Striatal action-value neurons reconsidered, Elber-Dorozko, Lotem, and Loewenstein Yonatan, Submitted, (Submitted)

It is generally believed that during economic decisions, striatal neurons represent the values associated with different actions. This hypothesis is based on a large number of electrophysiological studies, in which the neural activity of striatal neurons was measured while the subject was learning to prefer the more rewarding action. Here we present an alternative interpretation of the electrophysiological findings. We show that the standard statistical methods that were used to identify action-value neurons in the striatum erroneously detect the same action-value representations in unrelated neuronal recordings. This is due to temporal correlations in the neuronal data. We propose an alternative statistical method for identifying action-value representations that is not subject to this caveat. We apply it to previously identified action-value neurons in the basal ganglia and fail to detect action-value representations. In conclusion, we argue that there is no conclusive evidence for the generally accepted hypothesis that striatal neurons encode action-values.


According to the synaptic trace theory of memory, activity-induced changes in the pattern of synaptic connections underlie the storage of information for long periods. In this framework, the stability of memory critically depends on the stability of the underlying synaptic connections. Surprisingly however, synaptic connections in the living brain are highly volatile, which poses a fundamental challenge to the synaptic trace theory. Here we review recent experimental evidence that link the initial formation of a memory with changes in the pattern of connectivity, but also evidence that synaptic connections are considerably volatile even in the absence of learning. Then we consider different theoretical models that have been put forward to explain how memory can be maintained with such volatile building blocks.


It has long been known that we subjectively experience longer stimuli as being more intense. A recent study sheds light on the neural mechanisms underlying this bias by tracking the formation of a percept of intensity in the rat brain.
The selection and timing of actions are subject to determinate influences such as sensory cues and internal state as well as to effectively stochastic variability. Although stochastic choice mechanisms are assumed by many theoretical models, their origin and mechanisms remain poorly understood. Here we investigated this issue by studying how neural circuits in the frontal cortex determine action timing in rats performing a waiting task. Electrophysiological recordings from two regions necessary for this behavior, medial prefrontal cortex (mPFC) and secondary motor cortex (M2), revealed an unexpected functional dissociation. Both areas encoded deterministic biases in action timing, but only M2 neurons reflected stochastic trial-by-trial fluctuations. This differential coding was reflected in distinct timescales of neural dynamics in the two frontal cortical areas. These results suggest a two-stage model in which stochastic components of action timing decisions are injected by circuits downstream of those carrying deterministic bias signals.
There is a long history of experiments in which participants are instructed to generate a long sequence of binary random numbers. The scope of this line of research has shifted over the years from identifying the basic psychological principles and/or the heuristics that lead to deviations from randomness, to one of predicting future choices. In this paper, we used generalized linear regression and the framework of Reinforcement Learning in order to address both points. In particular, we used logistic regression analysis in order to characterize the temporal sequence of participants’ choices. Surprisingly, a population analysis indicated that the contribution of the most recent trial has only a weak effect on behavior, compared to more preceding trials, a result that seems irreconcilable with standard sequential effects that decay monotonously with the delay. However, when considering each participant separately, we found that the magnitudes of the sequential effect are a monotonous decreasing function of the delay, yet these individual sequential effects are largely averaged out in a population analysis because of heterogeneity. The substantial behavioral heterogeneity in this task is further demonstrated quantitatively by considering the predictive power of the model. We show that a heterogeneous model of sequential dependencies captures the structure available in random sequence generation. Finally, we show that the results of the logistic regression analysis can be interpreted in the framework of reinforcement learning, allowing us to compare the sequential effects in the random sequence generation task to those in an operant learning task. We show that in contrast to the random sequence generation task, sequential effects in operant learning are far more homogenous across the population. These results suggest that in the random sequence generation task, different participants adopt different cognitive strategies to suppress sequential dependencies when generating the random sequences.

**2015**


Dyslexics are diagnosed for their poor reading skills, yet they characteristically also suffer from poor verbal memory and often from poor auditory skills. To date, this combined profile has been accounted for in broad cognitive terms. Here we hypothesize that the perceptual deficits associated with dyslexia can be understood computationally as a deficit in integrating prior information with noisy observations. To test this hypothesis we analyzed the performance of human participants in an auditory discrimination task using a two-parameter computational model. One parameter captures the internal noise in representing the current event, and the other captures the impact of recently acquired prior information. Our findings show that dyslexics' perceptual deficit can be accounted for by inadequate adjustment of these components; namely, low weighting of their implicit memory of past trials relative to their internal noise. Underweighting the stimulus statistics decreased dyslexics’ ability to compensate for noisy observations. ERP measurements (P2 component) while participants watched a silent movie indicated that dyslexics' perceptual deficiency may stem from poor automatic integration of stimulus statistics. This study provides the first description of a specific computational deficit associated with dyslexia. SIGNIFICANCE STATEMENT This study presents the first attempt to specify the mechanisms underlying dyslexics' perceptual difficulties computationally by applying a specific model, inspired by the Bayesian framework. This model dissociates between the contribution of sensory noise and that of the prior statistics in an auditory perceptual decision task. We show that dyslexics cannot compensate for their perceptual noise by incorporating prior information. By contrast, adequately reading controls' usage of previous information is often close to optimal. We used ERP measurements to assess the neuronal stage of this deficit. We found that unlike their peers, dyslexics’ ERP responses are not sensitive to the relations between the current observation
and the prior observation, indicating that they cannot establish a reliable prior.


Dynamic remodeling of connectivity is a fundamental feature of neocortical circuits. Unraveling the principles underlying these dynamics is essential for the understanding of how neuronal circuits give rise to computations. Moreover, as complete descriptions of the wiring diagram in cortical tissues are becoming available, deciphering the dynamic elements in these diagrams is crucial for relating them to cortical function. Here, we used chronic in vivo two-photon imaging to longitudinally follow a few thousand dendritic spines in the mouse auditory cortex to study the determinants of these spines’ lifetimes. We applied nonlinear regression to quantify the independent contribution of spine age and several morphological parameters to the prediction of the future survival of a spine. We show that spine age, size, and geometry are parameters that can provide independent contributions to the prediction of the longevity of a synaptic connection. In addition, we use this framework to emulate a serial sectioning electron microscopy experiment and demonstrate how incorporation of morphological information of dendritic spines from a single time-point allows estimation of future connectivity states. The distinction between predictable and nonpredictable connectivity changes may be used in the future to identify the specific adaptations of neuronal circuits to environmental changes. The full dataset is publicly available for further analysis.

SIGNIFICANCE STATEMENT The neural architecture in the neocortex exhibits constant remodeling. The functional consequences of these modifications are poorly understood, in particular because the determinants of these changes are largely unknown. Here, we aimed to identify those modifications that are predictable from current network state. To that goal, we repeatedly imaged thousands of dendritic spines in the auditory cortex of mice to assess the morphology and lifetimes of synaptic connections. We developed models based on morphological features of dendritic spines that allow predicting future turnover of synaptic connections. The dynamic models presented in this paper provide a quantitative framework for adding putative temporal dynamics to the static description of a neuronal circuit from single time-point connectomics experiments.


We investigated the course of language processing in the context of a verification task that required numerical estimation and comparison. Participants listened to sentences with complex quantifiers that contrasted in Polarity, a logical property (e.g., more-than-half, less-than-half), and then performed speeded verification on visual scenarios that displayed a proportion between 2 discrete quantities. We varied systematically not only the sentences, but also the visual materials, in order to study their effect on the verification process. Next, we used the same visual scenarios with analogous non-verbal probes that featured arithmetical inequality symbols (). This manipulation enabled us to measure not only Polarity effects, but also, to compare the effect of different probe types (linguistic, non-linguistic) on processing. Like many previous studies, our results demonstrate that perceptual difficulty affects error rate and reaction time in keeping with Weber’s Law. Interestingly, these performance parameters are also affected by the Polarity of the quantifiers used, despite the fact that sentences had the exact same meaning, sentence structure, number of words, syllables, and temporal structure. Moreover, an analogous contrast between the non-linguistic probes () had no effect on performance. Finally, we observed no interaction between performance parameters governed by Weber’s Law and those affected by Polarity. We consider 4
Probability estimation is an essential cognitive function in perception, motor control, and decision making. Many studies have shown that when making decisions in a stochastic operant conditioning task, people and animals behave as if they underestimate the probability of rare events. It is commonly assumed that this behavior is a natural consequence of estimating a probability from a small sample, also known as sampling bias. The objective of this article is to challenge this common lore. We show that, in fact, probabilities estimated from a small sample can lead to behaviors that will be interpreted as underestimating or as overestimating the probability of rare events, depending on the cognitive strategy used. Moreover, this sampling bias hypothesis makes an implausible prediction that minute differences in the values of the sample size or the underlying probability will determine whether rare events will be underweighted or overweighed. We discuss the implications of this sensitivity for the design and interpretation of experiments. Finally, we propose an alternative sequential learning model with a resetting of initial conditions for probability estimation and show that this model predicts the experimentally observed robust underweighting of rare events.


Biases such as the preference of a particular response for no obvious reason, are an integral part of psychophysics. Such biases have been reported in the common two-alternative forced choice (2AFC) experiments, where participants are instructed to compare two consecutively presented stimuli. However, the principles underlying these biases are largely unknown and previous studies have typically used ad-hoc explanations to account for them. Here we consider human performance in the 2AFC tone frequency discrimination task, utilizing two standard protocols. In both protocols, each trial contains a reference stimulus. In one (Reference-Lower protocol), the frequency of the reference stimulus is always lower than that of the comparison stimulus, whereas in the other (Reference protocol), the frequency of the reference stimulus is either lower or higher than that of the comparison stimulus. We find substantial interval biases. Namely, participants perform better when the reference is in a specific interval. Surprisingly, the biases in the two experiments are opposite: performance is better when the reference is in the first interval in the Reference protocol, but is better when the reference is second in the Reference-Lower protocol. This inconsistency refutes previous accounts of the interval bias, and is resolved when experiments statistics is considered. Viewing perception as incorporation of sensory input with prior knowledge accumulated during the experiment accounts for the seemingly contradictory biases both qualitatively and quantitatively. The success of this account implies that even simple discriminations reflect a combination of sensory limitations, memory limitations, and the ability to utilize stimuli statistics.

Spatial Generalization in Operant Learning: Lessons from Professional Basketball, Tal, Neiman, and Yonatan Loewenstein, PLoS Computational Biology, 05/2014, Volume 10, Issue 5, p.e1003623,
In operant learning, behaviors are reinforced or inhibited in response to the consequences of similar actions taken in the past. However, because in natural environments the same situation never recurs, it is essential for the learner to decide what similar is so that he can generalize from experience in one state of the world to future actions in different states of the world. The computational principles underlying this generalization are poorly understood, in particular because natural environments are typically too complex to study quantitatively. In this paper we study the principles underlying generalization in operant learning of professional basketball players. In particular, we utilize detailed information about the spatial organization of shot locations to study how players adapt their attacking strategy in real time according to recent events in the game. To quantify this learning, we study how a make \ miss from one location in the court affects the probabilities of shooting from different locations. We show that generalization is not a spatially-local process, nor is governed by the difficulty of the shot. Rather, to a first approximation, players use a simplified binary representation of the court into 2 pt and 3 pt zones. This result indicates that rather than using low-level features, generalization is determined by high-level cognitive processes that incorporate the abstract rules of the game.


Organisms modify their behavior in response to its consequences, a phenomenon referred to as operant learning. The computational principles and neural mechanisms underlying operant learning are a subject of extensive experimental and theoretical investigations. Theoretical approaches largely rely on concepts and algorithms from reinforcement learning. The dominant view is that organisms maintain a value function, that is, a set of estimates of the cumulative future rewards associated with the different behavioral options. These values are then used to select actions. Learning in this framework results from the update of these values depending on experience of the consequences of past actions. An alternative view questions the applicability of such a computational scheme to many real-life situations. Instead, it posits that organisms exploit the intrinsic variability in their action?selection mechanism(s) to modify their behavior, e.g., via stochastic gradient ascent, without the need of an explicit representation of values. In this review, we compare these two approaches in terms of their computational power and flexibility, their putative neural correlates, and, finally, in terms of their ability to account for behavior as observed in repeated-choice experiments. We discuss the successes and failures of these alternative approaches in explaining the observed patterns of choice behavior. We conclude by identifying some of the important challenges to a comprehensive theory of operant learning.

an explicit world model in terms of state-action pairs.

Complex Population Response of Dorsal Putamen Neurons Predicts the Ability to Learn


Day-to-day variability in performance is a common experience. We investigated its neural correlate by studying learning behavior of monkeys in a two-alternative forced choice task, the two-armed bandit task. We found substantial session-to-session variability in the monkeys' learning behavior. Recording the activity of single dorsal putamen neurons we uncovered a dual function of this structure. It has been previously shown that a population of neurons in the DLP exhibits firing activity sensitive to the reward value of chosen actions. Here, we identify putative medium spiny neurons in the dorsal putamen that are cue-selective and whose activity builds up with learning. Remarkably we show that session-to-session changes in the size of this population and in the intensity with which this population encodes cue-selectivity is correlated with session-to-session changes in the ability to learn the task. Moreover, at the population level, dorsal putamen activity in the very beginning of the session is correlated with the performance at the end of the session, thus predicting whether the monkey will have a "good" or "bad" learning day. These results provide important insights on the neural basis of inter-temporal performance variability.

Stochasticity, bistability and the wisdom of crowds: a model for associative learning in genetic regulatory networks


It is generally believed that associative memory in the brain depends on multistable synaptic dynamics, which enable the synapses to maintain their value for extended periods of time. However, multistable dynamics are not restricted to synapses. In particular, the dynamics of some genetic regulatory networks are multistable, raising the possibility that even single cells, in the absence of a nervous system, are capable of learning associations. Here we study a standard genetic regulatory network model with bistable elements and stochastic dynamics. We demonstrate that such a genetic regulatory network model is capable of learning multiple, general, overlapping associations. The capacity of the network, defined as the number of associations that can be simultaneously stored and retrieved, is proportional to the square root of the number of bistable elements in the genetic regulatory network. Moreover, we compute the capacity of a clonal population of cells, such as in a colony of bacteria or a tissue, to store associations. We show that even if the cells do not interact, the capacity of the population to store associations substantially exceeds that of a single cell and is proportional to the number of bistable elements. Thus, we show that even single cells are endowed with the computational power to learn associations, a power that is substantially enhanced when these cells form a population.

The Role of First Impression in Operant Learning


We quantified the effect of first experience on behavior in operant learning and studied its underlying computational principles. To that goal, we analyzed more than 200,000 choices in a repeated-choice experiment. We found that the outcome of the first experience has a substantial and lasting effect on
participants' subsequent behavior, which we term outcome primacy. We found that this outcome primacy can account for much of the underweighting of rare events, where participants apparently underestimate small probabilities. We modeled behavior in this task using a standard, model-free reinforcement learning algorithm. In this model, the values of the different actions are learned over time and are used to determine the next action according to a predefined action-selection rule. We used a novel nonparametric method to characterize this action-selection rule and showed that the substantial effect of first experience on behavior is consistent with the reinforcement learning model if we assume that the outcome of first experience resets the values of the experienced actions, but not if we assume arbitrary initial conditions. Moreover, the predictive power of our resetting model outperforms previously published models regarding the aggregate choice behavior. These findings suggest that first experience has a disproportionately large effect on subsequent actions, similar to primacy effects in other fields of cognitive psychology. The mechanism of resetting of the initial conditions that underlies outcome primacy may thus also account for other forms of primacy. (PsycINFO Database Record (c) 2012 APA, all rights reserved)

Covariance-Based Synaptic Plasticity in an Attractor Network Model Accounts for Fast Adaptation in Free Operant Learning

In free operant experiments, subjects alternate at will between targets that yield rewards stochastically. Behavior in these experiments is typically characterized by (1) an exponential distribution of stay durations, (2) matching of the relative time spent at a target to its relative share of the total number of rewards, and (3) adaptation after a change in the reward rates that can be very fast. The neural mechanism underlying these regularities is largely unknown. Moreover, current decision-making neural network models typically aim at explaining behavior in discrete-time experiments in which a single decision is made once in every trial, making these models hard to extend to the more natural case of free operant decisions. Here we show that a model based on attractor dynamics, in which transitions are induced by noise and preference is formed via covariance-based synaptic plasticity, can account for the characteristics of behavior in free operant experiments. We compare a specific instance of such a model, in which two recurrently excited populations of neurons compete for higher activity, to the behavior of rats responding on two levers for rewarding brain stimulation on a concurrent variable interval reward schedule (Gallistel et al., 2001). We show that the model is consistent with the rats' behavior, and in particular, with the observed fast adaptation to matching behavior. Further, we show that the neural model can be reduced to a behavioral model, and we use this model to deduce a novel ?conservation law,? which is consistent with the behavior of the rats.

2012

How Recent History Affects Perception: The Normative Approach and Its Heuristic Approximation

In this paper we study how history affects perception using an auditory delayed comparison task, in which human participants repeatedly compare the frequencies of two, temporally-separated pure tones. We demonstrate that the history of the experiment has a substantial effect on participants' performance: when both tones are high relative to past stimuli, people tend to report that the 2nd tone was higher, and when
they are relatively low, they tend to report that the 1st tone was higher. Interestingly, only the most recent trials bias performance, which can be interpreted as if the participants assume that the statistics of stimuli in the experiment is highly volatile. Moreover, this bias persists even in settings, in which it is detrimental to performance. These results demonstrate the abilities, as well as limitations, of the cognitive system when incorporating expectations in perception.

**2011**


Reinforcement learning in complex natural environments is a challenging task because the agent should generalize from the outcomes of actions taken in one state of the world to future actions in different states of the world. The extent to which human experts find the proper level of generalization is unclear. Here we show, using the sequences of field goal attempts made by professional basketball players, that the outcome of even a single field goal attempt has a considerable effect on the rate of subsequent 3 point shot attempts, in line with standard models of reinforcement learning. However, this change in behaviour is associated with negative correlations between the outcomes of successive field goal attempts. These results indicate that despite years of experience and high motivation, professional players overgeneralize from the outcomes of their most recent actions, which leads to decreased performance.


BACKGROUND: To what extent are the properties of neuronal networks constrained by computational considerations? Comparative analysis of the vertical lobe (VL) system, a brain structure involved in learning and memory, in two phylogenetically close cephalopod mollusks, Octopus vulgaris and the cuttlefish Sepia officinalis, provides a surprising answer to this question. RESULTS: We show that in both the octopus and the cuttlefish the VL is characterized by the same simple fan-out fan-in connectivity architecture, composed of the same three neuron types. Yet, the sites of short- and long-term synaptic plasticity and neuromodulation are different. In the octopus, synaptic plasticity occurs at the fan-out glutamatergic synaptic layer, whereas in the cuttlefish plasticity is found at the fan-in cholinergic synaptic layer. CONCLUSIONS: Does this dramatic difference in physiology imply a difference in function? Not necessarily. We show that the physiological properties of the VL neurons, particularly the linear input-output relations of the intermediate layer neurons, allow the two different networks to perform the same computation. The convergence of different networks to the same computational capacity indicates that it is the computation, not the specific properties of the network, that is self-organized or selected for by evolutionary pressure.

What fundamental properties of synaptic connectivity in the neocortex stem from the ongoing dynamics of synaptic changes? In this study, we seek to find the rules shaping the stationary distribution of synaptic efficacies in the cortex. To address this question, we combined chronic imaging of hundreds of spines in the auditory cortex of mice in vivo over weeks with modeling techniques to quantitatively study the dynamics of spines, the morphological correlates of excitatory synapses in the neocortex. We found that the stationary distribution of spine sizes of individual neurons can be exceptionally well described by a log-normal function. We furthermore show that spines exhibit substantial volatility in their sizes at timescales that range from days to months. Interestingly, the magnitude of changes in spine sizes is proportional to the size of the spine. Such multiplicative dynamics are in contrast with conventional models of synaptic plasticity, learning, and memory, which typically assume additive dynamics. Moreover, we show that the ongoing dynamics of spine sizes can be captured by a simple phenomenological model that operates at two timescales of days and months. This model converges to a log-normal distribution, bridging the gap between synaptic dynamics and the stationary distribution of synaptic efficacies.


Delayed comparison tasks are widely used in the study of working memory and perception in psychology and neuroscience. It has long been known, however, that decisions in these tasks are biased. When the two stimuli in a delayed comparison trial are small in magnitude, subjects tend to report that the first stimulus is larger than the second stimulus. In contrast, subjects tend to report that the second stimulus is larger than the first when the stimuli are relatively large. Here we study the computational principles underlying this bias, also known as the contraction bias. We propose that the contraction bias results from a Bayesian computation in which a noisy representation of a magnitude is combined with a-priori information about the distribution of magnitudes to optimize performance. We test our hypothesis on choice behavior in a visual delayed comparison experiment by studying the effect of (i) changing the prior distribution and (ii) changing the uncertainty in the memorized stimulus. We show that choice behavior in both manipulations is consistent with the performance of an observer who uses a Bayesian inference in order to improve performance. Moreover, our results suggest that the contraction bias arises during memory retrieval/decision making and not during memory encoding. These results support the notion that the contraction bias illusion can be understood as resulting from optimality considerations.


According to the theory of Melioration, organisms in repeated choice settings shift their choice preference in favor of the alternative that provides the highest return. The goal of this paper is to explain how this learning behavior can emerge from microscopic changes in the efficacies of synapses, in the context of a two-alternative repeated-choice experiment. I consider a large family of synaptic plasticity rules in which changes in synaptic efficacies are driven by the covariance between reward and neural activity. I construct a general framework that predicts the learning dynamics of any decision-making neural network that implements this synaptic plasticity rule and show that melioration naturally emerges in such networks. Moreover, the resultant learning dynamics follows the Replicator equation which is commonly used to
phenomenologically describe changes in behavior in operant conditioning experiments. Several examples demonstrate how the learning rate of the network is affected by its properties and by the specifics of the plasticity rule. These results help bridge the gap between cellular physiology and learning behavior.

2009

**Learning reward timing in cortex through reward dependent expression of synaptic plasticity**


The ability to represent time is an essential component of cognition but its neural basis is unknown. Although extensively studied both behaviorally and electrophysiologically, a general theoretical framework describing the elementary neural mechanisms used by the brain to learn temporal representations is lacking. It is commonly believed that the underlying cellular mechanisms reside in high order cortical regions but recent studies show sustained neural activity in primary sensory cortices that can represent the timing of expected reward. Here, we show that local cortical networks can learn temporal representations through a simple framework predicated on reward dependent expression of synaptic plasticity. We assert that temporal representations are stored in the lateral synaptic connections between neurons and demonstrate that reward-modulated plasticity is sufficient to learn these representations. We implement our model numerically to explain reward-time learning in the primary visual cortex (V1), demonstrate experimental support, and suggest additional experimentally verifiable predictions.
The Jerusalem Brain Sciences Building will provide a state-of-the-art research and teaching facility for the Edmond and Lily Safra Center for Brain Sciences.

read more

ELSC Media Channel

Get into our media channel and investigate ELSC's latest videos: seminars, public lectures, courses and video articles.

read more

Source URL: https://elsc.huji.ac.il/loewenstein/science/publications

Links:
[1] https://elsc.huji.ac.il/loewenstein/publications/export/tagged
[12] https://elsc.huji.ac.il/loewenstein/publications/export/xml/5432
[14] http://scholar.google.com/scholar?btnG=Search%2BScholar&as_q=%22Striatal%2Baction-value%2Bneurons%2Breconsidered%22&as_sauthors=Elber-Dorozko&as_occt=any&as_epq=&as_oq=&as_eq=&as_publication=&as_ylo=&as_yhi=&as_sdtp=1&as_sdtAAP=1&as_sdtp=1
[18] https://elsc.huji.ac.il/loewenstein/publications/export/tagged/5494
[19] https://elsc.huji.ac.il/loewenstein/publications/export/xml/5494
[22] https://elsc.huji.ac.il/loewenstein/publications/view/5469
[23] https://elsc.huji.ac.il/loewenstein/publications/neuroscience-formation-percept-rat-cortex
[25] https://elsc.huji.ac.il/loewenstein/publications/export/tagged/5469
[26] https://elsc.huji.ac.il/loewenstein/publications/export/xml/5469
[27] https://elsc.huji.ac.il/loewenstein/publications/export/bibtex/5469
[28] http://scholar.google.com/scholar?btnG=Search%2BScholar&as_q=%22Neuroscience%3A%2BFormation%2Bof%2BPercept%2Bin%2Bthe%2BRat%2BCortex%22&as_sauthors=Mongillo&as_occt=any&as_epq=&as_oq=&as_eq=&as_publication=&as_ylo=&as_yhi=&as_sdtAAP=1&as_sdtp=1
[29] https://elsc.huji.ac.il/loewenstein/publications/view/5450