Model for Inferring the Value of Rewarding Brain Stimulation

Brain Stimulation as a Reward

A rat sits in an operant box, grooming itself. Its paws graze over the small mound of acrylic on his head, easily incorporating the foreign object as part of its own body. The location's odours and appearance are familiar; it has been here before. Suddenly, the rat's head perks up. It darts to the lever that has emerged from the wall. The rat holds the bar down impatiently, periodically tapping it. Finally, its work pays off: the animal is rewarded. This rat is neither hungry, nor thirsty, nor socially isolated, nor drug-addicted. To an ignorant onlooker, this animal wastes its time bar pressing in exchange for nothing. In effect, its reward is relatively intangible. This rat is harvesting trains of electrical pulses delivered to his brain: a brain stimulation reward. The phenomenon of brain stimulation reward is fascinating. The animal (whose biological needs have seemingly all been met) will work endlessly to administer electrical stimulation of a subset of the thousands of axons coursing through the medial forebrain bundle.

Brain stimulation is a powerful enough reward to motivate animals to go to great lengths to obtain a single train of pulses. The nervous system did not evolve for scientists to probe it with electrical stimulators. The firing resulting from a successful harvest presumably produces the illusion of reward by mimicking neural activation induced by natural goal objects such as food, water and sex. For proper survival, animals must opt for alternatives that will provide it with the greatest subjective reward at the smallest subjective price. Selection among competing goals requires that animals gauge the subjective payoffs of the alternatives. An animal electing to perform some task must necessarily trade off time it could spend performing other tasks. Since appetitive motivation organizes and directs our behaviour in such a direct way, studying the circuitry that underlies rewards helps us understand how decisions are evaluated and the subsequent choices generated.

Brain stimulation rewards and natural rewards must be attributed a common currency at some point in their processing for animals to be capable of comparing the rewards arising from electrical brain stimulation and the rewards derived from food and water. An experiment conducted by Green & Rachlin (1991) partly confirms this hypothesis. Hungry and thirsty rats were given the opportunity to press for brain stimulation, water or food in a two-choice operant situation. Reinforcement on either of two levers was delivered on a variable ratio schedule. The total number of presses the rats made was constrained in a session; if an animal allocated all of its presses to a single alternative, there was no way for it to respond on the other. Such a procedure imposed a budgetary constraint on the animal that made it possible to evaluate the economic substitutability of brain stimulation with water and food rewards. Highly substitutable rewards result in a high sensitivity to relative prices (the average numbers of lever presses required to earn a reward): an individual will adjust allocation of responses on the two levers according to the ratio of the two prices. At the limit, if an individual is responding for perfect substitutes—for instance, different brands of cola—then an increase in the price of one relative to the other will result in the complete abandonment of its pursuit. In the case of very non-substitutable (complementary) goods, for example right and left shoes, changes in the relative prices do not alter the quantities of the two goods that are sought. Green and Rachlin (1991) attempted to establish the uniqueness of the reinforcing properties of brain stimulation by testing its economic substitutability with natural rewards. If electrical brain stimulation was substitutable for natural goal objects that are rather unsubstitutable for each other, it would follow that the brain stimulation partly mimics their common, rewarding properties. If it was high substitutable for only a specific natural reward, the brain stimulation would highly mimic its unique properties. Poor substitutability between brain stimulation and either food or water would indicate a functionally independent action from natural rewards. Substitutability estimates can range from large, negative values in the case of complements, to 1 for perfect substitutes. The results were clear: whereas substitutability estimates between food and water were close to 0 or negative (indicating that these are complements), the substitutability between brain stimulation and either food or water was above 0.6 in all cases. The cross-modal substitutability was in the same range or greater than that of the brain stimulation for itself. This compellingly suggests that the stimulation mimics rewarding properties common to both food and water.

In addition, Conover, Woodside & Shizgal (1992) have described rat operant responding to brain stimulation and natural (saline) rewards in a forced-choice setting. Sodium-depleted subjects chose the larger reward when one was pitted against the other: they chose brain stimulation when the pulse frequency was high and saline when it was weak. Moreover, saline could summate with brain stimulation, that is, the effect of both together was equivalent to the sum of each. Similar results were found by Conover & Shizgal (1994) with sucrose rewards. It therefore appears that some aspects of the rewarding effects of brain stimulation are shared with a variety of natural reinforcers (water, food pellets, saline and sucrose) and evaluated on some common dimension. Taken together, the results from Green and Rachlin as well as Conover et al. suggest that brain stimulation reward and natural reinforcers are compared at some point downstream from where the rewarding effect of stimulation begins.

Electrical stimulation is not the only artificial means of high-jacking the reward system. Drugs of abuse are rewarding to users because of the activity they induce in reward-relevant parts of the brain. In addition to addiction, many impulse-control syndromes such as pathological gambling, morbid obesity, obsessive compulsive disorder and eating disorders may reflect an underlying malfunction in the circuitry designed for goal-selection and survival. Research on the biological basis brain stimulation reward may well provide insight into the aetiology, mechanism of action and treatment of pathological impulse-control.

When a rat harvests a brain stimulation reward, pulses cause a volley of action potentials to propagate from the tip of an electrode often implanted in the medial forebrain bundle. Activity reaches synaptic terminals over an unknown number of relays until the signal that carries reward-relevant information is spatio-temporally integrated. The result of the integration of reward signals over space in a given time window is a stored record of reward. It is this reward signal that forms the basis of a behavioural decision, rather than the first-stage neurons directly activated by the electrode. A decision is made by combining the strength of rewards with their price: the final choice is presumably that which provides the greatest benefit/cost ratio.

The parameters of the stimulation—the current that specifies the radius of excitation, the frequency of pulses that specifies the firing rate of the excited substrate, the price of the reward that specifies the rate of reinforcement—are all experimentally controlled. The behavioural output can be directly observed. However, the intermediate steps that translate the firing rate of first-stage neurons into a subjective reward signal, and the translation of that subjective reward signal into a behavioural decision are completely hidden from the experimenter.

The subjective reward value of a train of electrical stimulation pulses is related to the number of pulses delivered in a time window by the reward-growth function. The intensity of a brain stimulation reward increases rapidly with firing rate (as a power function) and levels off to asymptote at sufficiently high stimulation frequencies (Gallistel & Leon, 1991, Simmons & Gallistel, 1994). When given the option to respond on one of two levers for brain stimulation reward, rats were insensitive to differences among frequencies much higher than those required to produce asymptotic responding in single-operant situations. Furthermore, the effectiveness of subsequent increases in train duration to give rise to a subjective reward intensity decreases (in a stimulation strength-dependent manner) largely within first two seconds. These temporal characteristics make it unlikely that the first-stage neuron firing elicited by the electrical stimulation directly give rise to a subjective reward signal. If the primary reward signal were contained only in those neurons that are directly activated by the injected electrical pulses, the intensity would continue to grow for as long as the neurons still fire. That increases in duration fail to be effective in increasing reward intensity in a stimulation strength-dependent manner suggests that the directly activated substrate does not give rise to the reward intensity of the stimulation. Instead, activity is summated across these neurons over a short interval at some site downstream from the stimulation. It is the action of this process—the integrator—that outputs a subjective reward intensity.

An animal will base its decisions on the combination of the subjective intensity of the rewards it will earn from each alternative with the cost of those rewards. This cost comprises both an effort cost, the exertion required to earn each reward, and an opportunity cost, the price of foregoing other alternatives while pursuing the reward. In the case of a rat pressing for electrical stimulation, it will decide to allocate its behaviour between bar pressing and all other activities it could engage in while inside the operant box. The behavioural allocation function, then, models the relationship between the payoff from brain stimulation and the animal's actual allocation to activities that will procure that brain stimulation reward. If the reward it receives in exchange for work (bar-pressing) is subjectively weak enough or expensive enough, the animal will opt to engage in grooming, sleeping and exploration. If the reward is subjectively intense or easily procured, the animal will prefer spending its time bar pressing to any other activity it can perform in the operant box.

There exists, therefore, a black box within the self-stimulating rat's mind that we cannot directly observe. Many of the parameters that determine the subjective reward intensity of brain stimulation can be independently manipulated, and the animal's behavioural allocation to obtaining brain stimulation can be directly observed. In between the manipulable inputs and the observable output is a black box: those processes responsible for translating the injected impulse flow into a stored record of reward, translating opportunity and effort costs into a subjective price, combining subjective reward intensities and prices into payoffs, as well as generating behaviour from this subjective payoff. The transformations occurring within this black box have been hitherto immeasurable in a practical sense and have been equally difficult to infer without the assumptions of a strong version of the matching law.

Two-dimensional representations are misleading

Changes to the reward circuitry are commonly inferred from shifts in two-dimensional psychometric functions that assess an animal's behaviour in response to either the frequency of brain stimulation or its price (the time and effort required to earn a reward). Two 2D functions are in common use: rate-frequency measures and progressive-ratio measures. Rate-frequency measures assess how many times an animal responds for rewarding brain stimulation of a given intensity. At non-saturating frequencies, the more stimulation pulses are delivered in a time window, the greater the subjective reward value of that stimulation (as given by their relationship in the reward-growth function) and therefore, the more an animal presses for the reward. On successive trials, the frequency of the stimulation—the number of electrical pulses injected per time window—is decreased and the animal's rate of responding is measured. As the frequency decreases logarithmically, the animal's response rate decreases sigmoidally. Progressive-ratio measures assess how many rewards an animal harvests for rewards of a given price. The greater the number of presses the animal must make to earn a reward, the less an animal is willing to expend effort to obtain a reward. On successive trials, the number of presses required to obtain a reward is increased and the number of harvested rewards is measured. As the ratio requirement increases logarithmically, the number of rewards harvested decreases sigmoidally.

The two above behavioural assessments are used to infer the effect of manipulations to the reward pathway. However, fundamentally different changes in the circuitry underlying decision making can produce indistinguishable changes in either rate-frequency or progressive-ratio measures. If a manipulation is made to the reward circuitry that acts before the signal is spatio-temporally integrated, we observe a shift in the position of the rate-frequency curves collected. For instance, if some pharmacological agent changes the impact of each stimulation pulse arriving at the integrator, thereby increasing the subjective intensity of that train of pulses, fewer pulses are required to obtain a similar subjective reward value. If this occurs, one would expect to see a leftward shift in rate-frequency measures: the animal harvests a given number of rewards when fewer pulses are injected into the system. On the other hand, if some pharmacological agent facilitates an animal's willingness to work for a given subjective reward, each pulse contributes to the subjective reward signal in the same way, but the animal will invest more time into collecting those rewards. As a result, the thinking goes, we observe a rightward shift in progressive-ratio measures: the animal maintains a high level of responding even when it must expend a large amount of time and energy responding for a single reward. However, an increased willingness to work can produce a leftward shift in a rate-frequency curve, and increased reward intensity can produce a rightward shift in a progressive-ratio curve. Therefore we can not infer the subjective reward intensity of brain stimulation from either of these measures, and therein lays their fundamental flaw (Mullett, 2005; Breton, 2006). The following section illustrates this point more fully.

Manipulations to components of the reward circuitry do not transparently manifest themselves by traditional behavioural measures. When a particular structure is lesioned or its activity amplified pharmacologically, it is often useful to assess its effects on self-stimulation behaviour. The following two paragraphs elaborate on the conclusions that have come from manipulation of two structures of interest: the ventral tegmental dopaminergic neurons and the central extended amygdala. When the activity of ventral tegmental dopaminergic neurons or central extended amygdala neurons is manipulated, it is not possible on the basis of two-dimensional representations alone to determine whether the change has altered the intensity of the rewarding effect of a given reinforcer or has affected the animal's willingness to respond for the intensity of the reward it is to receive.

It is widely known that cocaine makes animals press more vigorously for less intense stimulation, shifting rate-frequency curves leftwards (Bauco & Wise, 1997). Conversely, Salamone, Wisniecki, Carlson & Correa (2001) have reported that animals working on various fixed-ratio schedules of reinforcement respond fewer times at a high fixed ratio when dopamine is depleted in the nucleus accumbens by 6-hydroxydopamine lesions. However, dopamine depletion in the nucleus accumbens does not appear to affect the response rate for a low (1 press per reward) response requirement. On the basis of this and other work (2001), Salamone suggests that these animals are less willing to work for a given reward because dopamine modulates the effortful mobilization of resources toward producing motivated behaviour. The conclusions of Bauco & Wise (1997) and Salamone et al. (2001) rely mainly on interpretations of two-dimensional representations considering operant responding as a function of only stimulation strength (in the case of Bauco & Wise) or price (in the case of Salamone et al.).

A structure of interest to neurobiologists studying the reward circuitry is the central extended amygdala. Temporary anaesthesia of the central amygdala by lidocaine infusion produces a rightward shift of rate-frequency curves (Waraczynski, 2003), indicating that animals will work less for stimulation of a given intensity if these amygdala neurons or fibres of passage do not function properly. The interpretation resulting from this two-dimensional analysis is that central amygdala neurons are necessary for the neural computation of reward intensity.

The traditional methods (rate-frequency and progressive-ratio measures) often claim to dissociate components of the circuitry that compute reward intensity from those that combine it with other decision-making variables (Franklin, 1978; Clark & Bernstein, 2006). They do not. As explained above and illustrated more comprehensively below, making an inference about changes to subjective reward value based only on the frequency of the electrical pulses delivered or the stimulation price alone is misleading. Manipulating both the strength of the stimulation and its price while assessing behaviour in three dimensions can disambiguate shifts that reflect pre-integration processing from post-integration processing. In fact, consideration of only reinforcer strength or willingness to pay has led to many controversies in elucidating the role of various neural substrates in addiction and decision-making processes.

A three-dimensional model: the Mountain

A more sophisticated view of decision-making considers both the strength of the stimulation and its price. When an animal responds for brain stimulation, it must constantly decide between bar-pressing to obtain a reward and reaping the rewards that come from doing anything else it could do in the operant chamber. A model of choice developed by Shizgal (1999, 2004) takes into account multiple components of decision-making including how rewarding brain stimulation is with how costly it is. The scalar combination of subjective reward value and price yields the utility of brain stimulation reward. The proportion of time an animal invests in harvesting brain stimulation is then related to the utility of brain stimulation reward (U_{bsr}) and the utility of everything else the animal could be doing in the operant chamber (U_{ee}). The equation, shown below (equation 1), is a simple modification of Herrnstein's single-operant matching law (1970,

$$TA = \frac{U^{a}{}_{bsr}}{U^{a}{}_{bsr} + U^{a}{}_{ee}} \tag{1}$$

1974) that takes into account how well brain stimulation reward substitutes for all the other activities that can be performed in the box (a). A plot of the animal's allocation to bar-pressing resulting from stimulation of a given frequency and price in semi-logarithmic space yields a mountain-like structure (see figure 1A) whose shape recaptures the shape of the reward-growth function.[1]

The utility of brain stimulation reward depends on the subjective reward value of the train of pulses (*I*, as determined by the reward growth function) and the subjective price the animal must pay to obtain the reward (P_s). The rat's responding is presumably the result of some comparison between the utility of self stimulation and the utility of all other possible activities (everything else). There is a special case that allows us to express the utility of everything else in terms of variables that we can control and infer, the subjective intensity and price of the reward. When the animal spends as much time at the bar as it does doing everything else, the utilities of either set of activities are inferred to be equal. If this situation occurs when the stimulation is maximally intense, then it must be because the stimulation is so costly for the animal to procure that it equally prefers

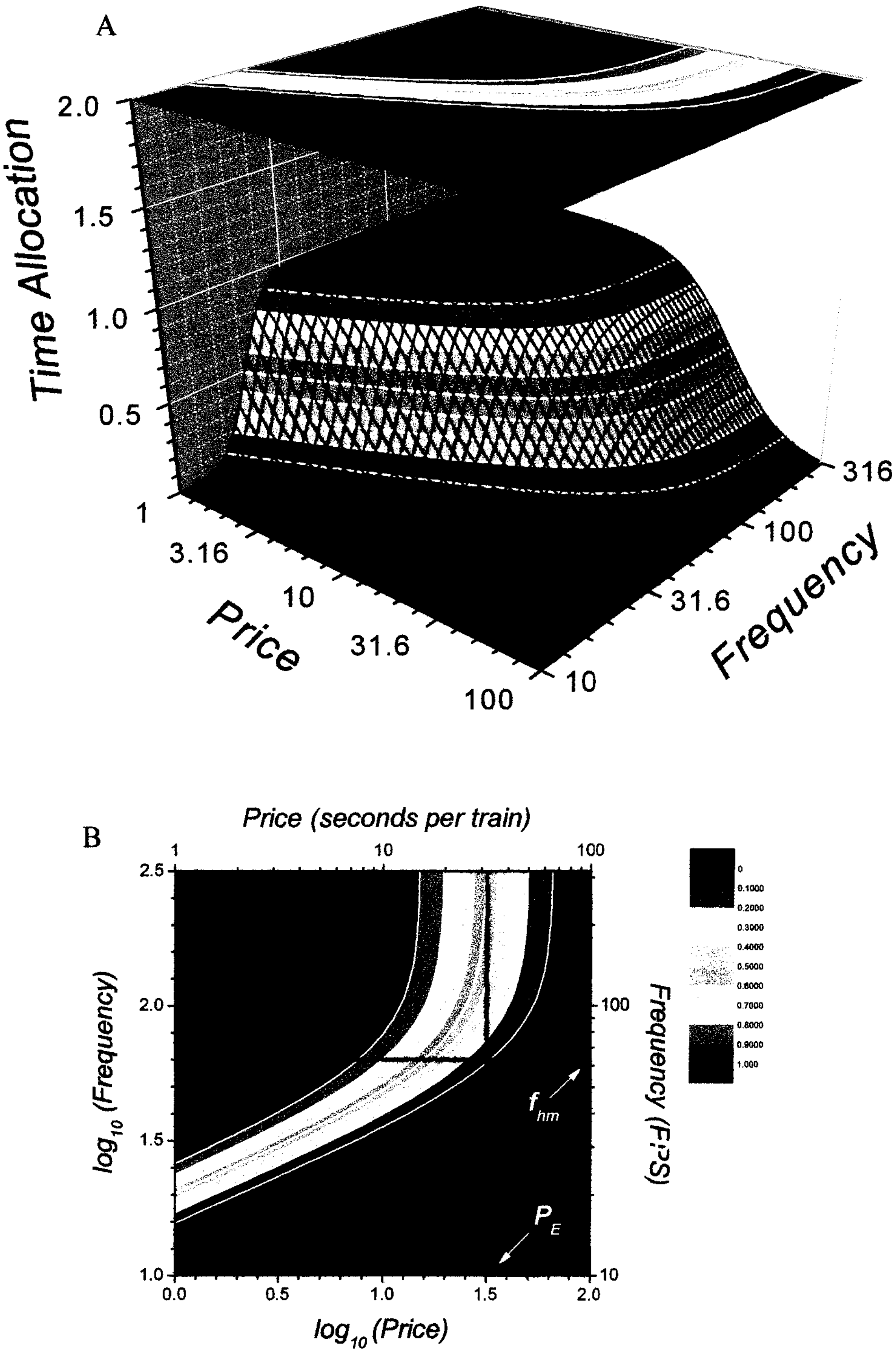


Figure 1. Structure representing the relationship between pulse frequency, price, and performance for rewarding brain stimulation. The top view depicts a contour graph of this 3D structure by representing successive heights (deciles of time allocation) with different colours delineated by white contours. Panel B is a contour plot equivalent to the top view of panel A, with the values of fhm and Pe plotted as black lines. (Taken from Shizgal, 2004)

working for the stimulation and doing anything else it can do. As a result, the utility of everything else is equal to the maximum possible reward (I_{max}) divided by the price at which a maximal reward produces 50% time allocation to bar pressing (P_e), scaled by the exertion required to earn each reward (ξ). The animal's time allocation to activities that will procure it brain stimulation reward therefore depends on 6 parameters, since the required exertion factors out because it appears in both the numerator and denominator of the P_s/P_e ratio. Two parameters specify the location of the three-dimensional function in space (f_{hm} and P_e), two parameters specify the rate of reward growth with stimulation and the increase in behavioural allocation with subjective reward (g and a) and two parameters specify the ceiling and floor of the structure (TAX and TIN). In panel B of figure 1, dark black lines indicate the values of f_{hm} and P_e on the contour plot. This contour plot is identical to the one plotted above the 3D structure in panel A.

It is possible, then, to disambiguate the pre- and post-integration effects of manipulations based on changes in the location of this three-dimensional structure (the mountain). Only those manipulations that alter the circuitry before the output of the spatio-temporal integrator will change the location of the mountain along the frequency axis, producing a change in the f_{hm} parameter. Manipulations that alter the circuitry after spatio-temporal integration will move the mountain along the price axis, producing a change in the P_e parameter.

The dissociation is possible because events prior to the output of the integrator alter the trade-off between stimulation current and pulse frequency. As a result, interruption or amplification of those events will change only how stimulation frequency is translated into subjective reward intensity and therefore only change the value of f_{hm}.

Events beyond the output of the integrator will not be able to take into account the electrically-evoked impulse flow to the integrator, and thus can only take into account the subjective reward intensity that results from the spatio-temporal integration process. Rearranging terms in equation 2, one can see that P_e—which sets the position of the

$$U_{ee} = \frac{I_{\max}}{P_e \times \xi} \quad (2)$$

mountain along the price axis—depends on the utility of everything else, the effort cost of self-stimulation activities, and the maximal intensity of brain stimulation reward. None of these three terms depend directly on the injected impulse flow, but rather, on processes occurring beyond the output of the integrator. As a result, alterations to events occurring beyond the output of the integrator will change only the translation of subjective reward intensity into behaviour and thus only change the value of P_e.

Why the two-dimensional psychometric functions are insufficient

It is clear, when considering this comprehensive model of choice in its three-dimensional space, why shifts in the traditional two-dimensional psychometric functions are misleading. Say a manipulation has shifted rate-frequency measures rightward resulting in an inference that the manipulation has made each stimulation pulse less rewarding. A silhouette of the mountain along the frequency axis taken at a very low, negligible price is analogous to this kind of measure. Although a shift in f_{hm} could account for a shift in this two-dimensional representation, so could a shift in P_e, because of the curved nature of the structure at low prices. Similarly, a silhouette of the mountain along the price axis taken at a high price, analogous to progressive-ratio measures, can misleadingly lead someone to believe that P_e has changed, but an identical shift can be produced by changing f_{hm}. In both cases, neither measure truly captures the change in the animal's decision policy, so neither measure can, alone, disambiguate shifts in f_{hm} (related to pre-integration effects) from those in P_e (related to post-integration effects) (Mullett, Conover & Shizgal, 2004).

Construction of Preference in Rat Responding

The methods by which brain stimulation reward experiments are conducted all fundamentally borrow from the psychophysical tradition of the method of limits. Early auditory psychophysicists (Titchener, 1905, cited in Kling & Riggs, 1971) would measure detection thresholds of sound pitch intensity by decreasing and increasing the pitch of their stimulus in small logarithmic steps until the subject reliably stopped or started reporting hearing the tone. Similarly, BSR experiments measure the number of presses an animal makes (in the case of this thesis, the proportion of time spent holding down the active lever) in response to stimulation of incrementally decreasing reward intensity (Miliaressis, Rompré, Laviolette, Philippe & Coulombe, 1986). An auditory psychophysics manipulation, such as changing the intensity of the tone presented, would shift the perception curve; a BSR manipulation, such as changing the intensity of the electrical current, would shift the response curve.

Since the curve-shift method began widespread use following criticisms of inferring changes to the reward value of BSR from simple changes in response rate, an important factor has never truly been addressed. Animals quickly learn the structure of the experiment: for a fixed number of trials, each delineated by a cue-signalled inter-trial interval, the stimulation frequency will decrease in logarithmic steps, and that sequence will repeat until the experimenter stops the experiment. For nearly 40 years, experimenters never expressed concern that this would cause a problem with data collection, because they tested only increasing or decreasing stimulation frequencies at a negligeable price.

The problem is not unique to the rate-frequency curve shift method. In progressive ratio schedules, the animal's task is to respond for increasingly expensive rewards. The sequence becomes predictable and can possibly be learned given enough experience in the operant chamber. Since only a very valuable reward is used in describing the relationship between responding and the rate of reinforcement, it is impossible to know whether the predictability of this sequence of trials has any impact on the animal's performance. Assessing behaviour on both the pulse frequency and price axes makes possible the detection of changes in an animal's time allocation to working for brain stimulation reward depending on whether a given trial is approached by incremental changes in pulse frequency or price. This thesis will provide evidence that an animal's behaviour is constructed by its expectancy for future payoffs. Most importantly, it will provide evidence that expectancy for decreasing stimulation frequency or increasing price will alter the proportion of time rats spend harvesting rewards. This thesis will also provide evidence for a solution that minimizes systematic biases that the rat may have that result from expectations about parameter values on future trials.

A Validation of the Model

Although the model can differentiate between pre- and post-integration manipulations in principle, it remains to be seen whether it is a valid model of choice in practice. In order to do so, it would be necessary to demonstrate whether each prediction of the model is supported. Most notably, producing changes that are known to affect only the translation of stimulation into subjective reward value, like shortening the duration of the train, would be expected to change f_{hm}. Producing changes that are known to affect only the translation of subjective reward value into behaviour, like delaying the reward by a few seconds, would be expected to change P_e.

It follows that for the mountain to be valid, we must be capable of manipulating each variable modeled and empirically demonstrate that their alteration corresponds to a predicted shift in the relevant location parameters. This thesis will provide evidence that a manipulation clearly acting prior to the output of the spatio-temporal integrator resulted in a shift of the mountain primarily along the frequency axis.

The case of altering train duration to shift the mountain along the frequency axis has already been documented (Arvanitogiannis, 1997; Mullett, 2004). These experiments established that the mountain model provides a relatively accurate general description of how reward strength and cost combine to produce performance for brain stimulation. However, some problems were left unsolved by these trailblazing studies of the reward circuitry that will be addressed in this thesis. There is good reason to believe the methods employed by the authors of both these validation experiments biased the animal's own choice behaviour in a systematic way. Systematic bias in the data collected by both Arvanitogiannis (1997) and Mullett (2005) would not only detract from the conclusiveness of their findings, but would also make the analysis of other manipulations altering P_e (such as leisure-contingent stimulation and delay) difficult.

In the first validation experiment performed by Arvanitogiannis (1997), a traditional variable-interval schedule of reinforcement was used to alter the experimenter-controlled rate of reinforcement, or inversely, the price of stimulation. This type of schedule is inadequate because the animal's harvest of brain stimulation rewards is not proportional to their investment in self-stimulation activities. Since, in this schedule, the lever remains armed at the end of the variable interval, animals could determine the time at which they could expect to be rewarded and only begin pressing at that point. As a result, rats could "cheat the system" by spending most of their time in leisure activities at the cost of foregoing only a few rewards. Since it is advantageous for the animal to wait for the bar to be armed before responding, the schedule does not provide a tight control of the opportunity cost of brain stimulation rewards. Furthermore, Arvanitogiannis (1997) employed a strict version of the single-operant matching law that assumes the asymptotic rate of responding throughout a trial is constant. This assumption has since been shown to be untenable for brain stimulation rewards (Conover, Fulton & Shizgal, 2001). In particular, it was shown that asymptotic responding slowed as the variable interval was lengthened, suggesting that employing single-operant matching under a variable interval schedule to scale the payoff derived from brain stimulation is problematic.

The variable interval schedule of reinforcement that was used is not analogous to many natural foraging situations because prey don't tend to wait for their predators to consume them. One can imagine that a bear can forage a patch of berries (leisure activities in the rat scenario) or fish for salmon in a nearby stream (bar pressing in the rat scenario). If salmon passed by the bear as a traditional variable-interval schedule, they would wait for the bear to be done picking berries until he fished them out. If the bear can reliably time the average arrival times of salmon, he is free to engage in berry-picking until he's almost sure there will be salmon waiting. Even if some of the salmon will have been waiting to be harvested for a while, the bear can still collect most of the fish while spending most of its time eating berries. In effect, the bear has cheated the system.

In a latter validation experiment by Mullett (2005), a free-running variable interval (first used in pilot experiments conducted by Arvanitogiannis) was used to prevent animals from cheating the system. In this schedule of reinforcement, if the animal isn't working at the time the interval elapses, a new interval is generated and the rat has missed that opportunity to collect a brain stimulation reward. This free-running variable interval schedule is analogous to the situation in which the above salmon swim down the stream without stopping to wait for the bear to return. If the bear isn't in the stream when the salmon come by, he misses the opportunity to eat them. Just like the bear making a choice between fishing and berry-picking, the rat's choice is exclusive between bar-pressing and leisure. Its brain stimulation harvest is directly proportional to the amount of time it invests in bar-pressing activities. An animal's opportunity cost can be defined as the subjective cost of foregoing the opportunity to pursue one alternative and its rewards in favour of pursuing another. The free-running variable interval schedule provides a tight control over the opportunity cost by allowing experimenters to control the average amount of work an animal must invest in self-stimulation in order for it to harvest brain stimulation rewards. While the animal is holding down the bar, it cannot engage in other rewarding activities such as grooming, exploring and sleeping. Bar-pressing for a given number of seconds presents a lost opportunity for the rat to engage in other activities. This is not true for traditional variable interval schedules, where the animal's opportunity cost is not controlled because the animal is free to engage in other activities before the required interval elapses. The animal need not forego many leisure-related rewards in favour of the rewarding brain stimulation derived from work.

However, changing the schedule of reinforcement to a zero-hold free-running variable interval to evade assumptions shown to be untenable for brain stimulation reward and to tightly control the opportunity cost of self-stimulation may not be sufficient to demonstrate the capacity of the Mountain to model performance. The method of descending limits borrowed from classical psychophysics may have altered bar-pressing behaviour by instilling expectancy for future rewards based on previous reinforcement history. Two problems were encountered by Mullett (2005) in analyzing the performance for rewarding brain stimulation. It appeared as though the locus of rise of pulse frequency-performance functions at high and low prices were fixed together. The slope of the curve obtained at a high price was also decreased, resulting in a rotation rather than a shift in some, but not all, cases. This is a problematic observation, because logically, the injection of more action potentials should be necessary to motivate an animal to work for stimulation requiring a greater opportunity cost. A second problem was encountered by Mullett (2005): the proportion of time allocated to working for high frequency brain stimulation rewards taken at a high price appeared to depend on whether that combination was approached by varying price from trial to trial or by varying the pulse frequency. It was therefore necessary to document the instances in which our method for eliciting choice changed bar-pressing behaviour. A proper understanding of those processes by which animal behaviour depends on the testing situation facilitates an understanding of how to minimize the influence of the testing situation on the psychometric functions obtained. As a result, we set out both to document the instances of construction of preference in animals responding for rewarding brain stimulation and to develop the solutions that minimized the systematic construction of preference in self-stimulating rats. With evidence confirming that animal behaviour was biased by the testing situation and support for methods that would reduce the influence of these biases, we attempted to replicate the findings of Arvanitogiannis (1997) and Mullett (2005), thereby providing further validation of the mountain model.

Objectives of the Thesis

The experiments presented in this thesis were designed to accomplish three goals. First, I assessed whether psychometric functions relating time allocation to the strength or cost of brain stimulation reward depended on the structure of the test sessions. Second, an improved data collection methodology was developed to minimize the biases that had been detected. Finally, the validity of a mathematical model of choice was empirically considered with regard to its capacity to disambiguate the effects of pre-integration manipulations from post-integration manipulations. Experiment 1 deals with the first two objectives of this thesis, whereas experiment 2 deals with the last. This thesis is also the first and only replication of a validation of the mountain model to use re-sampling, an advanced statistical method, to improve the accuracy and power of the curve-fitting process. Taken as a whole, the results of the experiments described here depict a validation of the mountain model using an improved methodology and statistical technique. The collection method could then be used in a plethora of future experiments aimed at further validating the model. A validated model could be subsequently used to investigate the contributions of different components of the reward system.

To accomplish these objectives, experiment 1 will provide evidence that the data collection methodology that is traditionally used (the sweep method) is fundamentally flawed; a means of controlling this systematic bias will be proposed. This experiment demonstrates that animals form a higher-level representation of the structure of the testing session that influences the time they allocate to holding down a bar for stimulation. Well-trained animals were subjected to frequency sweeps, in which pulse frequency decreased within sessions from trial to trial with the reinforcement density kept constant within sessions. Animals were also subjected to sweeps in which the variable interval was increased from trial to trial within a session (price sweeps). Finally, their behaviour was assessed when the trials of descending frequency sweeps and ascending price sweeps were presented at random. The differences in their allocation of time to self-stimulation activities among all types of sweeps and randomization conditions were compared in normatively equivalent trials. Experiment 2 will provide evidence of the validity of the model in its capacity to discern pre-integration effects using the methodology elaborated in experiment 1. Bar-pressing for stimulation bursts of 0.25 seconds was compared to responding for stimulation bursts of 1 second. Following these two experiments, a broad discussion will elaborate on the conclusions of experiments 1 and 2.

EXPERIMENT 1

Construction of preference in human beings

The curve-shift method assesses animal self-stimulation performance for a repeating series of decreasing stimulation frequencies. The series is presumably quickly learned: it is not uncommon for an animal in training to begin responding during the last low-frequency trials of such a series in anticipation of the high-frequency trials to come. The sequence is presented in the same order until the animal's rate of responding on trials in which the rat is presented a given pulse frequency is consistent with its rate of responding on other trials of the same pulse frequency.

Progressive-ratio schedules are no more immune to these considerations. Over the course of an operant-response session, the animal is presented with a predictable series of increasing work requirements.

Methods used by both Arvanitogiannis (1997) and Mullett (2005) to accurately assess an animal's decision to press for rewarding stimulation used variations on the above two paradigms. In both cases, logarithmically decreasing frequencies (frequency sweeps) were presented with a low response requirement, or highly rewarding stimulation frequencies were presented at increasing response requirements (price sweeps).

The idea that human preferences are constructed, not revealed, has long been prominent in the study of economics and choice, best documented by Kahneman and Tversky (1979). They cite a number of situations in which humans' preferences do not simply depend on the payoff they receive. Although a person's preference for a given option should be independent of its description (descriptive invariance) and the method used to elicit the preference (procedural invariance), these two factors appear to heavily influence the final decision. This assumption—a key component for the "rational-man" model in the social sciences —has often been shown to be invalid.

The phenomenon of descriptive and procedural dependency is made obvious by a study by Tversky and Thaler (1990). When presented with different gambles with equivalent expected values, individuals overwhelmingly chose the one in which the probability of winning was higher, despite the low financial gain produced by such a gamble. Nevertheless, their risk-aversion did not translate into their subjective pricing of those gambles. When they were subsequently asked the lowest price for which they would sell the gambles, individuals put a higher premium on the low-probability, high-remuneration gamble. On the basis of this data, Thaler and Tversky concluded that although a person's preference is influenced by the probability of winning (exhibiting risk aversion), their evaluation of price is dictated in part by the money they would receive from the option. Preferences are unduly influenced by probabilities whereas pricing is unduly influenced by the amount at play.

Ariely, Loewenstein and Prelec (2003) recently explored the notion that preferences are constructed by presenting MBA students with various valuable consumer goods worth on average 70$. Students were simply asked if they would purchase each product for a dollar amount equal to the last two digits of their social security number. Following the simple yes/no response, students were asked the maximum amount they were willing to pay for each good. To ensure honesty, students were obliged to purchase at least one item, with the price of that item being a combination of their answer on the initial accept/reject decision and their subsequent maximal willingness to pay. Since the last two digits of the social security number are random and meaningless, they shouldn't alter someone's willingness to pay for a product. The experimenters found that when subsequently asked the maximum they would pay, students with higher final digits were willing to pay a greater amount than those with lower digits. Individuals' assessment of the value of consumer products appeared to be anchored by an arbitrary number.

In a study by Simonsohn & Loewenstein (2006), it was found that individuals moving from a city where housing was expensive tended to overpay in a less expensive city, whereas those moving from a city where housing was inexpensive attempted to pay less in a more expensive city. Clearly, the individuals in the study did not base their willingness to pay on only the value of the house and its price; their expectancies played a large role in their decisions. That these individuals later re-adjusted their willingness to pay implies that those coming from expensive cities were not simply wealthier. Instead, their anchor for the scale of housing values began at some different initial condition.

The body of Tversky's work (aided by Nobel prize-winning colleague Kahneman) led to the now law-like conjecture that preferences are not merely revealed by the experimenter: individuals' choices are *constructed* by the elicitation process. This postulate permeates all areas interested in the study of human decision-making, from economics to psychology.

Construction of preference in non-human animals

The application of these ideas to operant responding is not usually entertained. Nonetheless, Fantino (Navarro & Fantino 2005; O'Daly, Case & Fantino, 2006) has recently investigated some instances in which pigeon performance for food and water rewards is not completely economically rational. Using concurrent fixed ratio schedules, Navarro & Fantino (2005) has reproduced the "sunk cost" effect in pigeons. Classical microeconomics assumes that an individual's choice will not depend on "sunk costs", or costs that have already been incurred. If an individual has bought a movie ticket, the price of the ticket is a sunk cost. If the individual later regrets buying the movie ticket—for example because it is revealed that the movie is not entertaining—a decision about whether to see the movie should not include considerations about the "sunk cost" of the ticket. An individual can watch a bad movie, or do something else that would be more enjoyable. The rational choice, according to a classical microeconomist, is to do something else, since the money has already been spent whether one goes to the movie or not and one will get more enjoyment from not watching the movie. Clearly, this is not what individuals tend to do. A situation such as this one, in which sunk costs factor into an individual's later decisions, is referred to as the "sunk cost effect." Navarro and Fantino (2005) established that, just like human beings, pigeons are susceptible to this sunk cost effect. When presented with high- and low-fixed ratio schedules that can be aborted, food harvest can be maximized by responding the number of times required for the low-ratio schedule and aborting the trial if no reinforcement is given. Pigeons that are given the opportunity to escape a high fixed ratio schedule will not do so if the change in reinforcement schedule is not highly salient. In addition to this sunk cost effect, Fantino (Goldshmidt & Fantino, 2004; O'Daly, Case & Fantino, 2006) has described pigeon responding under risk conditions, and has illustrated a few cases in which economic context influences animal behaviour (Goldshmidt & Fantino, 2004; Fantino & Stolarz-Fantino, 2005).

Inconsistencies in preference have sometimes been studied in the case of brain stimulation reward. Konkle et al. (2001) published findings that animals responded at very similar rates for frequency sweeps regardless of whether they were presented with stimulation frequencies of increasing or decreasing frequency. However, the step size used was 0.1 common logarithmic units, with a majority of trials in which the stimulation was sufficiently strong to produce asymptotic responding. Such a series corresponds to an approximately 25% reduction or augmentation of the stimulation pulse frequency from trial to trial, resulting in a rather insensitive measure of changes in threshold. This kind of step size would be able to detect, at best, a 25% reduction of the animal's threshold responding level. Some laboratories have reported statistically significant changes in threshold of approximately 10%. Indeed, Phillips & Lepiane (1986) found modest but statistically reliable differences between ascending, descending and randomized presentations of pulse frequencies. Clearly if there had been a modest effect of directionality, this experiment would not have been able to detect it. Furthermore, animals still knew the large-scale structure of the experiment: on any given trial, the animal could predict whether the next trial would offer him the opportunity to bar press for electrical stimulation of greater or lesser frequency.

Evidence from Gallistel, Mark, King & Latham (2001) has demonstrated that an animal is capable of adjusting its behaviour to the stationarity of the reinforcement density. In effect, their results suggest that animals behave as ideal detectors of change when the rate of reinforcement is changed over long or short periods of time. In particular, if animals are exposed to long periods of stationary rates of reinforcement, they adjust slowly to changes in that rate. If animals are exposed to periods in which the rate of reinforcement is changed very quickly, they adjust quickly to alterations in that rate. It seems clear from these results that animals must have some representation of the structure of the testing situation for them to behave as Bayesian statisticians.

It is therefore possible that the highly predictable structure of the Arvanitogiannis and Mullett validation experiments altered the animals' response patterns. Rats in these experiments are exposed for long periods of time to periods of relatively stationary conditions. A frequency sweep is first collected at a negligible price until a sufficient amount of reliably predictable performance is recorded. A price sweep is then collected at a highly rewarding stimulation pulse frequency until similarly reliable performance is recorded. A final frequency sweep is collected at a high price. In both the Arvanitogiannis (1997) and Mullett (2005) cases, the animal is afforded the luxury of highly stable, highly stationary conditions. It is altogether likely that these situations considerably bias the animal's decision to press for rewarding stimulation in a systematic manner. In conditions of long-term stability in price, it may be possible that the internal anchor for a rat's scale of rewards shifts to adapt to this long-term rate of reinforcement. As a result of this anchor shift, the animal's performance for rewarding brain stimulation is different according to the long-term average price. When testing is carried out at high prices for long periods of time, the anchor that serves as an internal reference point may therefore be different from when it is carried out at low prices.

It has been hitherto nearly impossible to study the biasing effect of the testing situation on self-stimulation performance because experiments on brain stimulation reward have rarely been set up in such as way as to make it possible to detect biases, such as shifting anchor points. Even in cases when both rate-frequency and rate-price curves were collected in succession (Arvanitogiannis, 1997; Mullett 2005), the only price-frequency pairs which overlapped were those where the frequency of stimulation was relatively high and whose price was relatively low. As a result, only the very high payoff trials overlapped. It is likely the ceiling effect hides reinforcement history-dependent changes in performance, since when the animal is presented with a very high payoff its performance is roughly asymptotic. The experiment elaborated below demonstrates that there is no discrepancy between these price-frequency pairs in the case when the payoff on a given trial is very high or very low. In contrast, inconsistencies emerge when the payoff on a given trial is some intermediate value. Thus, seldom has the biasing effect of the testing situation on self-stimulation performance been observable since the key price-frequency pairs have rarely been compared previously. Mullett (2005) pointed out inconsistencies similar to those described in this thesis; however, the author did not formally assess the observation statistically and did not provide a solution to reduce them.

The present experiment

Prior work by Breton (2004) has already demonstrated the possibility that the means by which traditional psychometric functions of self-stimulation behaviour are collected can bias the animal's consequent allocation of time to self-stimulation. It was observed that by collecting much more data than in the Arvanitogiannis (1997) and Mullett (2005) experiments, smooth 3D surface fits to price and frequency "sweeps" poorly captured the pattern of responding for brain stimulation. Another key observation was made by Breton (2004): high-priced frequency "sweeps" appeared to change over time, with animals allocating increasing proportions of time to stimulation for which they had initially refused to work. In effect, there was a leftward shift in the threshold for performance for rewarding brain stimulation when comparing the first session of frequency sweeps taken at a 16-second price and the last session of these sweeps. It follows from these observations that the standard with which animals evaluated the stimulation had changed as a result of long-term adaptation to new testing conditions. Some anchor for scaling rewards had drifted as a result of the method of elicitation of choice. Later pilot work by Marcus (2005) provided further evidence that rats' preferences with regard to brain stimulation rewards was constructed by the testing situation rather than simply revealed.

If the testing situation had no impact on a rat's bar-pressing behaviour, that is, if the principle of procedural invariance held for self-stimulation, normatively equivalent methods of eliciting bar-pressing behaviour resulting from different pulse frequencies and prices of stimulation would not alter time allocation. It should not matter whether the amount of time required in harvesting a reward changes over a long period of time, nor should it matter whether the animal is experiencing increasingly costly or decreasingly rewarding stimulation. Time allocation, if it is procedurally invariant, should not differ statistically from one condition to the next: all that should matter, in an economic sense, is the scalar combination of the intensity of the stimulation with the perceived cost of acquiring it.

The present experiment tests this assumption and demonstrates one method to prevent procedural dependence from introducing inconsistencies into the dataset. Animals were presented with frequency "sweeps" in which the price was held constant throughout a single testing session, but stimulation got increasingly costly over the course of multiple sessions. After two price increases, the stimulation got equivalently less costly over the course of multiple sessions. Preceding and following this procedure, animals were presented with price “sweeps” in which the frequency was held constant throughout the testing session. The price-frequency pairs thus tested were chosen so that many would overlap: prices tested in the frequency sweep domain were equal or close to three of those used in the price sweep domain.

If procedural invariance held, then animals would be expected to allocate equivalent amounts of time to harvesting trains of the same frequency and price, regardless of the long-term increasing or decreasing price trend and regardless of the variable being changed from trial to trial. However, if rats’ expectancies (based on their previous experience) biased their self-stimulation performance, performance would depend on price history, changes in independent variable, or both.

Furthermore, making the rats’ long-term evaluations of price stable and removing any signal about future trial prices would be expected to remove these biases entirely. Animals will not be biased by their trial history if that trial history is statistically random and stable. It follows that rats with context-dependent preferences of price-frequency pairs will not show that context dependence if prices and frequencies are presented randomly across trials. This improved consistency, if found, would also be expected to produce a better-fitting 3D surface to price and frequency data. One such smooth-fitting 3D surface is the Mountain model under validation in experiment 2. If the sweep method produces large inconsistencies in self-stimulation performance for frequency “sweeps” and price “sweeps”, the model would be less capable to account for the variance in the data by producing systematic deviations of the data points from the fitted surface. If randomizing the presentation of trials of different prices and frequencies is a solution to the inconsistency, data points would not deviate in a systematic fashion from the fitted surface.

The goal of this experiment is to reproduce and statistically test for the internal inconsistencies that have already been reported informally and documented incompletely, and elaborate a methodology to reduce inconsistencies. In other words, this experiment will establish that in the case of rats, an internal anchor for the scale of rewards appears to change according to different long-term average prices. The inconsistencies that result from this change in anchoring reference point will then be lessened to nearly irrelevant simply by presenting the animals with those same prices and frequencies in random order. If it is possible to lessen or eliminate the inconsistencies that traditional methods impart to performance, this finding would suggest that animals learn much more than response-reinforcement associations in an operant experiment. It also suggests that experiments in which the long-term price changes phase-by-phase are tainted by an assumption that is untenable. The experiment demonstrates that large discrepancies in performance during trials in which the price and frequency are the same but the long-term average price and the session structure differ can be made concordant again if trial presentations are randomized, making the price stable over a long period of time.

Method

Subjects

Four experimentally-naïve, male Long-Evans rats (Charles River, St-Constant, Quebec) were used. Each weighed approximately 450 grams at time of surgery. Animals were housed in pairs on a 12h-12h reversed light-dark phase cycle until surgical implantation of the stimulation electrode. Food and water was made available *ad libitum* except during experimental testing. Surgically implanted animals were only considered for the experiment if one of the bilaterally-implanted electrodes evoked approach behaviours following electrical priming stimulation, without any accompanied vocalization or adverse motor effects. Any animals experiencing seizures, showing obvious signs of distress such as vocalizing or leaping, or displaying disruptive motor reactions in response to a half-second train of 78 Hz (40 pulses with a 12.2 millisecond inter-pulse period) stimulation at 200 microamperes were deemed ineligible for the study. Current was subsequently titrated with pulse frequency so animals self-stimulated at the highest current they would tolerate that still yielded a wide range of pulse frequencies for which performance was asymptotic.

Materials & Apparatus

Electrodes

Electrodes were made by coating 00 gauge stainless steel insect pins with Formvar enamel to within 0.5mm of the tip. The insect pin was soldered to a flexible insulated wire, which was itself soldered and crimped to a male amphenol pin. The amphenol pin was inserted into a 9-socket cylindrical connector (Ginder Scientific, Ottawa, ON) at the time of surgery. Two electrodes were glued together with dental acrylic 3.6mm (medio-laterally) and 0.5mm (dorso-ventrally) apart. The entire electrode assembly was implated at least 24 hours following construction.

Operant Chambers

Operant chambers were made of opaque Plexiglas walls and measured 12.75 inches long, 8.75 inches wide and 26.00 inches high. A wide wire-mesh served as a floor. A retractable lever (MedAssociates, VT) was mounted 3.00 inches above the floor on left and right walls with a key light 1.25 inches above each lever. Trial contingencies and the exact structure of each delivered pulse train were controlled by a computer, which also registered the time of each transition in the state of the lever (custom programming by Steve Cabilio, hardware designed by David Munro, 1998).

A large house light located 5.75 inches from the floor on the back wall flashed for the 10 second duration of the inter-trial interval. The key light above the active lever was illuminated throughout the trial; when a reward was delivered, the lever retracted for 2 seconds and the key light was turned off. A single train of priming stimulation indicating the frequency of the stimulation delivered on the subsequent trial was delivered during the inter-trial interval 8 seconds after the onset of the interval and ending 1.5 seconds before the end of the interval.

The price of the stimulation was controlled by a free-running variable interval schedule of reinforcement (FVI) in which the animal was rewarded if it was holding the bar down at the end of a randomly-drawn latency interval. The intervals were drawn from an exponential distribution with mean and variance set to the experimenter-controlled price. If the price was set at 4 seconds, for instance, an animal would get rewarded on average every four seconds if it held the bar down continuously. If the bar was not depressed at the end of the interval, a new interval was drawn from the exponential distribution. This reinforcement contingency rewards the animal in proportion to the time it invests in self-stimulating, since the reward is not held for the animal at the end of the latency as in traditional variable interval schedules. As a result, the rat's average harvest is proportional to the time it allocates to self-stimulation.

Surgical Procedure

The above-described electrode assemblies were aimed at the lateral hypothalamic level of the medial forebrain bundle. Stereotaxic coordinates for the deeper of the two electrodes were 2.8mm posterior to bregma, 1.7mm lateral to the saggital sinus and 8.3mm ventral to the dura mater. Surgery was performed under sodium pentobarbital (Somnotol) anesthesia at a dose of 60 mg/Kg, administered intraperitoneally. Atropine sulfate (0.5mg/Kg) was administered 15 minutes prior to pentobarbital injection to reduce mucous bronchial secretions throughout the procedure. Return wires were fastened to the 2 most rostral of the 6 inserted jeweler's screws; the entire electrode assembly was anchored to the jeweler's screws by dental acrylic. A low dose (0.05mg/Kg) of buprenorphine was administered 15 minutes after surgery to minimize discomfort.

Behavioural Testing

Training

Following Skinnerian shaping of the lever-press behaviour, animals were trained on the experimental protocol over multiple daily sessions in which a 10-trial sequence was run repeatedly (24 times); the price of the stimulation throughout these sessions was 1 s. After two identical trials of high pulse frequency stimulation, the pulse frequency on the subsequent 8 trials was decreased by a logarithmic step size so as to produce approximately 3 trials of asymptotically high responding, 3 trials of descending time allocation, and 3 trials of asymptotically low responding. The first trial of such a series was considered a warm-up and the results of this trial were not included in any analysis. Different animals stepped through a different range of pulse frequencies at different intervals. The ten-trial sequence was repeated 24 times in a session, and sessions were repeated until animals responded similarly across determinations and sessions for each pulse frequency tested.

Trials were separated by a ten second inter-trial interval during which the house light flashed continuously and the computer delivered a 500 millisecond burst starting 8 seconds into the interval and ending 1.5 seconds before the end of the interval. The non-contingent priming stimulation delivered during this time was exactly the same as the stimulation the animal would receive during the subsequent trial. Throughout the trial, any successful harvest triggered a lever retraction, accompanied by a two second blackout delay during which the trial clock paused. The length of each trial, excluding the blackout delay periods, was set to 25 times the price, allowing the rat to earn, on average, a maximum of 25 rewards on any given trial.

Following the collection of the one second frequency sweep condition, animals were presented with price "sweeps", in which the price on every non-warm up trial was increased from 1 second in 0.185 logarithmic steps to 30.2 seconds. The pulse frequency presented during all trials was as high as the animal would tolerate. Such a procedure served two purposes: first, to expose rats to increasingly expensive rewards, and second, to determine the price at which they would perform at a half-maximal level.

Procedure

Following training, animals were again exposed to a frequency sweep carried out with the price set to 1 second. The price at which the frequency "sweeps" were collected was systematically increased and then decreased in 2 equal logarithmic steps, starting at 1 second and ending at the price which supported half-maximal performance, rounded to the nearest 0.1 logarithmic unit. Animals were therefore presented with a set of frequency "sweeps" taken at a low price, another set taken at a medium price (logarithmically half-way between the low and high prices), another set taken at a high price and a final set taken at the medium price. For 2 animals, another low-price frequency sweep was also collected. A minimum of 20 determinations of each frequency and each price sweep was considered for analysis, excluding the first determination of every session and the first session of a new condition.

The first session of a condition (training price, low-price frequency, medium-price frequency, high-price frequency, return to medium-price frequency, and return to low-price frequency sweeps) was excluded. The first determination of each condition was also excluded. These were deemed adjustment periods for the rat, and likely represent a comparison of the current session structure with the previous structure. One rat was excluded from analysis because of a consistent rightward shift that began to manifest itself late in the experimental protocol; this shift may be symptomatic of electrode displacement by skull growth or pathological changes in the brain.

At the end of the sweep portion of the experiment, a final price sweep was taken using the fixed cumulative handling time schedule (FCHT) to at least minimally train the animal for the schedule of reinforcement that would be used in the randomized portion of the experiment. In this case, the animal was rewarded after holding the bar for a fixed, cumulative amount of time. If the price was set at 4 seconds, for instance, an animal would get rewarded as soon as it had spent a total 4 seconds holding the bar. In such a case, an animal could hold for one second, stop, hold another 2 seconds, stop, and would be rewarded once the computer recorded another 1 second of depression. This reinforcement contingency removes reward-to-reward variability in the proportion of time required to hold down the bar in order to harvest a reward. Given the random nature of the following experimental session structure, it was necessary to provide the animal with as much information as possible in order to accurately evaluate the payoff it would receive on a particular trial.

In the randomized portion of the experiment, the same 36 price-frequency pairs that were tested as in the sweep portion of the experiment (including the price sweep) were placed in a list. This list was constructed so that the central three points of each sweep were represented twice as many times as the upper and lower extremes. This list was then randomized into a new list in which each price-frequency pair that would be tested was bracketed by a leading trial carried out using a high pulse frequency and a 1 second price and by a trailing trial carried out using a low pulse frequency and a 1 second price. These brackets allowed the experimenters to track any long-term changes in the effectiveness of the electrode or anomalous behaviour of the rat, in addition to possibly providing stable anchors for the rat's evaluative scale. A new list of randomized test points was generated every day using custom-programmed software in the MATLAB programming language.

The list of bracketed, randomized (or re-randomized) test points was presented to the rats daily until their behaviour became stable. Stability was indicated by steady, asymptotically-high time allocation on leading bracket trials and steady, asymptotically-low time allocation on trailing trials. Once animals were deemed to have learned the randomized nature of the task by these criteria, 12 passes through re-randomized lists were analyzed for consistency and compared to the sweep data. As in the sweep condition, data from the first pass through the randomized list and performance on leading and trailing bracket trials were not included in the data analysis.

Data Collection and Analysis

Every change in the state of the lever was recorded and time-stamped by computer at an accuracy of 0.1 seconds. Bar releases not exceeding one second were included in work time, since in those short intervals it is unlikely the animals engaged in activity other than self-stimulation. On the basis of casual observation, the rat typically is standing at the lever during these brief pauses and is often touching the lever lightly. Conover (personal communication) has performed analyses on the types of lever holds and releases on uncorrected data. Specifically, they have identified multiple components in the distribution of hold and release times. Their results suggest that work time is spent in bouts of bar-pressing behaviour consisting of bar holds and brief (less than one second) bar releases. The "tapping" correction increases the dynamic range of the collected data and compensates for the underestimation of animals' true allocation of time to self-stimulation activities.

Results

The following presentation of the results of this experiment will begin with a description of the analysis method for detecting internal inconsistencies in the data set. The curve-fitting technique will be elaborated on, as well as the logic of the statistical tests that will be used to assess performance, on different trials, for stimulation trains of the same strength and price. The analysis of whether or not performance depended on whether the price-frequency pair was approached from a frequency sweep (or pseudo-sweep) or a price sweep (or pseudo-sweep) will follow. Finally, the results section will consider the 3D surface fits to each animal's performance to further assess the relative prevalence and severity of internal inconsistency in data collected during sweep conditions and randomized conditions.

Fitting Strategy for Assessing Internal Consistency

A sigmoid-appearing curve (see equation A1 in appendix) consisting of a dual-quadratic spline function was fit to the data from the frequency sweep taken at the highest price and the initial price "sweep." This spline function consisted of a straight line at the lower asymptote joined to a positive quadratic, itself joined to a negative quadratic, joined in turn to a straight line (see figure 2). The spline function smoothly captures the s-shape form of the data: stable, asymptotic responding when the frequency or price is very high or low with a transition region between the two asymptotes. The advantage of this piecewise-defined curve over other sigmoid-shaped curves is that the parameters defining the function do not tend to be as correlated as those defining a continuous function, leading to a well-fitting function that does not require a long time for an iterative least-squares fitter to come to a local minimum (Conover, personal communication).

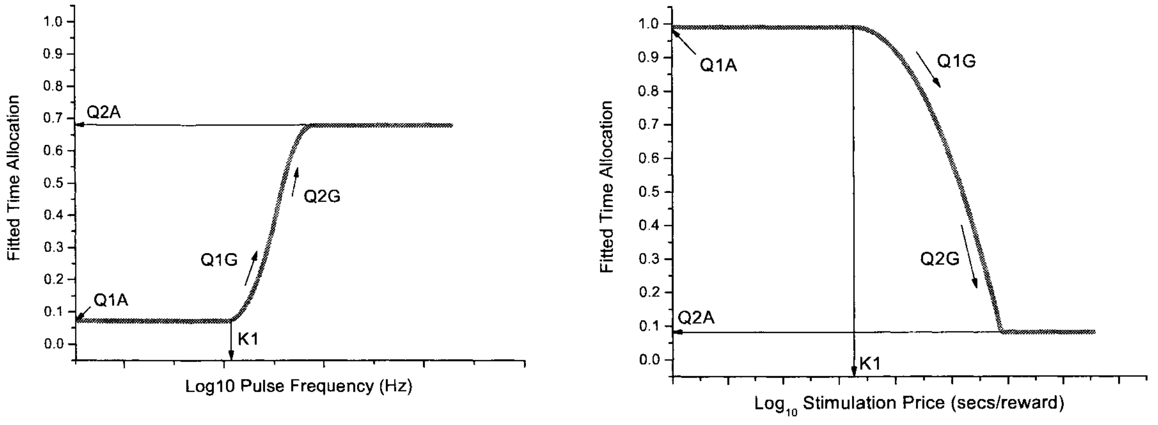


Figure 2. Example pseudo-sigmoids fit to frequency (left) and price (right) sweeps and pseudo-sweeps. The components of the pseudo-sigmoids are indicated by arrows on the two panels. Q1 and Q2 refer to the two component quadratics of the spline function. A refers to the asymptotic time allocation. G refers to the growth of the quadratic. K1 refers to the first point (knot) along the abscissa at which the asymptote of the spline function encounters a piece-wise change; it is the position along the abscissa at which the pseudo-sigmoid stops describing a straight asymptote and begins to describe a quadratic.

Confidence intervals around each fitted multi-segment sigmoid were then estimated by means of a bootstrapping procedure. The frequency and price values tested were chosen so as to cause the frequency and price sweeps to intersect. It was therefore possible to interpolate with 95% confidence whether performance for a train of a particular frequency and price differed depending on the type of sweep being presented. A similar procedure was conducted for data points collected during the randomized phase of the experiment. Figure 3 shows, for each rat, the points presented in random order during the randomized portion of the experiment. The circles represent the sampled price-frequency pairs; the lines represent the pseudo-sweeps in which the points were grouped. The independent variables (price and frequency) were re-ordered to produce a series of re-constructed pseudo-sweeps. In other words, the time allocations collected at each point along each line in figure 3 were grouped together for each rat as though group of points on a line had been collected as a sweep. Comparisons were therefore made between the fitted pseudo-sigmoid curves for frequency sweeps taken at low, medium and high prices and the projected curve fit for the price sweep at each point of intersection, in sweep and randomized experimental phases. At the point in the parameter space where the price and frequency sweeps intersect, overlap of confidence intervals around the observed time allocation signifies that the two estimates are consistent. In contrast, non-overlap of confidence intervals signifies that it can not be said with 95% confidence that the animal's behaviour is consistent. Thus, if they do not, one can presume that the animal's behaviour is statistically different in the frequency and price sweep cases. If consistency between price sweeps and frequency sweeps is restored by the randomization process, it

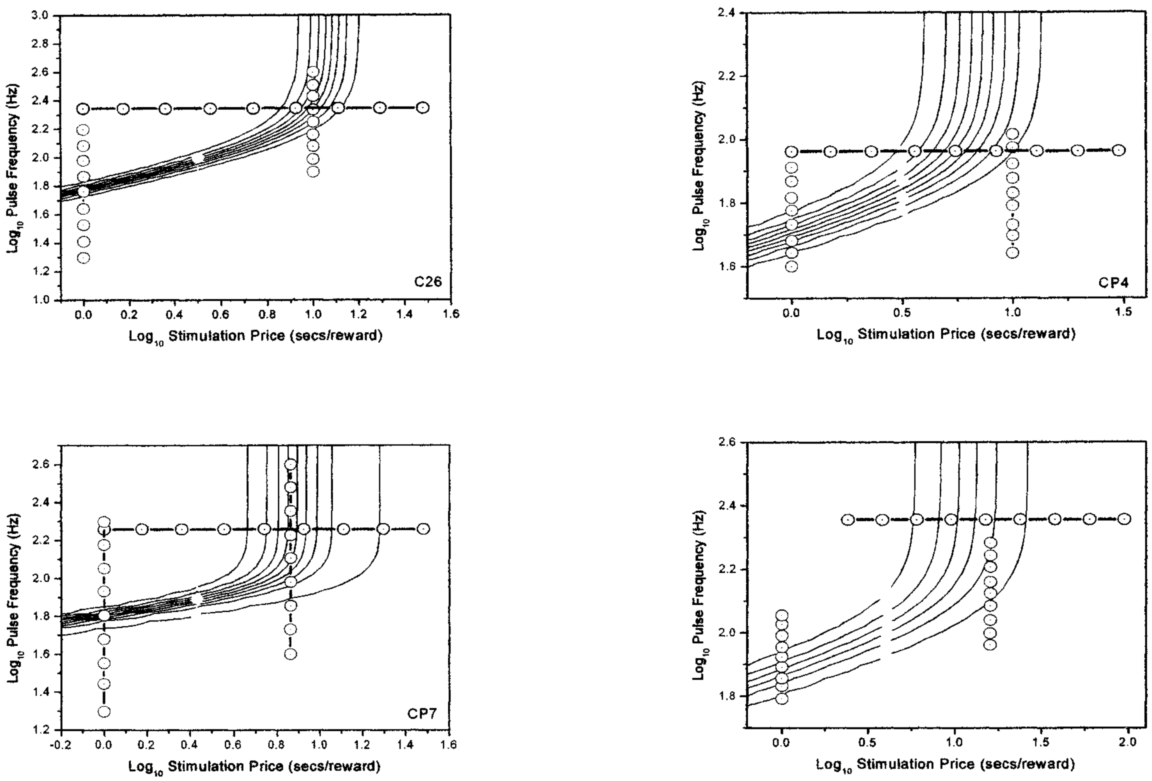


Figure 3. Position of collected frequency and price pseudo-sweeps in the randomized condition of the experiment. Price-frequency pairs belonging to a given pseudo-sweep are plotted in the same colour and connected by a line. Overlain with them is a contour plot of successive deciles of time allocation predicted by a surface fit to the data points.

follows that the structure of the sweep portion of the experiment biases the animal's behaviour.

The fitted multi-segment sigmoid functions comprised 5 separate parameters (see figure 2). Arranged in order of the position along the abscissa of the portion of the function they affect, the parameter names are: Q1A, K1, Q1G, Q2G, Q2A. The parameter Q1A (first asymptote) is the estimated time allocation along the first asymptote. Parameter K1 (first knot) is the position along the abscissa of the junction between the end-point of the first asymptotic segment and the start-point of the first quadratic. The value Q1G determines the growth rate of the first quadratic function whereas Q2G determines the growth of the second. Parameter Q2A (second asymptote) is the estimated time allocation along the second asymptote. The same functional form was fit to frequency sweeps and price sweeps; initial seeds for the asymptote values were exchanged because time allocation grows as the frequency is increased but decreases as the price is increased.

The model under validation in experiment 2 (the Mountain Model) was subsequently fit to the data collected in the sweep portion of the experiment and to the randomized portion of the experiment separately in order to assess the goodness-of-fit of each dataset to the model. This fit was based on all frequency sweeps taken at all prices, rather than only the frequency sweep taken at the highest price and the price sweep.

The purpose of this fit was not to validate the model, but rather, to compare the goodness of fit of the model to data collected using traditional sweep methodology and data collected using randomized sampling of the parameter space. If there existed large inconsistencies in the data collected under different conditions of the sweep portion, there would be large, systematic deviations of data points from the fitted surface Furthermore, if the two portions of the experiments differed drastically in the model's parameter estimates and goodness-of-fit statistics, it would suggest that some psychological process had contributed differentially to performance during the sweep and randomized portions of the experiment. Using the mountain model has the advantage of assessing whether a data collection method that would remove systematic biases provides a smoother fit of the model used later on in this thesis than traditional methods.

The data were fit according to a non-linear estimation routine coded in MATLAB (Conover, 2006). Weights for each time-allocation value obtained at a given price and frequency were computed using Tukey's bi-square estimator, with the tuning constant set to 6. All confidence intervals surrounding estimates derived from the fitting procedure, including mean time allocations, fitted surface estimates, and parameter shifts, were calculated based on a bootstrapping procedure. Details of the bootstrap re-sampling procedure, justification for the use of Tukey's bi-square estimator, and a description of this robust statistical technique are included in appendix A.

Internal Consistency

Sweep Portion

Table 1 shows, for all rats, the estimated parameters of the multi-segment sigmoids fitted to price and frequency sweep data. Also included are the 95% confidence intervals surrounding these parameter estimates, as determined by the bootstrapping procedure.

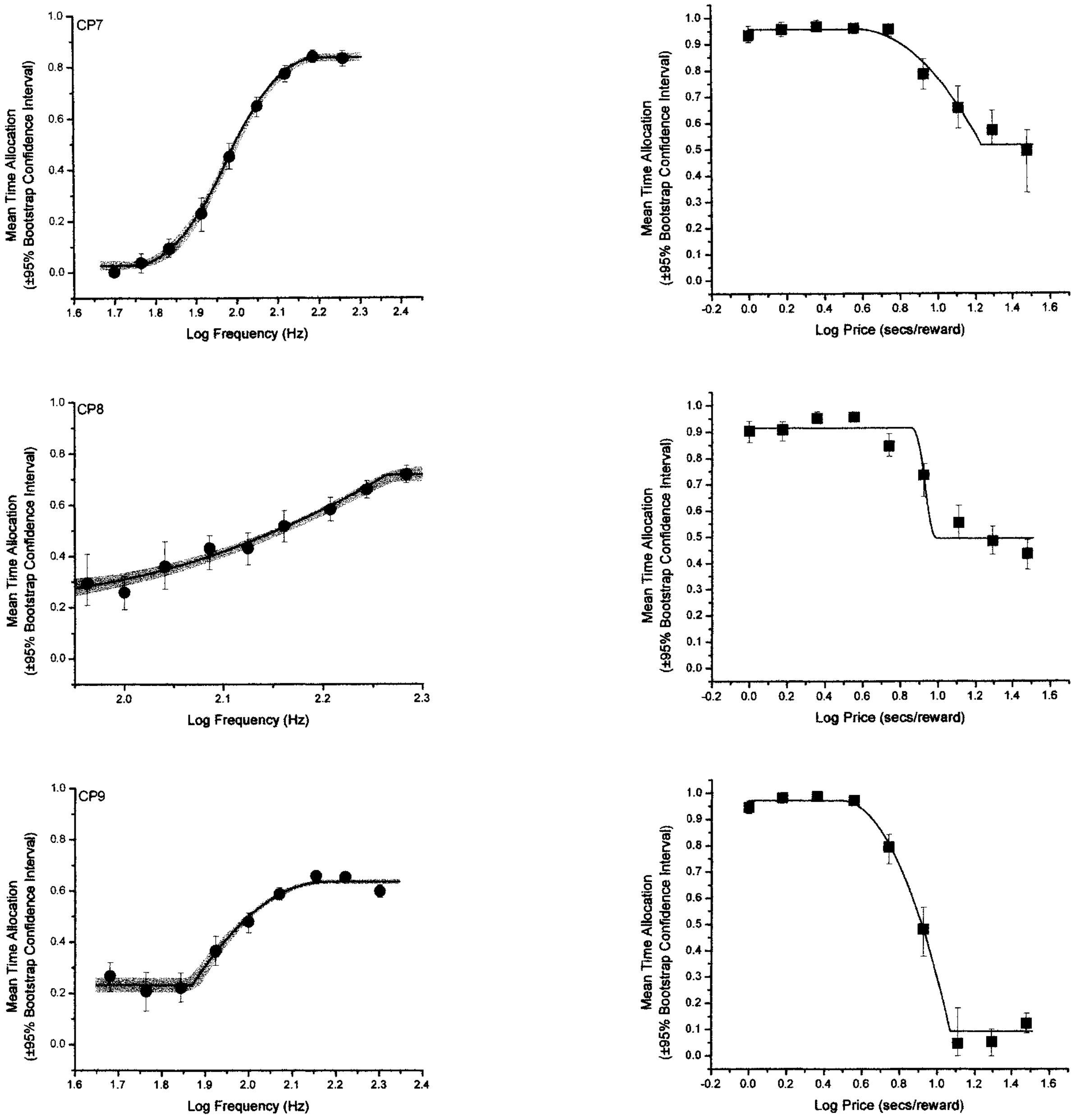
Figures 4 and 5 plot the sigmoid fitted to the data from each price sweep (in dark blue) and frequency sweep (in dark red) for each rat. Coincident with these fits are

Table 1

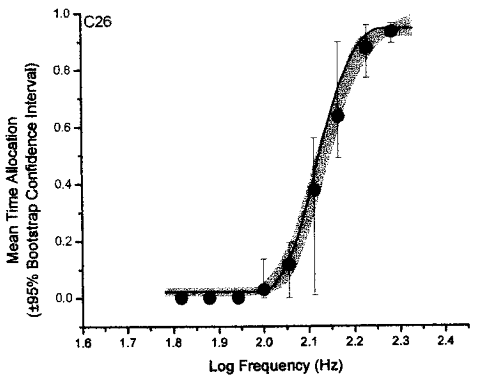
Estimated Parameters of Pseudo-Sigmoid Fits

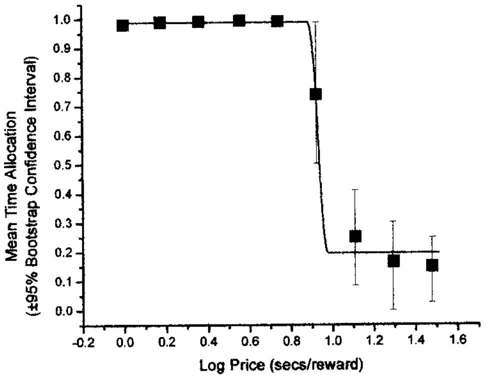
Frequency sweeps conducted in "sweep" condition				Price sweeps conducted in "sweep" condition			
			C26				
Parameter	Estimated value	Lower bound of confidence interval	Upper bound of confidence interval	Parameter	Estimated value	Lower bound of confidence interval	Upper bound of confidence interval
Q1A	0.0210	0.0066	0.0368	**Q1A**	0.9889	0.9870	0.9905
Q2A	0.9427	0.9136	0.9985	**Q2A**	0.1964	0.1513	0.2449
K1	1.9886	1.9401	2.0279	**K1**	0.9820	0.9713	0.9936
Q1G	29.0730	13.3570	54.4040	**Q1G**	210.8800	201.2300	219.8500
Q2G	26.6910	10.2280	88.3470	**Q2G**	179.3400	170.8600	187.5100
			CP2				
Q1A	0.0278	0.0147	0.0441	**Q1A**	0.9133	0.8995	0.9258
Q2A	0.7278	0.6575	0.7995	**Q2A**	0.1388	0.1137	0.1638
K1	2.1890	2.1329	2.2587	**K1**	0.7946	0.7884	0.8019
Q1G	9.5544	6.5596	16.5160	**Q1G**	181.2800	177.3400	185.3000
Q2G	958.5900	917.0900	988.8800	**Q2G**	69.9030	59.4630	79.8160
			CP3				
Q1A	0.1336	0.1100	0.1579	**Q1A**	0.8965	0.8800	0.9155
Q2A	0.5228	0.5041	0.5425	**Q2A**	0.1707	0.1411	0.2020
K1	1.6866	1.6383	1.7180	**K1**	1.0220	0.9692	1.0793
Q1G	328.0100	310.6900	337.2000	**Q1G**	320.5100	305.7300	334.8400
Q2G	17.5610	7.0564	50.3770	**Q2G**	1.0199	0.7824	1.2686
			CP4				
Q1A	0.3441	0.3107	0.3726	**Q1A**	0.7991	0.7798	0.8168
Q2A	0.6892	0.6777	0.6988	**Q2A**	0.1155	0.0966	0.1349
K1	1.7215	1.6858	1.7613	**K1**	1.2094	1.1671	1.2596
Q1G	22.9050	15.7200	40.0280	**Q1G**	360.8600	340.7100	378.8000
Q2G	77.7490	31.4790	139.2700	**Q2G**	2.6597	1.8825	3.8104
			CP7				
Q1A	0.0272	0.0120	0.0453	**Q1A**	0.9573	0.9501	0.9655
Q2A	0.8373	0.8215	0.8523	**Q2A**	0.5128	0.4278	0.5618
K1	1.7416	1.7010	1.7829	**K1**	1.2458	1.1547	1.4069
Q1G	9.2419	6.3866	13.2930	**Q1G**	436.7300	435.7700	442.3900
Q2G	9.7250	7.0894	13.2200	**Q2G**	2.1829	0.9416	3.5250
			CP8				
Q1A	0.2491	0.1945	0.3056	**Q1A**	0.9139	0.9043	0.9236
Q2A	0.7188	0.6904	0.7465	**Q2A**	0.4945	0.4717	0.5170
K1	1.8483	1.7830	1.9294	**K1**	0.9937	0.9869	1.0000
Q1G	5.7484	4.1750	8.9659	**Q1G**	119.6100	111.2000	127.4100
Q2G	302.5000	237.4800	374.3200	**Q2G**	105.5900	99.1160	112.6400
			CP9				
Q1A	0.2333	0.2057	0.2610	**Q1A**	0.9714	0.9661	0.9765
Q2A	0.6363	0.6278	0.6445	**Q2A**	0.0939	0.0752	0.1143
K1	1.8646	1.8444	1.8847	**K1**	1.0742	1.0390	1.1123
Q1G	242.4600	226.7900	256.7100	**Q1G**	291.6200	271.3200	309.6400
Q2G	10.5820	8.3666	13.1610	**Q2G**	2.9542	2.2859	3.7791

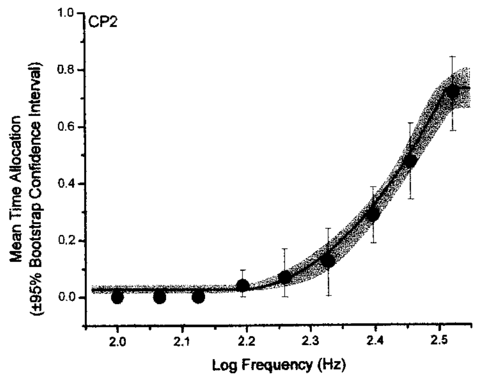
Frequency sweep reconstructions conducted in randomized condition				Price sweep reconstructions conducted in randomized condition			
			C26				
Parameter	Estimated value	Lower bound of confidence interval	Upper bound of confidence interval	Parameter	Estimated value	Lower bound of confidence interval	Upper bound of confidence interval
Q1A	0.1150	0.0968	0.1374	**Q1A**	0.9584	0.9555	0.9615
Q2A	0.6808	0.6269	0.7502	**Q2A**	0.1188	0.1027	0.1364
K1	2.2103	2.1779	2.2419	**K1**	1.1933	1.1704	1.2242
Q1G	61.3720	35.9690	94.5670	**Q1G**	61.4810	21.5590	97.2330
Q2G	96.9680	9.0092	194.9600	**Q2G**	4.1651	3.4875	4.9824
			CP4				
Q1A	0.2053	0.1344	0.2251	**Q1A**	0.9216	0.9065	0.9352
Q2A	0.2584	0.2163	0.3722	**Q2A**	0.2614	0.2383	0.2852
K1	1.9831	1.6389	2.0378	**K1**	0.8362	0.8005	0.8785
Q1G	178.7400	99.5460	183.5500	**Q1G**	219.3800	200.9900	232.6000
Q2G	145.1600	125.8300	299.7800	**Q2G**	6.3616	4.0940	9.3370
			CP7				
Q1A	0.1035	0.0799	0.1279	**Q1A**	0.8992	0.8906	0.9082
Q2A	0.5501	0.5121	0.5870	**Q2A**	0.1461	0.1313	0.1601
K1	1.9058	1.8804	1.9218	**K1**	1.0814	1.0271	1.1368
Q1G	44.5150	25.0880	59.6000	**Q1G**	100.1700	20.6730	167.5100
Q2G	177.2900	154.0800	178.3400	**Q2G**	3.6806	2.4010	5.5660
			CP8				
Q1A	0.3986	0.3728	0.4167	**Q1A**	1.0000	1.0000	1.0000
Q2A	0.5204	0.4827	0.5542	**Q2A**	0.2835	0.2521	0.3210
K1	2.0930	2.0356	2.1176	**K1**	1.6749	1.6050	1.7473
Q1G	194.0800	94.5480	220.2300	**Q1G**	416.4900	326.9800	494.8500
Q2G	361.3200	99.7600	521.5400	**Q2G**	0.2857	0.2435	0.3314
			CP9				
Q1A	0.1545	0.1336	0.1753	**Q1A**	0.8133	0.7851	0.8475
Q2A	0.5210	0.5004	0.5410	**Q2A**	0.2407	0.1914	0.2689
K1	1.8224	1.7698	1.8596	**K1**	1.0522	0.9633	1.2174
Q1G	82.9250	20.0720	138.8000	**Q1G**	113.7000	23.4820	180.9700
Q2G	7.8196	4.8629	11.7580	**Q2G**	1.3840	0.5503	2.4452

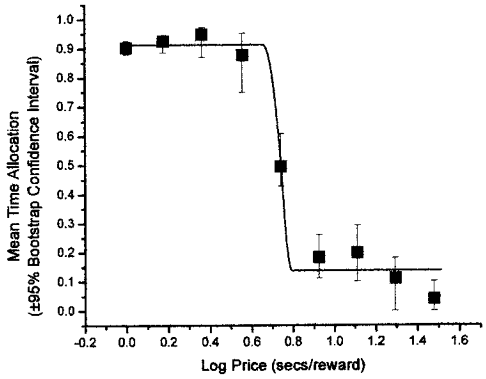


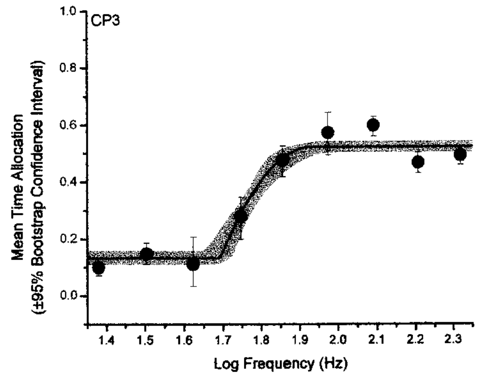
Figures 4 and 5. Observed mean corrected time allocation to self-stimulation activities in the "sweep" portion of the experiment. Solid lines depict the best-fitting dual-quadratic spline function. Error bars represent 95% confidence intervals (derived by bootstrap-resampling) surrounding the mean corrected time allocation estimates; envelopes surrounding pseudo-sigmoids represent the 95% confidence region of the function.

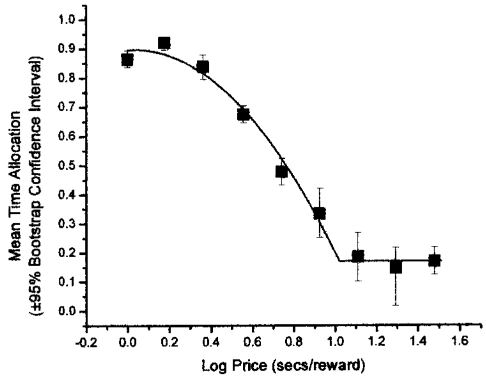

C26
Mean Time Allocation
(±95% Bootstrap Confidence Interval)
0.0
0.2
0.4
0.6
0.8
1.0
1.6
1.7
1.8
1.9
2.0
2.1
2.2
2.3
2.4
Log Frequency (Hz)

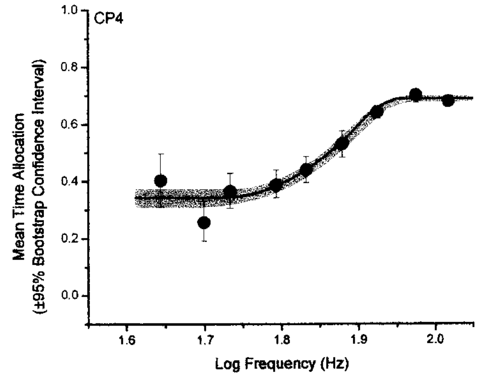

Mean Time Allocation
(±95% Bootstrap Confidence Interval)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
Log Price (secs/reward)

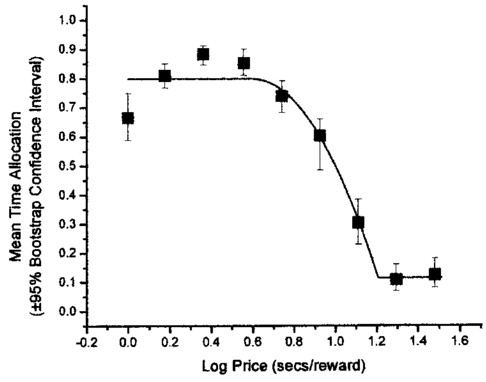

CP2
Mean Time Allocation
(±95% Bootstrap Confidence Interval)
0.0
0.2
0.4
0.6
0.8
1.0
2.0
2.1
2.2
2.3
2.4
2.5
Log Frequency (Hz)


Mean Time Allocation
(±95% Bootstrap Confidence Interval)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
Log Price (secs/reward)


CP3
Mean Time Allocation
(±95% Bootstrap Confidence Interval)
0.0
0.2
0.4
0.6
0.8
1.0
1.4
1.5
1.6
1.7
1.8
1.9
2.0
2.1
2.2
2.3
Log Frequency (Hz)


Mean Time Allocation
(±95% Bootstrap Confidence Interval)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
Log Price (secs/reward)


CP4
Mean Time Allocation
(±95% Bootstrap Confidence Interval)
0.0
0.2
0.4
0.6
0.8
1.0
1.6
1.7
1.8
1.9
2.0
Log Frequency (Hz)

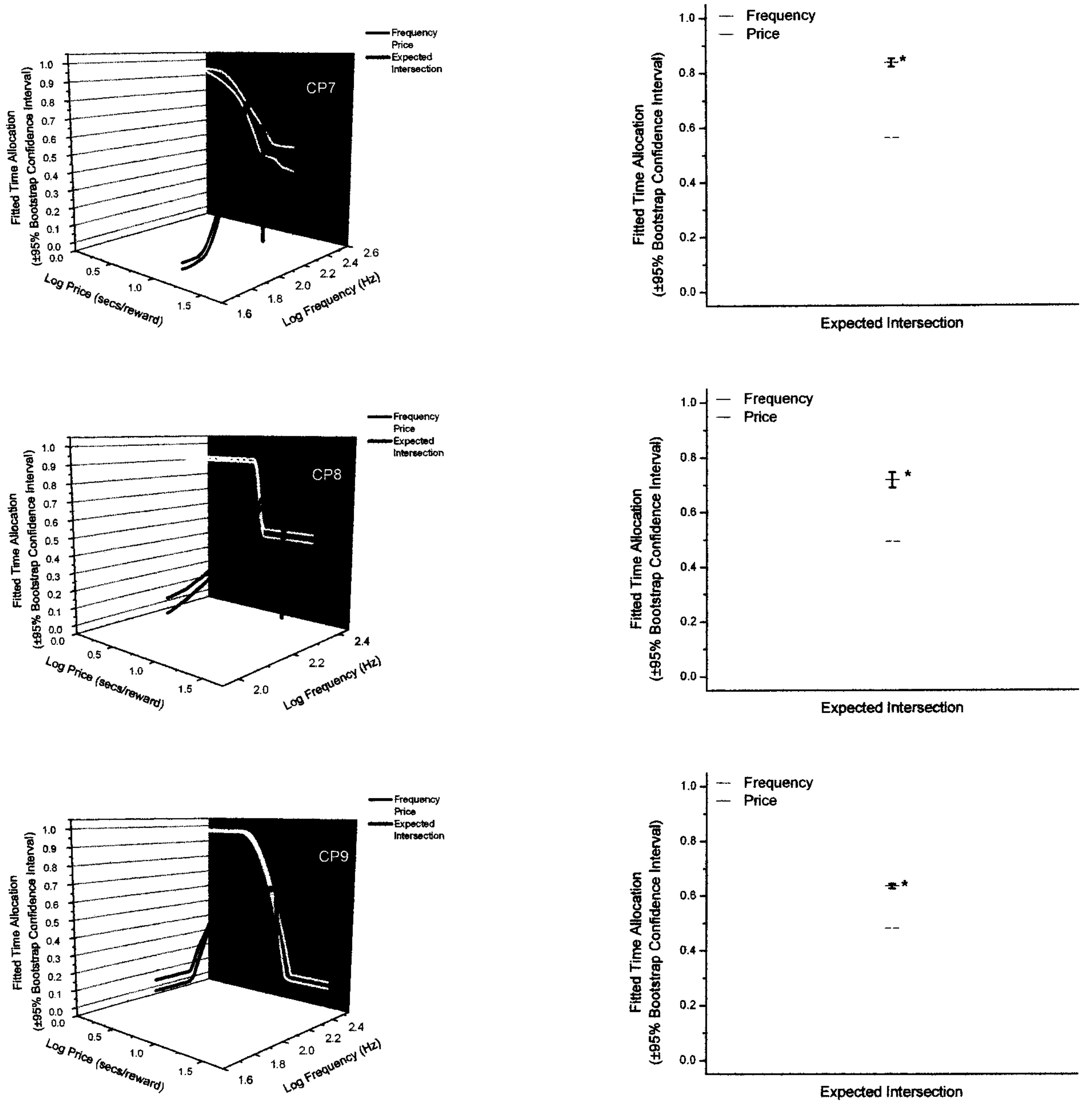

Mean Time Allocation
(±95% Bootstrap Confidence Interval)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
Log Price (secs/reward)

ribbons corresponding to the 95% confidence interval (determined by the bootstrapping procedure) surrounding the fits. The cyan ribbon represents the 95% confidence surrounding the fitted price sweep sigmoid; the magenta ribbon represents the same level of confidence surrounding the frequency sweep. The animal's bi-square-weighted mean time allocation and associated confidence intervals derived by means of resampling are shown in the same panel; in almost all cases, the fitted sigmoid provides a close approximation to the data.

The right column of figures 6 and 7 presents the fitted time allocations for the expected intersection point; the vertical black line in the 3D plot in the left column designates the point in the parameter space where the sigmoids fitted to the frequency- and price-sweep data intersect. Coincident with this left-hand 3D plot are the 95% confidence intervals surrounding the curves fitted to frequency sweep (magenta) or price sweep (cyan) data. The extrapolations of time allocation at the intersection point were estimated from the sigmoidal functions fitted to the frequency-sweep and price-sweep data. In all 7 cases, different time allocations were obtained at the point of intersection in the parameter space when that point was sampled during a frequency sweep or a price sweep.

Randomized Portion

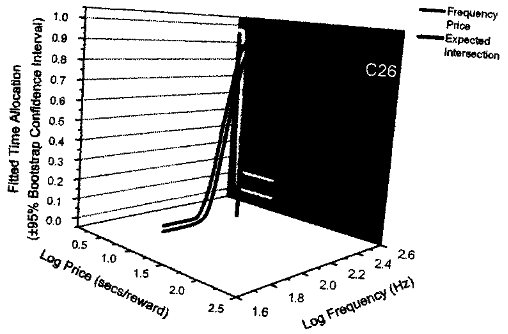
Figures 8 and 9 show the fitted ribbon plots as in figures 4 and 5, for the data collected in the randomized portion of the experiment. The right sides of figures 10 and 11 present the time allocations and 95% confidence estimated by fits to the randomized price sweep data and the randomized frequency sweep data; the left side contains 3D plots similar to those in figures 6 and 7. In 3 of 5 cases, the 95% confidence surrounding

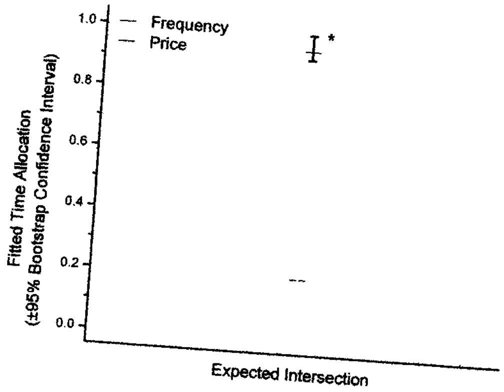


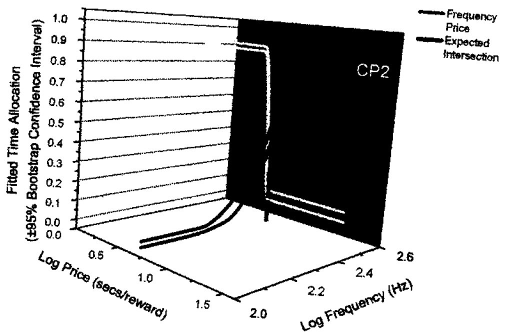
Figures 6 and 7. 95% confidence bounds of the dual-quadratic spline function fitted to data collected in the "sweep" portion of the experiment.

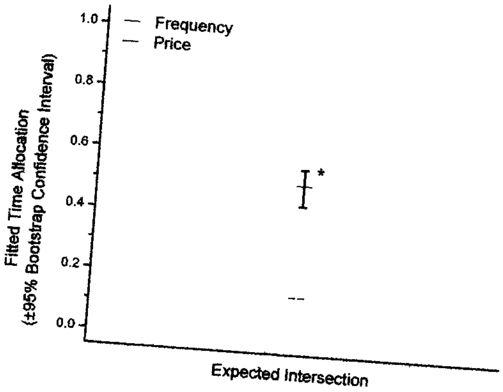
The left panel shows upper and lower bounds in a 3D space in which time allocation is plotted as a function of pulse frequency and stimulation price. A vertical line is traced at the point at which one would expect the intersection of the frequency and price "sweeps" to occur.

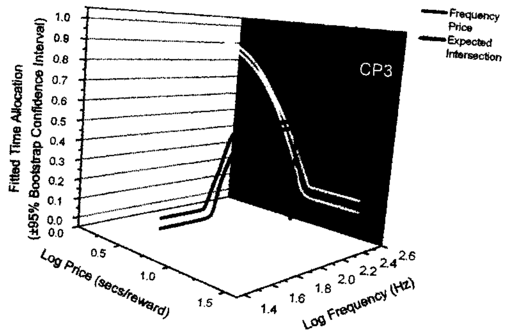
The right panel shows the time allocation predicted by the fitted function according to interpolations of the price "sweep" or extrapolations from the frequency "sweep". Error bars represent the 95% confidence surrounding the time allocation predicted by the fit.

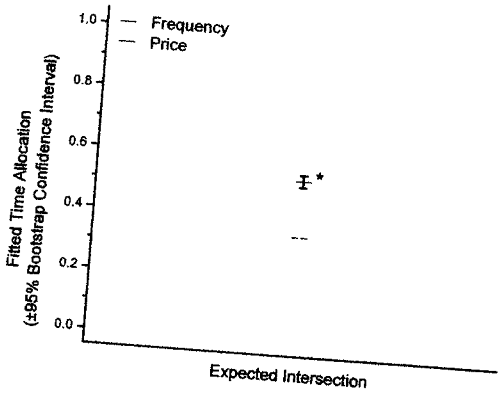

C26
Frequency
Price
Expected Intersection
Fitted Time Allocation (±95% Bootstrap Confidence Interval)
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Log Price (secs/reward)
0.5 1.0 1.5 2.0 2.5
Log Frequency (Hz)
1.6 1.8 2.0 2.2 2.4 2.6

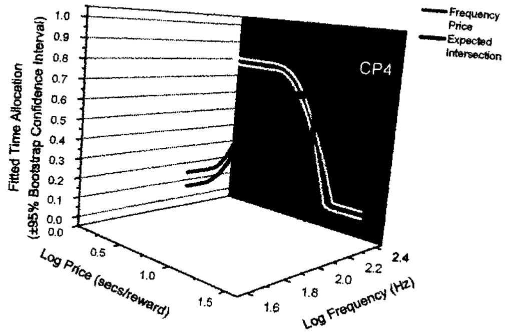

Frequency
Price
*
Fitted Time Allocation (±95% Bootstrap Confidence Interval)
0.0 0.2 0.4 0.6 0.8 1.0
Expected Intersection

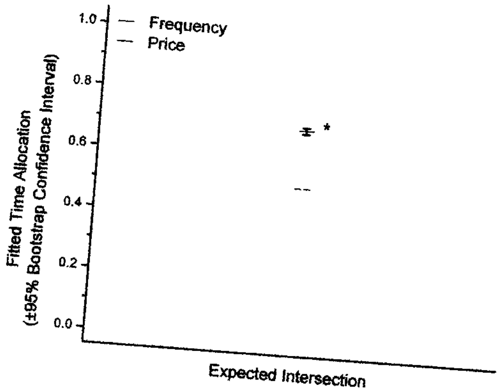

CP2
Frequency
Price
Expected Intersection
Fitted Time Allocation (±95% Bootstrap Confidence Interval)
0.0 0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Log Price (secs/reward)
0.5 1.0 1.5
Log Frequency (Hz)
2.0 2.2 2.4 2.6

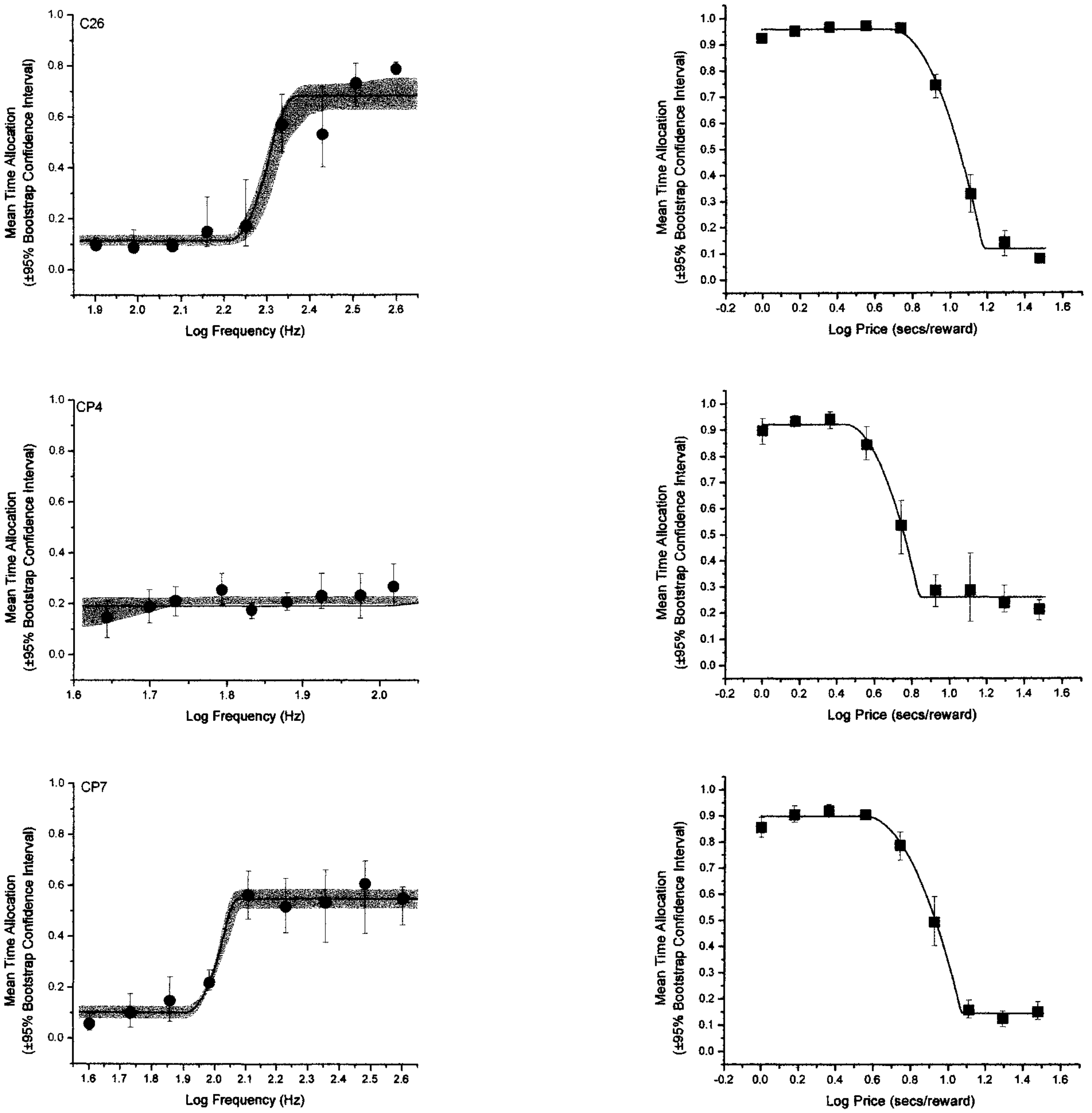

Frequency
Price
*
Fitted Time Allocation (±95% Bootstrap Confidence Interval)
0.0 0.2 0.4 0.6 0.8 1.0
Expected Intersection


CP3
Frequency
Price
Expected Intersection
Fitted Time Allocation (±95% Bootstrap Confidence Interval)
0.0 0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Log Price (secs/reward)
0.5 1.0 1.5
Log Frequency (Hz)
1.4 1.6 1.8 2.0 2.2 2.4 2.6

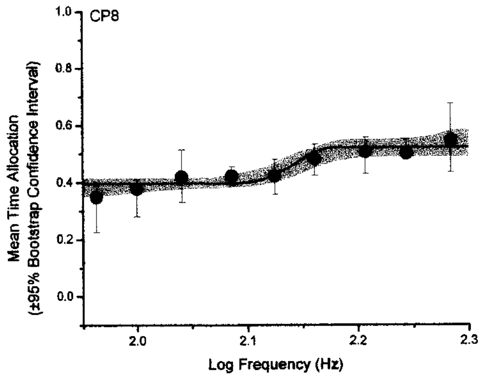

Frequency
Price
*
Fitted Time Allocation (±95% Bootstrap Confidence Interval)
0.0 0.2 0.4 0.6 0.8 1.0
Expected Intersection

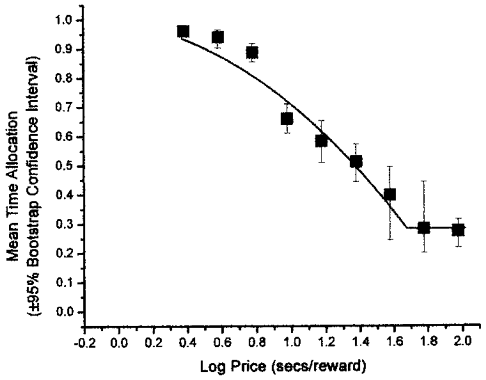

CP4
Frequency
Price
Expected Intersection
Fitted Time Allocation (±95% Bootstrap Confidence Interval)
0.0 0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Log Price (secs/reward)
0.5 1.0 1.5
Log Frequency (Hz)
1.6 1.8 2.0 2.2 2.4

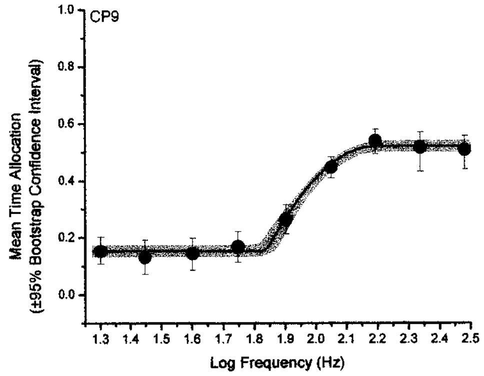

Frequency
Price
*
Fitted Time Allocation (±95% Bootstrap Confidence Interval)
0.0 0.2 0.4 0.6 0.8 1.0
Expected Intersection

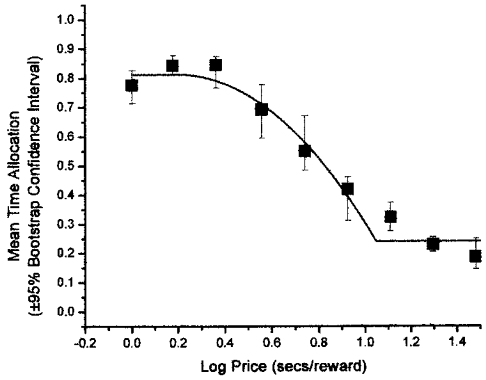


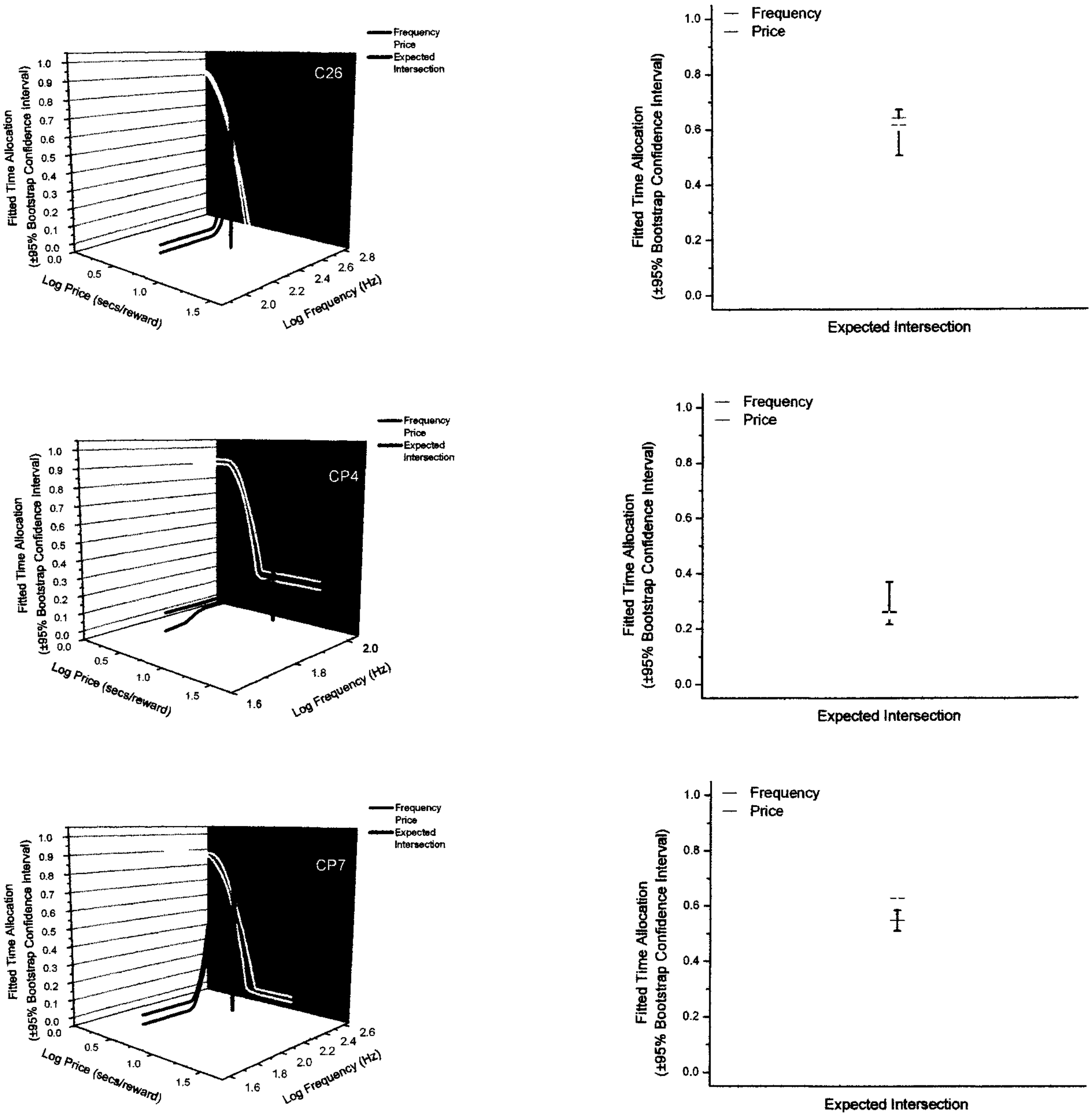
Figures 8 and 9. Observed mean corrected time allocation to self-stimulation activities in the randomized portion of the experiment. Solid lines depict the best-fitting dual-quadratic spline function. Error bars represent 95% confidence intervals (derived by bootstrap-resampling) surrounding the mean corrected time allocation estimates; envelopes surrounding pseudo-sigmoids represent the 95% confidence region of the function.


CP8
Mean Time Allocation
(±95% Bootstrap Confidence Interval)
0.0
0.2
0.4
0.6
0.8
1.0
2.0
2.1
2.2
2.3
Log Frequency (Hz)


Mean Time Allocation
(±95% Bootstrap Confidence Interval)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
Log Price (secs/reward)


CP9
Mean Time Allocation
(±95% Bootstrap Confidence Interval)
0.0
0.2
0.4
0.6
0.8
1.0
1.3
1.4
1.5
1.6
1.7
1.8
1.9
2.0
2.1
2.2
2.3
2.4
2.5
Log Frequency (Hz)

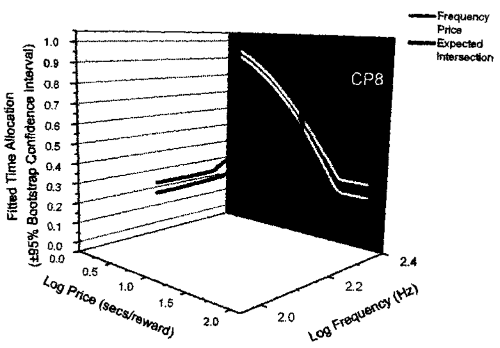

Mean Time Allocation
(±95% Bootstrap Confidence Interval)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
Log Price (secs/reward)

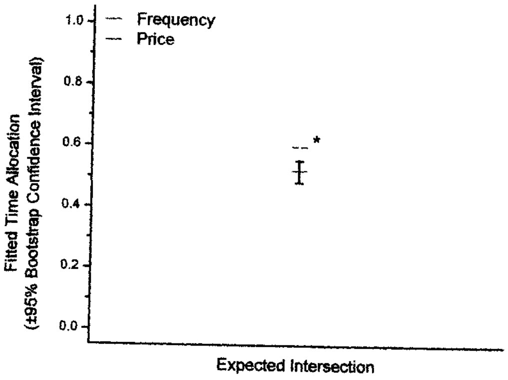


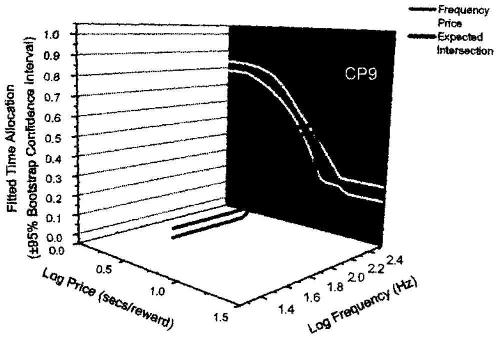
Figures 10 and 11. 95% confidence bounds of the dual-quadratic spline function fitted to randomized data.

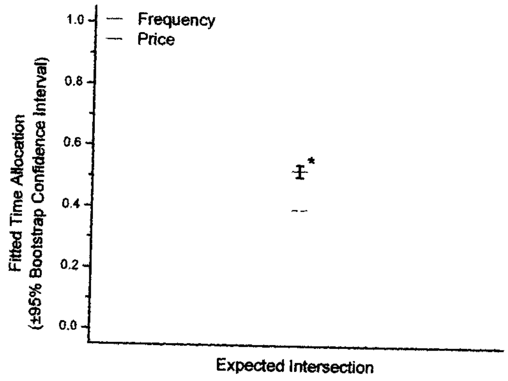
The left panel shows upper and lower bounds in a 3D space in which time allocation is plotted as a function of pulse frequency and stimulation price. A vertical line is traced at the point at which one would expect the intersection of the frequency and price "sweeps" to occur.

The right panel shows the time allocation predicted by the fitted function according to interpolations of the price "sweep" or extrapolations from the frequency "sweep". Error bars represent the 95% confidence surrounding the time allocation predicted by the fit.


Frequency
Price
Expected Intersection
CP8
Fitted Time Allocation (±95% Bootstrap Confidence Interval)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.0
0.5
1.0
1.5
2.0
Log Price (secs/reward)
2.0
2.2
2.4
Log Frequency (Hz)


Frequency
Price
Fitted Time Allocation (±95% Bootstrap Confidence Interval)
1.0
0.8
0.6
0.4
0.2
0.0
*
Expected Intersection


Frequency
Price
Expected Intersection
CP9
Fitted Time Allocation (±95% Bootstrap Confidence Interval)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.0
0.5
1.0
1.5
Log Price (secs/reward)
1.4
1.6
1.8
2.0
2.2
2.4
Log Frequency (Hz)


Frequency
Price
Fitted Time Allocation (±95% Bootstrap Confidence Interval)
1.0
0.8
0.6
0.4
0.2
0.0
*
Expected Intersection

estimated time allocations under price sweep and frequency sweep reconstructions overlap almost completely, and in the remaining two cases, the gap between them is small.

Surface Fits to Sweep and Random Data Sets

Table 2 shows the fitted surface parameters and the associated 95% confidence bounds as estimated by bootstrapping, for the sweep and randomized portion of the experiment for all rats. In certain cases, it can not be said with 95% confidence that the estimated parameter under sweep conditions of the experiment is the same under randomized conditions. Notably, in all cases, the growth of the rewarding effect with pulse frequency is steeper when animals are tested in traditional, sweep procedures than when test trials are randomized. As a result, the curves appear to drop in a near step-like fashion from maximal time allocation to minimal time allocation when pulse frequency crosses some threshold level during sweep conditions.

Figures 12 through 18 depict the observed bi-square-weighted mean time allocations for each rat as well as multiple projections of the fitted surface, for sweep and randomized data. In many cases, the surface fit to sweep condition data is reliably biased in different directions upon repetition of the same sweep. In the upper left corner, there is a comparison of the fitted surface with sweeps obtained when the price of the frequency sweep was medium and the previous frequency sweep was taken at a low (upward-pointing triangles) or high (downward-pointing triangles) price. In those cases where a second frequency sweep was obtained at a low price, there is a similar comparison of the fitted surface with sweeps obtained when the previous sweep was a price sweep (upward-pointing triangles) or a frequency sweep obtained at a medium price (downward-pointing

Table 2

Estimated Parameters of Surface Fit

"Sweep" Portion

Randomized Portion

CP2

Parameter	Estimated Value	Lower confidence bound	Upper confidence bound	Confidence Interval Width
A	1.7836	1.2502	2.7330	1.4828
Fhm	2.5253	2.5131	2.5375	0.0244
G	14.5998	12.6720	16.1449	3.4728
Pe	1.8518	1.6842	1.9875	0.3033
TAX	0.8293	0.7895	0.8702	0.0808
TIN	0.0480	0.0373	0.0601	0.0227

CP3

Parameter	Estimated Value	Lower confidence bound	Upper confidence bound	Confidence Interval Width
A	1.0208	0.9604	1.0816	0.1213
Fhm	1.8982	1.8757	1.9190	0.0434
G	6.5358	5.7898	7.4836	1.6939
Pe	0.6787	0.6559	0.7028	0.0469
TAX	1.0000	0.9999	1.0000	0.0001
TIN	0.1173	0.1050	0.1301	0.0251

C26

Parameter	Estimated Value	Lower confidence bound	Upper confidence bound	Confidence Interval Width	Parameter	Estimated Value	Lower confidence bound	Upper confidence bound	Confidence Interval Width
A	17.2128	8.7556	19.9999	11.2443	**A**	7.2551	5.7988	9.0735	3.2747
Fhm	[illegible]	2.0304	2.0348	0.0044	**Fhm**	2.1141	2.0938	2.1399	0.0461
G	11.9416	11.5378	12.7844	1.2466	**G**	3.1046	2.8602	3.3948	0.5347
Pe	1.0439	1.0329	1.0646	0.0317	**Pe**	1.0732	1.0516	1.1034	0.0518
TAX	0.9557	0.9499	0.9614	0.0115	**TAX**	0.9556	0.9501	0.9613	0.0112
TIN	0.0395	0.0293	0.0483	0.0190	**TIN**	0.1322	0.1094	0.1570	0.0477

CP4

Parameter	Estimated Value	Lower confidence bound	Upper confidence bound	Confidence Interval Width	Parameter	Estimated Value	Lower confidence bound	Upper confidence bound	Confidence Interval Width
A	2.1960	1.9619	2.4620	0.5001	**A**	3.8955	3.2452	4.7735	1.5283
Fhm	1.8416	1.8355	1.8476	0.0121	**Fhm**	1.8685	1.8336	1.9073	0.0737
G	13.0614	12.0492	14.2284	2.1793	**G**	5.0025	4.1928	6.1805	1.9877
Pe	0.9618	0.9358	0.9847	0.0489	**Pe**	0.8793	0.8091	0.9545	0.1455
TAX	0.9046	0.8918	0.9176	0.0258	**TAX**	0.8937	0.8738	0.9126	0.0389
TIN	0.2039	0.1930	0.2145	0.0216	**TIN**	0.2083	0.1916	0.2252	0.0336

CP7

Parameter	Estimated Value	Lower confidence bound	Upper confidence bound	Confidence Interval Width	Parameter	Estimated Value	Lower confidence bound	Upper confidence bound	Confidence Interval Width
A	1.1575	1.0729	1.2502	0.1774	**A**	5.2870	4.6114	6.1296	1.5183
Fhm	2.0192	2.0129	2.0275	0.0146	**Fhm**	1.9257	1.9156	1.9374	0.0218
G	14.4655	13.3598	15.5914	2.2316	**G**	7.2158	6.5001	7.9345	1.4344
Pe	1.6100	1.5700	1.6516	0.0816	**Pe**	0.8910	0.8758	0.9073	0.0315
TAX	0.9976	0.9908	1.0000	0.0092	**TAX**	0.9063	0.8995	0.9129	0.0134
TIN	0.0360	0.0310	0.0411	0.0101	**TIN**	0.1211	0.1093	0.1323	0.0230

CP8

Parameter	Estimated Value	Lower confidence bound	Upper confidence bound	Confidence Interval Width	Parameter	Estimated Value	Lower confidence bound	Upper confidence bound	Confidence Interval Width
A	1.3140	1.1926	1.5111	0.3185	**A**	2.5837	2.1612	3.1210	0.9598
Fhm	2.0953	2.0848	2.1191	0.0342	**Fhm**	2.0659	2.0436	2.0916	0.0480
G	9.7888	8.2333	10.8540	2.6207	**G**	5.5832	4.8583	6.4583	1.6001
Pe	1.2667	1.2298	1.3155	0.0858	**Pe**	1.0687	1.0208	1.1191	0.0983
TAX	1.0000	0.9996	1.0000	0.0004	**TAX**	0.9728	0.9521	0.9949	0.0428
TIN	0.1731	0.1554	0.1917	0.0363	**TIN**	0.3559	0.3241	0.3886	0.0645

CP9

Parameter	Estimated Value	Lower confidence bound	Upper confidence bound	Confidence Interval Width	Parameter	Estimated Value	Lower confidence bound	Upper confidence bound	Confidence Interval Width
A	2.8126	2.5552	3.1406	0.5854	**A**	2.7317	2.4084	3.0626	0.6542
Fhm	1.9385	1.9323	1.9449	0.0126	**Fhm**	1.8605	1.8459	1.8745	0.0287
G	8.7394	8.2052	9.3352	1.1300	**G**	5.6507	5.1127	6.2617	1.1490
Pe	0.9380	0.9283	0.9480	0.0197	**Pe**	0.9137	0.8920	0.9365	0.0445
TAX	0.9594	0.9507	0.9680	0.0173	**TAX**	0.8170	0.8001	0.8337	0.0336
TIN	0.1312	0.1222	0.1401	0.0179	**TIN**	0.1572	0.1445	0.1698	0.0253

Note. *Values in red demonstrate parameter estimates that are statistically different (a = 0.05 by bootstrap-resampling) between "sweep" and randomized portions of the experiment.*

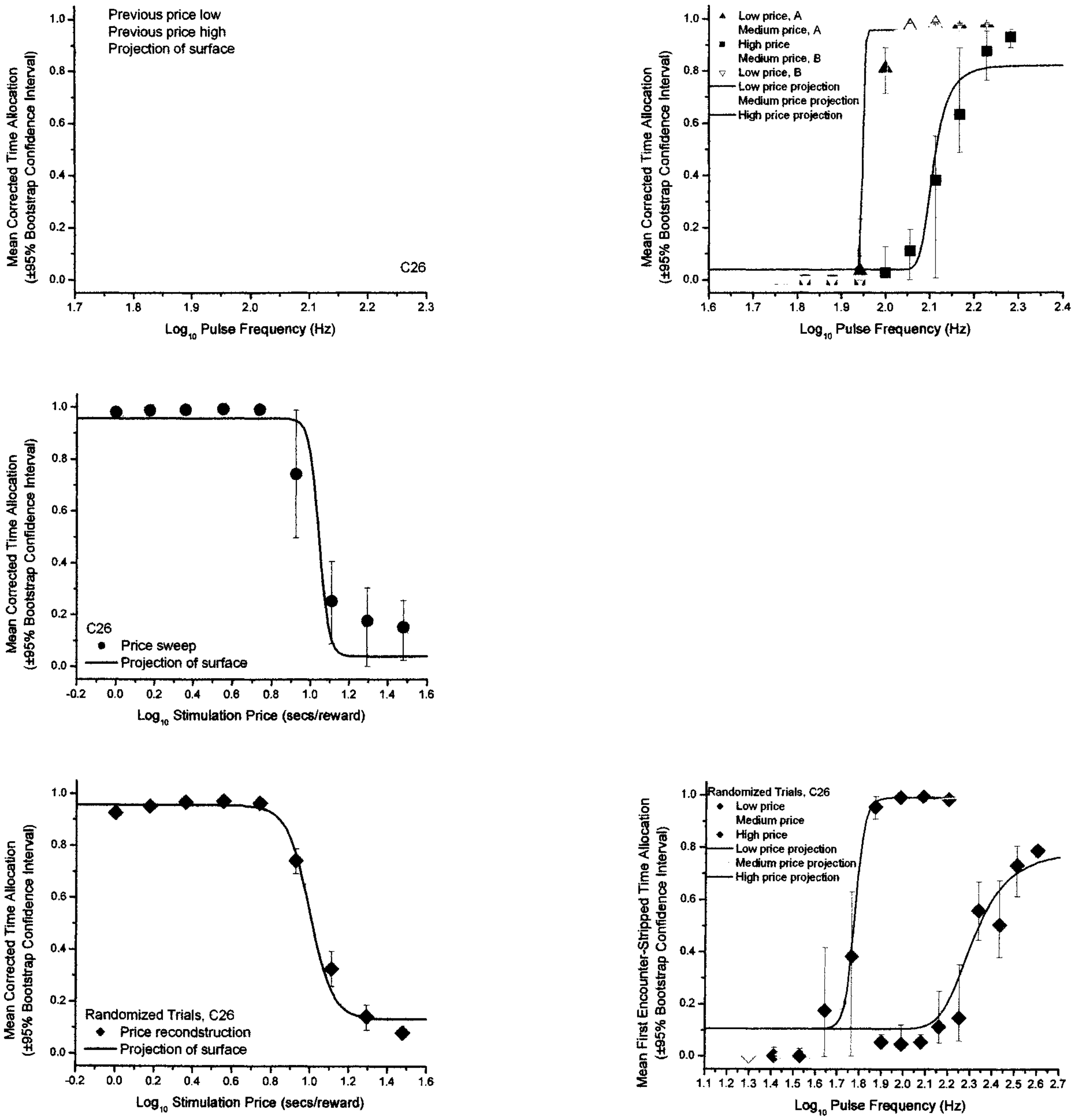


Figure 12. Observed mean corrected time allocation at each condition in the experiment contrasted with a projection of the fitted 3D surface, for rat C26. Error bars represent 95% confidence intervals (derived by bootstrap-resampling) about the estimated mean.

Upper left: Contrast of the medium-price frequency “sweep” obtained when the previous frequency sweep was obtained at high or low price with the fitted surface.

Upper right: Contrast of all frequency “sweeps” with the fitted surface.

Middle left: Contrast of price “sweep” with the fitted surface.

Lower left: Contrast of reconstructed price pseudo-``sweep” with the fitted surface.

Lower right: Contrast of reconstructed frequency pseudo-”sweeps” with the fitted surface.

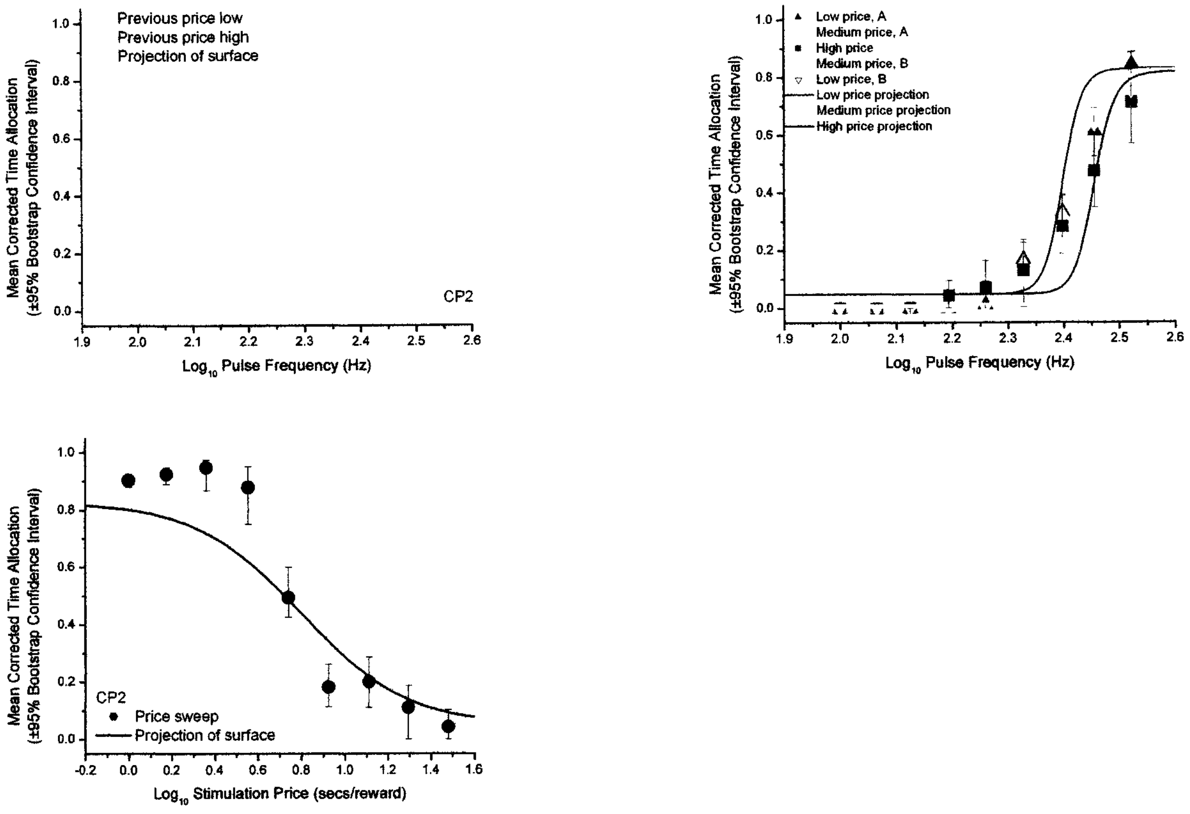


Figure 13. Observed mean corrected time allocation at each condition in the experiment contrasted with a projection of the fitted 3D surface, for rat CP2. Error bars represent 95% confidence intervals (derived by bootstrap-resampling) about the estimated mean.

Upper left: Contrast of the medium-price frequency "sweep" obtained when the previous frequency sweep was obtained at high or low price with the fitted surface.

Upper right: Contrast of all frequency "sweeps" with the fitted surface.

Middle left: Contrast of price "sweep" with the fitted surface.

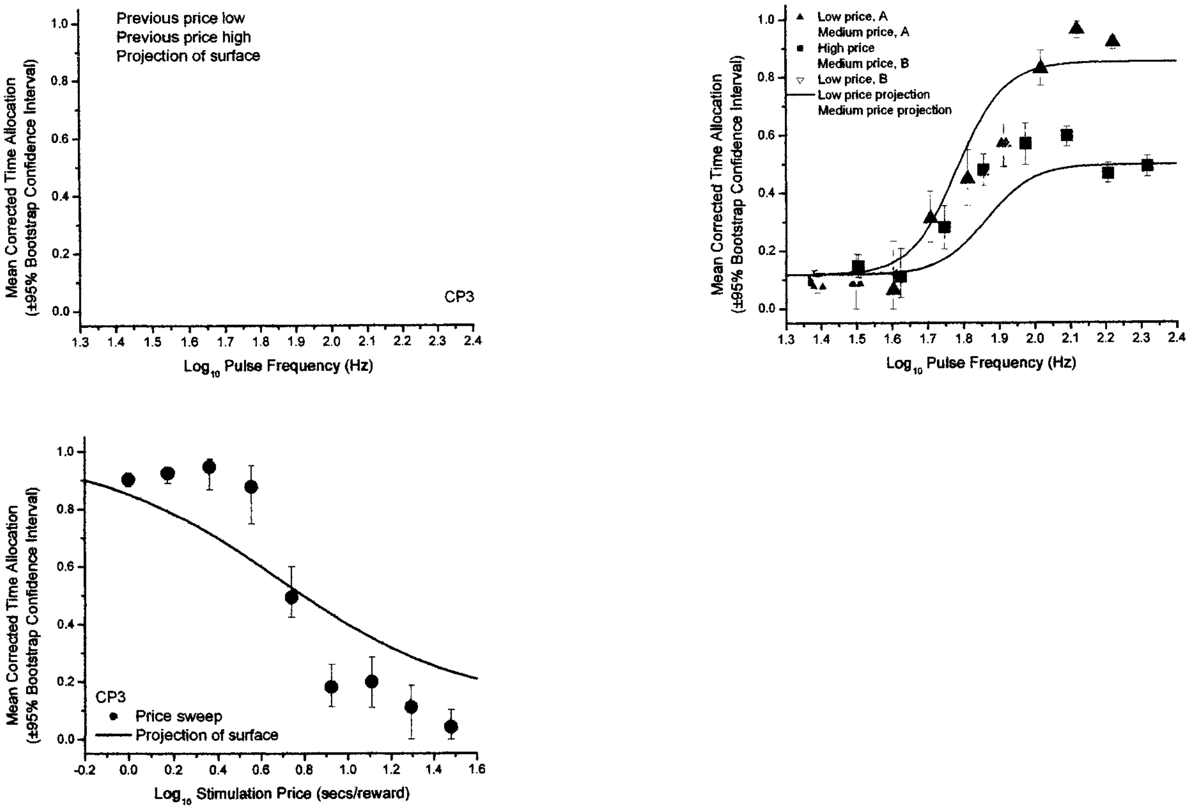


Figure 14. Observed mean corrected time allocation at each condition in the experiment contrasted with a projection of the fitted 3D surface, for rat CP3. Error bars represent 95% confidence intervals (derived by bootstrap-resampling) about the estimated mean.

Upper left: Contrast of the medium-price frequency "sweep" obtained when the previous frequency sweep was obtained at high or low price with the fitted surface.

Upper right: Contrast of all frequency "sweeps" with the fitted surface.

Middle left: Contrast of price "sweep" with the fitted surface.

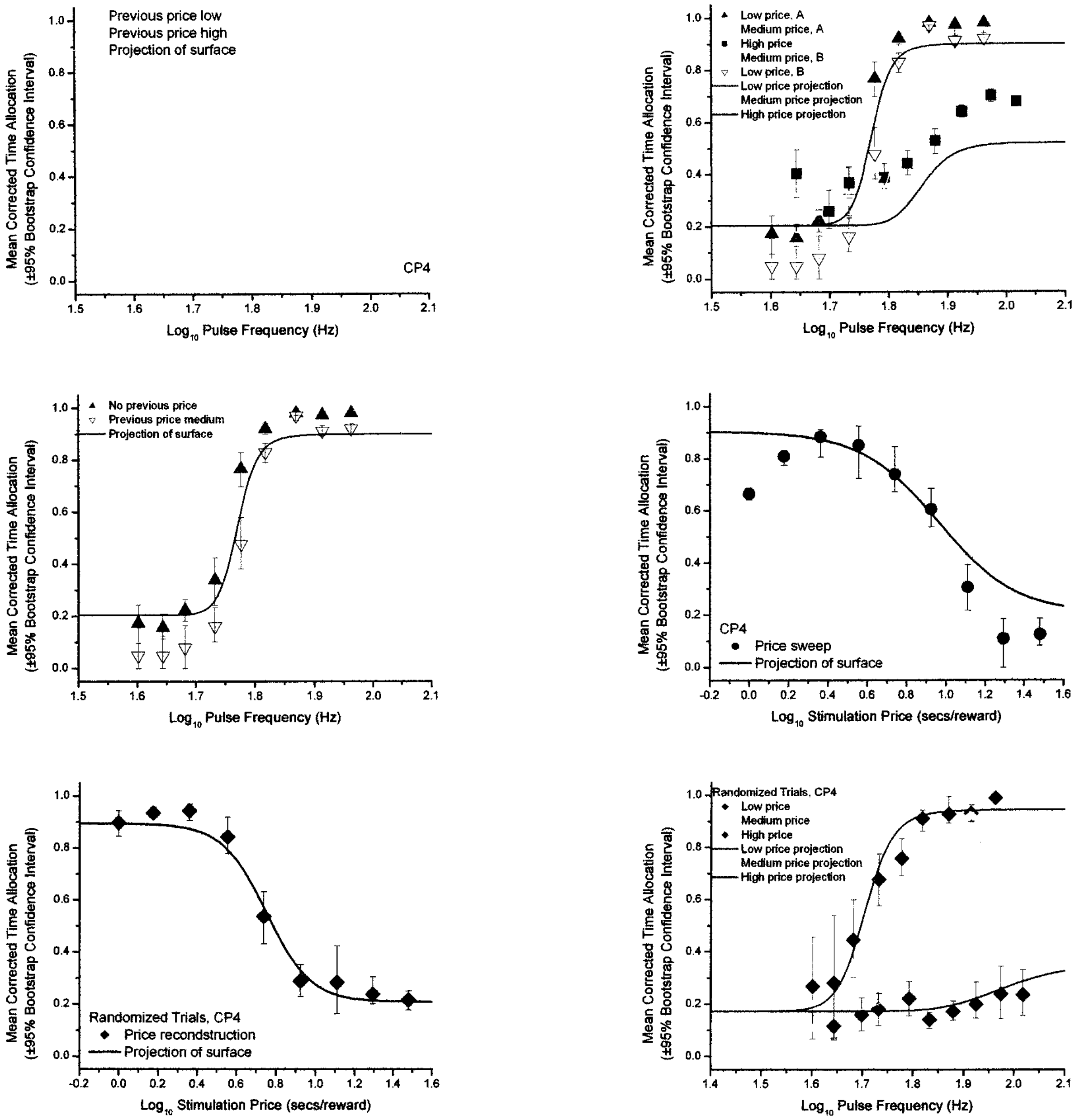


Figure 15. Observed mean corrected time allocation at each condition in the experiment contrasted with a projection of the fitted 3D surface, for rat CP4. Error bars represent 95% confidence intervals (derived by bootstrap-resampling) about the estimated mean.

Upper left: Contrast of the medium-price frequency “sweep” obtained when the previous frequency sweep was obtained at high or low price with the fitted surface.

Upper right: Contrast of all frequency “sweeps” with the fitted surface.

Middle left: Contrast of price “sweep” with the fitted surface.

Lower left: Contrast of reconstructed price pseudo-``sweep” with the fitted surface.

Lower right: Contrast of reconstructed frequency pseudo-”sweeps” with the fitted surface.

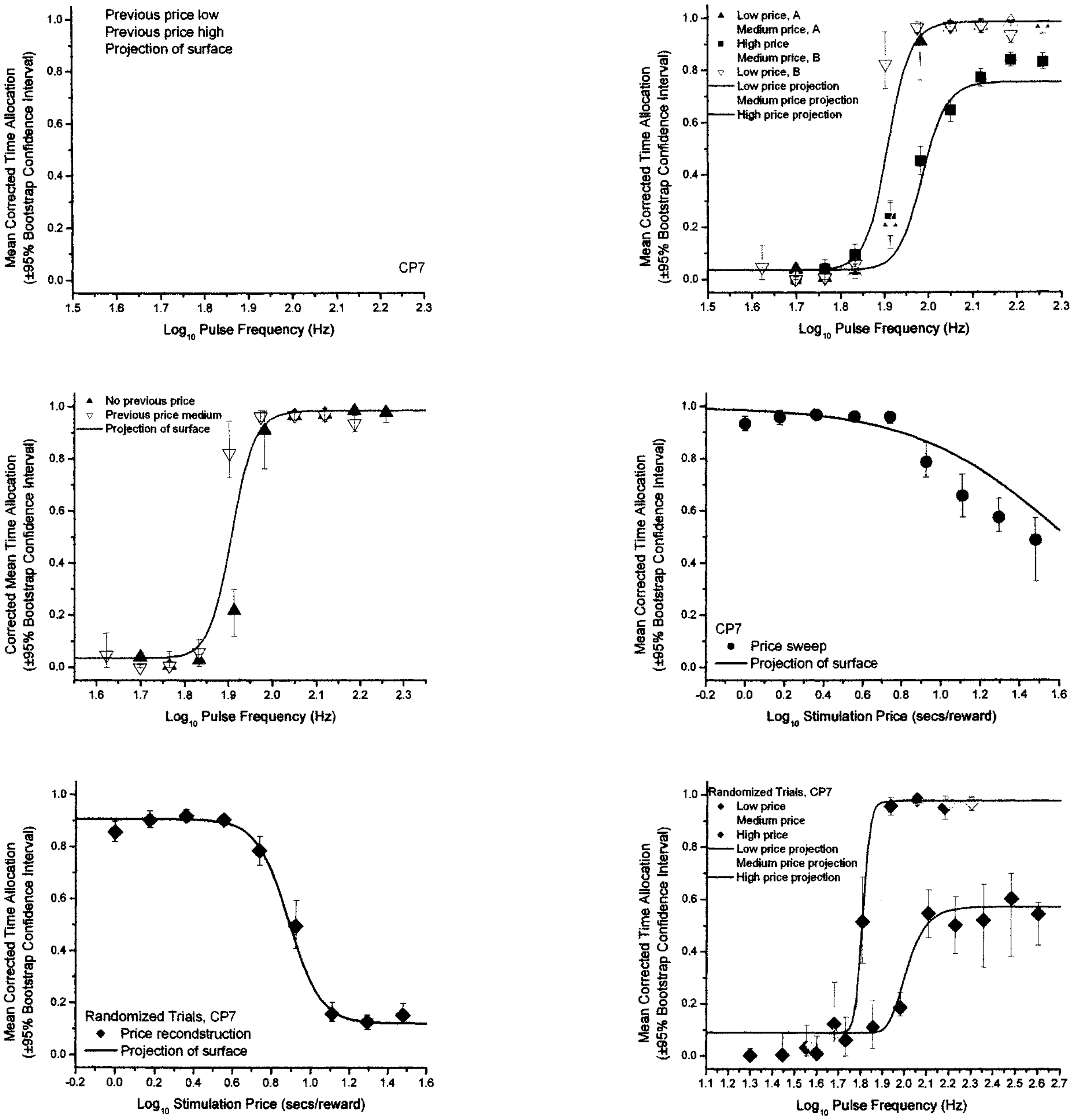


Figure 16. Observed mean corrected time allocation at each condition in the experiment contrasted with a projection of the fitted 3D surface, for rat CP7. Error bars represent 95% confidence intervals (derived by bootstrap-resampling) about the estimated mean.

Upper left: Contrast of the medium-price frequency “sweep” obtained when the previous frequency sweep was obtained at high or low price with the fitted surface.

Upper right: Contrast of all frequency “sweeps” with the fitted surface.

Middle left: Contrast of price “sweep” with the fitted surface.

Lower left: Contrast of reconstructed price pseudo-``sweep” with the fitted surface.

Lower right: Contrast of reconstructed frequency pseudo-”sweeps” with the fitted surface.

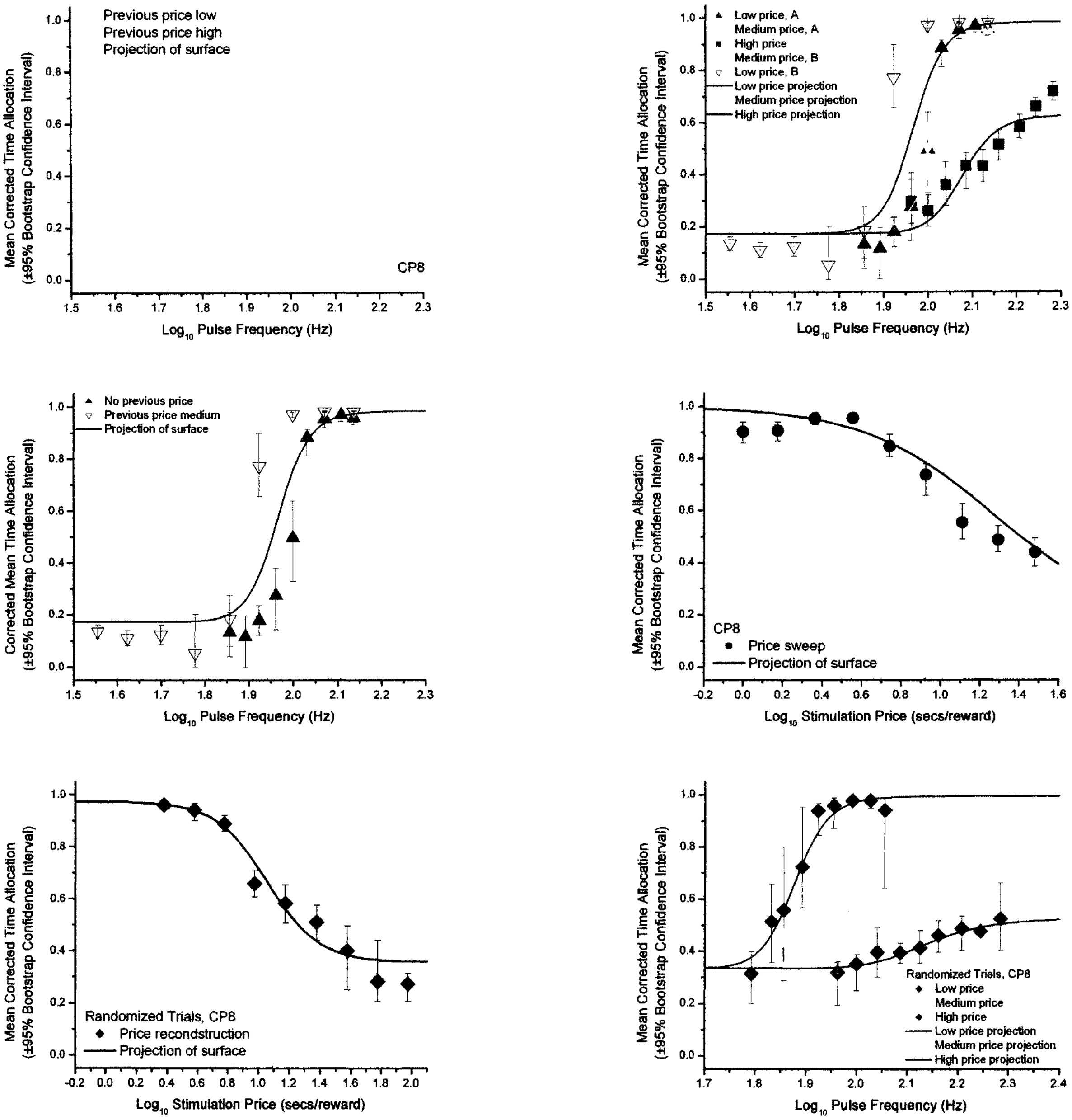


Figure 17. Observed mean corrected time allocation at each condition in the experiment contrasted with a projection of the fitted 3D surface, for rat CP8. Error bars represent 95% confidence intervals (derived by bootstrap-resampling) about the estimated mean.

Upper left: Contrast of the medium-price frequency "sweep" obtained when the previous frequency sweep was obtained at high or low price with the fitted surface.

Upper right: Contrast of all frequency "sweeps" with the fitted surface.

Middle left: Contrast of price "sweep" with the fitted surface.

Lower left: Contrast of reconstructed price pseudo-``sweep" with the fitted surface.

Lower right: Contrast of reconstructed frequency pseudo-"sweeps" with the fitted surface.

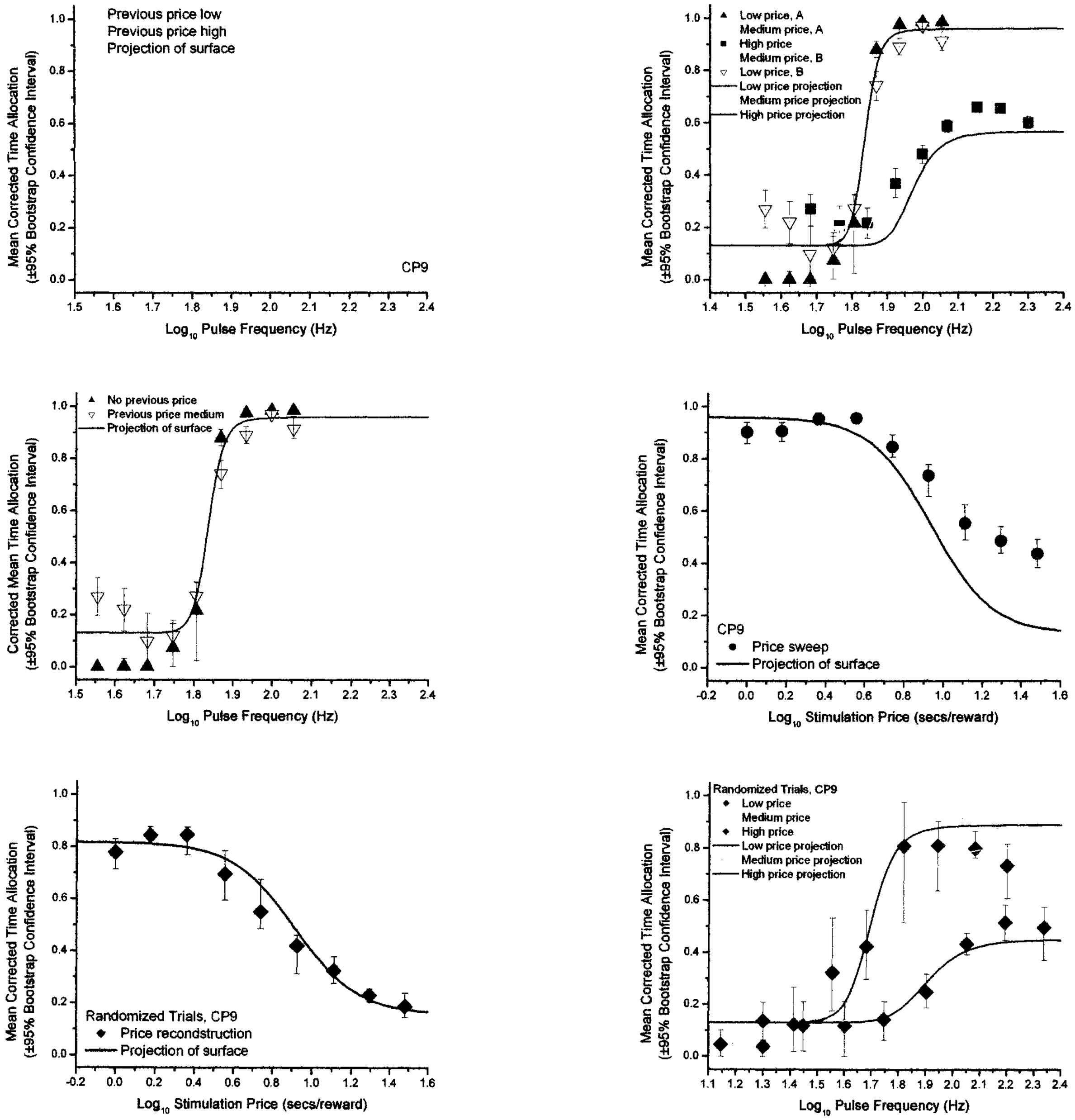


Figure 18. Observed mean corrected time allocation at each condition in the experiment contrasted with a projection of the fitted 3D surface, for rat CP9. Error bars represent 95% confidence intervals (derived by bootstrap-resampling) about the estimated mean.

Upper left: Contrast of the medium-price frequency "sweep" obtained when the previous frequency sweep was obtained at high or low price with the fitted surface.

Upper right: Contrast of all frequency "sweeps" with the fitted surface.

Middle left: Contrast of price "sweep" with the fitted surface.

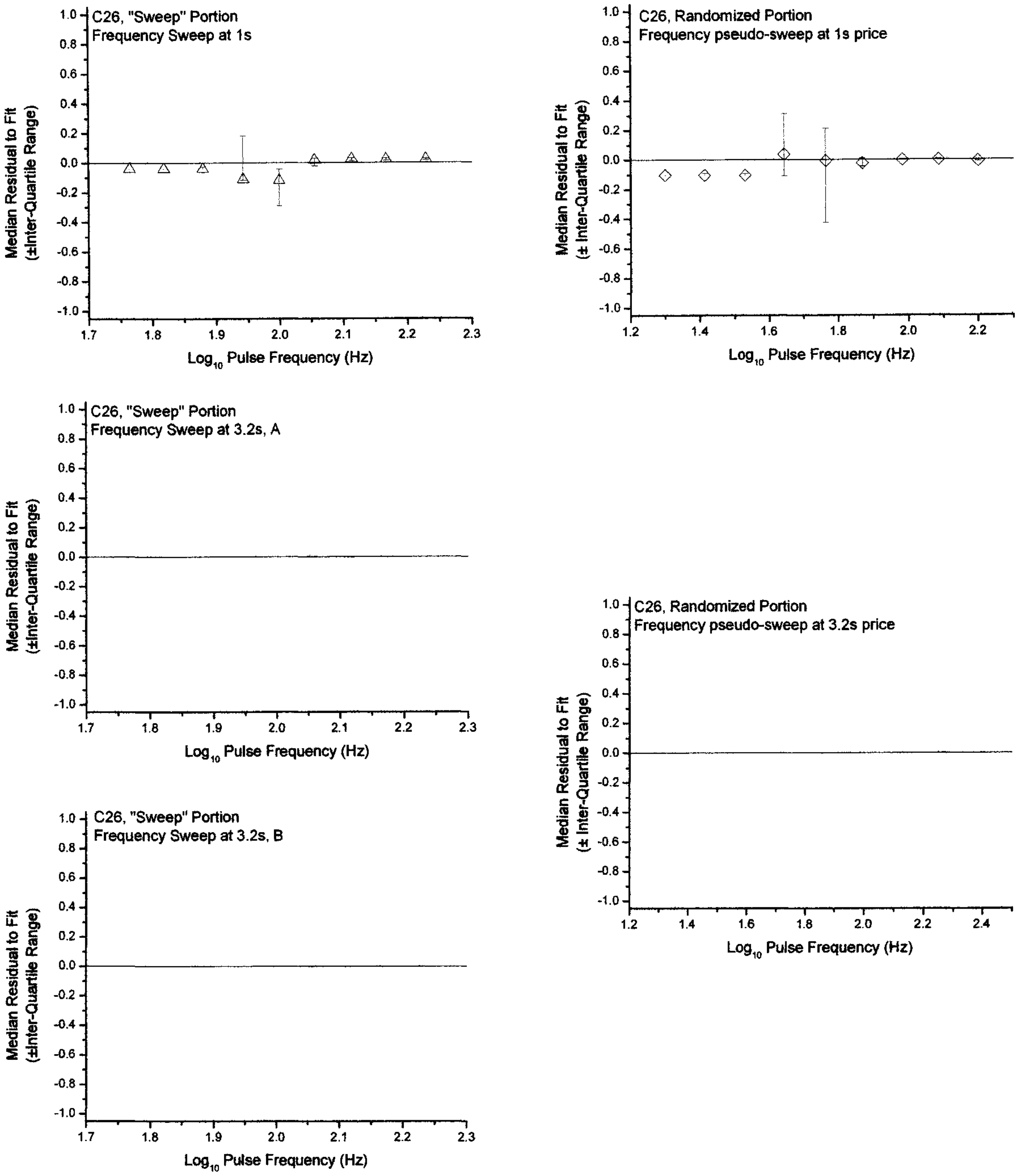
Lower left: Contrast of reconstructed price pseudo-"sweep" with the fitted surface.

Lower right: Contrast of reconstructed frequency pseudo-"sweeps" with the fitted surface.

triangles). Notice how, often, there is a shift between the curves obtained during these two phases of the sweep portion of the experiment. In the middle left (when no second low-priced frequency sweep was collected) or middle right (when such a condition was run) panels depict the price sweeps collected prior to the experiment. The curve often does not track the points very well, as a result of the compromise that is made by the fitter to reduce the residual sum of squares between the data points and the fitted surface. Since a compromise must be made, neither the projection of the surface to all the frequency sweep data (upper right panels) nor to the price sweep data provides a convincing fit overall. In other words, although the fit may have converged on a set of parameters that describes the data set, the pattern of responding as a whole does not appear to be well-captured by the surface. This is typical of a surface that does not provide a smooth, coherent fit to the underlying data. Projections of surfaces fitted to data from randomized conditions (bottom panels) are generally better-behaved: there does not appear to be a sweep-specific, systematic and reliable deviation of points from associated curves in any of the 5 rats considered.

It is true that the proportion of variance accounted for by fits to sweep condition surfaces ranges from 0.69 to 0.86, while that of fits to "random" condition surfaces ranges from 0.64 to 0.88. This may simply be due to the greater variability in responses when trials are presented in random order. The decreased variability of the animals' behaviour during the sweep portion of the experiment may overshadow the actual fit of the projected surface to the data points. Since variability of performance during sweep conditions is much less than in randomized conditions (compare the upper right panels of figures 12 through 18 to the bottom right panels), it would be disingenuous to conclude that the sweep conditions provide a better fit. Increased error variance, as is seen in the randomized portion of the experiment, tends to mask the decrease in systematic deviations from the surface. This section will therefore conclude with an analysis of the systematic deviations from the fitted surface. Indeed, the purpose of the 3D surface fit is to determine whether the general pattern of responding, as a whole, is consistent across different long-term average prices and session structures. If the general pattern of responding is not consistent, it will produce systematic deviations of the data to the fitted surface.

Figures 19 through 30 depict the median residual and inter-quartile range of the residuals of each data point to the surface fit (the differences between the data points and the fitted surface). On the left hand side, the median residual is plotted as a function of the frequencies or prices of the corresponding sweeps during the sweep portion of the experiment. On the right hand side, the median residual is plotted as a function of the frequencies or prices of the corresponding pseudo-sweeps constructed from the data obtained during the randomized portion of the experiment. When frequency sweeps were taken at the same price in two different phases of the sweep portion of the experiment, these plots are juxtaposed on top of one another. In most cases, the residuals from the sweep condition tend to fall on either side of the zero mark, indicating regions in which data points are systematically above or below the surface. Visual inspection of the residuals from the randomized condition shows, by and large, fewer regions indicating systematic deviation from the fitted surface. The median residual tends to be on or close to the zero mark (indicating a perfect regression). When the median residuals stray from zero, they tend to do so in a non-systematic fashion. Overall, there appears to be less



C26, "Sweep" Portion
Frequency Sweep at 1s

C26, Randomized Portion
Frequency pseudo-sweep at 1s price

C26, "Sweep" Portion
Frequency Sweep at 3.2s, A

C26, Randomized Portion
Frequency pseudo-sweep at 3.2s price

C26, "Sweep" Portion
Frequency Sweep at 3.2s, B

Figure 19. Contrast of residuals in "sweep" and randomized portions of the experiment for rat C26. Left panel contains median residuals at each frequency used in each frequency sweep condition; right panel contains the same for the randomized portion of the experiment. Error bars represent inter-quartile range.

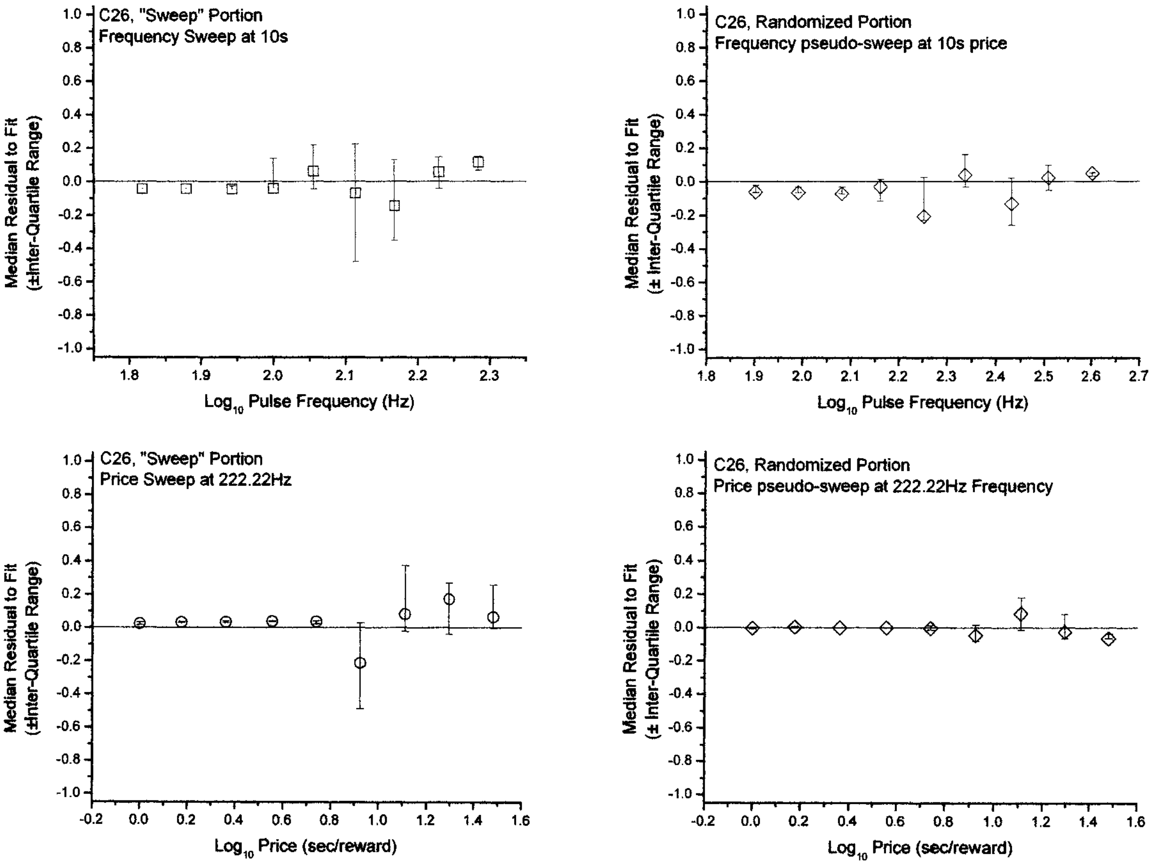


Figure 20. Contrast of residuals in "sweep" and randomized portions of the experiment for rat C26. Left panel contains median residuals at each price and frequency used in each price and frequency sweep condition; right panel contains the same for the randomized portion of the experiment. Error bars represent inter-quartile range.

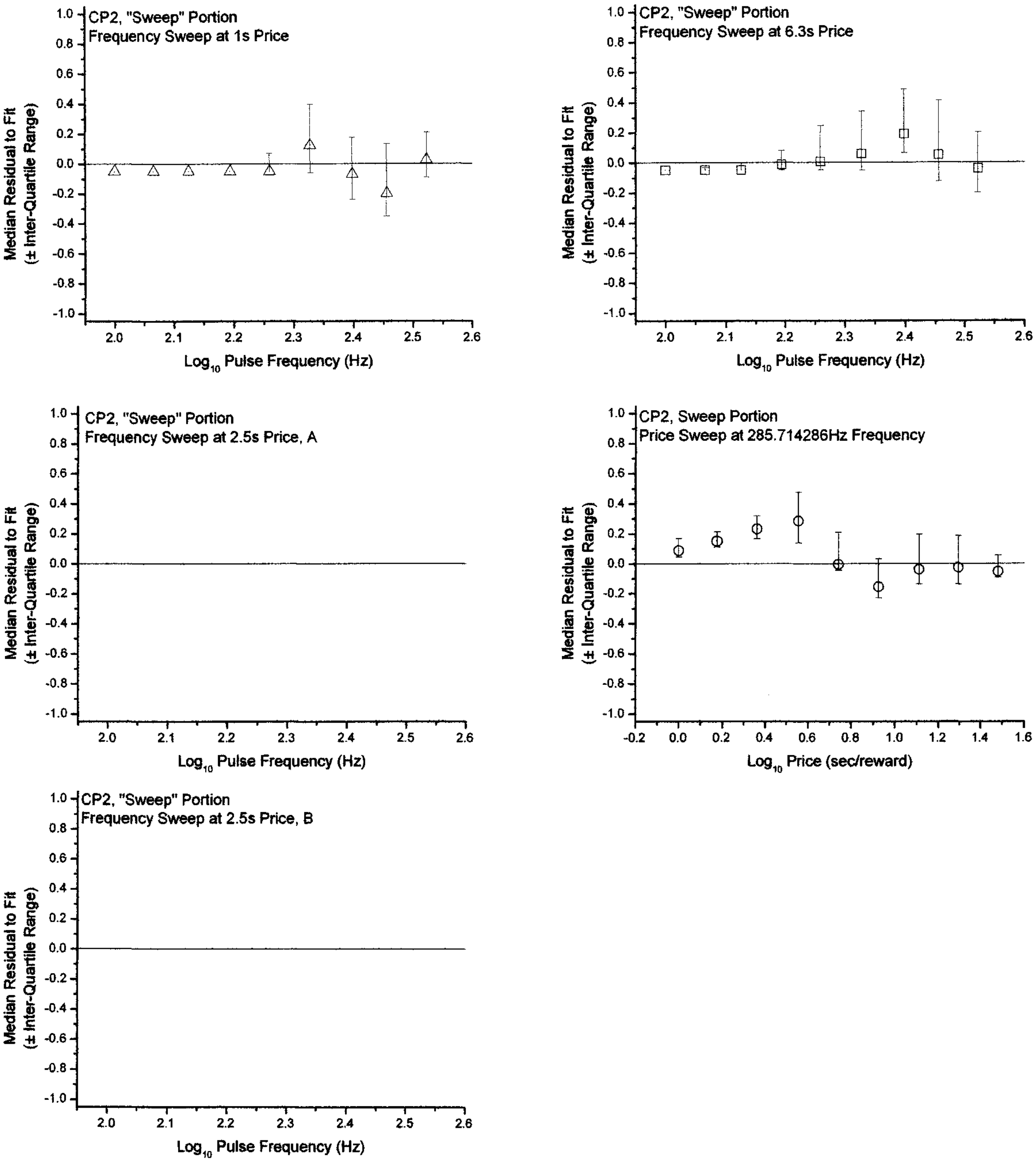


Figure 21. Median residual to fitted surface in in "sweep" portion of the experiment for rat CP2, at each price and frequency used in each price and frequency sweep condition. Error bars represent the inter-quartile range.

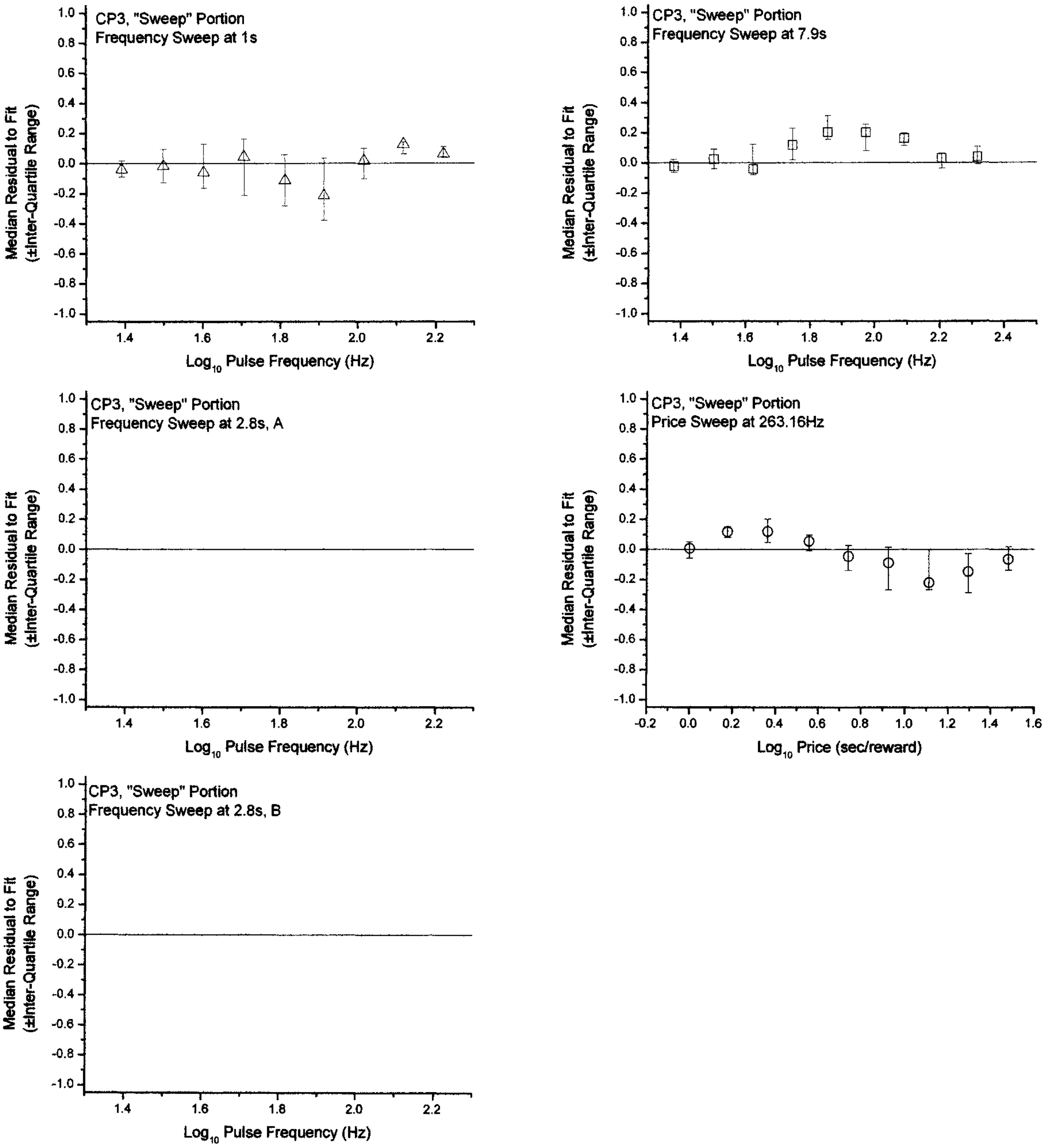


Figure 22. Median residual to fitted surface in in "sweep" portion of the experiment for rat CP2, at each price and frequency used in each price and frequency sweep condition. Error bars represent the inter-quartile range.

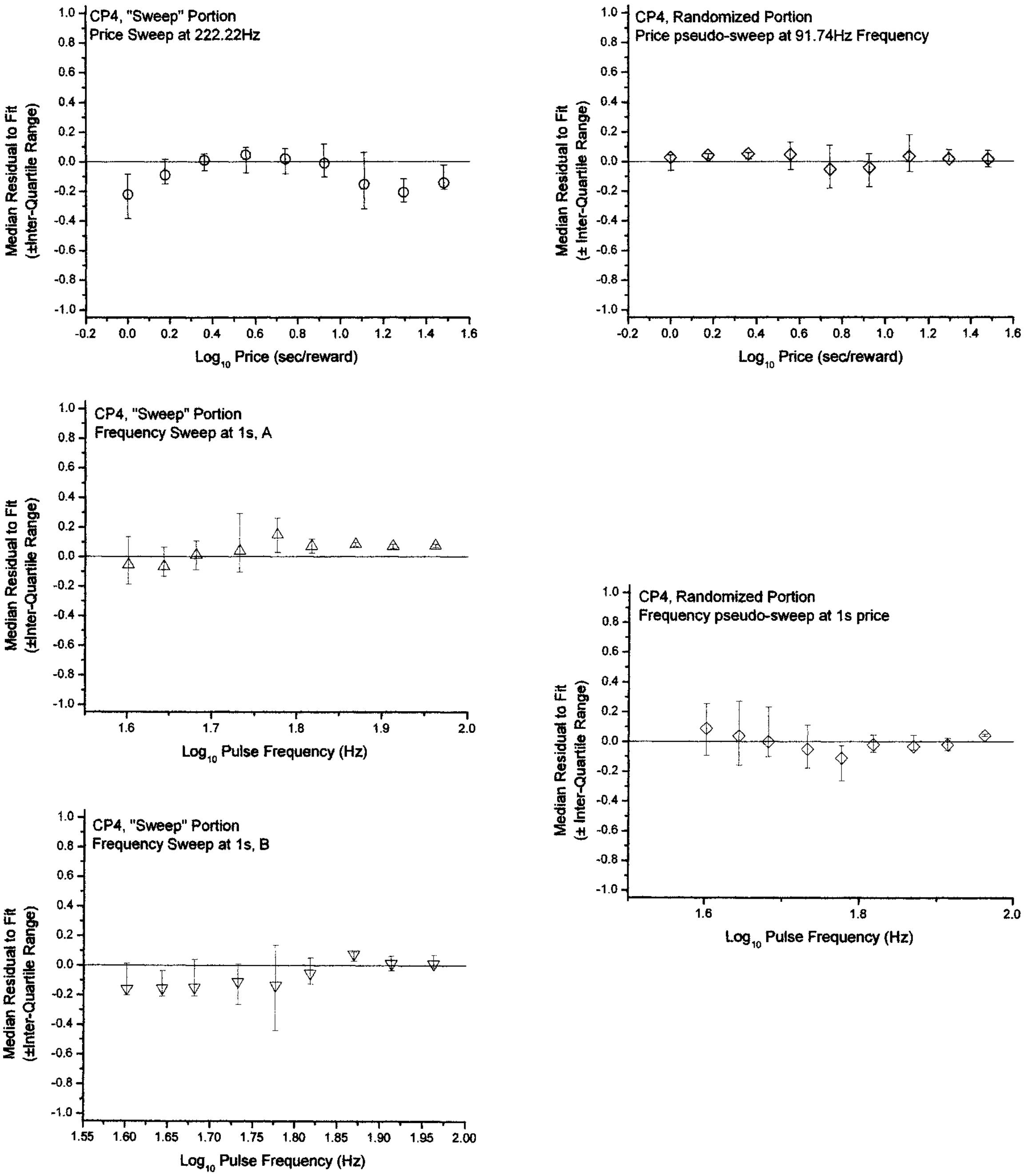


Figure 23. Contrast of residuals in "sweep" and randomized portions of the experiment for rat CP4. Left panel contains median residuals at each price and frequency used in each price and frequency sweep condition; right panel contains the same for the randomized portion of the experiment. Error bars represent inter-quartile range.

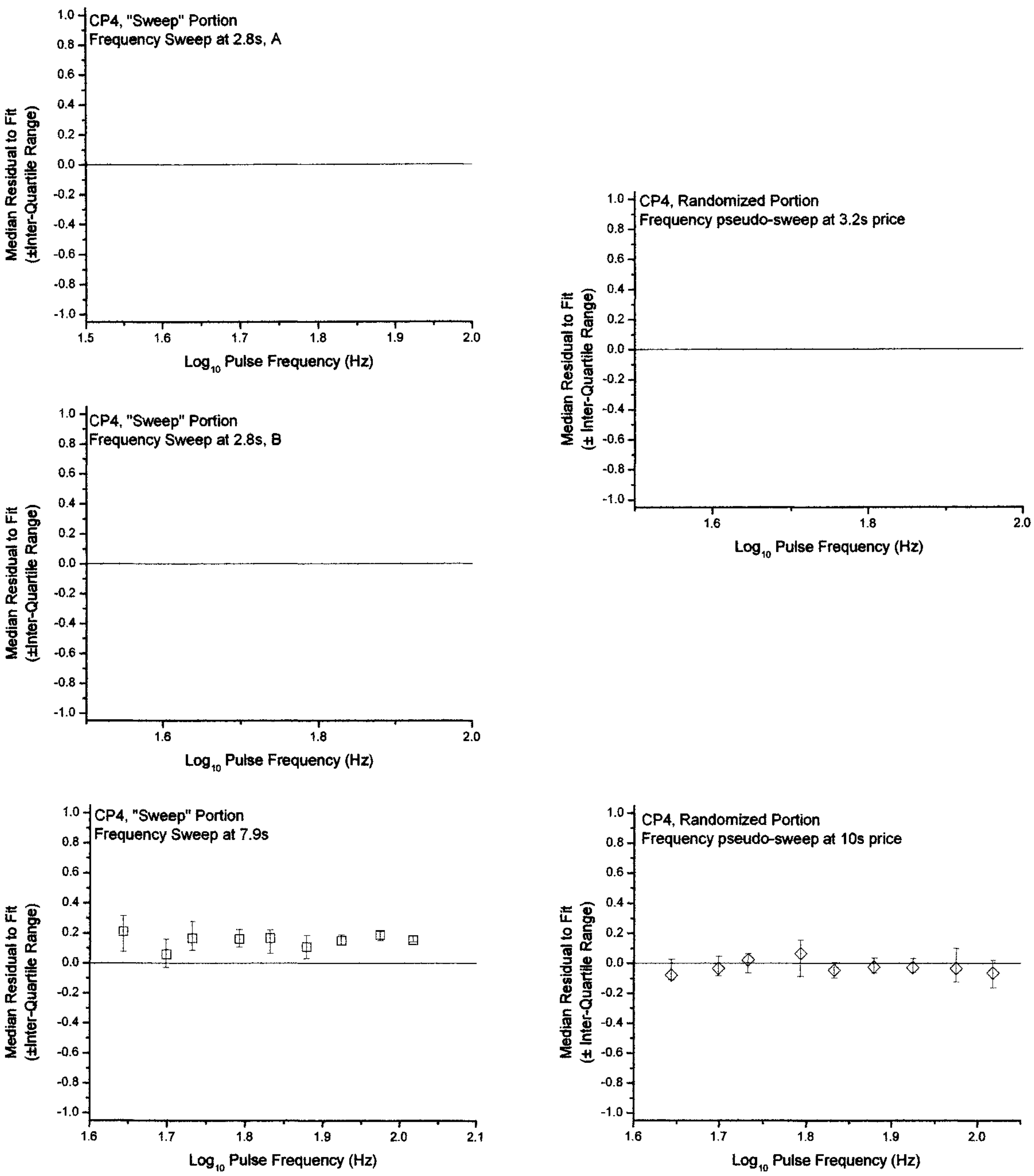


Figure 24. Contrast of residuals in "sweep" and randomized portions of the experiment for rat CP4. Left panel contains median residuals at each price and frequency used in each price and frequency sweep condition; right panel contains the same for the randomized portion of the experiment. Error bars represent inter-quartile range.

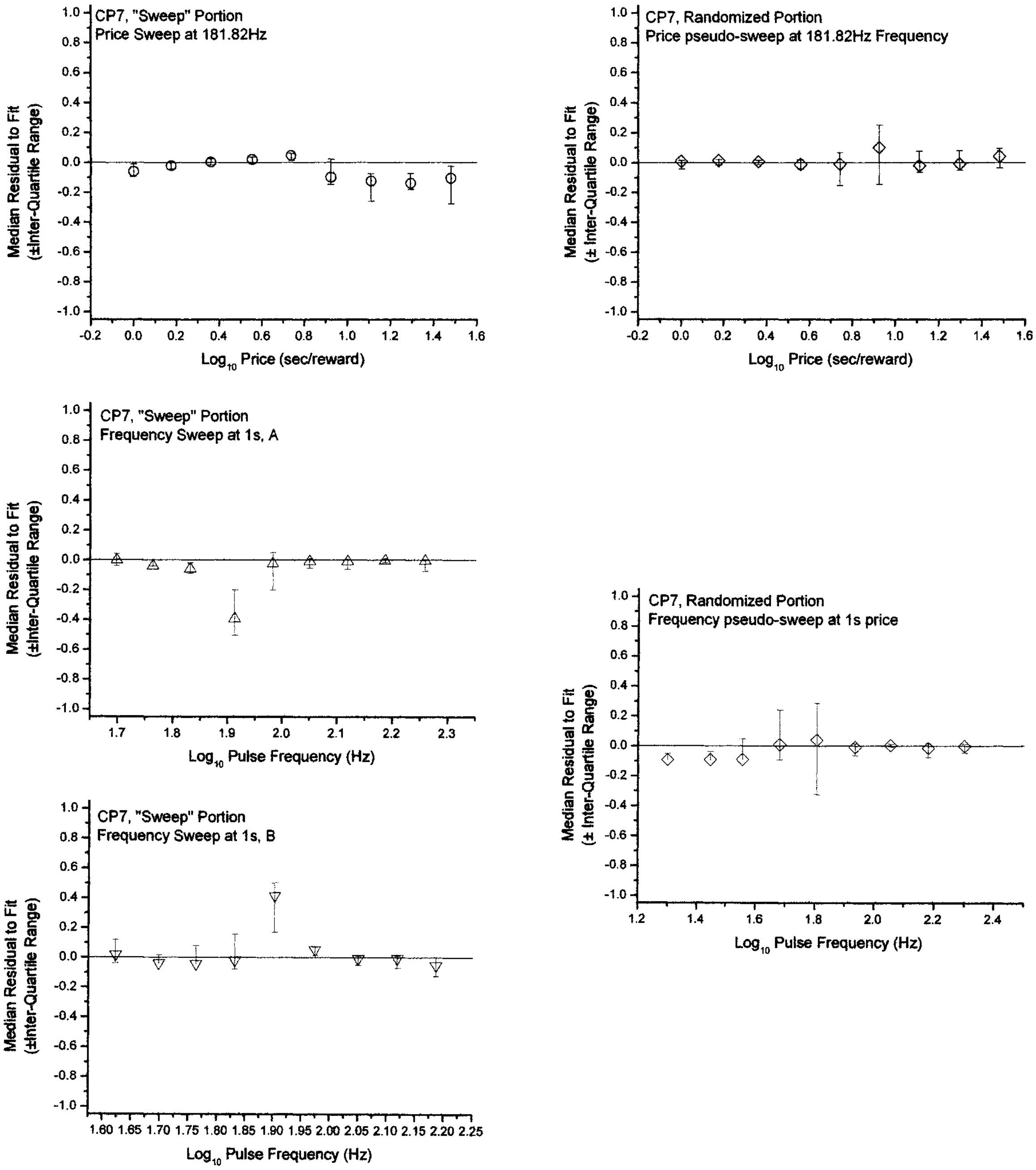


Figure 25. Contrast of residuals in "sweep" and randomized portions of the experiment for rat CP7. Left panel contains median residuals at each price and frequency used in each price and frequency sweep condition; right panel contains the same for the randomized portion of the experiment. Error bars represent inter-quartile range.

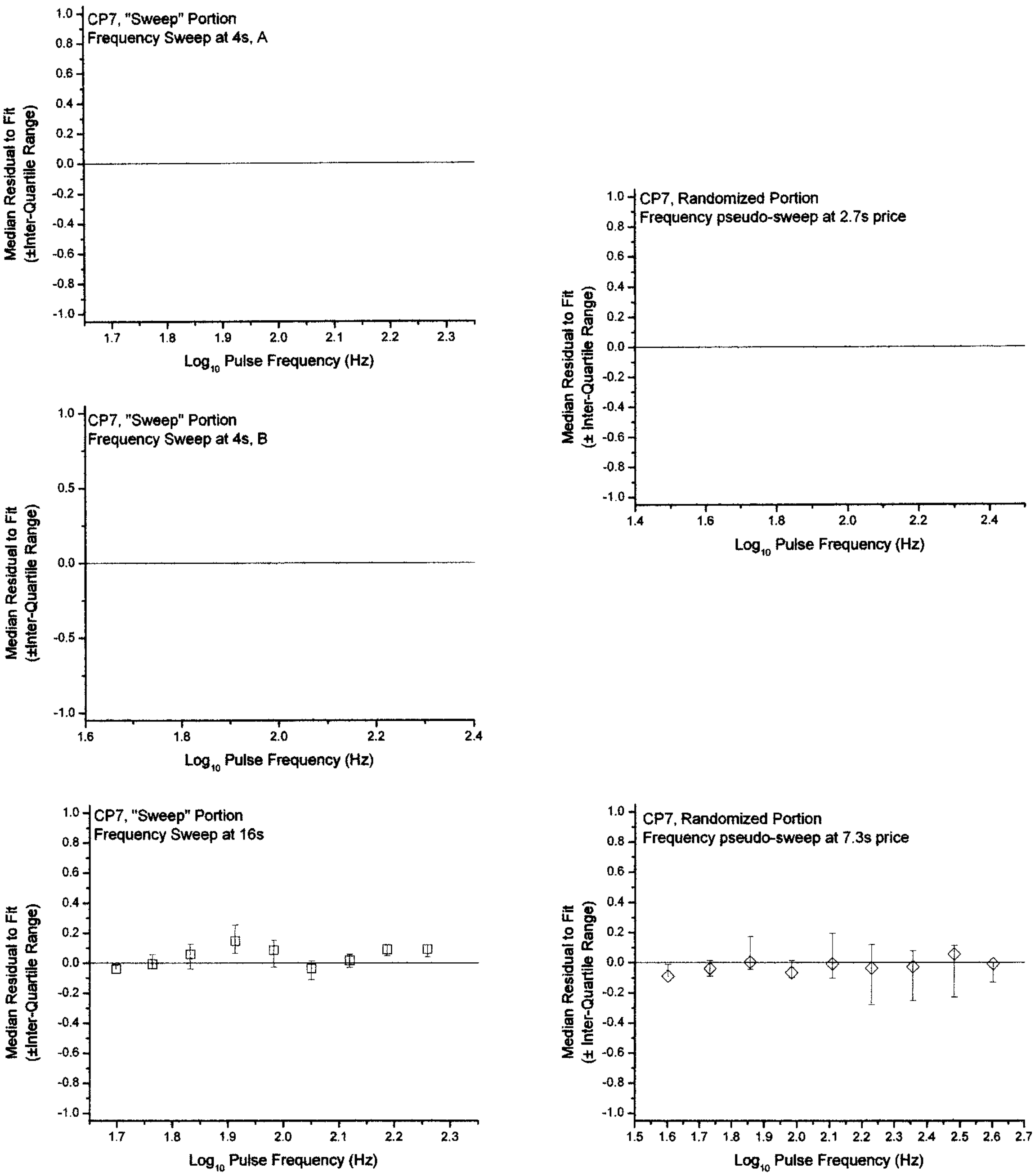


Figure 26. Contrast of residuals in "sweep" and randomized portions of the experiment for rat CP7. Left panel contains median residuals at each price and frequency used in each price and frequency sweep condition; right panel contains the same for the randomized portion of the experiment. Error bars represent inter-quartile range.

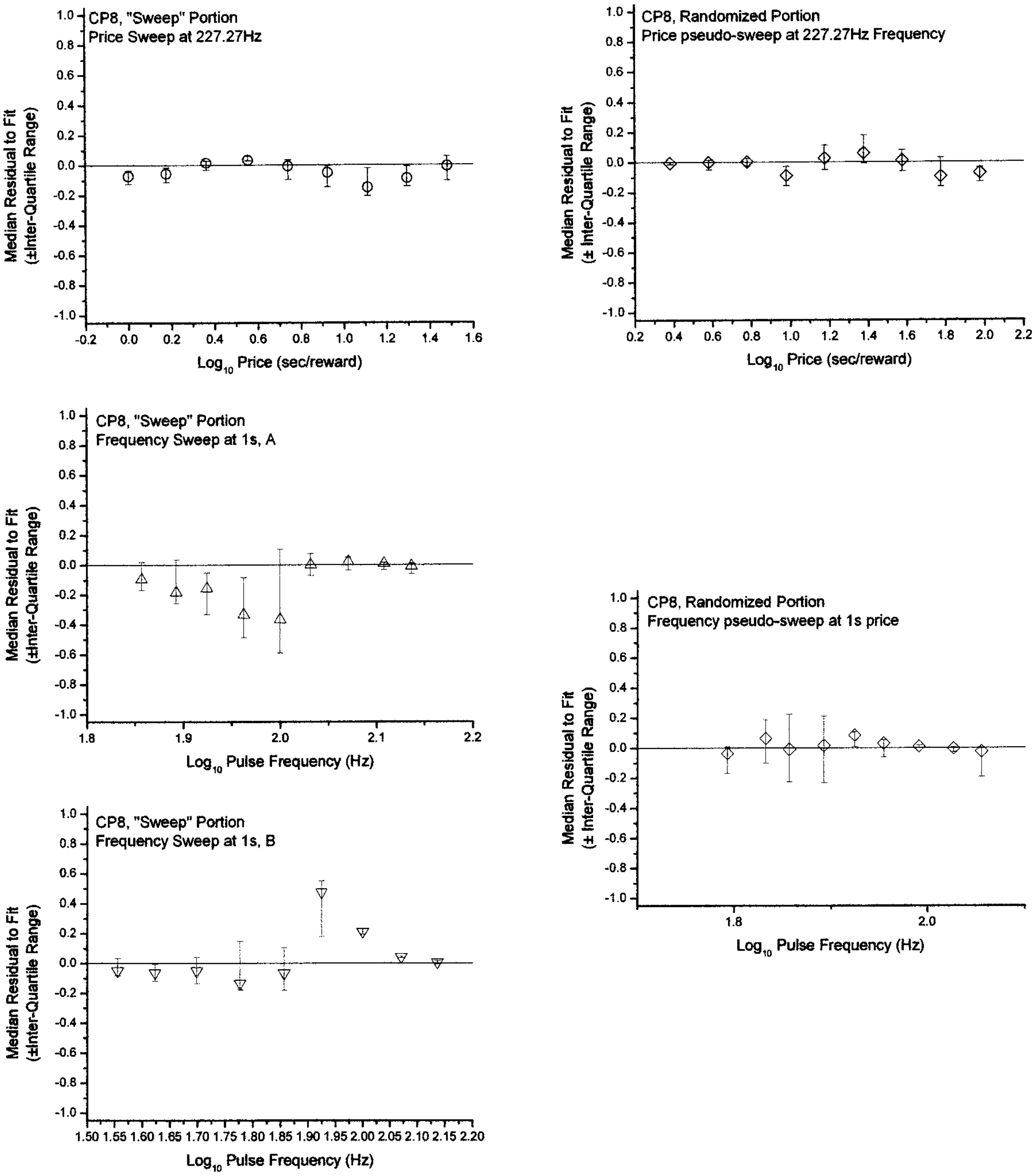


Figure 27. Contrast of residuals in "sweep" and randomized portions of the experiment for rat CP8. Left panel contains median residuals at each price and frequency used in each price and frequency sweep condition; right panel contains the same for the randomized portion of the experiment. Error bars represent inter-quartile range.

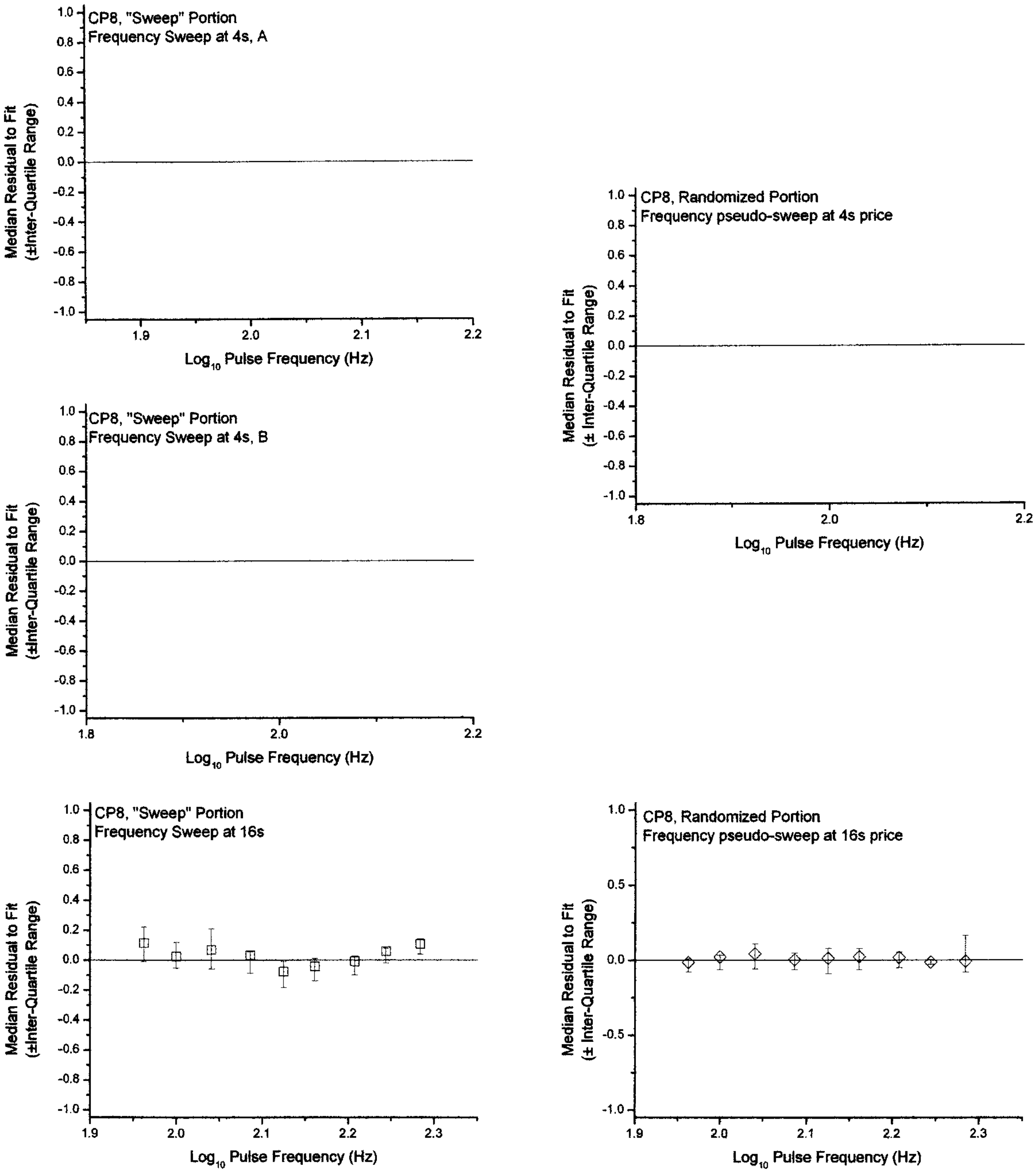


Figure 28. Contrast of residuals in "sweep" and randomized portions of the experiment for rat CP8. Left panel contains median residuals at each price and frequency used in each price and frequency sweep condition; right panel contains the same for the randomized portion of the experiment. Error bars represent inter-quartile range.

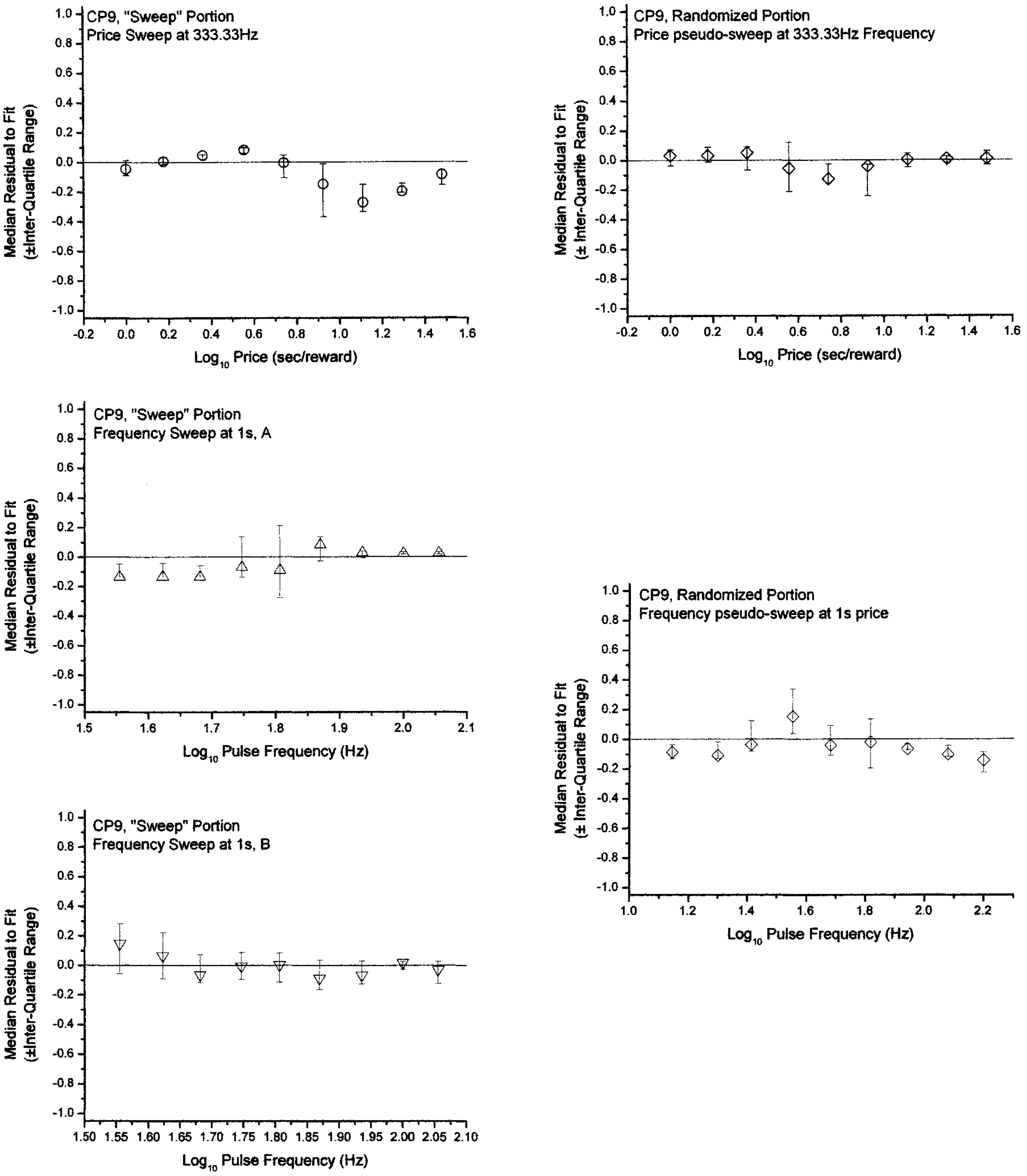


Figure 29. Contrast of residuals in "sweep" and randomized portions of the experiment for rat CP9. Left panel contains median residuals at each price and frequency used in each price and frequency sweep condition; right panel contains the same for the randomized portion of the experiment. Error bars represent inter-quartile range.

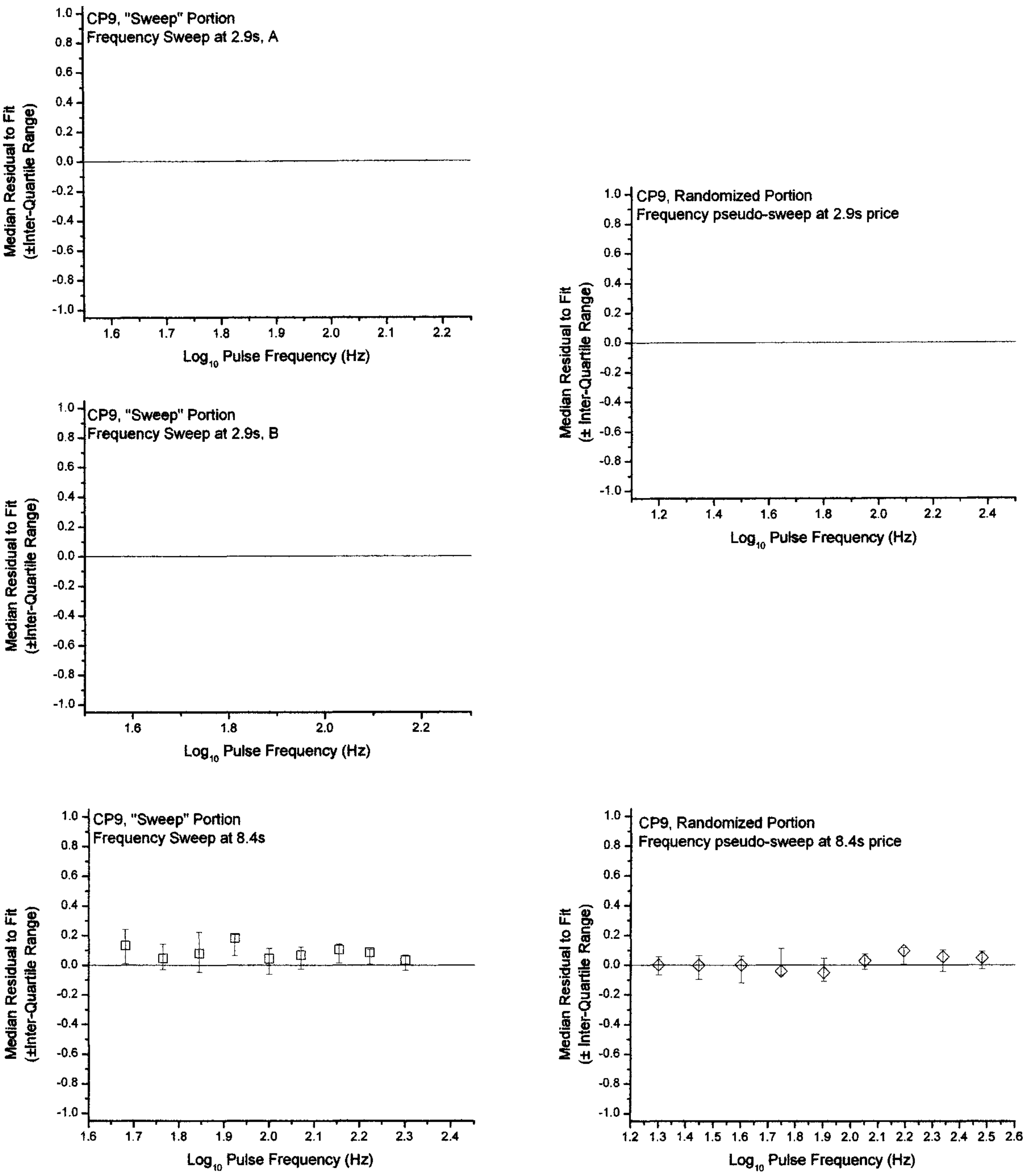


Figure 30. Contrast of residuals in “sweep” and randomized portions of the experiment for rat CP9. Left panel contains median residuals at each price and frequency used in each price and frequency sweep condition; right panel contains the same for the randomized portion of the experiment. Error bars represent inter-quartile range.

systematic deviation of the data from the fitted surface in the case of the randomized than the sweep condition.

Discussion

It is important to describe the relationship between time allocation, stimulation strength and price if we are to tie different manipulations, such as alterations in dopamine tone, to particular stages in the circuitry underlying the evaluation of payoff and the allocation of behaviour towards the procurement of rewards. Indeed, a good model of performance would incorporate the strength of the stimulation and its cost; without such a model, it would not be possible to determine how performance has changed as the result of a manipulation. The Mountain model is capable of providing such information. However, inconsistency problems have been encountered when attempting to validate its capacity to describe behaviour. In many cases, time allocation for a given price-frequency pair depends on whether that pair is approached from a frequency sweep or a price sweep (documented in this experiment). The position of frequency sweeps in many cases appears to depend on whether the previous frequency sweep was obtained at a high or low price (Marcus, 2005). Finally, in some cases, the locus of rise of frequency sweeps obtained at high prices does not appear shifted from that of a frequency sweep obtained at a low price (Mullett, 2005). These inconsistencies stand in the way of determining the relationship between performance, pulse frequency and price because they systematically distort the performance we are trying to measure. Eliminating these inconsistencies makes possible a clear picture of the relationship we are trying to describe.

The present study attempted to reliably reproduce context-dependent inconsistencies in time allocation and develop a data collection method to correct for them. Although a few experiments have demonstrated that anchor points of subjective scales can drift in rats (Breton, 2004; Marcus, 2005), this experiment provides both substantial evidence for abandoning the use of the method of descending limits as well as a remedy for the alterations in performance it produces. When animals are exposed to predictably high prices for a long period of time, there appears to be a kind of drift in their internal anchor for the scale of rewards that changes their performance for rewarding brain stimulation. This occurs when the price of rewarding brain stimulation is increased from a negligible level in frequency sweeps and when the average price is changed again during price sweeps.

The time the rat spends working for a stimulation train of a given strength and price varies as a function of whether this point in the parameter space (made up of the frequencies and prices an animal will encounter) is approached along a frequency sweep or a price sweep. When animals are exposed to high average prices over long periods of time, their behaviour shifts: it is as though the underlying process evaluating the payoff of stimulation activities changes through the very action of attempting to quantify it experimentally. In randomized versions of the experiment, the average price of the stimulation is constant over long periods of time, ostensibly making the animal's behaviour more consistent even if more variable.

The idea of randomizing trials is not new. Valenstein (1964, cited in Konkle, Bielajew & Fouriezos, 2001) first considered the possibility, but not in the context of simultaneously manipulating both the strength and price of reward. Clearly, a simple randomization of frequency-manipulated trials at a constant price would not solve the problem of internal consistency if the source of the inconsistencies was adaptation to a long-term price. Pilot work (Smith, 2005) attempting to solve the inconsistency problem without randomization has also failed. Smith (2005) showed that providing a stable anchor point by running a single trial of high-frequency, negligible-price stimulation at the beginning of each set of nine test trials does not appear to reduce inconsistency or produce smoother data, suggesting that the changes occur over long periods of time. The present thesis is the first study to reproduce the problem and offer a solution. By virtue of this, it is also apparent that either the highly-predictable nature of the sweep method, the animal's adjustment to non-negligible average prices, or some combination thereof, biases the animal's decision to press for rewarding brain stimulation.

This bias is all the more evident when considering the systematic shifts of data collected by orderly sweep data compared to random-order trial presentation. Performance for medium- or low-priced frequencies often depended on whether the animal had previously experienced subjectively high or low reward costs. This results in a poor surface fit because the common surface must accommodate two distinct populations of data values at a given point in the parameter space. Under sweep conditions, the data deviate systematically from a smooth fitted surface. Under randomized conditions there is more error variance but little systematic deviation from a fitted surface. If an animal's reinforcement history alters its performance on any given trial, then the phase-by-phase changes in average long-term price would produce the kinds of systematic deviations that were recorded in the sweep portion of the experiment. In contrast, the average long-term price the animal sees is stable when test trials are presented in random order. The animal is likely still basing its performance at least to some extent on its reinforcement history; nonetheless, the influence of high-priced trials is interspersed with the influence of low-priced trials. As a result, the effect of reinforcement history on performance becomes non-systematic, because the influence of past test trials on performance eventually cancels out.

The balance of the discussion of experiment I will focus on the implications of these results and the application of the ideas they may spawn. In particular, I will elaborate on how these findings establish that the preferences of non-human animals are constructed rather than revealed by the elicitation process. Then, I will propose one means by which the observed inconsistencies may come about. Following this, I will discuss some of the implications of the changing anchor for the scale of rewards for the human decision-making literature. I will then propose one means of determining whether differences in long-term average price or differences in predictability in the sweep conditions leads to internal inconsistencies. I will conclude with the implications of the results for the design of future experiments.

Construction of preference: rats do it, too

Decision theorists studying human beings have long known that we are not rational machines with fixed internal anchors. Multiple psychological processes intervene to alter how one will respond to a specific set of circumstances. Eliciting preferences in humans alters those preferences: the means by which we are asked to make a decision and the description of the alternatives both change what we ultimately choose. As such, human preferences are constructed rather than simply revealed by the elicitation process.

This experiment provides early evidence that the same is true in rats. The method of elicitation certainly influences the animal's bar-pressing performance. It is not, however, the only result suggesting construction of preference in rats. Sonnenschein (2004) found differences in performance under single- and dual-operant methods of elicitation, suggesting that alternatives presented one at a time differ from those presented in pairs with regard to self-stimulation performance.

That the inconsistency findings have never been documented or resolved in any systematic way is not surprising if we consider that rarely have experiments been designed to detect those inconsistencies. Previous experiments combining measurements of performance for stimulation of different pulse frequencies and prices (Mullett, 2005) used stable, predictable average prices for the frequency sweeps over long periods of time. A frequency sweep would be obtained at a low price for a long period of time. A price sweep would then be obtained at a different, higher average price for a long period of time. A final frequency sweep would then be obtained at a different, high average price for a long period of time. The results of this experiment suggest that this phase-by-phase experimental structure leads to inconsistencies between price and frequency sweeps that can be eliminated when test trials are presented in random order. When the payoff of self-stimulation is on average sufficiently low over long periods of time, the animal adjusts its behaviour, perhaps by changing the anchor for the scale of rewards or by changing its subjective evaluation of price. What is it about the pattern of responding for rewarding brain stimulation that makes the data set inconsistent? In other words, why will an animal spend a greater proportion of time bar-pressing for a trial part of a high-priced frequency sweep than it will for an equivalent trial that is part of a high-frequency price sweep?

The source of the inconsistency

In understanding the neural basis of decision-making, it would well be useful—if not crucial—to determine in which ways the sweep method biases performance for rewarding brain stimulation and at which stage of processing such a bias acts. These two points are both difficult problems to study, since previous work (Marcus, 2005; Breton, 2004) has shown multiple types of discrepancy and, more problematically, inconsistent discrepancies across animals. In particular, Marcus (2005) found that the position of frequency sweep data collected at a medium price sometimes changed depending on whether the frequency sweep in the previous experimental phase was taken at a low price or a high price. That is, a lower frequency was required to perform a criterion level of responding during the medium-priced frequency sweep when the rats had just experienced a high-priced frequency sweep. Despite this, many animals showed the effect only when the price of the frequency sweep was dropped again. Taken together, the results from Marcus (2005) suggest that although animals do not consistently show a single, unified data distortion, virtually all animals tested exhibit at least one type of inconsistency.

Breton (2004) also found more nuanced inconsistencies. Specifically, the point at which frequency sweeps collected at a high price begin to rise (which is often used as a measure of the stimulation threshold) does not appear to shift from that of frequency sweeps collected at a low price. Instead, with this locus of rise firmly attached to that of steeply-rising low-priced frequency sweeps, the curves appear to climb much more gradually, in some cases to a lower asymptote. Although the decreased slope and lower asymptote is predicted by the model under validation in experiment 2, the model also predicts that the position of the curve should shift: a greater number of action potentials should be required to compensate for the increased price of the stimulation. One solution to the distortion is to assume that animals cannot detect differences in prices below the highest price used in the Breton (2004) experiment. In this case, animals would be indifferent to a two-fold difference in price between 16 seconds and 8 seconds. If these particular animals were indifferent to changes in price below 16 seconds, their price sweep curves would be asymptotic up to 16 second prices. Breton (2004) clearly showed this not to be the case.

The problem is striking when revisiting the results of an experiment by Mullett (2005). The loci of rise of the frequency sweeps taken at a high price were not considerably displaced from those taken at a low price for some subjects. Indeed, these high-priced frequency sweeps rose more gradually than low-priced sweeps but started from a similar locus of rise. In part to account for this, the author fit a version of the model under validation in experiment 2 that considers that there may be some minimum subjective price involved in stopping leisure activities to bar-press. It is ridiculous to believe that animals would be as sensitive to a two-fold change in price from 0.01 to 0.02 seconds as they would be to a two-fold change from 20 to 40. Surfaces could hence be fit reasonably well to this pattern of gradual rise from an identical locus if a parameter existed that specified a price below which animals became insensitive to further price decreases. The estimated value of this parameter was approximately 11 seconds. Such a large estimate of the minimum subjective price is necessary if the locus of rise of high-priced frequency sweep curves is not significantly shifted from low-priced frequency sweep data. The smaller the shift in frequency sweeps obtained at different prices, the more similar their subjective interpretation must be. If the pattern of responding for high-priced stimulation is almost identical to the pattern of responding for low-priced stimulation, due to the fixed locus of rise, the best-fitting model will incorporate a minimum subjective price that is almost or just as large as the price used for the high-priced frequency sweep. Although surfaces could be made to fit the data extremely well, the derived values for the minimum subjective price are suspect; instead of reflecting the curvature of the subjective-price function, the fixed locus of rise could well have resulted from the shift in anchor point suggested by the results of the present study.

Rat cognition?

The results of the present experiment raise an important issue. An assumption of the Matching Law is that the testing situation merely reveals underlying fixed criteria for evaluating reward intensities and costs. I demonstrate here that instead of rigid, fixed criteria, animals may employ flexible heuristics that reflect their expectations of future payoff based on previous experience. It is possible that the scale of rewards is shifted by long-term adaptations in a way that is similar to the visual system shifting its light-sensitivity as a function of ambient luminance. In other words, lean rates of reinforcement over long periods of time may shift the reward-sensitivity curve.

In this case, it is not surprising that randomizing the order in which test trials are presented to the animals would reduce or eliminate the adaptation. Once the animal has completed relatively few determinations in a randomized presentation, the average price of the stimulation stabilizes. There is no way for a cognitive anchor of the scale for reward evaluation to change systematically from experiment phase-to-experiment phase. Since the animal simply can not know whether a given point is part of a frequency sweep or a price sweep, its behaviour in the randomized condition is a much more reliable estimate of an "ideal" revelation of preference than traditional sweep methods.

That the randomization process restores the internal consistency of most data sets considered is telling of the underlying process by which animals evaluate and decide. The results of the present experiment suggest that construction of preference in this context need not occur in verbal animals and be directly related to linguistic limitations. Indeed, the proportion of brain matter devoted to frontal cortex—the presumed seat of rational deliberation—is quite small in the rat. Nevertheless, it appears that rats' preferences are experientially constructed by the testing situation, not merely revealed. It is all the more striking that evidence of a decision anomaly such as sensitivity to a long-term average price can be obtained in a non-verbal subject.

It is possible that it is merely the predictable structure of the experimental sessions that contributes to an animal's construction of preference. If this were the case, animals would perform consistently for randomized price-frequency pairs presented in two phases of an experiment in which the long-term average price is changed in each phase. If the inconsistency effects were due to the long-term average price, making that price different in two separate phases of a randomized experiment would restore the inconsistencies. The test of such a hypothesis is simple. Over multiple sessions, an animal could be run in a frequency sweep at a low price, randomized with a price sweep. Then, conditions would change for another long period of time, and the price of the randomized frequency sweep would increase. If the animal were simply reacting to the predictability of the sequence, all datasets in all conditions would be expected to be consistent with one another. That is, if the predictability of the price and frequency sweeps were the cause of the inconsistency, it would not matter whether you randomized a frequency sweep taken at a low price or one taken at a high price with a price sweep. All four would be expected to be consistent with one another. If an anchor for the scale of reward evaluation were being pulled by the long-term rate of reinforcement, the frequency sweeps would not be consistent with price sweeps collected under different conditions. In other words, if the long-term average price was the cause of the inconsistency, a frequency sweep taken at a low price and randomized with a price sweep would be consistent with that price sweep but inconsistent with a subsequent price sweep randomized with a frequency sweep taken at a high price.

The measurement problem

The present experiment reproduces the most consistent discrepancy observed in previous studies and probes the usefulness of a method that would reduce incongruities. In a majority of cases, the act of assessing an animal's performance for rewarding brain stimulation alters the performance the investigators are interested in measuring. This has serious implications for the design of future experiments that assume a simple multiplicative combination of price and reward intensity. If animals experience long periods of average high-priced stimulation in a predictable session structure, it is likely the scale of rewards will have changed from its state when stimulation could be harvested at a negligible price.

As such, previous experiments in which both the pulse frequency and price of stimulation were manipulated are tainted by the highly-reproducible internal inconsistencies. These inconsistencies are the product of an as-yet unknown process of adaptation that the animal undergoes resulting either from highly-predictable session structures, high average prices over long periods of time, or both. Previous findings, however revealing they may be, must be re-evaluated in order for the validation of a model of self-stimulation performance to be truly meaningful and acceptable. Experiment 2 therefore provides a validation of a model of performance for rewarding brain stimulation using a randomized-trials approach to reduce potential systematic biases caused by long-term adaptations.

EXPERIMENT 2

In the early years following the discovery of brain stimulation reward following the initial report by Olds & Milner (1954), the effect of alterations to reward-relevant parts of the brain was inferred from changes in the rate of responding for highly rewarding stimulation on a continuous reinforcement schedule. If an animal stopped pressing following a lesion to a part of the brain, it was inferred that the lesion disrupted the rewarding properties of the brain stimulation. Criticisms of this method were not widely acknowledged until Valenstein & Hodos (1962) raised their concerns with the numerous problems of the maximum rate of response measure. Performance variables in this unidimensional view were inevitably confounded with reward-relevant variables. There was no way of knowing whether the animal was simply tired faster, whether forced movements, tremors and seizures were disrupting self-stimulation, whether the animal perceived pursuing the stimulation to be more subjectively costly, or the subjective reward intensity had decreased.

Such criticisms led to the curve-shift method for linking the effect of alterations to neural mechanisms of reward. The method was essentially borrowed from visual psychophysics; better known as the method of limits. An individual's lower threshold for pitch perception, for example, can be inferred by decreasing the frequency of a stimulus in logarithmic steps until the subject no longer reliably responds that they hear the tone (Titchener, 1905, cited in Kling & Riggs, 1972). If one were to hypothetically lesion a certain portion of their brain and this threshold increased, the inference would be that the lesion has impaired the subject's capacity for lower pitch detection. A higher pitch would be required for the subject to detect it. A similar line of thinking has been the dominant view in understanding brain reward manipulations. A shift in the two-dimensional curve relating the rate of responding to the intensity of the stimulation (either the current, specifying the extent of excitation, or the pulse frequency, specifying the number of action potentials induced per activated neuron) has traditionally been interpreted as a change in reward value of the stimulation. It is argued if the subject requires more stimulation to perform a given criterion level of responding, the manipulation has rendered the stimulation less rewarding. If the maximum rate of responding has substantially decreased, the reasoning goes, then the manipulation has altered performance without altering reward value.

The number of scholarly articles claiming to have dissociated performance and reward effects on the basis of two-dimensional psychometric functions is large. Franklin (1978) claimed to have demonstrated dissociation between reward and performance in evaluating the role of catecholamines in brain stimulation reward. This experiment used a version of Gallistel's runway methodology in which the speed at which an animal runs to a goal box for an electrical brain stimulation is varied independently of the constant priming stimulation the animal receives in its start box. Such a procedure produces a threshold curve very much like a rate-frequency curve: the animal's running speed increases rapidly to asymptote as the goal box stimulation frequency is increased. Changes in the maximum speed are interpreted as changes in the rat's performance capacity, whereas alterations in the locus of rise of the curve (the threshold) are interpreted as changes to the reward value of the stimulation.

A few dissenters have shown this to be a false assumption. Fouriezos, Bielajew & Pagotto (1990) showed that increasing the effort required to respond by weighting the lever resulted in a rightward shift of the rate-frequency curve. It is unlikely that adding weight to the bar has somehow changed the reward intensity of the electrical stimulation delivered to the animal's brain. Yet, the uninformed conclusion of an individual blind to the nature of the manipulation using an orthodox interpretation of curve-shift data would be that it has affected how the stimulation translates into reward. Clearly, it has not. This finding alone would appear to contradict the claim that the two-dimensional rate-frequency or rate-intensity curve can disambiguate changes in reward intensity (by changes in threshold) from changes in performance variables (by changes in maximum response rate).

Other individuals rely on the progressive-ratio schedule of reinforcement to either preclude or implicate a manipulation's effect on an animal's willingness to respond. The breakpoint is simply a different sort of threshold; one at which the animal is no longer willing to respond for stimulation. Changes in this breakpoint are traditionally interpreted as changes in the animal's willingness to respond for stimulation. For example, Salamone, Correa, Mingote & Weber (2005) have implicated dopaminergic tone in the cognitive mobilization of effortful responding by inference from changes in breakpoint. Animals with damage to their tegmental dopamine neurons by 6-hydroxydopamine lesions will respond fewer times than sham controls when the fixed ratio is high, but do not appear to differ in their responding when the fixed ratio is low. The inference is therefore that the lesioned animal is less willing to respond for food, despite the food's high reward value. An animal that has undergone such destruction of dopamine neurons will still consume large amounts of food but will not make a large number of responses in order to obtain it (Salamone, Wisniecki, Carlson & Correa, 2001).

These analytical methods are intrinsically flawed, because they confound the generation of reward value from the stimulation with the generation of behaviour from the reward value. Only a model considering at the same time the strength of the stimulation and the price of acquiring it can disambiguate those two processes. A shift in either the threshold of a rate-frequency curve or breakpoint of a progressive-ratio schedule can be produced by changing the way the stimulation translates into subjective reward value, or by changing the animal's willingness to work for a given reward value. The psychological process that determines how much time an animal will spend working for a reward presumably bases its decision on the intensity of the rewarding effect and the subjective costs the animal will incur. Manipulations to the brain may alter any of the variables that are involved in computing the payoff of pursuing a given alternative. Figures 31 and 32 (adapted from Shizgal, 2004) depict why it can not be said that a change in either the frequency threshold or progressive ratio breakpoint is a transparent manifestation of the psychological process that has been altered. The rate-frequency curve is equivalent to the silhouette of a mountain along the frequency axis taken at a low, negligible price. It is depicted in the upper left panel of figure 31. A progressive ratio measure is equivalent to the silhouette of a mountain along the price axis taken at a very high frequency. An example of this is depicted in the upper left panel of figure 32. Baseline silhouettes are plotted in these upper panels as cyan lines. If the mountain is shifted along the frequency axis—that is, a manipulation alters the reward circuitry before the output of the integrator—the projection of the mountain along the frequency axis will also displace (magenta lines). However, it is possible to produce an equivalent shift by moving the mountain along the price axis (black dotted lines); in other words, a

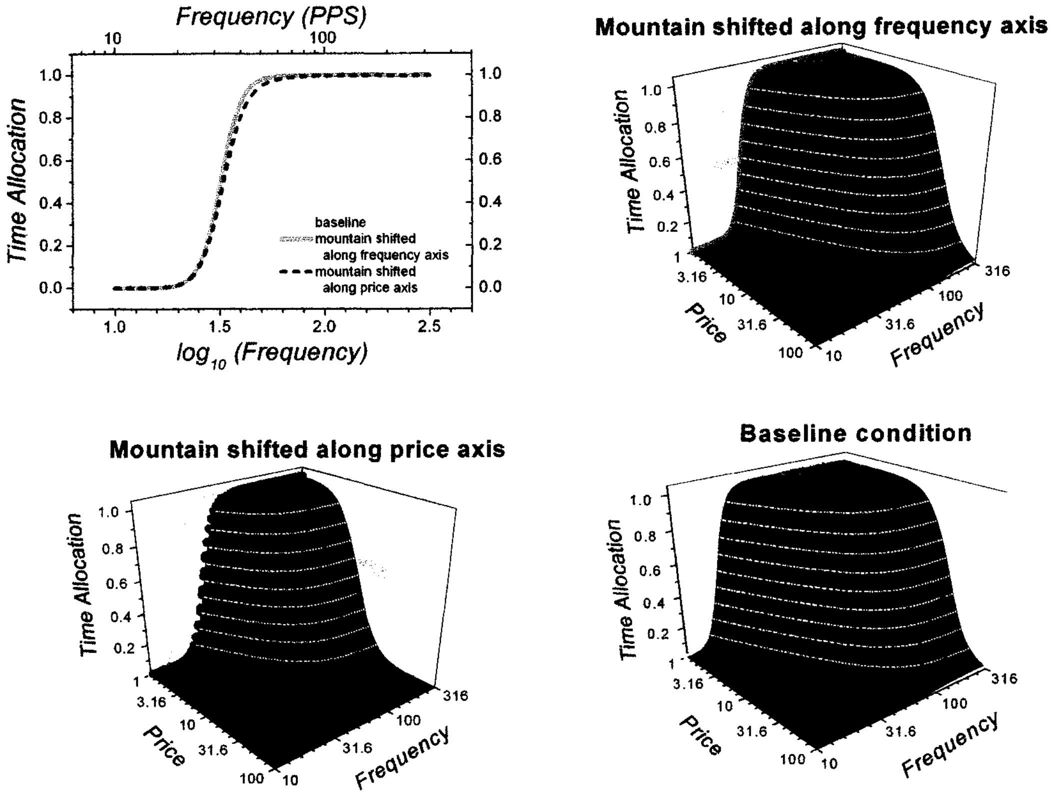


Figure 31. Manipulations altering the reward circuitry do not transparently manifest themselves in rate-frequency measures of performance. Rate-frequency measures of performance can be thought of as a series of frequencies obtained at a negligible price, and is thus analogous to a silhouette of this model of performance along the frequency axis at a low price. A manipulation that changes an animal's willingness to work for a given reward intensity can shift the rate-frequency curve in a way that is no different in two dimensions (upper left) from a anipulation that changes the reward produced by brain stimulation. (Taken from Shizgal, 2004)

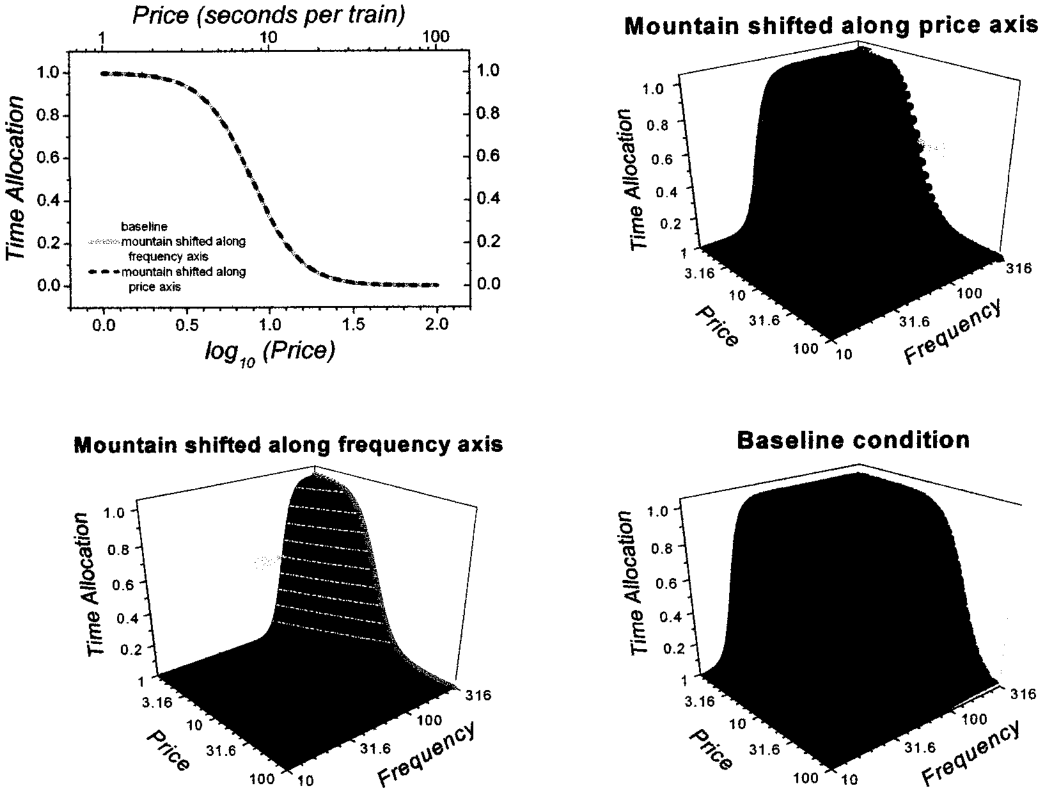


Figure 32. Manipulations altering the reward circuitry do not transparently manifest themselves in rate-frequency measures of performance. Rate-frequency measures of performance can be thought of as a series of frequencies obtained at a negligible price, and is thus analogous to a silhouette of this model of performance along the frequency axis at a low price. A manipulation that changes an animal's willingness to work for a given reward intensity can shift the rate-frequency curve in a way that is no different in two dimensions (upper left) from a manipulation that changes the reward produced by brain stimulation. (Taken from Shizgal, 2004)

manipulation altering reward processing beyond the output of the integrator can produce a shift in the two-dimensional rate-frequency representation. The same is true for the interpretation of the progressive ratio breakpoint, depicted in figure 32. Since the animal's behaviour, the pulse frequency of the stimulation and the schedule of reinforcement are the only decision-relevant variables available to the experimenter, behaviour can be thought of as resulting from an outer function within which is embedded an inner function: time allocation is the output of the outer function and hence, of the compound as well. The inner, reward-growth function (equation 3) specifies how

$$I = \frac{f^g}{f^g + {f_{hm}}^g} \tag{3}$$

where I is the intensity of the rewarding effect,

f is the pulse frequency of the stimulation,

f_{hm} is the frequency yielding half-maximal reward intensity, and

g specifies the rate at which reward intensity grows with frequency.

the stimulation gets translated into subjective reward value. The outer, behavioural-allocation function specifies how the reward value is mapped into behaviour by combining it with the price of self-stimulation. Whereas two-dimensional psychometric functions confound these two functions, a three-dimensional function can tease the two apart.

Building the mountain

When a rat harvests a brain stimulation reward, pulses delivered through the tip of the implanted electrode elicit a volley of action potentials in the underlying, reward-relevant substrate. This volley travels down the relevant axons to their terminals and across an unknown number of synaptic relays until its post-synaptic effects are spatio-temporally integrated. The integrator produces a pooled, subjective reward signal based on the total number of action potentials incident within a time window. The relationship between the frequency of stimulation and the subjective reward value it elicits was described by Gallistel using dual-operant methods as a power function when the frequency is low, and reaching asymptote at higher frequencies. Such a relationship can be approximated by a logistic of form $f^g/(f^g+f_{hm}{}^g)$, where f is the pulse frequency of the stimulation, f_{hm} is the frequency producing a reward of half maximal value and g is the rate at which reward grows with pulse frequency (Sonnenschein, Conover & Shizgal, 2003). This entire relationship is scaled by the maximum value of brain stimulation reward.

Once the stimulation has been transformed into a subjective reward value at some location in the brain, it is likely the animal's decision to respond depends on the multiplicative combination of that subjective reward value with the price of the reward. Prior work by Leon & Gallistel (1998) has provided evidence for the multiplicative combination of reward intensity and rate of reinforcement using assumptions of Herrnstein's Matching Law. Having elucidated the pulse frequency-subjective intensity relationship for each rat, the time allocation ratio of responding to each of two differentially priced levers was shown to be approximately the multiplicative combination of the subjective intensity of the stimulation with the rate of reinforcement. If the animal spends all its time working for the experimenter-controlled reward, price as is used in this thesis, is the inverse of the reward rate.

The subjective price of self-stimulation depends on both how much time must be diverted away from activities that are also rewarding such as grooming, exploring and sleeping. Free-running variable interval and fixed cumulative handling time schedules (described later) provide control over the opportunity costs to the rat, whereas ratio schedules provide control over the effort costs. The amount of time the animal is forced to spend holding a bar in order to harvest a reward is deemed the price of the stimulation. One can think of this time as the objective opportunity cost that is translated into the subjective price of the reward.

At any given moment, the animal's decision is dichotomous: either work for stimulation or engage in other possible activities in the operant box. If the rewards derived from self-stimulation were a perfect substitute for the rewards derived from all other activities, animals would work continuously for brain stimulation until the price was sufficiently high. At that point, animals would simply switch, engaging exclusively in leisure activities. All the data that have been collected so far demonstrates that the change from engaging exclusively in self-stimulation to engaging exclusively in leisure activities is gradual, suggesting the rewards derived from self-stimulation act as only partial economic substitutes for those derived from grooming, exploring and resting.

It stands to reason that if the payoff—the scalar combination of price and reward value—resulting from brain stimulation exceeds the payoff resulting from all other activities, the animal will prefer to self-stimulate rather than engage in unrelated activities. According to Herrnstein's Matching Law, an animal will allocate a proportion of its responses to an activity according to the proportion of the total payoff it represents. In mathematical terms, the strict single-operant version states that the response rate will be proportional to the ratio of the payoff it receives from responding to the sum of the payoff it receives from responding and the payoffs it receieves from all other activities. Herrnstein's version of the single-operant matching law was formulated to describe responding on variable interval schedules; the version presented below is intended to describe performance schedules on which payoff is directly proportional to time allocation. Since the rewards of self-stimulation act as partial substitutes for the rewards derived from alternate activities, a non-unity exponent has been added to the expression to account for this. The resulting expression that describes how an animal allocates its time to self-stimulation is $TA = U_{bsr}^{a}/(U_{bsr}^{a}+U_{ee}^{a})$, where TA is the proportion of time allocated to self-stimulation, U_{bsr} is the payoff from brain stimulation rewards, U_{ee} is the payoff from everything else, and a is an exponent reflecting the degree to which the two substitute.

If we know that U_{bsr} depends on the experimenter-set opportunity cost and the output of the reward growth function, there must be some way of quantifying the payoff of everything else. When the reward circuitry is saturated with action potentials and the integrator outputs a maximum reward value, increases in price will lead to drops in performance that cannot be compensated for by increases in pulse frequency. When this is true, the animal will allocate half of its time to self-stimulation when the payoffs from self-stimulation and everything else are equal. It therefore becomes easy to describe the payoff from everything else; it is equal to the payoff of brain stimulation reward when the animal allocates 50% of its time to self-stimulation. Recall that when time allocation is 50%, the payoff of brain stimulation reward must equal the payoff from everything else. The payoff from everything else can therefore be measured in terms of variables we can control, since it is possible for us to control the payoff from brain stimulation rewards. In other words, when the payoff from brain stimulation reward is equal to the payoff from everything else, the payoff of everything else can be described as a combination of the maximum intensity of the rewarding effect and the price of the stimulation. We therefore define a price, P_e, at which the payoff of a maximal reward will equal the payoff of everything else. The price at which the animal allocates half of its time to holding down the bar at a maximum reward value (P_e) is a benchmark for quantifying the payoff from everything else, because when the price is this high, a maximally intense brain stimulation reward will have the same payoff as all other activities the rat can engage in while in the operant chamber.

Manipulations that alter the reward circuitry prior to the output of the spatio-temporal integrator will alter the translation of the injected impulse flow into subjective reward intensity. This will change only the translation of stimulation into reward value, and will thus only change f_{hm}, the pulse frequency that generates a reward signal of half-maximal intensity. Manipulations that act after the signal is spatio-temporally integrated do not change how the injected signal translates into subjective reward intensity, rather, they alter how a given subjective reward intensity translates into behaviour. As a result, all that can possibly change within the confines of the model is the payoff of brain stimulation reward relative to everything else, and thus, will only change P_e.

When the inner, reward-growth function is embedded within the outer, neo-Herrnsteinian behavioural-allocation function and the animal's behaviour is plotted relative to the pulse frequency and price of the stimulation in three-dimensional semilogarithmic space, the structure resembles a mountain (figure 1A). Equation 4 shows

$$TA = (TAX - TIN) \times \frac{\left(\dfrac{f^g}{f^g + f_{hm}{}^g}\right)^a}{\left(\dfrac{f^g}{f^g + f_{hm}{}^g}\right)^a + \left(\dfrac{P}{P_e}\right)^a} + TIN \tag{4}$$

the resulting simplified combination of the two functions; this mathematical model will be referred to as the Mountain model from here on.

Proof of principle

Consider a manipulation such as by Waraczynski (2003), whereby a baseline frequency sweep is collected following saline injection and another is collected following the temporary inactivation of the central amygdala. In the upper left panel of figure 31, hypothetical results are shown. These results could be produced either by changing the translation of stimulation frequency into reward intensity (thereby changing f_{hm}, shown in the upper right panel) or by changing the animal's allocation of time for a given reward intensity (thereby changing P_e, shown in the lower left panel). There is no way to tell whether the effect occurs before the output of the spatio-temporal integration process (perhaps on the first-stage neurons themselves) or whether it occurs beyond that output. The 2D rate-frequency representations, which are essentially silhouettes of the Mountain along the frequency axis, provide a misleading and incomplete picture of the effect of a manipulation on self-stimulation.

Conversely, one could consider a manipulation such as demonstrated by Salamone (2001), inferring the involvement of dopamine tone in the effortful mobilization of resources by changes in breakpoint. In the upper left panel of figure 32, hypothetical results of a progressive ratio schedule are shown. These results could be produced either by changing the animal's willingness to work for a given reward intensity (thereby changing P_e) or by altering the translation of stimulation into subjective reward intensity (thereby changing f_{hm}). This 2D rate-ratio representation is essentially a silhouette of the Mountain taken along the price axis, and provides ambiguous information about a shift that could be due to an effect before the output of the integrator or to an effect beyond the output of the integrator (perhaps by changing the perceived harvesting effort required).

It is clear, then, that without a three-dimensional psychometric such as the Mountain, inferences from curve-shift methods (either rate-frequency or progressive-ratio) are relatively uninformative, insofar as they confound pre-integrator effects with post-integrator effects. Although two prior experiments (Arvanitogiannis, 1997; Mullett, 2005) have shown the Mountain Model to be a valid framework within which to understand brain reward manipulations, it had always been assumed that the testing situation had little impact on the animals' performance. The present experiment took into account the occurrence of preference construction elaborated on in experiment 1.

Strength-duration validation of the mountain model

If the Mountain Model is to be a solid framework upon which to interpret the effect of manipulations to the reward circuitry, one should be able to produce predictable changes known to affect only one parameter by virtue of their known effect on reward circuitry. A manipulation known to affect the temporal integration of the directly induced impulse flow should alter f_{hm} in a predictable way without altering P_e.

One such validation would involve changing the period during which neurons are stimulated during each train. If the train duration is short, there is less time for the reward signal to build up at the site of the integrator; as a result, the frequency of half maximal reward will be relatively high. Stimulation pulses will have to be injected at a higher rate into the substrate in order for a criterion level of reward to be reached. On the other hand, if the train duration is long, there is more time for the reward signal to build up at the site of the integrator. A slower rate will be necessary in order to elicit a similar level of subjective reward value: the frequency of half maximal reward will be relatively low.

Mark & Gallistel (1993) have already described the strength-duration relationship for stimulation trains. According to Gallistel, the frequency of half maximal reward at a given train duration is a hyperbolic function of the train duration taking into account the chronaxie of the relationship and scaling by the frequency of half-maximal reward for a train of infinite duration (see equation 5). Decreasing the train duration four-fold from 1 second to 0.25 seconds would be expected to shift f_{hm} by approximately 0.3 logarithmic units (Sonnenschein, Conover & Shizgal, 2003). Such a shift represents a two-fold change in the frequency of half maximal reward. One would not expect the output of the

$$f_{hm} = f_R \times \left(1 + \frac{C}{D}\right) \tag{5}$$

where f_{hm} is the frequency yielding half-maximal reward,

f_R is the frequency of the rheobase; the frequency yielding half-maximal reward when the duration of the train is infinitesimally large,

C is the chronaxie of the relationship, specifying how quickly it decays, and

D is the duration of the train.

reward-growth function to be rescaled, so one would not expect P_e to change significantly as a result of such a manipulation.

Two investigators have already reported validations of the model by assessing its capacity to identify the effect of altering train duration as occurring prior to the output of the integrator (Arvanitogiannis, 1997; Mullett, 2005). However, these studies were designed before unanticipated problems uncovered in experiment 1 had been described. I therefore attempted to partly validate the Mountain model's capacity to distinguish pre- from post-integration manipulations. As such, I attempted to verify whether shortening the train duration would produce a large, predictable increase in f_{hm} and a trivial or statistically unreliable shift in P_e, while removing the context-dependent biases described in Experiment 1. The present experiment further assesses the validity of the Mountain model.

Method

Subjects

Five male Long-Evans rats were used, from Charles River Breeding Farms (St-Constant, Quebec). Animals weighed a minimum of 350g before undergoing surgery. Testing always began during the dark phase of the cycle and never exceeded beyond 3 hours into the light phase of the cycle. Rats had *ad libitum* access to food and water throughout training and testing.

Materials & Apparatus

Electrodes

Electrodes used were similar to those in experiment 1. Electrodes were not glued into bilateral assemblies, but rather, a single electrode was implanted in the left hemisphere.

Operant Chambers

Operant chambers were identical to those used in experiment 1. The schedule of reinforcement used was a fixed cumulative handling time (FCHT) schedule, as discussed previously in experiment 1.

Surgical Procedure

Unipolar stimulation electrodes were implanted unilaterally into the lateral hypothalamic level of the medial forebrain bundle of subjects' left hemisphere. Stereotaxic coordinates for electrode implantation were 2.8mm posterior to bregma, 1.7mm lateral to the midline and 9mm ventral to the skull surface. Surgery was performed under isofluorane anesthesia at a maintenance concentration of 3% with a flow rate of 800mL/min. Atropine sulfate (0.5mg/Kg) was administered 15 minutes prior to surgery to reduce mucous secretions throughout the procedure. Return wires were fastened to the 2 most rostral of the 6 inserted jeweler's screws; the entire electrode assembly was anchored to the jeweler's screws by dental acrylic. A low dose (0.05mg/Kg) of buprenorphine was administered immediately after surgery, 24h after surgery and 48h after surgery, to minimize discomfort.

Behavioural Testing

Training

Animals were screened for the reward effectiveness of the implanted electrode a minimum of 4 days following surgical implantation. Animals were not included in the experiment if stimulation trains of 250 microampere intensity containing 45 pulses in half a second elicited motor stereotopy, aversive vocalizations, cowering, or any sign of discomfort on the rat. Following this confirmation, animals were screened for the highest current intensity and pulse frequency that did not produce an aversive effect. Using standard Skinnerian shaping techniques, animals were trained to press a bar to deliver the highest stimulation current and frequency. Animals were not included if they failed to learn an association between lever-pressing and the delivery of electrical stimulation.

Procedure

The experiment followed a multi-phase within-subject procedure. Animals were trained on descending pulse frequency sweeps (at a 1-second price) and increasing stimulation price sweeps (at the highest frequency the animal was willing to work for). Animals were then presented with a descending pulse frequency sweep in which each test trial was bracketed by a high-frequency, low price leading trial, and a low-frequency, low price trailing trial. A similarly bracketed trial structure was presented in an increasing stimulation price sweep taken at the highest frequency the animal would tolerate and another taken at the lowest frequency that produced maximal time allocation. In all cases (Y12, Y13, Y14 and C26), animals were presented with a bracketed descending pulse frequency sweep at train durations of 0.25 second and 1 second, bracketed by the same 0.5 second train duration trials.

Preliminary mountains were fit to the 0.5 second train duration data. The shift in the frequency sweep curves from 0.25 second and 1 second to 0.5 second train durations was used to estimate the approximate positions of mountains that would be fitted to eventual 0.25 and 1 second train durations.

Using these predicted mountain positions, points were chosen in 3 price-frequency sets to maximize the information provided by the data collected. Lists of price-frequency pairs were generated for each train duration. The lists comprised, for each train duration, a set of 9 frequencies at the same low price, a set of 9 prices at the highest frequency the animal could withstand, and a set of 9 price-frequency pairs arrayed along a radial axis through the estimated f_{hm} and P_e (see figure 33). Sets were chosen to maximize the likelihood of obtaining 3 points at the upper asymptote, 3 points on the lower asymptote, and 3 points along the rising portion of each set. Due to the expectation of higher variance, the five central price-frequency pairs in each set were sampled from twice as often as those in the upper and lower extremes.

To determine the effectiveness of the price-frequency pairs to produce asymptotically high, asymptotically low, and intermediate performance, animals were presented for five sessions the non-randomized, non-bracketed versions of the price-frequency sets. Thus, for five sessions, animals responded for a trial sequence in which

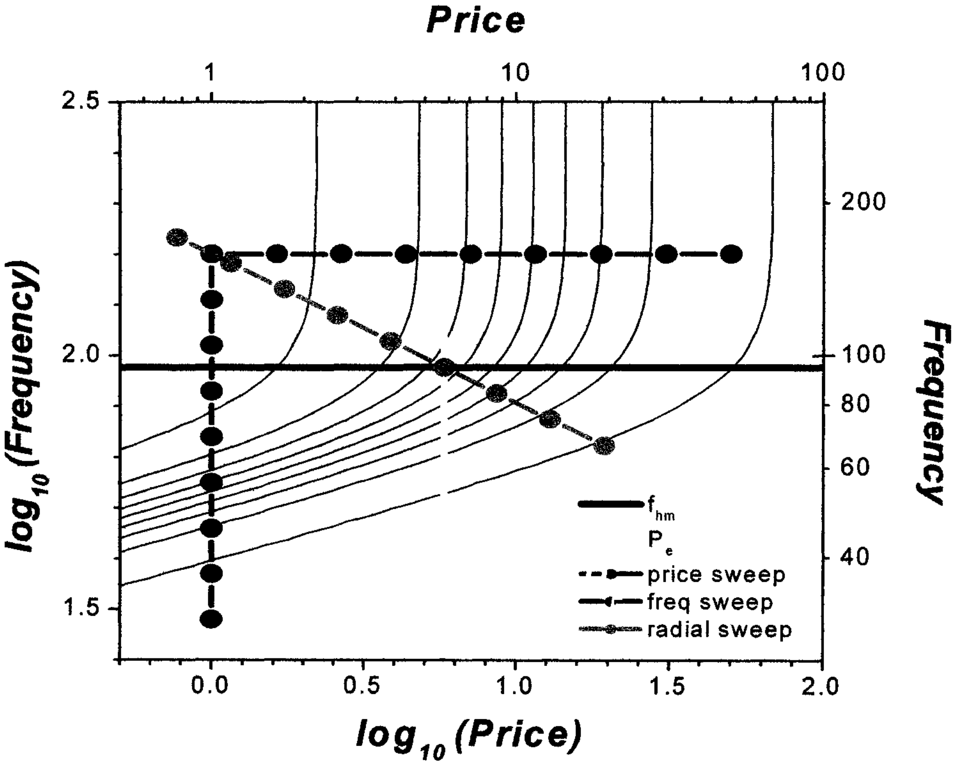


Figure 33. Example of a set of 3 pseudo-sweeps obtained at one of two train durations used in the experiment. Price-frequency pairs are represented by circles. They are grouped by pseudo-sweep for plotting purposes; pairs forming part of a pseudo-sweep are connected by a line and have identical colour. The presumed position of the mountain for this train duration is indicated by a magenta line representing fhm and a cyan line representing Pe.

the pulse frequency was logarithmically decreased, the price frequency was logarithmically increased, and both changed simultaneously at a 1 second train duration, and similarly for a 0.25 second train duration. Frequencies or prices were changed if it was deemed necessary, i.e., the set did not result in a sigmoidally-shaped curve or the animal exhibited adverse motor effects that compromised its capacity to respond.

In the final phase of the experiment, animals were presented with trial triads consisting of a test trial bracketed by a leading trial of high frequency, low price stimulation and by a trailing trial of low frequency, low price stimulation. Test trials were drawn without replacement from the two lists of price-frequency sets, thus merged into a large randomized list of triads. Animals were presented with re-randomized versions of these lists until their performance was stable over 8 lists. Stability was confirmed with consistently asymptotically high performance on leading bracket trials and low performance on trailing bracket trials. To avoid exhausting the rats, daily sessions were limited to 4 hours. As a result, a single pass through a list of price-frequency sets required 3 or 4 sessions to complete. Animals required approximately 13 passes through re-randomized lists to complete the final phase of the experiment.

Data Collection and Analysis

In addition to the tapping correction used in experiment 1, bar-pressing activity was only considered following the animals' first reward in a trial. If animals received only one or no rewards, the proportion of time allocated to self-stimulation was recorded as nil. Since animals encountered test trials in a random fashion, the criterion cumulative time required to trigger a reward on a given trial only became evident after the first reward was delivered. Time allocation was therefore calculated based on the proportion of time the bar was held down or undepressed for less than 1 second following the first reward encounter.

Results

Fitting Strategy

The data were fit according to the same non-linear estimation routine as in experiment 1. Different location parameters (f_{hm} and P_e) and maximum time allocations (TAX) were fit to the data from each train duration condition, but the slope and minimum time allocation estimates were fit to both simultaneously. Tukey's bi-square estimator was applied and the 95% confidence surrounding each parameter was calculated as in experiment 1. A description of Tukey's bi-square estimator, the bootstrapping approach, and the justification for these robust statistical methods can be found in appendix A. Final values resulting from the Mountain Model fit in experiment 1 did not appear to depend greatly on the initial seeds of the iterative fitting procedure, provided they were in the same range as the data collected. Initial seeds for the iterative least-squares fitting procedure were therefore kept constant across animals. Parameter A was seeded as 4, f_{hm} for the 1 second train duration as 1.7, f_{hm} for the 0.25 second train duration as 2.3, G as 4, P_e for both train durations as 1, TAX for both train durations as 0.9 and TIN as 0.1. Re-seeding with previously fit parameters did not change the fitted estimates.

In the case of fitted parameter estimates and differences in fitted estimates between train durations, confidence intervals were based on 1000 bi-square-weighted least-squares fits to re-sampled data. Thus, bi-square-weighted data points were sampled from each price-frequency pair (with replacement) in the same number as there were data points contributing to the mean time allocation for that pair. The confidence interval about each parameter estimate was based on those fit to 1000 such re-sampled datasets. The lower confidence bound was that below which 25 of the estimates were distributed;

the upper confidence bound was that below which 975 of the estimates were distributed. Confidence intervals about the differences in parameter estimates were similarly computed: they corresponded to the central 950 differences in re-sampled estimated parameters between 1 second train duration and 0.25 second train duration conditions.

Two-Dimensional Representations

Figures 34 through 37 provide a plot of the observed bi-square-weighted means and their 95% confidence intervals as calculated by the bootstrap re-sampling procedure, for both 1 second and 0.25 second train durations. Coincident with these are projected silhouettes of the 1 second and 0.25 fitted surfaces at the frequencies and prices used. In all cases, the projected curves provide a reasonable account of the animals' behaviour for stimulation of different pulse frequencies (panel A, frequency pseudo-sweeps), prices (panel B, price pseudo-sweeps) or both (panel C, radial pseudo-sweeps).

Three-dimensional surface fits

In figures 38 through 41, a plot of the observed, bi-square weighted, mean time allocations is depicted (on the Z-axis) as a function of pulse frequency and price. Draped over the observed points is the surface fitted by the iterative least-squares procedure, in black. The magenta lines on the floor of each graph indicate the fitted value of f_{hm} (intersecting the frequency axis) and P_e (intersecting the price axis). One second and 0.25 second data are shown in the same figure for each rat for comparison. In all cases, the fitted surface appears to track the animal's behaviour lawfully, with large changes in f_{hm} in all animals and small changes in P_e in 3 of 4 cases. In addition, the correlations between observed and predicted time allocations in all cases are high, with no fit

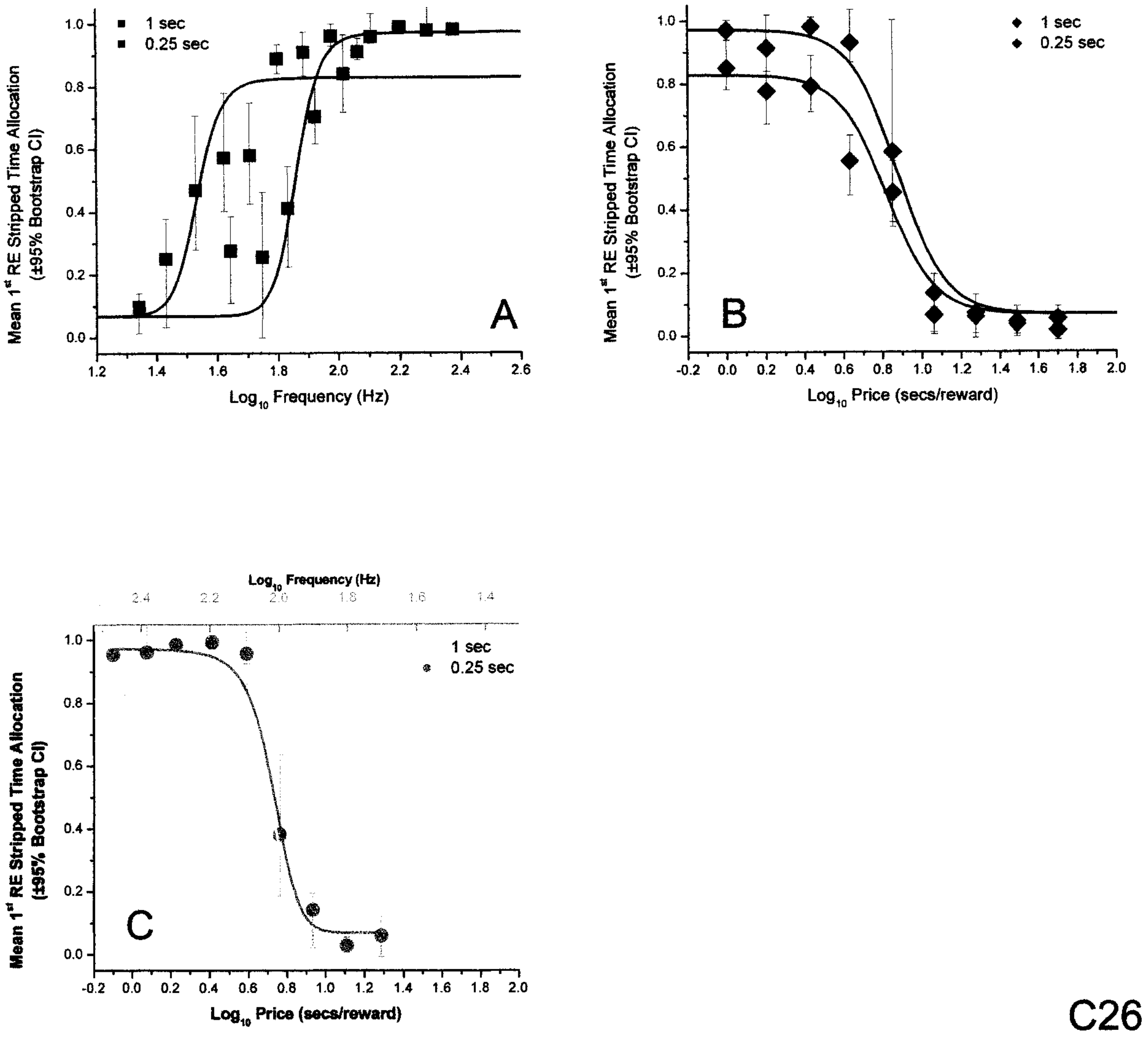


Figure 34. Observed, first-encounter stripped mean corrected time allocation compared to projections of the surfaces fit to the 1- and 0.25-second train duration data for rat C26.

Panel A: Reconstructed frequency axis-aligned pseudo-”sweeps” obtained at 1- and 0.25-second train durations.

Panel B: Reconstructed price axis-aligned pseudo-”sweeps” obtained at 1- and 0.25-second train durations.

Panel C: Reconstructed pseudo-”sweep” aligned with a radial passing through the intersection of the frequency and price pseudo-”sweeps”, and the point at which the frequency is fhm and the price is Pe. The bottom x-axis indicates the price of the stimulation, whereas the upper x-axes indicate, for each train duration, the pulse frequency.

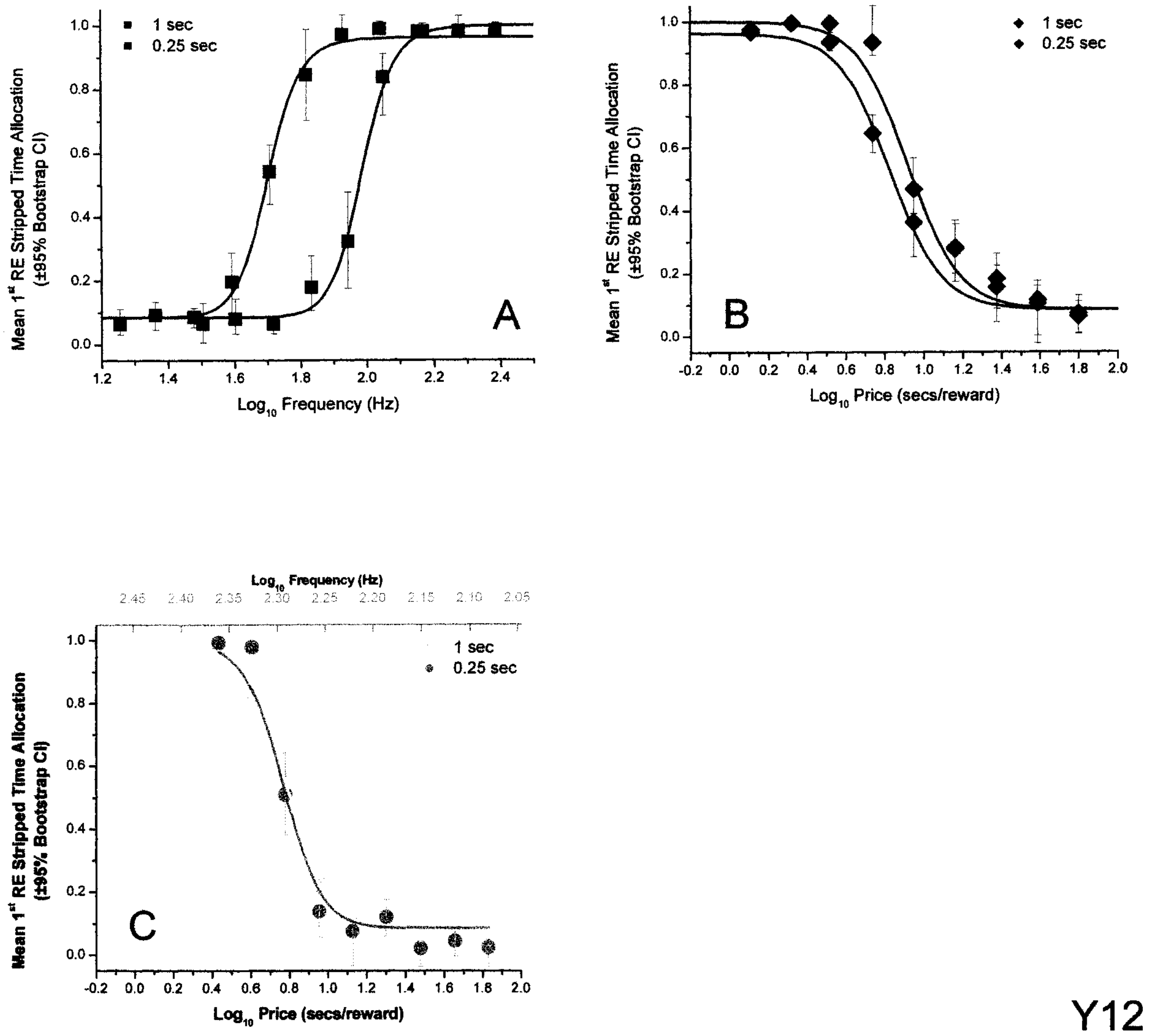


Figure 35. Observed, first-encounter stripped mean corrected time allocation compared to projections of the surfaces fit to the 1- and 0.25-second train duration data for rat Y12.

Panel A: Reconstructed frequency axis-aligned pseudo-"sweeps" obtained at 1- and 0.25-second train durations.

Panel B: Reconstructed price axis-aligned pseudo-"sweeps" obtained at 1- and 0.25-second train durations.

Panel C: Reconstructed pseudo-"sweep" aligned with a radial passing through the intersection of the frequency and price pseudo-"sweeps", and the point at which the frequency is fhm and the price is Pe. The bottom x-axis indicates the price of the stimulation, whereas the upper x-axes indicate, for each train duration, the pulse frequency.

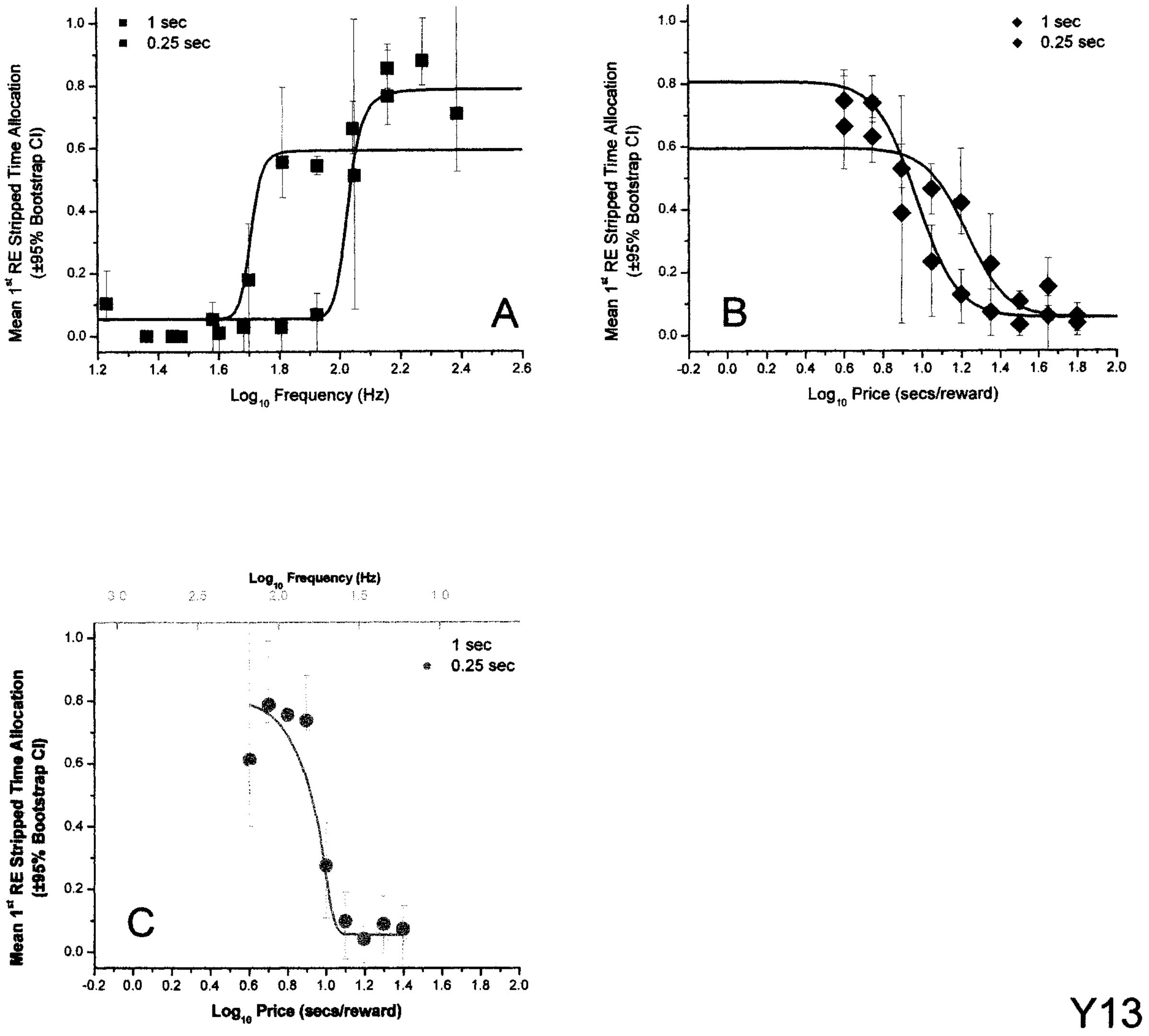


Figure 36. Observed, first-encounter stripped mean corrected time allocation compared to projections of the surfaces fit to the 1- and 0.25-second train duration data for rat Y13.

Panel A: Reconstructed frequency axis-aligned pseudo-"sweeps" obtained at 1- and 0.25-second train durations.

Panel B: Reconstructed price axis-aligned pseudo-"sweeps" obtained at 1- and 0.25-second train durations.

Panel C: Reconstructed pseudo-"sweep" aligned with a radial passing through the intersection of the frequency and price pseudo-"sweeps", and the point at which the frequency is fhm and the price is Pe. The bottom x-axis indicates the price of the stimulation, whereas the upper x-axes indicate, for each train duration, the pulse frequency.

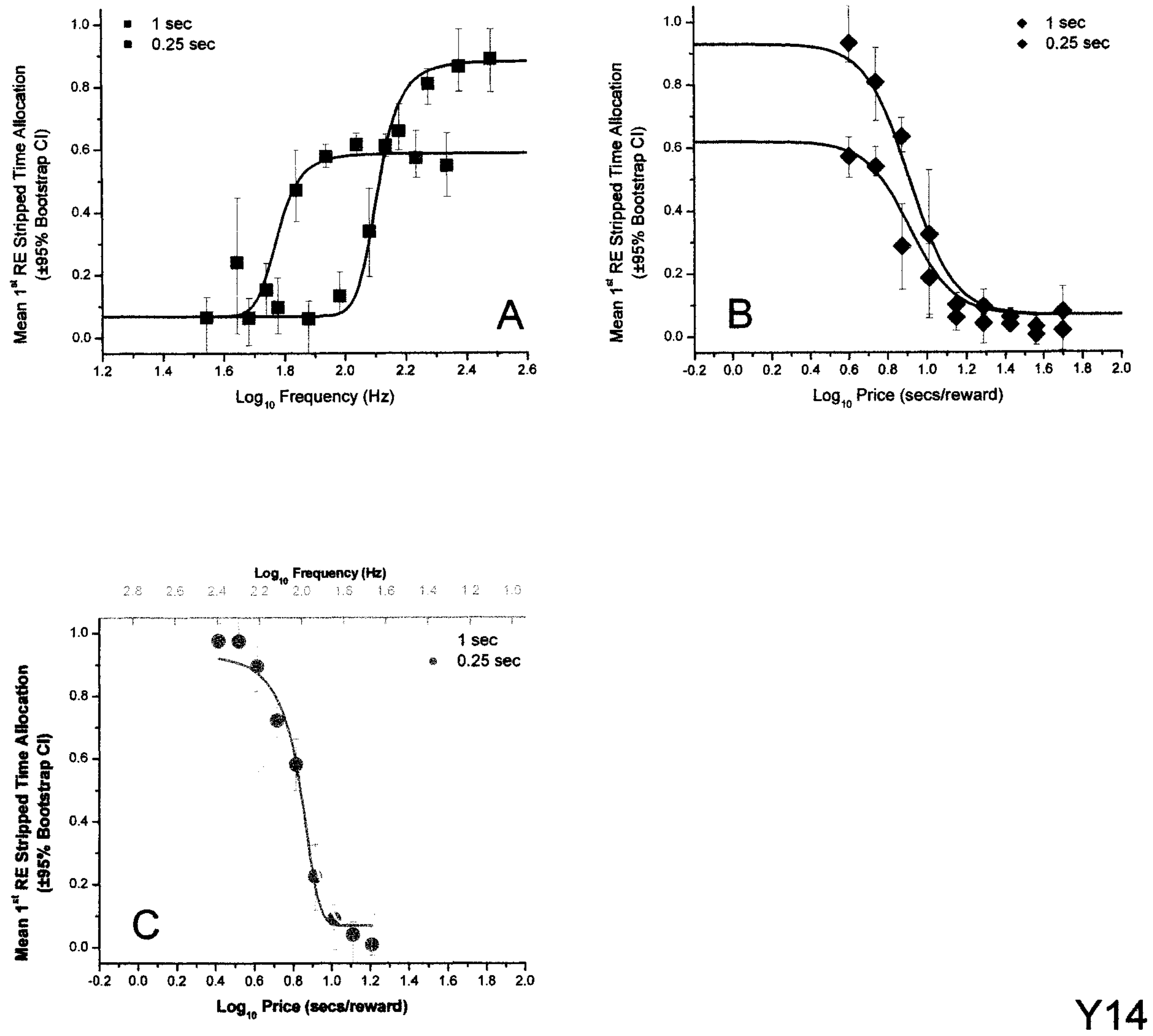


Figure 37. Observed, first-encounter stripped mean corrected time allocation compared to projections of the surfaces fit to the 1- and 0.25-second train duration data for rat Y14.

Panel A: Reconstructed frequency axis-aligned pseudo-"sweeps" obtained at 1- and 0.25-second train durations.

Panel B: Reconstructed price axis-aligned pseudo-"sweeps" obtained at 1- and 0.25-second train durations.

Panel C: Reconstructed pseudo-"sweep" aligned with a radial passing through the intersection of the frequency and price pseudo-"sweeps", and the point at which the frequency is fhm and the price is Pe. The bottom x-axis indicates the price of the stimulation, whereas the upper x-axes indicate, for each train duration, the pulse frequency.

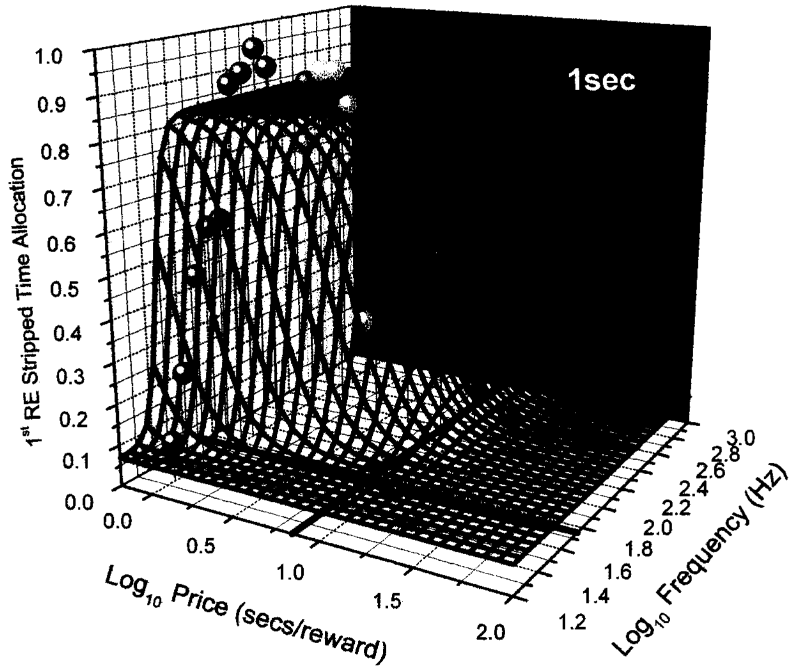
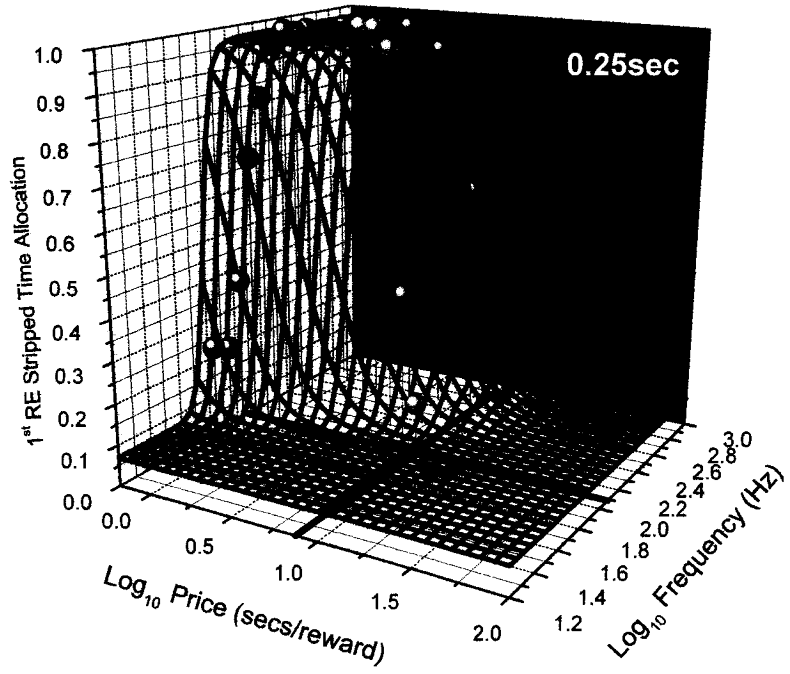


Figure 38. Comparison, for each train duration, of the observed time allocations and the fitted surface, for rat C26.



C26

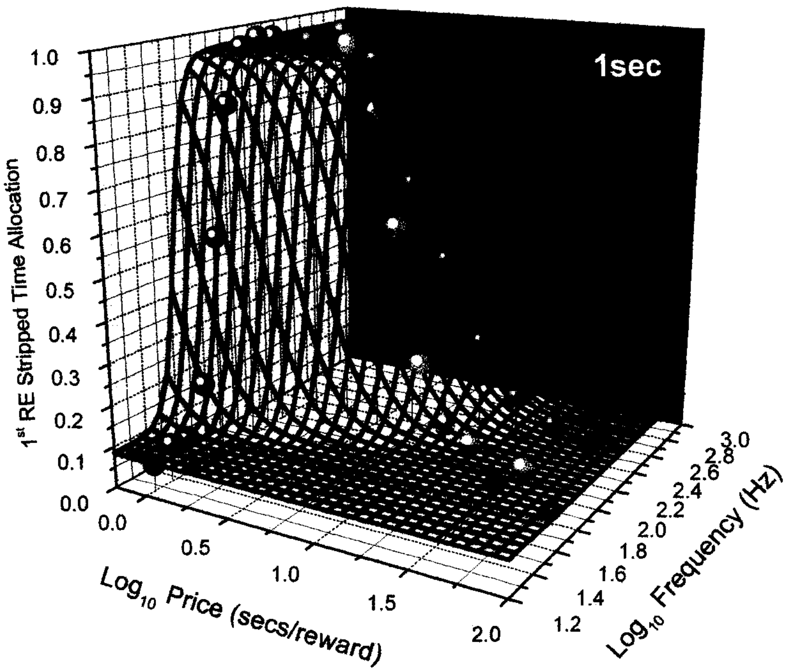
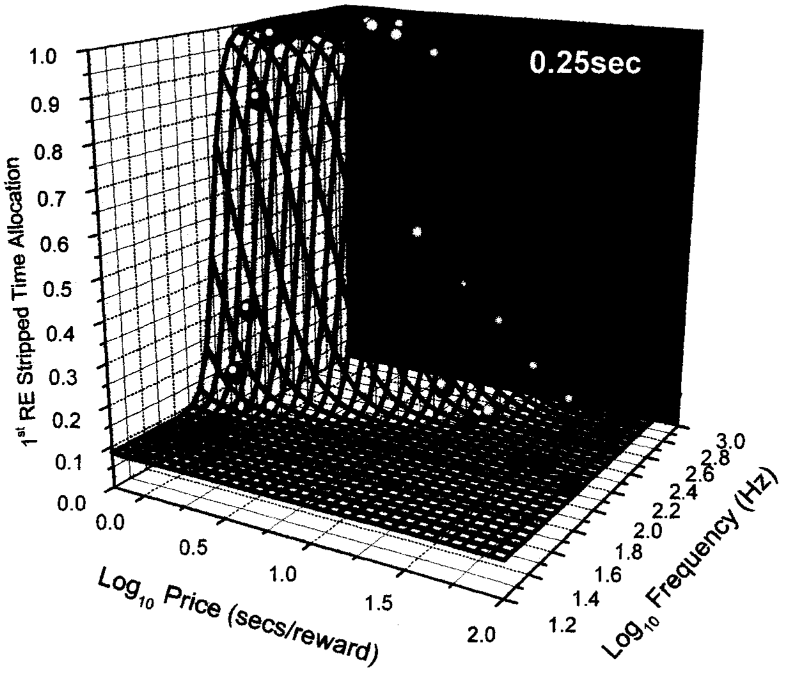


Figure 39. Comparison, for each train duration, of the observed time allocations and the fitted surface, for rat Y12.



Y12

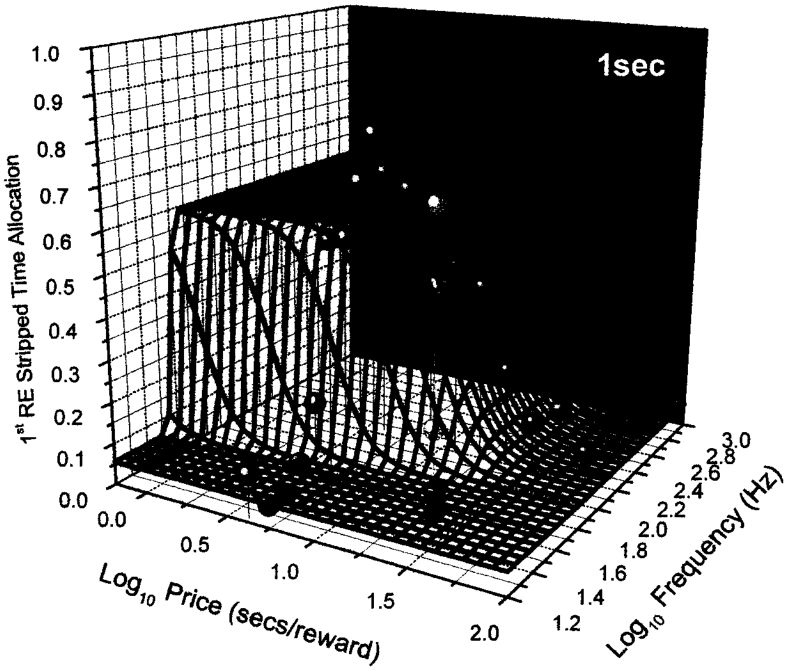
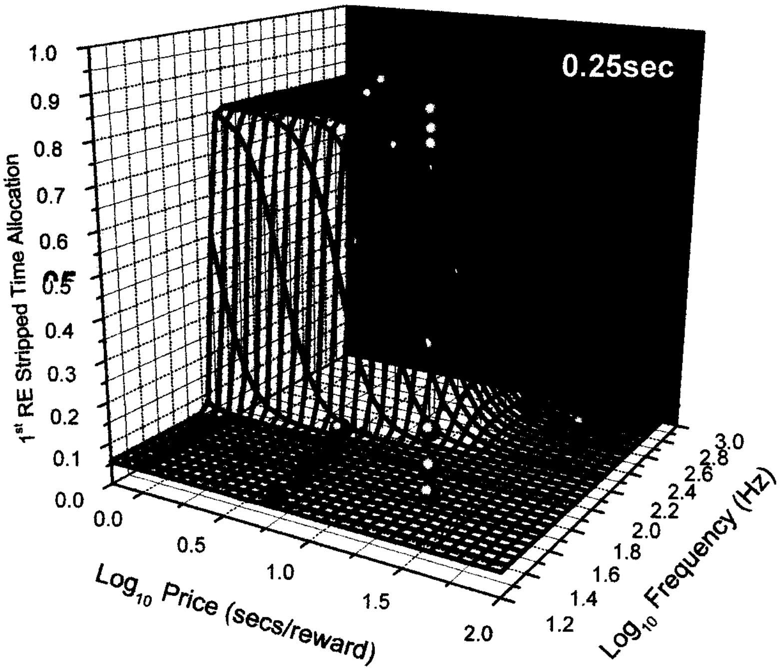


Figure 40. Comparison, for each train duration, of the observed time allocations and the fitted surface, for rat Y13.



Y13

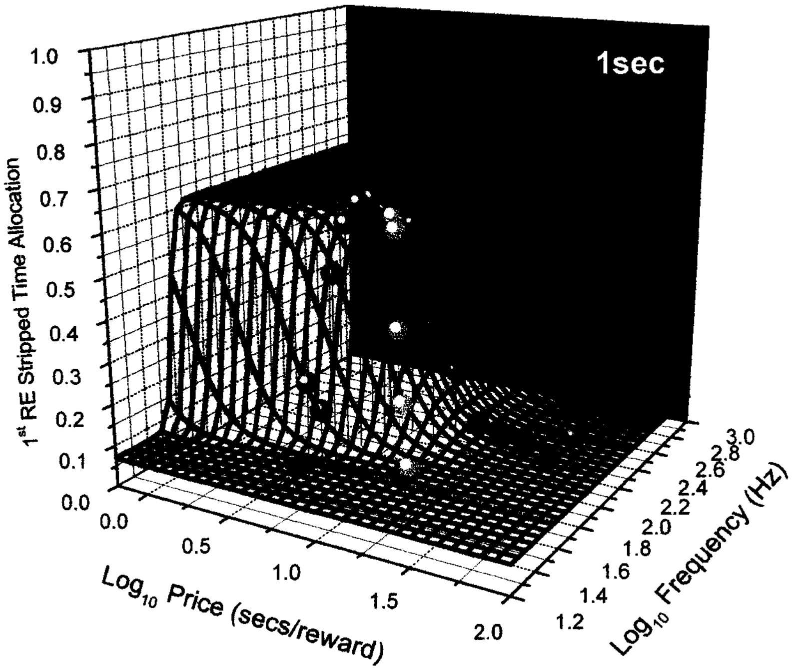
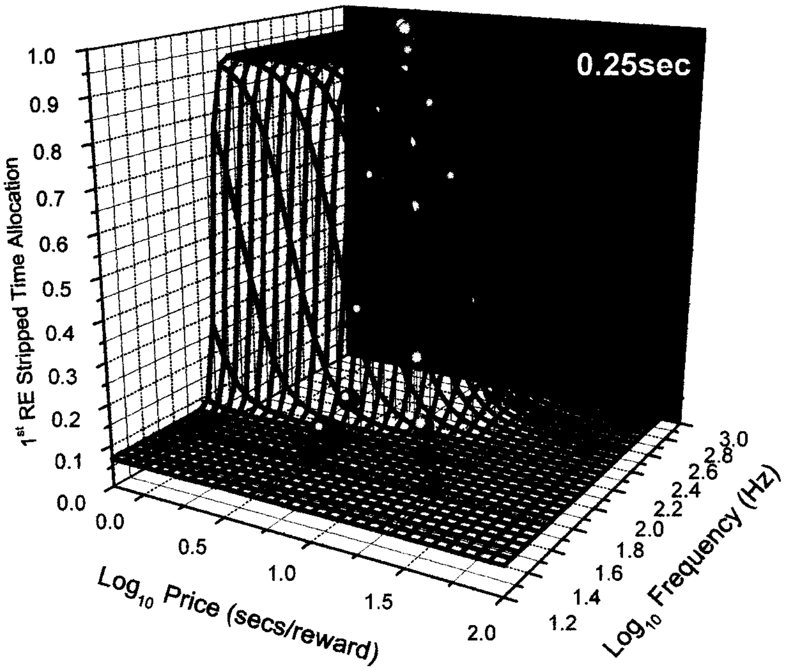


Figure 41. Comparison, for each train duration, of the observed time allocations and the fitted surface, for rat Y14.



Y14

accounting for less than 58% (the three others are 75%, 79% and 87%) of the variance in time allocation behaviour.

Parameter shifts

Figures 42 through 45 contain contour plots of the fitted mountain surfaces at each train duration as well as a bar graph of the estimated shift in TAX, f_{hm} and P_e with the associated 95% confidence intervals determined by the bootstrap re-sampling procedure. Contours of the mountain fitted to the data obtained at the 1-second train duration are reproduced once above the 0.25-second mountain fit (lining up on the price axis) and once to the right of it (lining up on the frequency axis). As is clear, the 95% confidence intervals of the estimated P_e parameter appear to overlap in 3 of the 4 animals, whereas the 95% confidence intervals of the estimated f_{hm} parameters do not in any animal.

Table 3 shows the estimated parameters, their upper and lower 95% confidence bounds and the width of the confidence interval surrounding parameter estimates of each fitted surface for all animals. Parameters without number suffixes were fit commonly to both train durations; those suffixed with 1 were fit only to 1-second train duration data while those suffixed with 2 were fit only to 0.25-second train duration data.

Table 4 shows the difference in estimated parameters TAX, f_{hm} and P_e along with 95% confidence bounds and the width of the confidence interval as calculated by the bootstrap procedure. In all animals, the difference in f_{hm} between 1- and 0.25-second train durations was approximately 0.3 logarithmic units, corresponding to a two-fold increase in frequency required to maintain an equivalent level of time allocation to self-stimulation activities at short train durations. In animals C26, Y12 and Y14, the

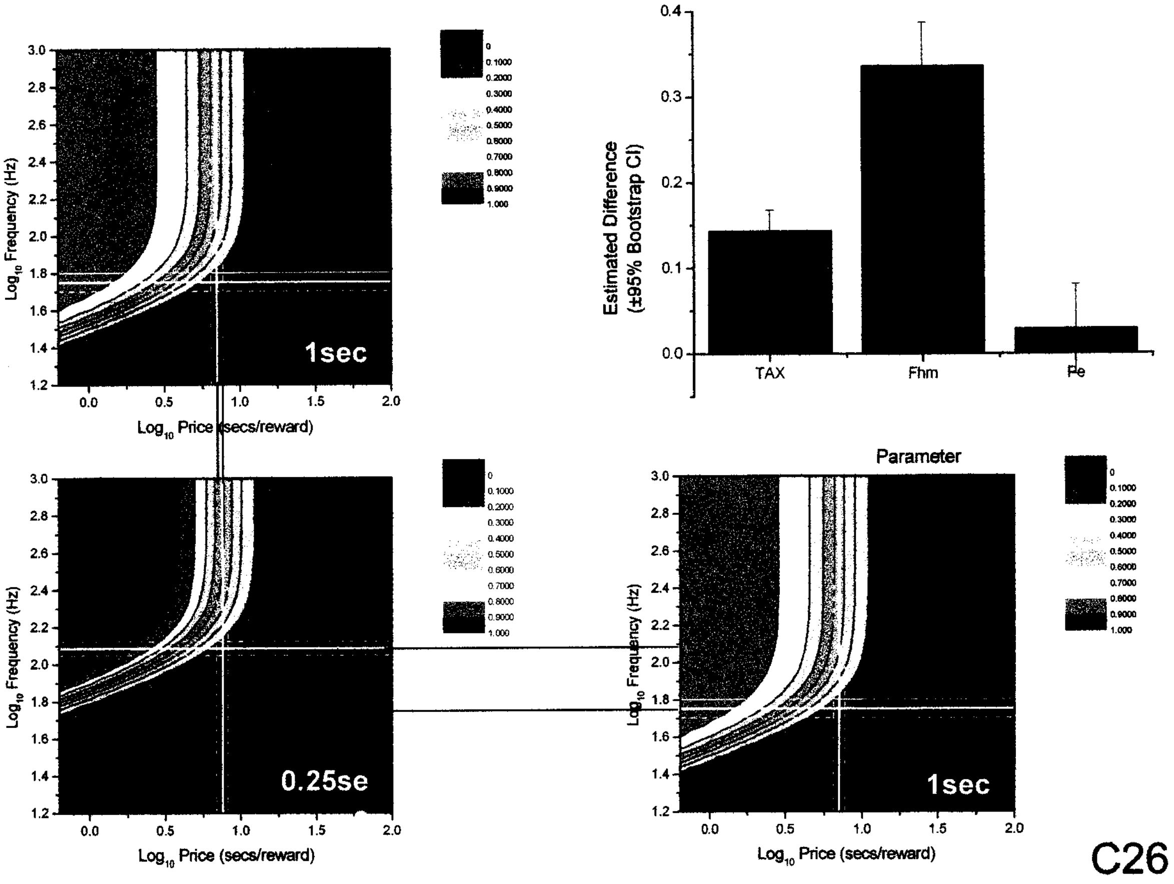


Figure 42. Contour plot of the fitted Mountain model at each train duration, demonstrating the train duration effect in C26. The upper left and lower right panels are identical representations of the 1-second train duration-fitted surface for contrast against the 0.25-second train duration-fitted surface. Purple lines project from the contour graph for better visualization of the position parameters fhm and Pe. The upper right panel is a bar graph representing the estimated difference in maximal time allocation (TAX), frequency of half-maximal reward (Fhm) and price at which a maximal reward yields half-maximal time allocation (Pe). There are statistically significant differences in TAX and Fhm, but the difference in Pe can not be said to differ statistically from 0.

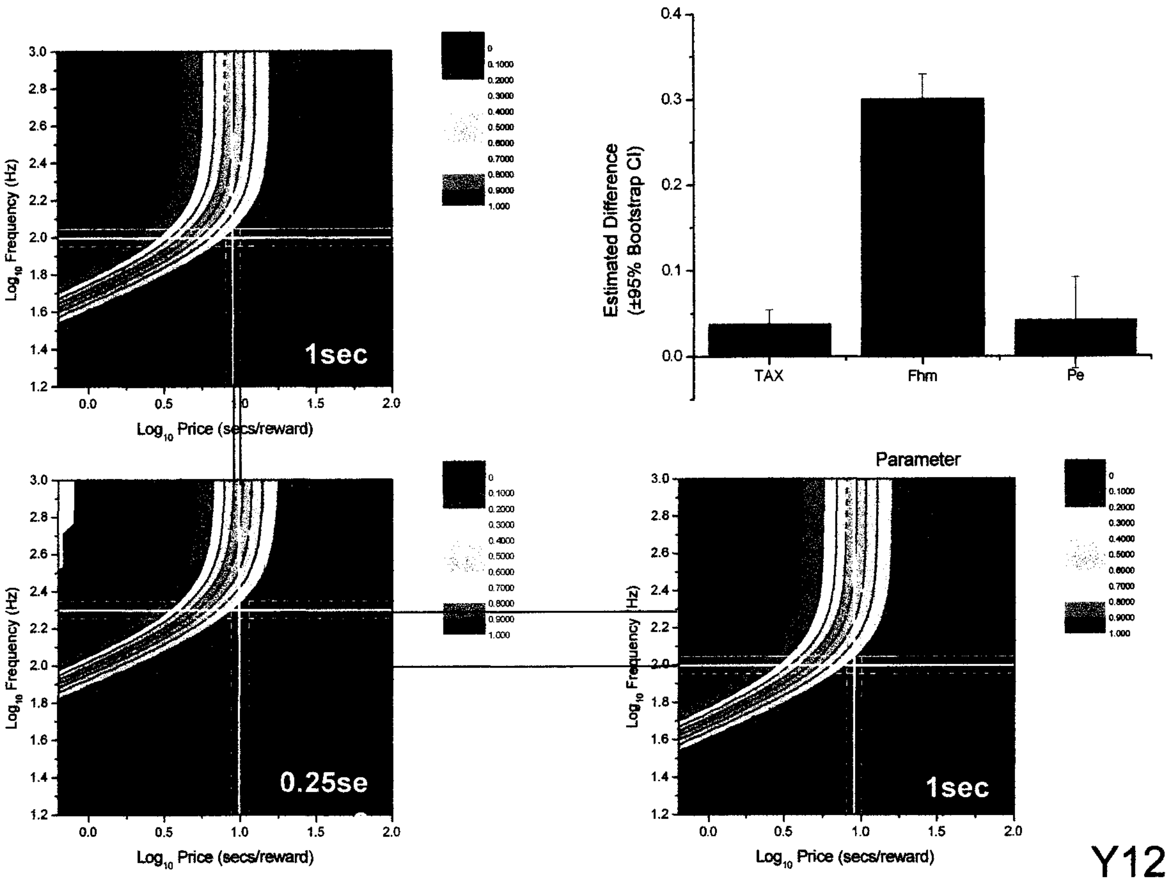


Figure 43. Contour plot of the fitted Mountain model at each train duration, demonstrating the train duration effect in Y12. The upper left and lower right panels are identical representations of the 1-second train duration-fitted surface for contrast against the 0.25-second train duration-fitted surface. Purple lines project from the contour graph for better visualization of the position parameters fhm and Pe. The upper right panel is a bar graph representing the estimated difference in maximal time allocation (TAX), frequency of half-maximal reward (Fhm) and price at which a maximal reward yields half-maximal time allocation (Pe). There are statistically significant differences in TAX and Fhm, but the difference in Pe can not be said to differ statistically from 0.

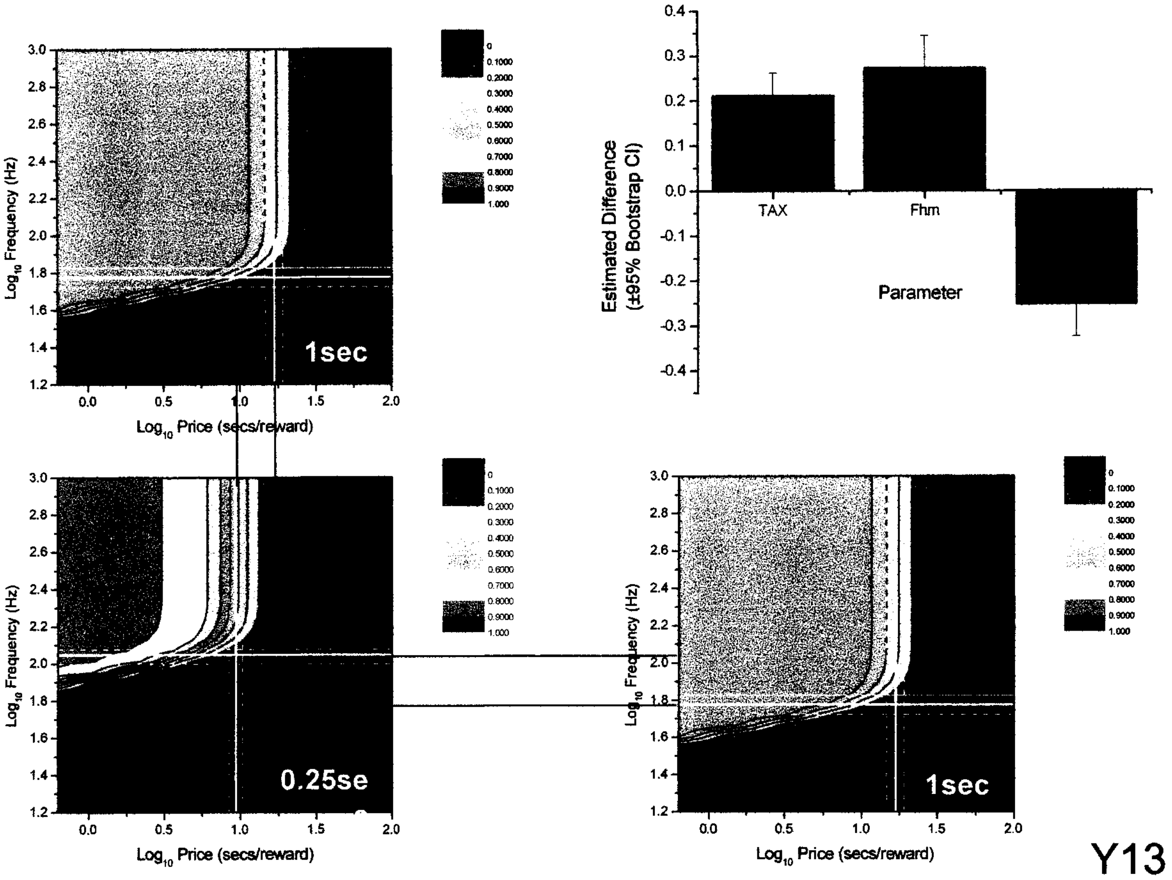


Figure 44. Contour plot of the fitted Mountain model at each train duration, demonstrating the train duration effect in Y13. The upper left and lower right panels are identical representations of the 1-second train duration-fitted surface for contrast against the 0.25-second train duration-fitted surface. Purple lines project from the contour graph for better visualization of the position parameters fhm and Pe. The upper right panel is a bar graph representing the estimated difference in maximal time allocation (TAX), frequency of half-maximal reward (Fhm) and price at which a maximal reward yields half-maximal time allocation (Pe). All three of the parameters that were allowed to vary with train duration condition are statistically different from 0.

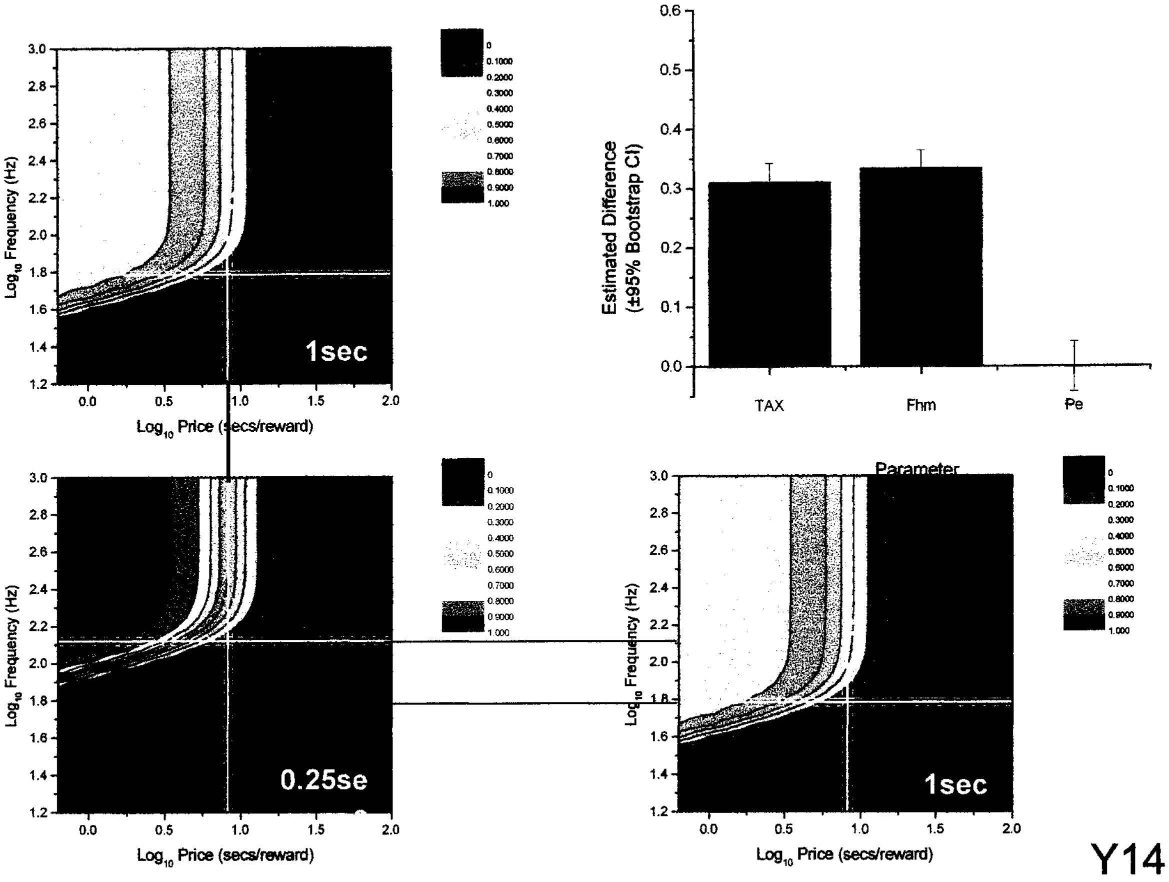


Figure 45. Contour plot of the fitted Mountain model at each train duration, demonstrating the train duration effect in Y14. The upper left and lower right panels are identical representations of the 1-second train duration-fitted surface for contrast against the 0.25-second train duration-fitted surface. Purple lines project from the contour graph for better visualization of the position parameters fhm and Pe. The upper right panel is a bar graph representing the estimated difference in maximal time allocation (TAX), frequency of half-maximal reward (Fhm) and price at which a maximal reward yields half-maximal time allocation (Pe). There are statistically significant differences in TAX and Fhm, but the difference in Pe can not be said to differ statistically from 0.

Table 3

Estimated Parameters of Mountain Model Simultaneously Fit to Both Train Durations

Parameter	Estimated Value	Lower Confidence Bound	Upper Confidence Bound	Confidence Interval Width
		C26		
A	3.5764	2.97549	4.28558	1.31009
Fh.1	1.75282	1.70395	1.80276	0.09881
Fh.2	2.08803	2.05048	2.12602	0.07554
G	3.60308	3.16556	4.10351	0.93795
Pe.1	0.84994	0.81242	0.88733	0.07491
Pe.2	0.87862	0.83778	0.91696	0.07918
TAX.1	0.82943	0.80258	0.85405	0.05147
TAX.2	0.97261	0.95739	0.98764	0.03025
TIN	0.06923	0.05449	0.08513	0.03064
		Y12		
A	3.29495	2.95096	3.69902	0.74806
Fh.1	1.99827	1.95318	2.04585	0.09267
Fh.2	2.29849	2.25492	2.3483	0.09338
G	3.04163	2.7493	3.38727	0.63797
Pe.1	0.95387	0.90679	1.00276	0.09597
Pe.2	0.99541	0.94217	1.06021	0.11804
TAX.1	0.96286	0.94236	0.97987	0.03751
TAX.2	1	0.99998	1	2E-05
TIN	0.08499	0.07368	0.09626	0.02258
		Y13		
A	4.23119	3.23513	5.50499	2.26986
Fh.1	1.7758	1.72362	1.8258	0.10218
Fh.2	2.04835	1.99918	2.07843	0.07925
G	7.27689	5.24886	12.28507	7.03621
Pe.1	1.22678	1.16725	1.28093	0.11368
Pe.2	0.97397	0.93133	1.0164	0.08507
TAX.1	0.59462	0.57321	0.61757	0.04436
TAX.2	0.80666	0.75427	0.86686	0.11259
TIN	0.05559	0.04306	0.06821	0.02515
		Y14		
A	3.86476	3.39017	4.45761	1.06744
Fh.1	1.78646	1.76384	1.80524	0.0414
Fh.2	2.11995	2.09837	2.14229	0.04392
G	5.46867	4.1961	6.91807	2.72197
Pe.1	0.91582	0.88146	0.94882	0.06736
Pe.2	0.91677	0.89212	0.94625	0.05413
TAX.1	0.61951	0.59969	0.64199	0.0423
TAX.2	0.92923	0.89375	0.96622	0.07247
TIN	0.06904	0.05729	0.08094	0.02365

Table 4

Estimated Differences in Parameters Fit Independently by Train Duration

Parameter	Estimated Difference	Lower Confidence Bound	Upper Confidence Bound
	C26		
TAX	0.14318	0.11862	0.16819
Fhm	0.33521	0.28424	0.38482
Pe	0.02868	-0.02311	0.08154
	Y12		
TAX	0.03714	0.02013	0.05764
Fhm	0.30022	0.27108	0.33441
Pe	0.04154	-0.0088	0.0977
	Y13		
TAX	0.21204	0.16208	0.26362
Fhm	0.27255	0.20107	0.33901
Pe	-0.25281	-0.31854	-0.1816
	Y14		
TAX	0.30972	0.27693	0.3422
Fhm	0.33349	0.30331	0.36632
Pe	9.47E-04	-0.03968	0.04367

Note. Expressed differences are (short train duration - long train duration).

Note. Values in red demonstrate statistical differences (a = 0.05, bootstrap) from 0.

difference in P_e was close to 0, with confidence bounds including 0 in these three cases. In animal Y13, P_e was approximately 0.3 logarithmic units smaller at the short train duration compared to the long train duration, corresponding to a two-fold increase in price required to maintain a level of time allocation that had been obtained at long train durations.

Discussion

This experiment tested a model of performance for rewarding brain stimulation. According to this model, when an animal harvests a brain stimulation reward, the electrical signal gives rise to a volley of action potentials that is subsequently converted into a reward intensity signal by a spatio-temporal integration process. On the basis of stored records of subjective reward intensity and price, the rat decides how to allocate its time between working for BSR and performing alternate activities. Figure 46 presents a diagram of presumed processes that intervene between the delivery of a brain stimulation reward and the animal's decision to press again. First, the elicited impulse flow is spatio-temporally integrated; the output of this integration process presumably leads to a stored record of reward within the brain, or a reward intensity signal. The animal's decision to press then depends on the intensity of the reward and the opportunity and effort costs it must incur in order to acquire the reward.

This is the model that has guided the study of brain stimulation reward for decades. Electrically evoked activity travels to an integrator, the activity of which reflects the total number of action potentials in the barrage produced by the electrode. Its peak is converted into a stored record of reward upon which future decisions can be based. The presumed existence of a stored record of reward that is dependent on the impulse flow requires that some process exist (the integrator) that is capable of counting the number of action potentials that the electrode injects into reward-relevant parts of the brain. There must exist some substrate in the brain which can carry out the spatio-temporal integration of the signal mimicked by electrical stimulation in order for current and frequency to trade off linearly (Gallistel, 1978). In other words, because it is possible to compensate

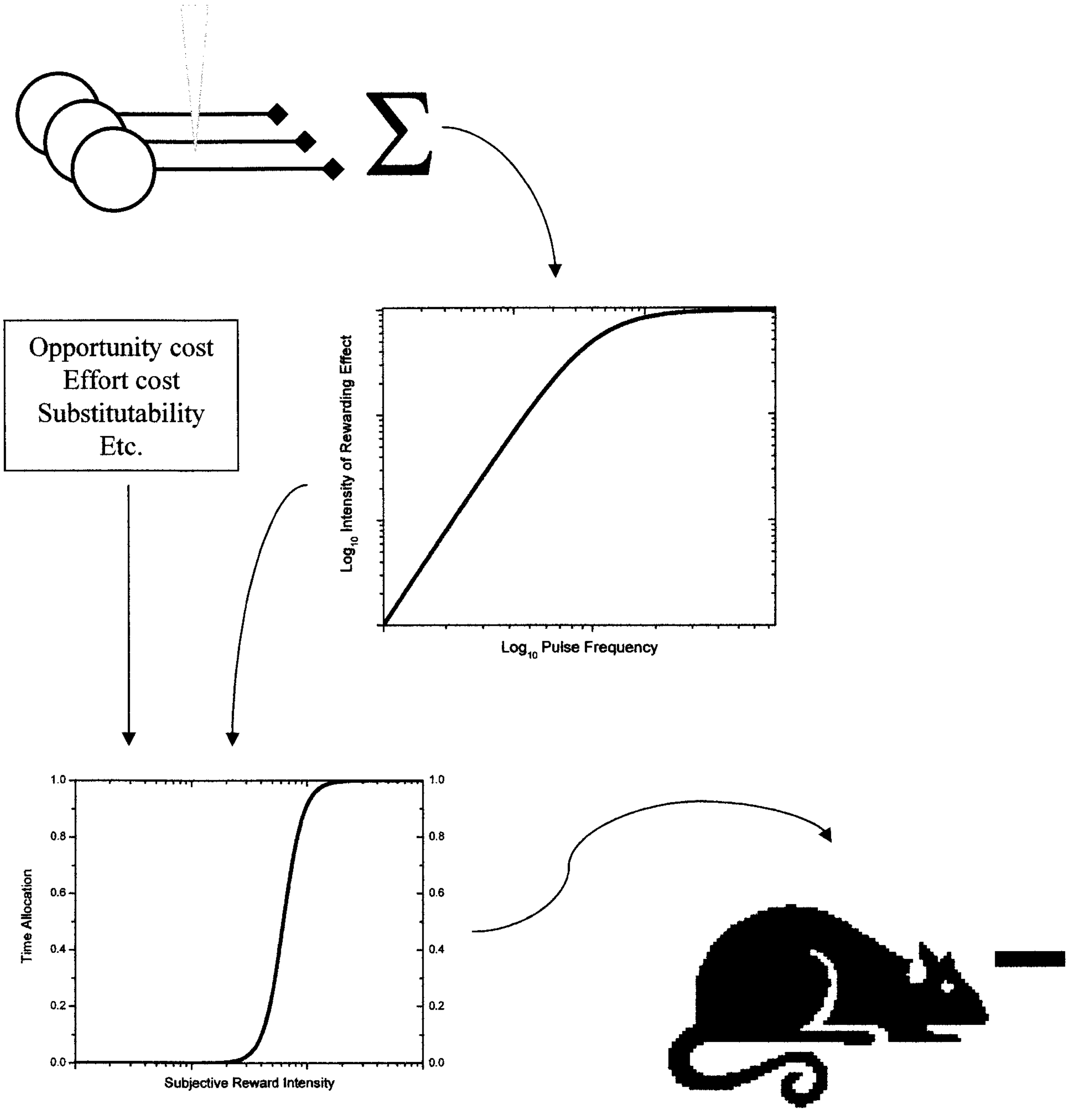


Figure 46. Psychological and neurobiological processing of rewards. Injected impulse flow is summated downstream by an integrator whose output determines subjective reward intensity, yielding a reward-growth relationship. This reward intensity is then combined with other factors that determine how much time the animal will allocate to harvesting brain stimulation rewards, yielding the behavioural-allocation function.

for a two-fold decrease in frequency by making a two-fold increase in current, it can be presumed that some integrator counts the number of action potentials that are incident within a time window. The peak of the integration process must in turn lead to a stored record of reward in order for the transient, half-second burst of electrical stimulation to be recorded in memory. Since animals readily learn to perform an operant task to deliver rewarding brain stimulation, there must be some conversion of the injected impulse flow into a stored record of reward.

Given that some process or set of processes trans-synaptically converts the injected impulse flow into a stored record of reward, the temporal summation exerted by the integrator will be leaky. In other words, the effects of action potentials that arrive late in the train will have to summate with the effects of action potentials that have arrived early and decayed. It follows that when the train duration is short, action potentials must be produced at a higher rate to drive the integrator to a given level of reward intensity. Longer train durations allow more time for temporal summation by the integrator, thereby requiring a lower rate to drive the integrator to a given level of reward. The result is that manipulations of train duration alter only the frequency of half-maximal reward because they alter the temporal window in which the integrator can summate action potentials.

Assuming that this model can, in theory, discern the stage of processing of manipulations, it should be possible to use the model to determine at what stage various pharmacological, neurotoxic and physiological manipulations act. As predicted, a manipulation that affects the process of temporal integration and thus acts prior to the output of the spatio-temporal integrator shifted the 3D surface benchmark for effects occurring prior to the output of the integration process. This shift was also, by and large, consistent with previous models of temporal integration (Gallistel, 1978; Sonnenschein et al., 2003). In effect, on the basis of chronaxie estimates (specifying the rate of decay of the strength-duration trade-off) described in previous experiments, one would expect a shift of the frequency of half-maximal reward intensity equivalent to approximately 0.3 common logarithmic units, which is within the range of all the shifts of f_{hm} observed.

Shifts in f_{hm} relative to P_e

In 3 of the 4 animals tested, the shift of f_{hm} along the frequency axis was large and reliable (approximately 0.3, corresponding to a two-fold change), and the shift of P_e along the price axis was 0 or negligible. Indeed, in those 3 cases (C26, Y12 and Y14), the 95% confidence interval around the difference in P_e estimates included zero. This suggests that the model is, in practice, capable of correctly detecting the effect of a manipulation acting prior to the output of the integrator. Although in one case (Y13) the shift along P_e was large, this finding is contrary to all similar validation experiments conducted to date (Arvanitogiannis, 1997; Mullett, 2005). A discussion of this result follows below.

The puzzling case of rat Y13

In the case of rat Y13, the shift along the price axis was even larger than that along the frequency axis. According to the model, for this result to be commensurate with the other 3 rats and the findings of Mullett (2005) and Arvanitogiannis (1997), it would be necessary for the 1-second train duration to entail, in addition to a longer integration window, a lower effort or subjective price of self-stimulation or an increase in the maximum reward intensity. The possibility of such effects is elaborated here.

Shizgal & Matthews (1977) established that although the reward-relevant signal injected into the brain grows and decays rapidly with time, the aversive side-effects of mixed appetitive-aversive stimulation appear to build and decay slowly. If there were a build-up of aversive stimulation in the case of Y13, the 1-second train duration would be much more likely to build up a behaviourally-relevant aversive signal than the 0.25-second train duration. This aversive build up would also take a longer time to decay at long train durations than shorter train durations. Since the build-up of an alleged aversive signal would occur primarily during 1-second train duration trials, all 1-second trains would be affected by this signal to a much greater extent than 0.25-second train duration test trials. Nonetheless, if the stimulation fires a mix of appetitive and aversive fibres, then increasing the duration for which the aversive-responsible population is fired would be expected to increase the aversive effect. If the longer train increased aversion, this would decrease I_{max} and hence P_e, which is opposite to the effect observed. Disruptive forced movements may well have been induced at the longer duration, which is consistent with the decrease in maximal time allocation. However, if these forced movements increased the effort cost, they would have decreased P_e, which is opposite to what was observed. Therefore, the more likely explanation for the increase in P_e is an increase in the maximal reward value.

If the maximal reward intensity is truly higher at the long train duration than it is at the short train duration, then an easy test of this would be to measure, in a dual-operant experiment, the animal's preference between maximally intense stimulation at each duration. If rat Y13 consistently prefers the maximally intense reward derived from the long train duration to the concurrently-presented maximally intense reward derived from the short train duration, it could be concluded that the maximum intensity of the 1-second train is greater than the maximum intensity of the 0.25-second train. In contrast, no such preference should be detected in the typical case of a subject in which increasing the train duration fails to alter P_e.

Other validations

The present experiment demonstrates that the model can, in practice, correctly detect the effect of manipulations altering the reward circuitry prior to the output of the integrator. It remains to be shown using the current method that a manipulation known to alter the reward circuitry beyond the output of the integrator can be correctly detected. Such a validation (or series of validations) would demonstrate, in conjunction with the current study, that the Mountain model is capable of discerning the stage at which a given manipulation alters the reward circuitry. A number of possible validation experiments that would be expected to shift the Mountain along the price axis are described below.

One experiment has already provided preliminary evidence that movement of the 3D structure along the price axis is possible. Arvanitogiannis (1997) provided background reinforcement contingent on the lever being released. This effectively increases the payoff that animals receive from leisure activities, which should require a corresponding decrease in the price of the stimulation. Since this sort of manipulation does not alter the translation of injected stimulation pulses into reward value, it should not alter the frequency of half-maximal reward. However, the addition of leisure-contingent, background reinforcement also makes the rewards derived from leisure activities more like the rewards the animals receive for lever-pressing. As a result, one would also expect the slope of the behavioural allocation function, A, to become steeper because the substitutability between the payoff from self-stimulation and leisure activities is closer to 1.

Another experiment could probe the effectiveness of the Mountain in discriminating between manipulations altering the reward circuitry prior to the output of the integrator from those altering the system beyond the output of the integrator. Delayed rewards are discounted with respect to immediate rewards (Mazur, 1984); the effect of delaying a brain stimulation reward by even a few seconds would therefore be expected to devalue the intensity of the resulting output of the integrator. In effect, delayed reinforcement should not alter the frequency of half-maximal reward because the integrator only considers the number of action potentials that have reached spatially-distributed input terminals in some temporal window. It is post-integrator output processes that would subsequently decrease the rat's willingness to pay for a given delayed reward intensity rather than an immediately-delivered reward of given intensity. An increase in the delay between completion of a successful harvest and reward delivery would decrease P_e because the entire reward-growth function is expected to shift downwards to a lower maximum as a result of delay discounting. Fouriezos & Randall (1997) have already described the effect of delay on brain stimulation reward; they found that the pulse frequency required to maintain a criterion level of responding had to be increased by approximately 10% for every second the reward was delayed.

The above is true for probability discounting, as well. A reward delivered with probability less-than-one upon completion of the required contingency will be less valuable than a reward delivered every time. In the simplest case, if a reward is delivered only on half the occasions in which an animal has completed the trial work requirement, it will have to devote twice as much time to working in order to harvest an equivalent number of rewards. Even if the animal does not realize the probabilistic nature of the rewards offered to him, the work requirement will have to be reduced by half for said rat to perceive the required performance for probabilistic rewards as equally effortful as the non-probabilistic one. Regardless of the functional form of the probability discounting relationship, it is the resulting reward intensity of the stimulation—as calculated by the integrator and output to the rest of the brain—that is discounted. The spatio-temporal integrator would simply not receive action potentials when the reward is not delivered; it would not receive more action potentials when it is. As a result, a decrease in the probability of reinforcement would shift the 3D structure along the price axis and not along the frequency axis.

In addition to manipulations of the circuitry beyond the output of the integrator by discounting, an equally simple validation would simply increase the work effort required in harvesting rewards. When an animal harvests rewards, it exerts some measurable effort on the lever by holding it down. If that effort requirement is increased by a behaviourally-relevant factor, the exertion-related effort costs increase. Thus, weighting the lever much like the Fouriezos et al. (1990) experiment would impose an additional exertion factor that is combined along with the stimulation price beyond the output of the integrator. Simply put, if it is harder for the rat to hold down a bar to self-administer brain stimulation, it will alter how willing it is to bar-press for that stimulation without changing the effectiveness of each stimulation pulse in giving rise to a subjective reward intensity. As a result, P_e would decrease when greater exertion is required of the rat. This experiment would have the additional benefit of pinning down the stage of processing responsible for the effects demonstrated in the Fouriezos et al. experiment. Recall from the introduction that the experiment produced shifts in the threshold of rate-frequency curves by making the lever increasingly more effortful to press. Using the Mountain model to replicate these findings would expose how powerful this paradigm is in disambiguating the effect of manipulations to reward circuitry.

A final, more conceptually-challenging validation involves directly altering the maximum reward value derived from self-stimulation. An increase in the maximum intensity of the reward circuitry would increase P_e. If P_e is the price at which a maximal intensity reward produces half-maximal time allocation, increasing the maximum intensity will increase P_e in direct proportion. Concurrent stimulation of lateral hypothalamic and prefrontal sites or bi-hemispheric lateral hypothalamic stimulation would produce such a change in maximal intensity. There is good evidence (Schenk & Shizgal, 1985) that the lateral hypothalamic level of the medial forebrain bundle and the prefrontal cortex project to different sites of integration with different growth properties. However, little is known about both the interaction between prefrontal and lateral hypothalamic stimulation and bilateral medial forebrain bundle sites of brain stimulation reward. The interpretation of these results would therefore be complex and require a greater understanding of the multiple integrators within the brain. Although investigators have provided evidence for interhemispheric summation- and collision-like effects (Miliaressis & Malette, 1995), it would be necessary to further characterize the interactions of multiple self-stimulation regions before designing an experiment targeted at changing the maximum reward intensity.

Why changing stimulation current is a challenging validation

The model validations described above would all be expected to shift P_e along the price axis. Although it would expose the true versatility of the model, I do not propose other validations of the model that would shift f_{hm}, as has been produced in this thesis and others (Arvanitogiannis, 1997; Mullett, 2005). The simple counter model posits that the integration process summates the action potentials that reach it in some temporal and spatial window. Would it not, therefore, be possible to produce a shift in f_{hm} by altering the number of primary reward fibers activated? In other words, is another perfectly acceptable validation of the model possible if one changes the stimulation current?

Indeed, if a single integrator were to count all the action potentials that arrive at the terminals of the primary reward cable, a two-fold decrease in current intensity should require a two-fold increase in pulse frequency. However, other evidence (Arvanitogiannis, Waraczynski & Shizgal, 1996; Murray & Shizgal, 1996) suggests that the reward cable reaches multiple integration processes; that it is, in essence, a heterogeneous neuronal population. In a homogeneous population with a unitary-integrator, a halving of the current would require a two-fold increase in pulse frequency, regardless of the rate at which reward intensity grows with stimulation strength. In contrast, manipulating stimulation current in a heterogeneous population subservient to multiple integrators with different reward-growth exponents would result in a much more complex current-frequency tradeoff. In this case, stimulation of a greater number of primary reward fibres could potentially activate multiple integrators with different integration characteristics whose interaction is as yet unknown. The lesion work produced and cited by Arvanitogiannis et al. (1996) appears to support the hypothesis that the primary reward substrate is heterogeneous and projects to multiple integrators. The largest shifts in required frequency—the position of the two-dimensional representation relating response rate to pulse frequency—were observed when the current was low. A low current is more likely to activate only one integrator in a heterogeneous pathway than a high current, and thus, compensation for the cell loss by other integration processes with different reward-growth rates is unlikely. Many investigators have argued for the presence of heterogeneity in the substrate responsible for self-stimulation in the MFB, including Murray & Shizgal (1996), Arvanitogiannis et al. (1996) and more recently Carr (2002), Waraczynski (2006), and Fulton, Woodside & Shizgal (in press). If the primary reward fibres indeed compose a heterogeneous substrate, changes in the intensity of the current are likely to affect not only the current-frequency trade-off, but would also recruit multiple integrators with different reward-growth properties. The inclusion of other recruited integrators by increasing the current intensity would alter the maximum reward intensity of electrical brain stimulation. In other words, it is possible that increasing the number of fibres activated by increasing the stimulation current would alter the maximal reward intensity by altering the number of different integrators recruited. An increase in the maximal reward intensity would change the price at which a maximal reward produces half-maximal time allocation; ergo, the increased current would also increase P_e. When Arvanitogiannis (1997) manipulated the intensity of the stimulation current, one subject showed a large (0.494 common logarithmic units) statistically significant increase in P_e following a two-fold increase in current.[2] This is consistent with the notion that the substrate for self-stimulation is heterogeneously distributed. According to the model, an increase in P_e would require either the perceived exertion to decrease, the payoff from everything else to become less attractive, or the maximum intensity of the rewarding effect to increase. Since the effort required did not change and the payoff of everything else was presumably constant across all conditions, the only variable that could change is the maximal intensity of the rewarding effect of brain stimulation. Manipulations of current have been attempted before (Arvanitogiannis, 1997), but the analysis of the results of such a manipulation would be complicated by heterogeneity in the reward pathway.

Re-interpreting the effect of drugs, lesions and physiological manipulations

The Mountain model provides a comprehensive and relatively parsimonious account of performance for rewarding brain stimulation. The purpose of such a model is not to simply account for performance, but rather, to infer the stage of processing at which various components of the reward circuitry contribute their effects. As such, a validated model would allow us to distinguish manipulations that alter the reward circuitry prior to the output of the integrator from those that affect the brain beyond the output of the summation process. It should therefore be possible to tie, albeit in a relatively crude way, the effect of drugs, lesions and physiological manipulations to two distinct stages of processing.

For example, it is known that drugs like cocaine and amphetamine increase dopaminergic tone. If administration of cocaine shifts the mountain along only the frequency axis, then the effect of cocaine is to alter neurotransmission prior to the output of the integrator. In this case, dopamine tone would modulate reward circuitry at the level of the integrator, possibly by increasing the effectiveness of each stimulation pulse in contributing to a given level of reward. If, on the other hand, dopaminergic tone affected perceived exertion costs, as is proposed by Salamone et al. (2005), cocaine administration would be expected to shift the mountain only along the price axis. A validated mountain model would, in theory, arbitrate the two opposing theories of the psychological functions of dopamine release.

Lesion studies would also benefit from a validated Mountain model. The medial forebrain bundle is a heterogeneous collection of axons that course rostro-caudally through the midbrain. The identity of the neurons primarily responsible for the rewarding effect of medial forebrain bundle self-stimulation is not yet known. A lesion to their cell bodies would deteriorate their projecting axons, resulting in a greatly reduced effectiveness of electrical stimulation. An additional criterion for the identity of the neurons whose axons comprise the primary reward cable would be that the destruction of their cell bodies—potentially in the sub-lenticular extended amygdala—should result in a shift of only the frequency of half-maximal reward. Conversely, structures contributing to reward processing beyond the output of the integrator would alter P_e.

Use of the Mountain model is not restricted to manipulations to the brain circuitry as overwhelming or devastating as drug administration and structure ablation. The circuitry responsible for brain stimulation reward has been implicated in the rewarding effects of natural goals. The stage of processing of natural rewards could be grossly identified by physiological manipulations. It is known that certain prosencephalic sites within the medial forebrain bundle are sensitive to the animal's long-term energy balance. These sites also likely project to an integration process distinct from the processes that integrate the rewarding effect of sites that are not modulated by long-term food restriction. The effect of the long-term depletion of energy stores could act on the integrator prior to its output, in which case, chronic food restriction would decrease f_{hm}. Fat depletion may also modulate the output of long-term energy store-specific integration processes, the willingness to engage in effortful performance, or the value of competing activites, in which case, long term food deprivation would increase P_e. The Mountain model would be capable of tying all kinds of physiological manipulations to the stage at which they contribute to reward processing.

General Discussion

The present thesis replicates findings from Mullett (2005) and Arvanitogiannis (1997) while correcting for biases that their work helped reveal. When animals are exposed to repeating trials in which the price is, on average, relatively high, their behaviour appears to change systematically over time. Although discrepant performance for combinations of price and pulse frequency has been documented before, this finding has always been inconsistent across rats: animals all appear to be crazy in their own way. This thesis is the first to uncover a finding that is virtually identical in all cases. In every rat tested, performance for equivalent combinations of price and frequency on trials in which the price is high and frequency is decreased from trial to trial is higher than conditions in which the frequency is high and price is increased from trial to trial.

This finding suggests that animals may adapt to the long-term reinforcement contingencies they face when the amount of time they must invest to earn a reward is relatively high. It is evident from this thesis and previous work (Breton, 2004; Marcus, 2005) that a rat's decision to press for rewarding brain stimulation depends not only on the pulse frequency and price it encounters, but also its long-term reinforcement history. The animal's preferences, just as has been known for some time in the human literature, are constructed during elicitation rather than simply revealed.

The true cost of high price

Experiment 1 demonstrates that exposure to long periods of high average prices alter the animal's entire pattern of responding for rewarding brain stimulation. Perhaps there exists some flexible anchor for the scale of rewards that alters the sensitivity of the system depending on the overall, long-term state of the world. Although it may seem counterintuitive for an animal to perform more vigorously for stimulation of a high price when it has been in such an environment for a long period of time, ecologically speaking, this result makes sense. In a world where the average opportunity cost of foraging is high over a long period of time and the food patch is relatively sparse, an animal cannot afford to eschew searching for food until the cost of acquiring it comes down. Since the animal will starve if it does not eat even high-cost foods, it may adjust its preference to include some of the lower-quality foods in its diet, thereby lowering the threshold at which it performs for food of various rewarding intensities. As a result, under these long-term conditions of high opportunity cost, an animal may become less sensitive to changes in food quality, thus allocating more time to foods it would ignore if their opportunity cost were low.

Compare the above scenario to the brain stimulation reward scenario in which animals have the opportunity to harvest decreasingly rewarding stimulation in conditions of long-term low reinforcement density. Stimulation of very low frequency remains unaffected by the increased price. Stimulation of intermediate frequency undergoes a bias in an upward direction: an intermediate pulse frequency may be acceptable if the price of a successful harvest will be highly costly for a long period of time. During price sweeps, the bias may be in an opposite direction: if highly appetitive stimulation becomes increasingly costly from trial to trial, it may result in an artificially lower breakpoint for self-stimulation performance. If the animal will be able to acquire the high-quality food at a low cost at some point in the series of increasing prices, it may reserve its responding for when the price is low and omit responding when it is higher and steadily escalating, thereby lowering the breakpoint. As a result, the pattern of responding for various pulse frequencies becomes inconsistent with the pattern of responding for various stimulation prices because the overall, long-term situation the rat is in is not the same in the two conditions.

Stability and variability

Randomized-trials designs for measuring self-stimulation performance have been shunned in the past because they lead to greater variability in behaviour. Indeed, the variability of the animal's time allocation to self-stimulation activities in sweep conditions of experiment 1 is much smaller than that of randomized conditions (compare upper and bottom right panels of figures 12 through 18). The stability of performance potentially reflects the animals' anticipation of future payoffs, which may ultimately alter performance for a particular type of sweep. In situations where the animal experiences a decreased pulse frequency from trial to trial and is well-informed about the price of this stimulation, it can come to predict the payoff it would receive on a subsequent trial from self-stimulation.

It is not uncommon for investigators to observe an increase in responding to the lowest pulse frequency used in a repeating series of frequency sweep trials, when animals are well-trained. This suggests an anticipation of the temporal distribution of the reward intensities derived from electrical stimulation: because animals tend to respond a certain way, the temporal distribution of the high pulse frequency trials appears at approximately similar times within a session. If animals evaluated only the stimulation price and pulse frequency on any given trial, the precise temporal distribution of high-frequency stimulation would not change the evaluation of the payoff on the current trial.

It is likely the combination of predictability and high price that produces the noteworthy inconsistencies in performance for stimulation trains of the same strength and price depending on whether this point in the parameter space is visited in the course of a price or frequency sweep. When the average price is high for a long period of time, and so predictably so that the animal cannot hope that the price will be lower considering the information available, the criterion for "acceptable" shifts to accommodate the leaner reinforcement density. When the average price is changed again throughout price sweeps, this criterion shifts with it. Although the long-term stability of the session structure may contribute to reducing variability in the data, one cannot assume that the rat fails to note this long-term stability and to use this information to adjust its strategy. The results of experiment 1 demonstrate that animals do learn, in fact, more than the simple response-reinforcement contingency. There is additional learning occurring on their part of the long-term conditions they are in, consequently resulting in a long-term adaptation of their behaviour to the average price. Randomized-trials designs result in more variable performance, but offer the advantage of presenting the rat with stable average conditions throughout the entire experiment.

Technical issues concerning the randomized-trials design

The randomized-trials design is ideal for studies in which both the strength and cost of the reward must be varied. The relationship between performance, stimulation strength and price is much more easily observed because the conditions throughout the entire experiment are relatively stationary over a long period of time. A randomized-trials approach removes systematic influences of reinforcement history on performance. As a result, the relationship between time allocation, pulse frequency and price is not influenced by other factors that may distort the relationship we are trying to describe.

Despite its many virtues, the particular implementation of the randomized-trials design used in the randomized portion of experiment 1 and all of experiment 2 is flawed in the sense that animals receive one piece of information before the trial even begins. Priming stimulation delivered 2 seconds before the onset of the trial informs the animal about the frequency of stimulation that will be delivered throughout the trial in question. Since the animal knows the pulse frequency of the stimulation before the trial begins, it is free to spend a large portion of the trial involved in leisure activities before sampling from the lever if the pulse frequency is sufficiently low. The only price information available that can truly guide its behaviour exists after the first reward encounter, since it is physically impossible for the animal to know the price of the stimulation before it has earned one reward. That the animal knows the frequency of stimulation before this first reward encounter makes analysis of the animal's interaction with the bar complex. It is possible that animals wait a very long period of time—potentially until the very end of the trial—before sampling the bar when the frequency is low. The pulse frequency may be so low it is barely perceptible, such that the animal waits a relatively long period of time before bar-pressing. Without taking too many liberties anthropomorphizing, the animal may sample after a long period to know whether the stimulation is truly as weak as it has been presented. This wait has been observed, and artificially inflates the resulting time allocation. If the animal harvests a reward requiring a 1 second hold when there are 2 seconds left to the trial, having waited 23 seconds before pressing, further presses will greatly inflate the resulting time allocation once performance during the first reward encounter has been removed. With 2 seconds remaining to the trial, a 1-second press will result in 50% time allocation to work activities when not counting the first reward encounter, but 8% (using the trial times used in experiments 1 and 2) time allocation overall.

Two simple corrections to this problem would involve simply eliminating the priming stimulation altogether and resetting the trial clock following the successful harvest of a single reward. The former correction would provide the animal with no information whatsoever before it harvests its first reward, rather than biasing the lever-pressing behaviour in complex ways. The latter would provide an equivalent time window for all first encounter-stripped estimates of time allocation. It would be impossible for time allocation to grow artificially as the result of the removal of irrelevant information.

A plot of the correlation between first encounter-stripped time allocation and overall corrected time allocation is shown in figure 47 for rat Y12 in the train duration experiment. The correlation points are additionally coded according to the proportion of trial time that has elapsed before the animal harvests its first reward. In the few instances in which low overall corrected time allocation corresponds to high first-encounter stripped time allocation, they result from the first reward being harvested after a large portion of the trial has already elapsed. Fortunately for the results of this thesis, the result is, by and large, an artefact of the leading and trailing bracket trials. On trailing bracket trials, in which stimulation frequency and price are low and produce floor responding, animals wait until the end of the trial to sample from the lever.

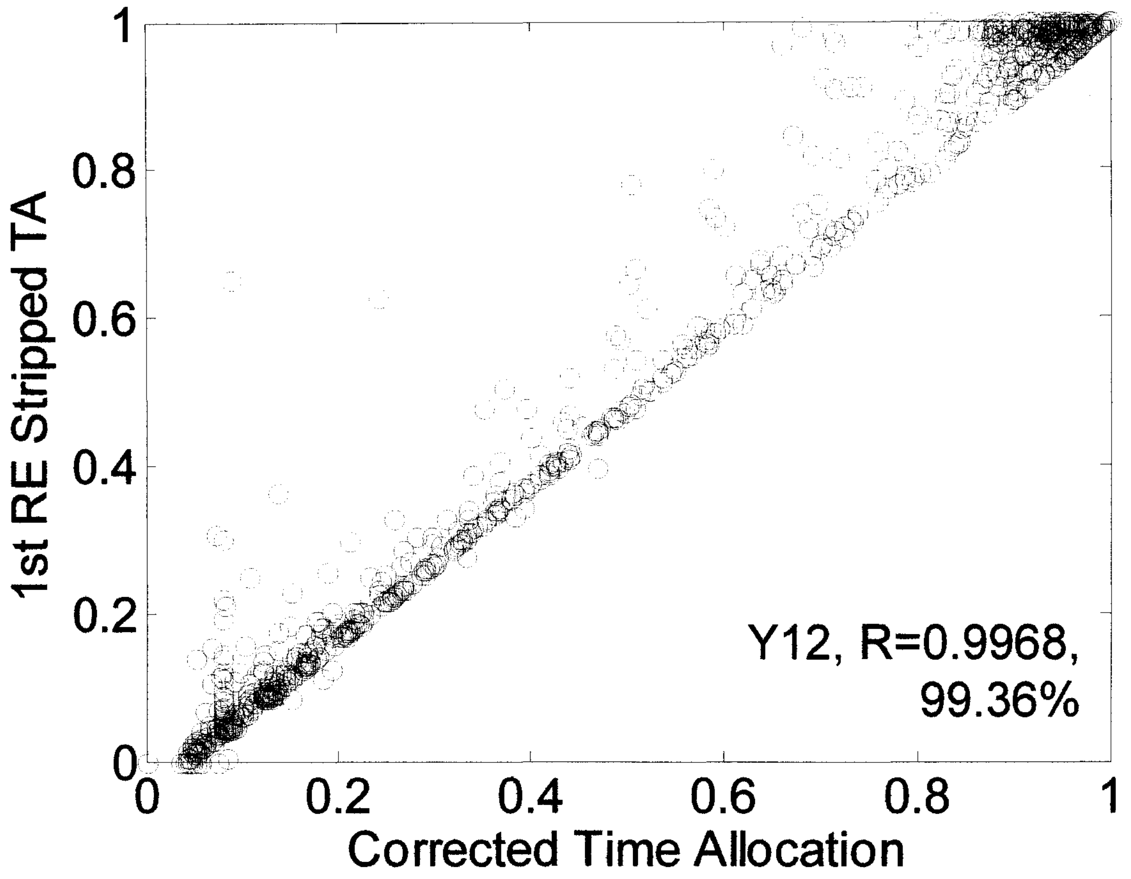


Figure 47. First reward encounter-stripped time allocation as a function of time allocation including first reward encounter. There is an almost-perfect relationship between the first encounter-stripped and first encounter-included time allocations. Deviations from this lawful relationship at the low end can be attributed to responses emitted at the end of the trial, when over 90% of the trial has elapsed before the first reward-encounter. Results are typical of all animals tested.

The additional advantage of removing the priming stimulation would be to further investigate the effect of predictably high- and low-payoff trials on performance for rewarding brain stimulation in the randomized-trials design. If every trial is preceded by a leading bracket trial of high frequency, low price rewards and followed by a trailing trial of low frequency low priced rewards, then it should be possible to determine whether animals can anticipate these trials. In effect, animals for which this trial structure would not substantially drive behaviour would necessarily have to sample the lever during the trailing bracket trials in order to derive the information they need to make a decision. If the rats can essentially count to three, their first encounter with the reward—indeed with the bar—will be much more protracted in the trial, if at all.

A robust model

That the Mountain model has yielded the consistent results that it has on a macro-level is quite telling. The model appears to be relatively robust, because even though certain assumptions used in the Arvanitogiannis (1997) and Mullett (2005) experiments have since been found to be untenable, the Mountain provided a reasonably good account of performance for BSR in these studies. Despite the unanticipated problems documented here, Arvanitogiannis (1997) and Mullett (2005) have nonetheless demonstrated the capacity of the Mountain model to discern the stage at which a manipulation affects the reward circuitry. The current thesis demonstrates that even when one removes assumptions about operant tempo and influences of long-term history on performance, the model is capable of correctly detecting the effect of a manipulation altering the circuitry prior to the output of the integrator. The randomized-trials design used in this thesis further frees the model from potential caveats about the inferred contribution of a manipulation to reward processing by reducing the influence of long-term reinforcement history.

Data from past validations of the model (Mullett, 2005) have been difficult to analyze. The sweep conditions inherent in the design this experiment resulted, in some cases, in frequency "sweeps" obtained at high prices that appeared un-shifted from those obtained at lower prices. Clear shifts were seen in some subjects; the failures were observed when rotations were seen rather than shifts. This effect was reproduced in many cases (Marcus, 2005): the locus of rise of frequency "sweeps" obtained at high prices was not substantially different from those obtained at lower prices. If a similar process of adaptation transpires in delayed rewards as takes place at high prices, a randomized-trials design should, in theory, resolve any issues related to correctly identifying this manipulation as acting beyond the output of the integration process.

The power of the model is also necessary for it to be a useful tool in tying manipulations to two distinct stages of processing. The Mountain model appears to be relatively powerful, because the accuracy of the estimates is directly related to the capacity to detect small effects. If statistical power is the probability of finding an effect provided there exists one in the population, very accurate estimates (i.e., estimates surrounded by very narrow confidence bands) would lead to the detection of very small effects. Considering that certain manipulations produce relatively small shifts in threshold in the rate-frequency realm, it is imperative that a replacement be capable of detecting these small changes. It is true that the confidence intervals surrounding position parameters in experiment 2 were relatively large (in some cases up to 0.14 common logarithmic units); the confidence intervals would almost inevitably shrink with a greater amount of data. In experiment 1, the large quantity of data that were collected resulted in small confidence intervals surrounding estimated position parameters. In other words, the randomized-trials design may not of its very nature degrade the capacity of the model to detect an effect. Instead, the increased variability of the data can easily be compensated for by obtaining more than 8 estimates of ceiling- and floor-asymptotic data and 16 estimates of sloping data on the 3D surface. An even more straightforward means of increasing the accuracy of fitted parameter estimates would be to obtain pseudo-sweeps at multiple angles to the principal axes of the parameter space (rather than the three used here). This increase would provide a great deal more information about the shape of the surface, and therefore decrease uncertainty about the position of the 3D surface along the price and frequency axes.

According to the model of performance considered here, time allocation reflects a molar estimate of payoff and costs. This estimate, in many cases, is an overall assessment of multiple interactions with the lever. If one were to consider the animal's behaviour on a much more molecular level, taking into account each reward encounter as a distinct event in the animal's history, it may be possible to increase the statistical power of the model without devoting one's lifetime to a single experiment. At the upper asymptotes, when the animal receives up to 25 rewards in a given trial, the amount of data contributing to a mean time allocation at that point on the 3D surface would increase by 25-fold. Along the sloping portions, when an animal harvests 10 or 15 rewards, the amount of data would increase proportionally. It is only at the lower asymptotic range, when animals harvest no rewards that the amount of data contributing to the mean time allocation would not change. It is highly probable that the drastic increase in sample size of most price-frequency pairs tested would more than compensate for encounter-to-encounter variability in behaviour. The increased power could potentially allow the model to be sensitive enough to detect even a 10% difference in the frequency of half-maximal reward or price at which a maximal reward produces half-maximal time allocation. Even considering all these issues, the confidence interval surrounding the estimated position parameters in some cases was rather small (as little as 0.06 common logarithmic units). The confidence interval width in these cases was smaller even than was found in Mullett (2005) when using the stabilizing sweep method of eliciting self-stimulation performance, suggesting that even without a molecular analysis of performance, the model is capable of detecting small shifts of the 3D surface along the frequency axis.

A framework for studying natural rewards

The Mountain model is not simply a tool for understanding how various components of brain reward circuitry contribute to the decision-making process. Although it will be extremely useful in tying the different experimenter-controlled manipulations to the stage of processing they contribute to, the model also opens the door to a discussion of which stage of processing natural states exert their effects. In essence, it provides a framework with which to study how animals' time allocation to natural rewards is affected by physiological and environmental states. For example, hunger could act prior to the output of an integrator computing the reward intensity of food. In this case, the effect of hunger is to alter the relative worth of foods of different quality, such that a simple pretzel is attributed the same reward intensity as a filet mignon. Foods that were attributed different reward intensities in satiety situations, along the sloping portion of the reward-growth function, may suddenly become equally intensely rewarding if hunger shifts the function leftward. Alternately or in addition, hunger could act beyond the output of an integrator computing the reward intensity of food. In this case, the effect of hunger is to alter one's willingness to pay for foods without altering the relative strengths of their rewarding effects: the pretzel may not be as satisfying as the steak, but you are more willing to acquire any food regardless of its rewarding properties.

This simple example is one of many that demonstrate that the Mountain model provides a comprehensive framework for understanding the ways in which choices are made in a natural setting. Changes to reward circuitry occurring prior to and beyond the output of the integrator alter all animals' behavioural allocation to a wide variety of natural goal objects. The phenomenon of brain stimulation reward provides direct access to reward-relevant circuitry in the brain, allowing researchers to directly study how the brain makes choices in the laboratory. Ultimately, the goal of research on appetitive motivation is to understand the processes by which the brain evaluates and decides outside the laboratory. The contribution of the Mountain model toward understanding those processes underlying this laboratory phenomenon—while taking into account that rats adapt to the long-term conditions of their environment—may well shed light on how animals and humans make decisions in natural contexts.

Endnotes

1. To visualize this, one can imagine a mirror image of the depiction in figure 1B (a theoretically-derived mountain) rotated 90 degrees clockwise.
2. The actual estimate obtained, Re_{min}, is the rate of reinforcement at which a maximally-intense brain stimulation reward yields 50% time allocation. Since the rate of reinforcement is inversely related to the price when an animal allocates all of its time to self-stimulation, the observed decrease in Re_{min} translates into an increase in P_e.

References

Ariely, D., Loewenstein, G. & Prelec, D. (2003). Coherent arbitrariness: Stable demand curves without stable preferences. *Quarterly Journal of Economics, 118*, 73-106.

Arvanitogiannis, A. (1997). *Mapping the substrate for brain stimulation reward: New approaches to an old problem*. Unpublished doctoral dissertation, Concordia University, Montreal, Quebec, Canada.

Arvanitogiannis, A., Waraczynski, M. & Shizgal, P. (1996). Effects of excitotoxic lesions of the basal forebrain on MFB self-stimulation. *Physiology & Behavior, 59*, 795-806.

Bauco, P. & Wise, R. (1997). Synergistic effects of cocaine with lateral hypothalamic brain stimulation reward: Lack of tolerance or sensitization. *The Journal of Pharmacology and Experimental Therapeutics, 283*, 1160-1167.

Breton, Y.A. (2004). *Using the reference-sweep method to infer the subjective value of rewarding brain stimulation in the rat*. Unpublished undergraduate honours thesis, Concordia University, Montreal, Quebec, Canada.

Breton, Y.A., Conover, K. & Shizgal, P. (2006). *Mountaineering in rats: A 3D model of performance for rewarding brain stimulation*. Poster session presented at the XXVIII international symposium on computational neuroscience, Université de Montréal, Montreal, Quebec, Canada.

Carr, K.D. (2002). Augmentation of drug reward by chronic food restriction: Behavioral evidence and underlying mechanisms. *Physiology & Behavior, 76*, 353-364.

Clark, J.J. & Bernstein, I.L. (2006). Sensitization of salt appetite is associated with increased "wanting" but not "liking" of a salt reward in the sodium-deplete rat. *Behavioral Neuroscience, 120*, 206-210.

Conover, K.L., Fulton S. & Shizgal, P. (2001). Operant tempo varies with reinforcement rate: implications for measurement of reward efficacy. *Behavioural Processes, 56*, 85-101.

Conover, K.L. & Shizgal, P. (1994). Differential effects of postingestive feedback on the reward value of sucrose and lateral hypothalamic stimulation in rats. *Behavioral Neuroscience, 108*, 559-572.

Conover, K.L. & Shizgal, P. (1994). Competiton and summation between rewarding effects of sucrose and lateral hypothalamic stimulation in the rat. *Behavioral Neuroscience, 108*, 537-548.

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*, 549-558.

Efron, B. & Tibshirani, R. (1993). *An Introduction to the Bootstrap.* New York: Chapman & Hall/CRC Press.

Fouriezos, G. & Randall, D. (1997). The cost of delaying rewarding brain stimulation. *Behavioural Brain Research, 87*, 111-113.

Fouriezos, G., Bielajew, C. & Pagotto, W. (1990). Task difficulty increases thresholds of rewarding brain stimulation. *Behavioural Brain Research, 37*, 1-7.

Franklin, K. (1978). Catecholamines and self-stimulation: Reward and performance effects dissociated. *Pharmacology, Biochemistry and Behavior, 9*, 813-820.

Fulton, S., Woodside, B. & Shizgal, P. (in press). Potentiation of brain stimulation reward by weight loss: Evidence for functional heterogeneity in brain reward circuitry. *Behavioural Brain Research.*

Gallistel, C.R. & Leon, M. (1991). Measuring the subjective magnitude of brain stimulation reward by titration with rate of reward. *Behavioral Neuroscience, 105*, 913-925.

Gallistel, C.R. (1978). Self-stimulation in the rat: Quantitative characteristics of the reward pathway. *Journal of Comparative and Physiological Psychology, 92*, 977-998.

Gallistel, C.R., Mark, T.A., King, A.P. & Latham, P.E. (2001). The rat approximates an ideal detector of changes in rates of reward: Implications for the law of effect. *Journal of Experimental Psychology: Animal Behavior Processes, 27*, 354-372.

Goldshmidt, J.N. & Fantino, E. (2004). Economic context and pigeons' risk-taking: an integrative approach. *Behavioral Processes, 65*, 133-154.

Green, L. & Rachlin, H. (1991). Economic substitutability of electrical brain stimulation, food and water. *Journal of the Experimental Analysis of Behavior, 55*, 133-143.

Herrnstein, R.J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior, 13*, 243-266.

Herrnstein, R.J. (1974). Formal properties of the matching law. *Journal of the Experimental Analysis of Behavior, 21*, 159-164.

Hoaglin, D.C. (1983). Understanding Robust and Exploratory Data Analysis. In D.C. Hoaglin, F. Mosteller & J.W. Tukey (Eds.), Understanding Robust and Exploratory Data Analysis. New York: Wiley.

Hodos, W. & Valenstein, E.S. (1962). An evaluation of response rate as a measure of rewarding intracranial stimulation. *Journal of Comparative and Physiological Psychology, 55*, 80-84.

Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263-292.

Konkle, A.T.M., Bielajew, C., Fouriezos, G. & Thrasher, A. (2001). Measuring threshold shifts for brain stimulation reward using the method of limits. *Canadian Journal of Experimental Psychology, 55*, 253-260.

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*, 265-277.

Marcus, J. (2005). *Does rattus economicus really exist? Eliciting preference reverals in rats*. Unpublished undergraduate honours thesis, Concordia University, Montreal, Quebec, Canada.

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*, 389-401.

Mazur, J.E. (1984). Fixed and variable ratios and delays: further tests of an equivalence rule. *Journal of Experimental Psychology: Animal Behavior Processes, 12*, 116-124.

Miliaressis, J. & Malette, E. (1995). Interhemispheric links in brain stimulation reward. *Behavioural Brain Research, 68*, 117-137.

Miliaressis, E., Rompré, P.P., Laviolette, P., Philippe, L. & Coulombe, D. (1986). The curve-shift paradigm in self-stimulation. *Physiology and Behavior, 37*, 85-91.

Mullett, A. (2005). *Linking changes in performance for brain stimulation reward to stages of neural processing.* Unpublished Master's thesis, Concordia University, Montreal, Quebec, Canada.

Mullett, A., Conover, K. & Shizgal, P. (2004). *Linking changes in performance for brain stimulation reward to stages of neural processing.* Poster session presented at the annual meeting of the Society for Neuroscience, San Diego, CA.

Murray, B. & Shizgal, P. (1996). Physiological measures of conduction velocity and refractory period for putative reward-relevant MFB axons arising in the rostral MFB. *Physiology and Behavior, 59*, 427-437.

Navarro, A.D. & Fantino, E. (2005). The sunk cost effect in pigeons and humans. *Journal of the Experimental Analysis of Behavior, 83*, 1-13.

O'Daly, M., Case, D.A. & Fantino, E. (in press). Influence of budget and reinforcement location on risk-sensitive preference. Behavioural Processes.

Olds, J. & Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. In Isaacson, R.L. (Ed.), *Basic readings in neuropsychology* (pp. 332-346). New York, Harper & Row Publishers.

Phillips, A.G. & LePiane, F.G. (1986). Effects of pimozide on positive and negative incentive contrast with rewarding brain stimulation. *Pharmacology, Biochemistry and Behavior, 24*, 1577-1582.

Salamone, J.D., Correa, M., Mingote, S.M. & Weber, S.M. (2005). Beyond the reward hypothesis: Alternative functions of nucleus accumbens dopamine. *Current Opinion in Pharmacology, 5*, 34-41.

Salamone, J.D., Wisniecki, A., Carlson, B.B & Correa, M. (2001). Nucleus accumbens dopamine depletions make animals highly sensitive to high fixed ratio requirements but do not impair primary food reinforcement. *Neuroscience, 105*, 863-870.

Schenk, S. & Shizgal, P. (1985). The substrates for self-stimulation of the lateral hypothalamus and medial prefrontal cortex: A comparison of strength-duration characteristics. *Physiology and Behavior, 34*, 943-949.

Shizgal, P. & Matthews, G. (1977). Electrical stimulation of the rat diencephalons: Differential effects of interrupted stimulation on on- and off-responding. *Brain Research, 129*, 319-333.

Shizgal, P. (1999). On the neural computation of utility: implications from studies of brain stimulation reward. In D. Kahneman, E. Diener & N. Schwarz (Eds.), *Well-Being: The Foundations of Hedonic Psychology*. New York: Russell Sage Foundation.

Shizgal, P. (2004). Proposal submitted to the Canadian Institute of Health Research (CIHR) on September 15.

Simmons, J.M. & Gallistel, C.R. (1994). Saturation of subjective reward magnitude as a function of current and pulse frequency. *Behavioral Neuroscience, 108*, 151-160.

Simonsohn, U. & Loewenstein, G. (2006). Mistake #37: The impact of previously observed prices on housing demand. *The Economic Journal, 116*, 175-199.

Smith, J. (2005). *A single reference point is insufficient at eliminating behavioural drift.* Unpublished undergraduate project, Concordia University, Montreal, Quebec, Canada.

Sonnenschein, B. (2004). *Evaluation of brain stimulation reward in rats: Heuristics, exemplars and context-dependency.* Unpublished doctoral dissertation, Concordia University, Montreal, Quebec, Canada.

Sonnenschein, B., Conover, K. & Shizgal, P. (2003). Growth of brain stimulation reward as a function of duration and stimulation strength. *Behavioral Neuroscience, 117*, 978-994.

The Mathworks. (2006). *Optimization Toolbox User's Guide* (Version 3). Natick, MA: The Mathworks.

Kling, J.W. & Riggs, L.A. (Eds.) (1971). *Experimental Psychology* (3rd Ed.). New York: Holt, Rinehart & Winston.

Tversky, A. & Thaler, R.H. (1990). Anomalies: Preference reversals. *Journal of Economic Perspectives, 4*, 201-211.

Waraczynski, M. (2003). Lidocaine inactivation demonstrates a stronger role for central versus medial extended amygdale in medial forebrain bundle self-stimulation. *Brain Research, 962*, 180-198.

Waraczynski, M.A. (2006). The central extended amygdale network as a proposed circuit underlying reward valuation. *Neuroscience & Biobehavioral Reviews, 30*, 472-496.

Appendices

Appendix A

Statistical Appendix

Tukey's bi-square estimator

Tukey's bi-square estimator was used because of the presence of heteroskedasticity, skew, and outliers. Heteroskedasticity is evident from the higher variance of data obtained along the sloping portions. Conditions throughout extended testing are also not likely completely stationary, which may account for the presence of outliers. The weighting procedure assigns lower weight to observations further away from the median; using a tuning constant of 6 results in an assignment of 0 weight to scores over 4 standard deviations above the mean in a normally-distributed dataset. A detailed description of this statistical technique can be found in Hoaglin (1983). A more rigorous argument for the use of bi-square-weighting is presented later.

Bootstrapping procedure

There is *a priori* reason to believe the data are not normally distributed. Time-allocation data are proportions and hence can vary only between 0 and 1. Thus, such data are not normally distributed. It is common for software packages to estimate confidence intervals for a parameter of a non-linear function from the steepness of the associated "loss function." (The Mathworks, 2006). Intercorrelation of parameter estimates is often ignored, resulting in unrealistically small confidence intervals. The fitting procedure poses a problematic distortion of the surface, because non-trivial correlations among the parameter estimates artificially reduces the width of 95% confidence intervals surrounding those estimates. These correlations are documented further for the current dataset (appendix B), and similar parameter inter-correlations were found in Mullett

(2005). The bootstrapping procedure was used in order to estimate the distribution of parameter values directly, forgoing assumptions of normality and absence of inter-correlation between parameter estimates. The 95% confidence surrounding mean time allocations at each price-frequency pair, estimated fitted mountain parameters and mountain parameter differences were empirically calculated with a bootstrapping procedure (a full treatment of this procedure and its theoretical foundations can be found in Efron & Tibshirani, 1993).

This procedure consisted of randomly sampling observations 1000 times with replacement from the data set, with a sample size equal to the number of observations that contributed to the original estimate of central tendency and spread. If 8 observations contributed to the mean at one price-frequency pair, the procedure would randomly sample 8 times with replacement from those 8 weighted observations and calculate the sample's mean, and repeat this process 1000 times. The confidence interval was then estimated by observing the points above and below which 2.5% of the means were distributed. Such a bootstrapping (re-sampling) procedure calculates confidence based on empirical observations, thus reducing the number of assumptions required in computing the interval. The confidence interval resulting from a bootstrapping procedure is free of distributional assumptions.

Since by definition loss functions of non-linear functions do not necessarily have global minima and maxima, initial seeds are required for least-squares fitting to find a local minimum in the loss function by iteration. The seeds used for the fitting of all pseudo-sigmoids were identical for all animals. In the case of price sweeps, the initial asymptote (Q1A) was entered as 0.95, the position of the first knot (K1) was entered as 1, the growth of the first (Q1G) and second (Q2G) quadratics were entered as 20 and the lower asymptote (Q2A) was entered as 0.1. In the case of frequency sweeps, the first asymptote (Q1A) was entered as 0.1, the position of the first knot was entered as 2, and the upper asymptote (Q2A) was entered as 0.95, while both growth parameters (Q1G and Q2G) were 20. The equation fit to the data is presented below, in equation A1. The seeds

$$TA = \begin{cases} x \leq K1 & Q1A \\ K1 \leq x < K2 & \mathrm{R \times (Q1G \times (x - K1)^2 - 0.5) + A_{min}} \\ K2 \leq x < K3 & \mathrm{R \times (0.5 - Q2G \times (x - K3)^2) + A_{min}} \\ K3 \leq x & Q2A \end{cases} \tag{A1}$$

$$K3 = \mathrm{K1} + \frac{\sqrt{\mathrm{Q1G \times (Q1G + Q2G)}}}{\mathrm{Q1G} \times \sqrt{\mathrm{Q2G}}}$$

$$K2 = \frac{\mathrm{K1 \times Q1G + K3 \times Q2G}}{\mathrm{Q1G + Q2G}}$$

$$R = |Q1A - Q2A|$$

$$A_{\min} = \frac{R}{2} + \min(Q1A, Q2A)$$

where TA is the fitted time allocation,

x is the independent variable, price or frequency,

K1, K2, K3 are the positions of the first, second and third knots,

Q1A, Q2A are the upper and lower asymptotic time allocations,

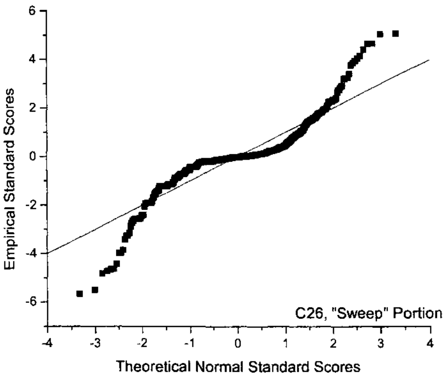
Q1G, Q2G specify the growth of the quadratic functions.

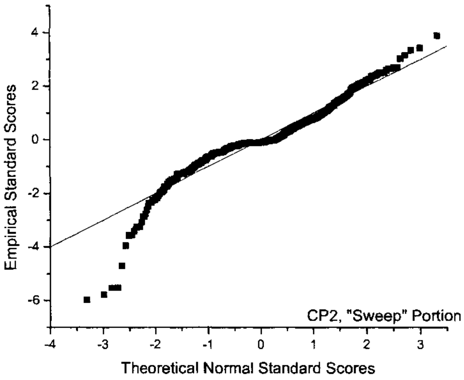
used for the fitting of all Mountain Model surfaces were estimated on the basis of the data taken along the frequency and price axes. The initial value for f_{hm} was input as the interpolation of the frequency of half-maximal performance for the price-frequency set lying along the frequency axis at the lowest price. The initial value for P_e was input as the interpolation of the price of half-maximal performance for the price-frequency set lying along the price axis at the highest frequency. Other parameters were kept constant across animals: the behavioural-allocation exponent (A) was seeded as 3, the reward-growth exponent (G) as 4, the minimum time allocation (TIN) as 0, and the maximum time allocation (TAX) as 1.

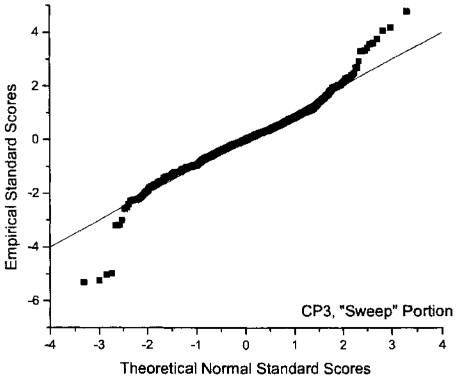
Violations of Normality, Experiment 1

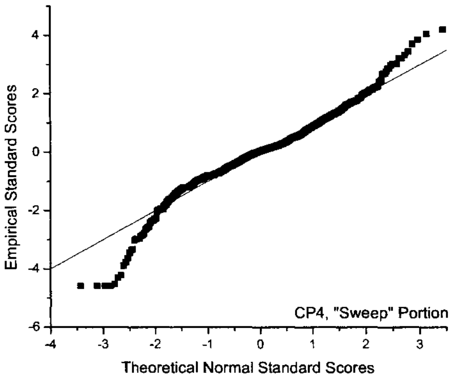
Normal probability plots of the data collected from each rat are presented in figures A1 and A2. The deviations of each score from the mean of each price-frequency pair in each condition were rank-ordered and normalized. The percentile rank of each Z-score was then computed and a theoretical Z-score was determined by computing the inverse normal cumulative density function for each percentile. As a result, each deviation from the mean has an associated theoretical Z-score (as determined by the inverse cumulative density function) and an empirical Z-score (as determined by the normalization of the deviations). The plots in figure A1 and A2 show the theoretical Z-score of all observed deviations from the mean for each rat in a normal distribution as a function of the observed Z-score.

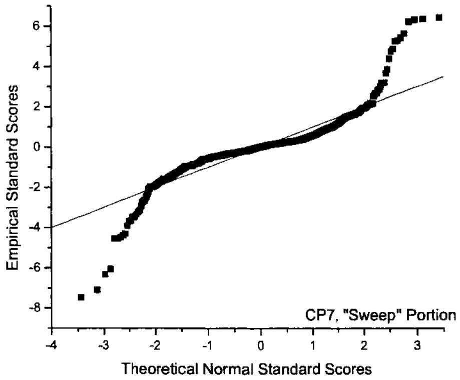
If the data were normally distributed about their respective means (price-frequency pairs at each sweep and randomized condition), the observed quantile and standardized score of each deviation would correspond approximately to the theoretical quantile and standardized score in a normal distribution; this would produce a straight line with a slope of 1 (thin line in figures A1 and A2). That is, deviations with a z-score of -1 should be greater than or equal to 16% of all the other deviations, those with a z-score of 0 should be greater than or equal to 50%, and so forth, assuming the distribution

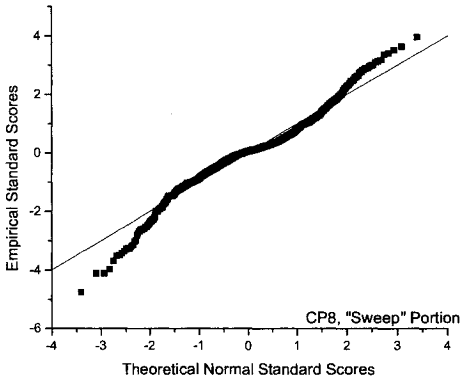












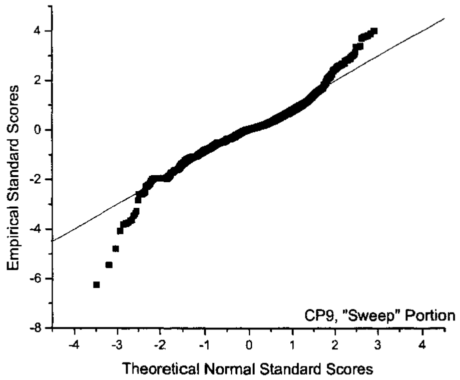


Figure A1. Normal probability plot for all rats in the sweep portion of experiment 1. The curves indicate that the data in every case deviate substantially from normality.

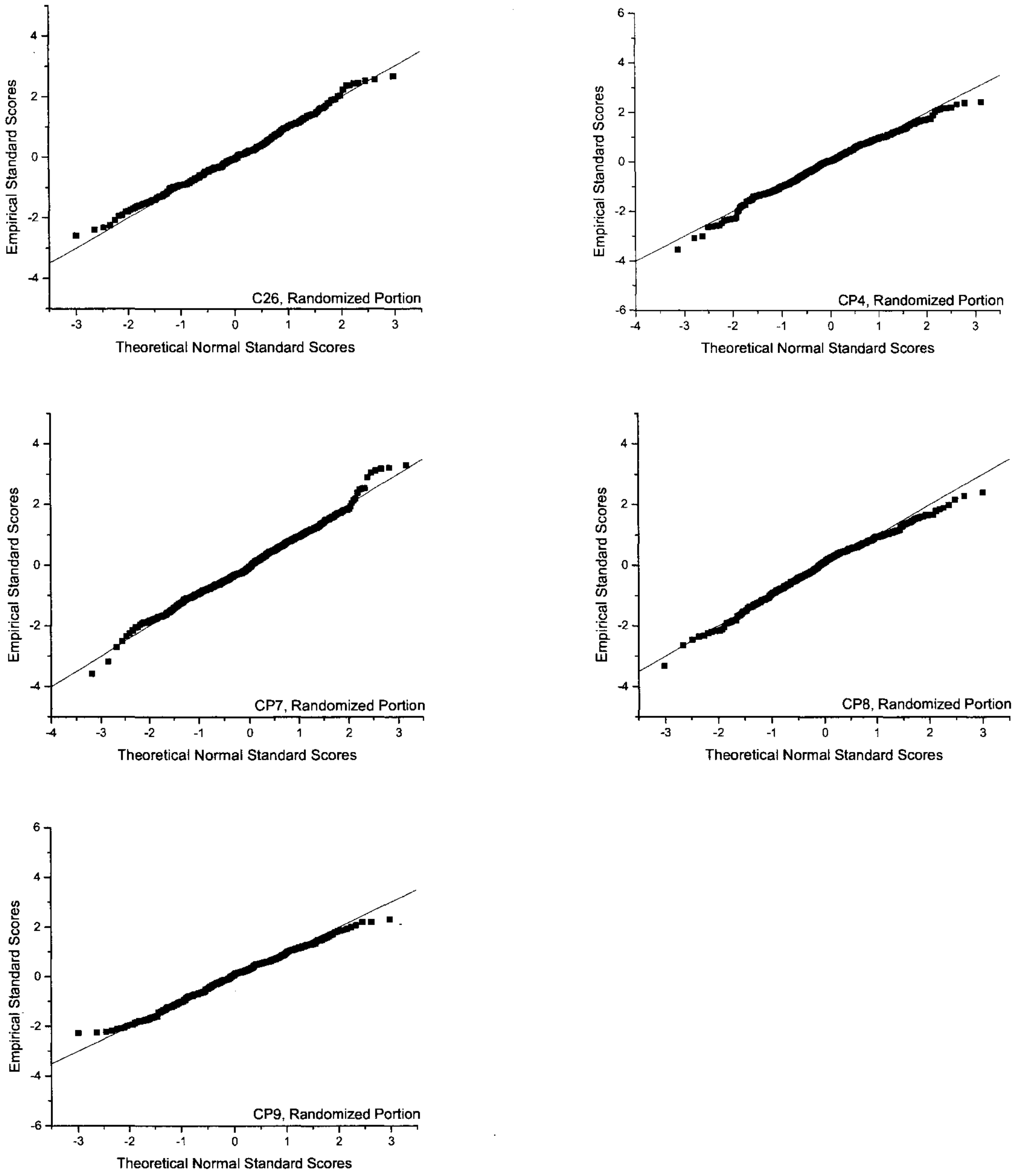


Figure A2. Normal probability plot for all rats in the randomized portion of experiment 1. The curves indicate that the data in every case deviate from normality.

is normal. Conversely, deviations that are greater than or equal to 16% of all other deviations should have a z-score of -1. A visual inspection of the plots of theoretical z-scores (as determined by the inverse normal cumulative density function) against the empirical z-scores indicates that the data deviate from normality. When the plot is curved downward at negative z-scores, the distribution has a heavier left-hand tail, since deviations have a more negative z-score than would be predicted by their percentile rank. Similarly, when it is curved upward at positive z-scores, the distribution has a heavier right-hand tail, since deviations have a more positive z-score than would be predicted by their percentile rank.

It is also clear from these results that the deviations are not simply due to a few single outliers and that the deviations depend on each rat in each condition. For example, while the data from the sweep portion of the experiment for rat CP7 do not appear to be normally distributed, the large, systematic deviations of the normal probability plot appear to be reduced in scope for the data collected in the randomized portion of the experiment for this rat.

The normal-probability plots clearly demonstrate the deviation of the data from normality and the presence of many outliers. Thus, Tukey's bi-square estimator was therefore applied to each data point, weighting the point based on the median absolute deviation from the median. Tukey's bi-square estimator rolls off the influence of outliers by reducing the contribution of observations that are very far from the median. This statistical procedure deals with the presence of outliers by robustly eliminating their presence without having to identify them by hand. As a result of this weighting procedure, problems of heteroskedasticity can also be partly dealt with, because the greatest source of variability in the regions where performance is non-asymptotic (and the data more variable) is the presence of outliers.

Violations of Normality, Experiment 2

As for experiment 1, figure A3 shows, for each rat, a normal probability plot of the data. The data from all rats deviated from normality with no across-subject pattern. This suggests a need for the analysis to be relatively free of distributional assumptions to avoid decreasing the statistical power of the analysis. As a result of this violation, it was deemed necessary to use Tukey's bi-square estimator to weight data points and use a bootstrapping procedure to assess the confidence surrounding parameters and data estimates

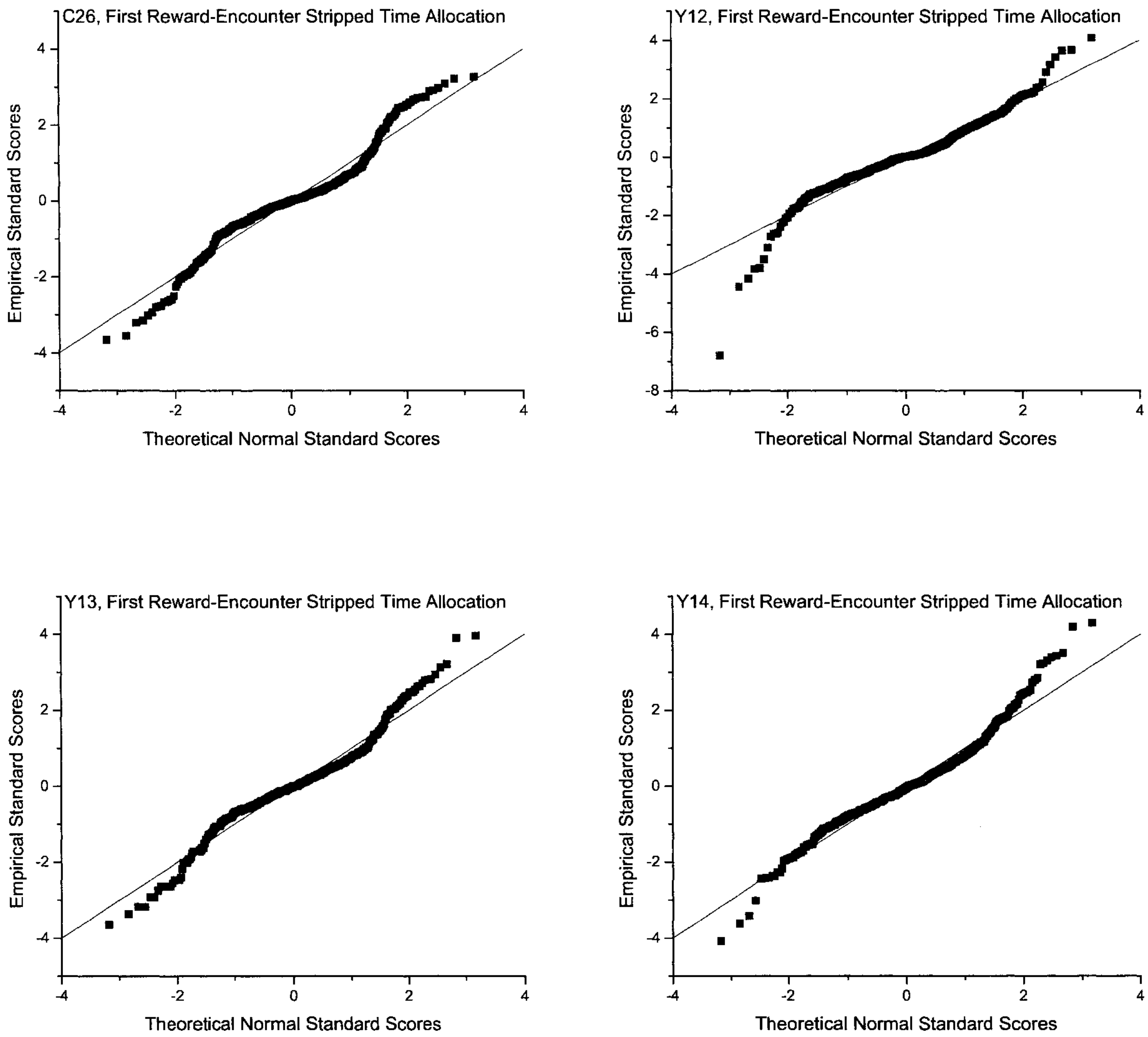


Figure A3. Normal probability plot for all rats in experiment 2. The curves indicate that the data in every case deviate substantially from normality.

Appendix B

Intercorrelations of Parameters in Mountain Model Fit

Table B1

Intercorrelations of Parameters Estimated by Surface Fit to Data of Sweep Portion in Experiment 1

			C26							CP2			
	A	Fh.1	G	Pe.1	Tax.1	Tin		A	Fh.1	G	Pe.1	Tax.1	Tin
A	-	-0.25	-0.73	-0.87	-0.2	0.59	A	-	-0.21	0.49	0.22	-0.71	0.75
Fh.1		-	-0.41	0.59	0.46	-0.23	Fh.1		-	-0.52	0.59	0.3	-0.26
G			-	0.44	0	-0.4	G			-	0.35	-0.68	0.71
Pe.1				-	0.31	-0.57	Pe.1				-	-0.37	0.36
Tax.1					-	-0.19	Tax.1					-	-0.75
Tin						-	Tin						-

			CP3							CP4			
	A	Fh.1	G	Pe.1	Tax.1	Tin		A	Fh.1	G	Pe.1	Tax.1	Tin
A	-	-0.14	-0.07	-0.54	0	0.27	A	-	0.4	-0.25	0.49	-0.76	0.37
Fh.1		-	-0.6	0.31	0.23	0.33	Fh.1		-	-0.66	0.41	-0.21	0.38
G			-	-0.29	-0.2	0.08	G			-	-0.14	0.1	-0.11
Pe.1				-	0	-0.47	Pe.1				-	-0.59	0
Tax.1					-	0.09	Tax.1					-	-0.22
Tin						-	Tin						-

			CP7							CP8			
	A	Fh.1	G	Pe.1	Tax.1	Tin		A	Fh.1	G	Pe.1	Tax.1	Tin
A	-	0	-0.58	-0.75	-0.67	0	A	-	0.71	-0.68	0.24	-0.65	0.25
Fh.1		-	-0.71	0.29	0.22	0	Fh.1		-	-0.91	0.77	-0.67	-0.08
G			-	0.28	0.25	0.17	G			-	-0.72	0.63	0.27
Pe.1				-	0.34	-0.17	Pe.1				-	-0.47	-0.52
Tax.1					-	0.1	Tax.1					-	0
Tin						-	Tin						-

			CP9			
	A	Fh.1	G	Pe.1	Tax.1	Tin
A	-	0.28	0	0.33	-0.86	0.44
Fh.1		-	-0.81	0.39	0	0.2
G			-	-0.15	-0.25	0
Pe.1				-	-0.39	-0.15
Tax.1					-	-0.31
Tin						-

Table B2

Intercorrelations of Parameters Estimated by Surface Fit to Data of Randomized Portion in Experiment 1

C26

	A	Fh.1	G	Pe.1	Tax.1	Tin
A	-	-0.71	0.7	-0.79	-0.91	0.75
Fh.1		-	-0.93	0.97	0.7	-0.61
G			-	-0.92	-0.72	0.57
Pe.1				-	0.75	-0.7
Tax.1					-	-0.71
Tin						-

CP4

	A	Fh.1	G	Pe.1	Tax.1	Tin
A	-	-0.37	0.32	-0.34	-0.68	0.57
Fh.1		-	-0.9	0.98	0.3	-0.29
G			-	-0.9	-0.14	0.38
Pe.1				-	0.19	-0.36
Tax.1					-	-0.19
Tin						-

CP7

	A	Fh.1	G	Pe.1	Tax.1	Tin
A	-	-0.58	0.62	-0.39	-0.63	0.37
Fh.1		-	-0.79	0.59	0.4	-0.15
G			-	-0.43	-0.51	0.2
Pe.1				-	0	-0.33
Tax.1					-	-0.19
Tin						-

CP8

	A	Fh.1	G	Pe.1	Tax.1	Tin
A	-	-0.09	0.11	-0.38	-0.52	0.73
Fh.1		-	-0.85	0.57	0.15	0
G			-	-0.39	-0.17	0
Pe.1				-	-0.13	-0.61
Tax.1					-	-0.16
Tin						-

CP9

	A	Fh.1	G	Pe.1	Tax.1	Tin
A	-	-0.43	0.63	0	-0.66	0.72
Fh.1		-	-0.75	0.48	0.24	-0.34
G			-	-0.34	-0.45	0.57
Pe.1				-	-0.29	-0.32
Tax.1					-	-0.49
Tin						-

Table B3

Intercorrelations of Parameters Estimated by Surface Fit to Data in Experiment 2

	C26										Y12								
	A	Fh.1	Fh.2	G	Pe.1	Pe.2	Tax.1	Tax.2	Tin		A	Fh.1	Fh.2	G	Pe.1	Pe.2	Tax.1	Tax.2	Tin
A	-	-0.35	0	0	0	0	-0.47	-0.63	0.38	A	-	-0.44	-0.43	0.42	-0.36	-0.55	-0.49	0	0.62
Fh.1		-	0.36	-0.37	0.41	0.08	0.37	0.5	0	Fh.1		-	0.77	-0.9	0.84	0.7	0.17	0	-0.4
Fh.2			-	-0.8	0.23	0.54	-0.12	0.24	0	Fh.2			-	-0.85	0.57	0.94	0.27	0	-0.36
G				-	-0.29	-0.28	0.12	-0.21	0.23	G				-	-0.67	-0.77	-0.27	0	0.38
Pe.1					-	0.12	-0.33	0.09	-0.21	Pe.1					-	0.54	-0.14	0	-0.44
Pe.2						-	-0.08	-0.15	-0.18	Pe.2						-	0.33	0	-0.48
Tax.1							-	0.35	0	Tax.1							-	-0.08	-0.23
Tax.2								-	-0.15	Tax.2								-	0
Tin									-	Tin									-

	Y13										Y14								
	A	Fh.1	Fh.2	G	Pe.1	Pe.2	Tax.1	Tax.2	Tin		A	Fh.1	Fh.2	G	Pe.1	Pe.2	Tax.1	Tax.2	Tin
A	-	-0.37	0.19	0.29	-0.27	0.21	-0.44	-0.6	0.56	A	-	0	-0.2	0.11	0.09	0	-0.62	-0.71	0.39
Fh.1		-	-0.3	-0.49	0.69	0	0	0.16	-0.35	Fh.1		-	-0.07	0.07	0.45	-0.09	-0.21	0	0.12
Fh.2			-	0.45	-0.25	0.17	0	0	0.38	Fh.2			-	-0.19	0	0.42	0.11	0.24	0.1
G				-	-0.59	-0.14	0.12	0	0.44	G				-	-0.22	-0.32	0	-0.27	0.24
Pe.1					-	0.09	-0.34	0	-0.42	Pe.1					-	0.13	-0.6	0	-0.17
Pe.2						-	-0.18	-0.57	-0.12	Pe.2						-	0	-0.37	-0.22
Tax.1							-	0.36	-0.14	Tax.1							-	0.45	-0.19
Tax.2								-	-0.24	Tax.2								-	-0.19
Tin									-	Tin									-

Appendix C

Glossary of terms

a	See behavioural growth exponent.
Behavioural growth exponent	The exponent of each term in the behavioural allocation function, indicating the rate of growth of allocation to bar-pressing as a function of the payoff.
Behavioural allocation function	The "outer" function translating the reward-growth function into observable behaviour.
Bootstrap re-sampling	A statistical procedure for estimating the confidence interval about a given statistic based on the observations themselves.
C	See chronaxie.
Chronaxie	A term in Gallistel's strength-duration relationship specifying the decay of the effectiveness of increases in train duration in eliciting a rewarding effect. It is the train duration at which the frequency yielding half-maximal reward intensity is twice the frequency of the rheobase.
Complementarity	A property of goods for which individuals' choices are insensitive to the relative pricing. Complementarity is the opposite of substitutability and represents the dissimilarity of two goods.
Exertion	The effort cost required for an animal to acquire a given reward.
f	The pulse frequency of a train of electrical stimulation.
FCHT	Fixed cumulative handling time schedule of reinforcement. The animal will only be rewarded when the bar has been held down for a total and interruptible number of seconds.
f_{hm}	See frequency of half-maximal reward.
First-encounter stripped time allocation	The time allocation when all interactions with the bar before the animal has harvested a reward are removed. If no rewards are earned, the resulting time allocation will be zero.
f_R	See rheobase.

Frequency of half-maximal reward	The pulse frequency yielding a reward intensity that is half-maximal.
FVI	Free-running variable interval schedule of reinforcement. The animal is only rewarded if it is holding the bar down at the end of a randomly-drawn latency.
g	See reward-growth exponent.
I	See intensity of rewarding effect.
I_{max}	See maximal intensity of rewarding effect.
Intensity of rewarding effect	The subjective intensity of the rewarding effect, determined by the aggregate firing rate in stimulated reward-relevant neurons.
Maximal intensity of rewarding effect	The subjective intensity of the rewarding effect at the point beyond which further increases in frequency fail to increase the subjective reward intensity.
Maximum time allocation	The maximum proportion of time an animal will spend holding down the bar.
Minimum time allocation	The minimum proportion of time an animal will spend holding down the bar.
Mountain model	A computational model that describes the psychological and neurobiological events that occur between the activation of reward-relevant neurons by the electrode and the animal's behaviour.
P_e	Price at which a stimulation train yielding a maximally intense subjective reward will produce only half-maximal time allocation.
Price	The average (in the case of FVI) or exact (in the case of FCHT) number of seconds the bar must be held in order for an animal to harvest a reward. It directly controls the opportunity cost of harvesting brain stimulation reward.
P_s	See subjective price.
Pseudo-sweep	The reconstruction of points that lie along the same ray in a randomized-trials experimental procedure.
Reward growth exponent	The exponent of each term in the reward-growth function, indicating the rate at which subjective reward intensity grows with pulse frequency.

Rheobase	A term in Gallistel's strength-duration relationship; the frequency of the rheobase. It is the frequency of half-maximal reward when the train duration is infinitesimally large.
Subjective price	The animal's subjective evaluation of the price of the stimulation.
Substitutability	A property of goods for which individuals' choices are sensitive to the relative pricing. Substitutability is the opposite of complementarity and represents the similarity of two goods.
Sweep	Refers to either to the structure of an experiment or the components of that structure. Sweeps comprise repeating trial sequences of decreasing pulse frequency or increasing price; sweep-structured experiments contain multiple series of these sequences.
TA	See time allocation.
TAX	See maximum time allocation.
Time allocation	Proportion of time the animal engages in self-stimulation in a trial. It is calculated as the amount of time spent holding down the bar (corrected for tapping) divided by the total trial time.
TIN	See minimum time allocation.
Tukey's bi-square estimator	A statistical procedure that assigns a weight to observations based on their deviation from the central tendency.
U_{bsr}	The payoff from brain stimulation reward resulting from the scalar combination of the intensity of brain stimulation reward and the subjective price of acquiring it.
U_{ee}	The payoff of rewards derived from all other activities the animal may engage in while in the operant chamber.
ξ	See exertion.