Computational Underpinnings of Neuromodulation in Humans

P. Read Montague[1,2,3] and Kenneth T. Kishida[4,5]

[1]*Department of Physics, Virginia Tech, Blacksburg, Virginia 24061, USA*
[2]*Wellcome Centre for Human Neuroimaging, University College London, London WC1N 3AR, United Kingdom*
[3]*Fralin Biomedical Research Institute, Virginia Tech, Roanoke, Virginia 24018, USA*
[4]*Department of Physiology and Pharmacology, Wake Forest School of Medicine, Winston-Salem, North Carolina 27101, USA*
[5]*Department of Neurosurgery, Wake Forest School of Medicine, Winston-Salem, North Carolina 27101, USA*
Correspondence: read@vt.edu; kkishida@wakehealth.edu

We summarize a new approach to neuromodulator detection that provides colocalized detection of dopamine, serotonin, and norepinephrine at subsecond timescales and promises to provide submillisecond estimates of the same. The methodology, elastic net electrochemistry, is used to estimate dopamine and serotonin in the striatum of conscious human subjects during active decision-making. We show a proof-of-principle example of the same method working on commercially available depth electrodes in common use for epilepsy monitoring and neurosurgical planning in humans, which further promises to make such electrodes sources of fast neuromodulator information never before available in human subjects. We discuss the implications of this methodology for making direct tests in humans of the computations carried by these three important neuromodulatory systems. The methods also promise great utility in model organisms, but this chapter focuses on the possibilities for human use.

INTRODUCTION

Neuromodulatory systems that deliver dopamine, serotonin, and norepinephrine to target neural regions are crucial for sustaining healthy mental function. Disturbances in these systems by injury or disease underlie a wide range of psychiatric and neurological dysfunction. Over the last two decades, these systems have been the focus of modeling that seeks to understand in computational terms their role in learning, memory, mood, and mental disorders. These systems are hypothesized to encode important learning signals about rewards, punishments, and attentional allocation as modulations in their spike rates. In principle, these modulations in spike rate translate into subsequent changes in the downstream delivery of their neuromodulators. As this modeling work progresses into its third decade, it is important to highlight some critical gaps in our understanding of diffuse neuromodulatory systems that modern methodologies stand poised to surmount. We focus here on two big gaps, but allowing that there are many others: (1) the neurophysiology of these systems in humans and (2) the feasibility of ultrafast neuromodulator measurements.

From a neurophysiological and signaling perspective, the vast majority of work on neuromodulatory systems has been in model organisms, which provide fantastically high-precision access and control. The caveat here, however, is in the difficulty of understanding the relationship of model organism behavior—alongside some interesting biological perturbation or measurement—to human behavior. This is simply a hard problem biologically and computationally. This kind of cross-species behavior gap is not easily bridged because it is difficult to know which behavioral primitives in rodents represent homologous behavioral capacities in humans. Moreover, experiments in model organisms must necessarily focus on simple behaviors (approach, avoidance, simple choices), and this leaves out the kind of important abstractions available to humans and that may be perturbed in humans by disease and injury. Human behavioral work—in the healthy and otherwise—brings its own face validity, but at a cost—the methodologies available for neural eavesdropping in humans have simply not been at the same level of granularity available in model systems.

A New Inferential Approach to Fast, Selective Neuromodulator Detection

We have recently developed new approaches that permit fast (subsecond), simultaneous, and colocalized detection of extracellular dopamine, serotonin, and norepinephrine and have extended the use of these tools for use in conscious human subjects (Kishida et al. 2011, 2016; Moran et al. 2018; also see Platt and Pearson 2016). Our electrochemical detection approaches require direct access to brain tissue, which in humans can only be gained by piggybacking on clinical procedures requiring neurosurgery.

Nevertheless, direct investigation of human brain function is requisite if we are to develop an understanding of how moment-to-moment fluctuations in dopamine, serotonin, and norepinephrine encode information that affects human behavior, thoughts, and feelings. We review the general approach and its connection to established machine learning techniques and we further point out how these methods can be implemented on electrodes in routine use in model organisms and during neurosurgical procedures in humans. These latter implementations have the potential to be transformative for our understanding of the computational underpinnings of neuromodulation (e.g., Dayan 2012; Sutton and Barto 2018) because they will make fast neuromodulator detection possible using off-the-shelf hardware and software.

Despite the recent revolution in methods to record and induce neural activity, there has been relatively less progress in making dynamic, chemically specific measurements of neurotransmitter fluctuations in the extracellular space. Fast-scan cyclic voltammetry has been the only rapid way to monitor subsecond neurochemical changes in neural tissue (Stamford et al. 1984; Kuhr and Wightman 1986; Mermet and Gonon 1986; Stamford 1990). However, the recent advent of an expressible dopamine sensor (DLIGHT; Patriarchi et al. 2018) should provide a host of new insights in model organisms where such an innovation can be expressed selectively in specific cells types. Cyclic voltammetry has been adapted for use in behaving animals over sufficiently long periods suitable for connecting neuromodulator fluctuations (e.g., typically dopamine) to behavior (Phillips et al. 2003; Robinson et al. 2008; Huffman and Venton 2009; Clark et al. 2010). The basic approach is to introduce a voltage sweep on a carbon fiber, record the measured currents, and take advantage of the fact that different oxidizable species react on the surface of the carbon fiber at different rates at different voltages. These rates of reaction are time and concentration dependent. In this fashion, the induced current time series potentially carries a "signature" for different important, oxidizable neurotransmitters that can be calibrated against known concentrations. These approaches contain pitfalls because of the potentially adulterating influences of compounds like ascorbate, pH, and other neurotransmitters with nearby oxidation peaks (e.g., norepinephrine vs. dopamine) as well as a number of other potential confounds.

The aim of deploying voltammetric approaches in humans inherits another challenge. In a human subject, because of the risk for contamination, it is not feasible to calibrate sensors beforehand and then introduce the calibrated electrode into the brain. Instead, a model for dopamine detection must be developed in an in vitro setting and used to infer concentrations on a similar, but distinct electrode to be deployed in vivo. Hence in vitro calibration models must be made stable to known influences such as pH and norepinephrine, both of which could confuse a putative dopamine measurement in vivo. The models also must be shown to generalize across electrodes and must be robust to dopamine levels on which the models were not trained. These challenges strongly suggested to us the use of a modern statistical inference method alongside very large training data sets typical for modern machine learning approaches. In our initial work (Kishida et al. 2016; Moran et al. 2018), we retained the voltage sweeps typical of prior fast-scan cyclic voltammetry work (here a 10-msec triangular sweep followed by a 90-msec waiting period—below we show in experiments that this waiting period is not necessary), but adopted a different approach to extracting a concentration-prediction model. As shown in Figure 1, after recording the current time series we computed a finite difference through time (Fig. 1C) and entered this into an "elastic net" regression (Zou and Hastie 2005; Kishida et al. 2016) on labeled data (here the "label" is the concentration). Each time step in the differentiated current time series is entered as an independent predictor. The concentration-prediction models were extracted using a standard cross-validation method ("glmnet.m" in MATLAB; Qian et al. 2013).

One motivation for this approach was the fact that current responses measured throughout nearly the entire voltammetric cycle provide an excellent encoding of the known analyte concentrations. Traditional analytic approaches focus the development of inference models on a single point in the voltammogram (e.g., typically the oxidation peak), which forces a loss of the information contained in the rest of the voltammetric measurement that is required to determine the chemical species identity. This is illustrated in Figure 1D where the color code shows the differentiated currents for different concentrations of dopamine and serotonin. Although each color traces out a "wiggly" line, these concentration-dependent traces remain visibly discriminable throughout almost the entire time series between the start and the capacitive transient at 5 msec. The same claim holds for the time series from 5 msec onward (not shown). Thus, information about a particular dopamine or serotonin level is not concentrated solely at the theoretically reported oxidation potential for each but is instead spread through a relatively broad region of the time series. This is visible for the ∼2-msec section shown (indicated by the rectangular box in Fig. 1C) but is present statistically for almost the entire time series. We believed that the highly distributed concentration information had not been exploited in the past; hence, we sought a way to "dig out" a wiggling but coherent representation of each concentration-dependent response.

Specifics of Elastic Net Electrochemistry

Figure 1A–C introduces the basic elastic net electrochemistry workflow (Kishida et al. 2016; Moran et al. 2018; also see Kishida et al. 2011). The basic approach is to train an N-fold (in our case, a 10-fold) cross-validated recognition model using data collected in a flow cell where mixtures of dopamine, serotonin, and other analytes (or contaminants) can be exactly controlled (and known), and then further validate this model using out-of-sample data in two ways. First, we test our within training set cross-validated models on measurements of dopamine and serotonin not used in building the cross-validated recognition model. Second, we test our models using out-of-probe data
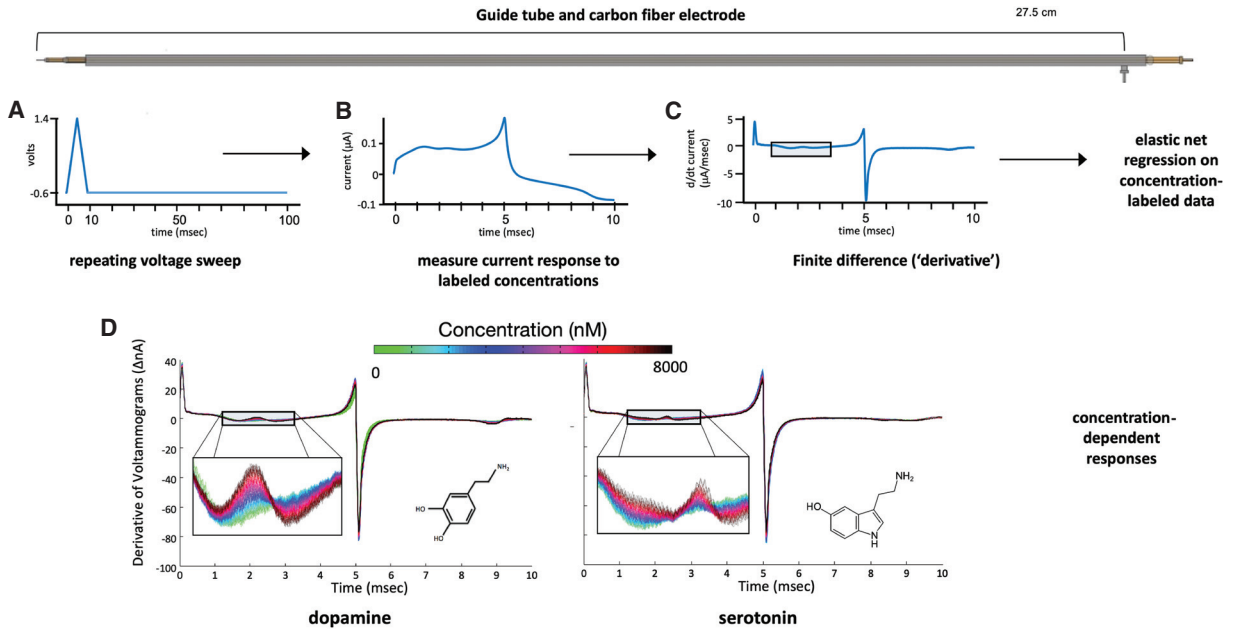
Figure 1. Elastic net electrochemistry. (*Top*) Diagram of the guide tube used during neurosurgery for DBS electrode implantation. The carbon fiber probe is inserted through the guide tube, and the stainless-steel pin at right acts as reference ground. This is the same ground used during training of models from flow cell data. Workflow for elastic net electrochemistry (see Kishida et al. 2016; Moran et al. 2018). (*A*) Voltage waveform on electrode. (*B*) Measured current time series during ~10-msec triangular waveform portion of the 100-msec duty cycle. (*C*) Finite time difference of current. (*D*) Carbon fiber electrode responses to dopamine and serotonin concentrations in format of finite time difference plot. Concentration-specific information about dopamine and serotonin is "wiggly," but this information is not simply concentrated at the theoretical oxidation potentials for both neuromodulators.

sets to show how well it generalizes to measurements made with completely naive electrodes. Figure 2 shows one example of how a serotonin model generalizes out of probe. This figure also shows that the model reports 0 serotonin (when that is actually the case) and reports no response to pH changes and dopamine changes over a large range. The same basic approach is taken for multianalyte mixtures (Fig. 3) and for depth electrodes used for epilepsy monitoring in humans (Fig. 4).

For purposes of discussion, here we focus the description of our approach on detecting dopamine, but the same basic principles apply to generating multianalyte models (like those shown in Figs. 3–6). To fit a model using known concentrations of dopamine in the context of varying pH, we use the "elastic net" to perform regularization and automatic variable selection to determine a good fit for a linear regression model

$$\hat{y} = \hat{\beta}_0 + x_1\hat{\beta}_1 + x_2\hat{\beta}_2 + \cdots + x_p\hat{\beta}_p$$

that predicts the concentration of dopamine (y) given a fast-scan cyclic voltammetry measurement. Note, one fast-scan cyclic voltammetry measurement is equal to the current measured during the application of a 10-msec triangular voltage sweep, as indicated in Figure 1A, followed by a 90-msec "wait period" for total of a 100-msec duty cycle. Here, we aim to estimate $\hat{y}$, the predicted concentration of dopamine for each 100-msec cycle (Fig. 1A) given the vector of parameters $x_1 \ldots x_p(\vec{x})$, which is the finite time derivative (dI/dt) of a single cyclic

voltammogram measurement. The betas $\vec{\beta}$ are regression weights. The elastic net procedure for linear regression models minimizes the residual sum of squares with an additional penalty term, $P_\alpha(\beta)$. The elastic net penalty

$$P_\alpha(\beta) = (1 - \alpha)(1/2)||\beta||^2_{\ell_2} + \alpha||\beta||_{\ell_1}$$

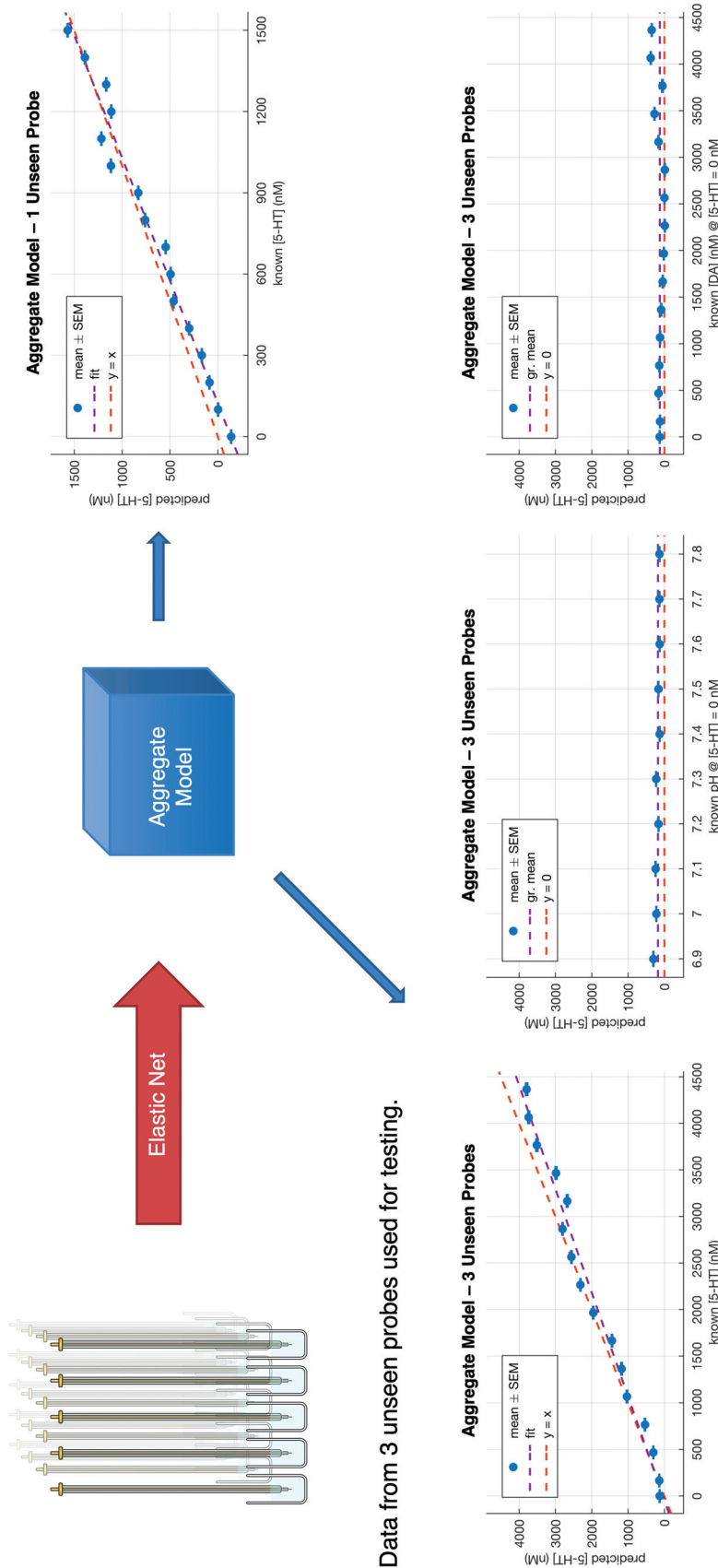is a mixture (convex hull) of the "ridge regression penalty"

$$\ell_2 - norm: (1/2)||\beta||^2_{\ell_2}$$

(Hoerl and Kennard 1970) and "lasso penalty"

$$\ell_1 - norm: ||\beta||_{\ell_1}$$

(Tibshirani 1996) parameterized by α, which takes a value between 0 and 1. To determine a best-fit linear regression model, we collect voltammetric measurements from samples of known concentrations of analyte in vitro and perform 10-fold cross-validation before further validating model performance on out-of-sample test cases (Kishida et al. 2016).

In one sense, our approach is not remarkable—we use off-the-shelf voltage clamp hardware (Molecular Devices) and off-the-shelf software (elastic net through calls to the glmnet toolbox in MATLAB). The model building approach, going from big data to a well-behaved model, uses standard principles in statistical learning methods. However, one major departure from prior voltammetric inference methods is that we train on large data sets (e.g.,

Figure 2. Out-of-probe and out-of-concentration generalization of elastic net electrochemistry approach. Twenty probes' responses were aggregated to form a composite model for predicting dopamine and serotonin that would not track pH and would generalize well out of probe. Single unseen probe predictions versus actual serotonin levels in flow cell (*upper right*). Aggregate model tested on three unseen probes (*bottom* row—these plots are averages over responses of the three naive probes). Here we show the serotonin model matching actual serotonin, not tracking pH, and not tracking dopamine. An average of 20,000 samples per probe were taken (this included 160 different concentration levels) for a grand total of ~400,000 samples for the entire 20-probe aggregate. The model so extracted generalizes well out of probe and out of sample (the concentrations in these plots are out of sample).
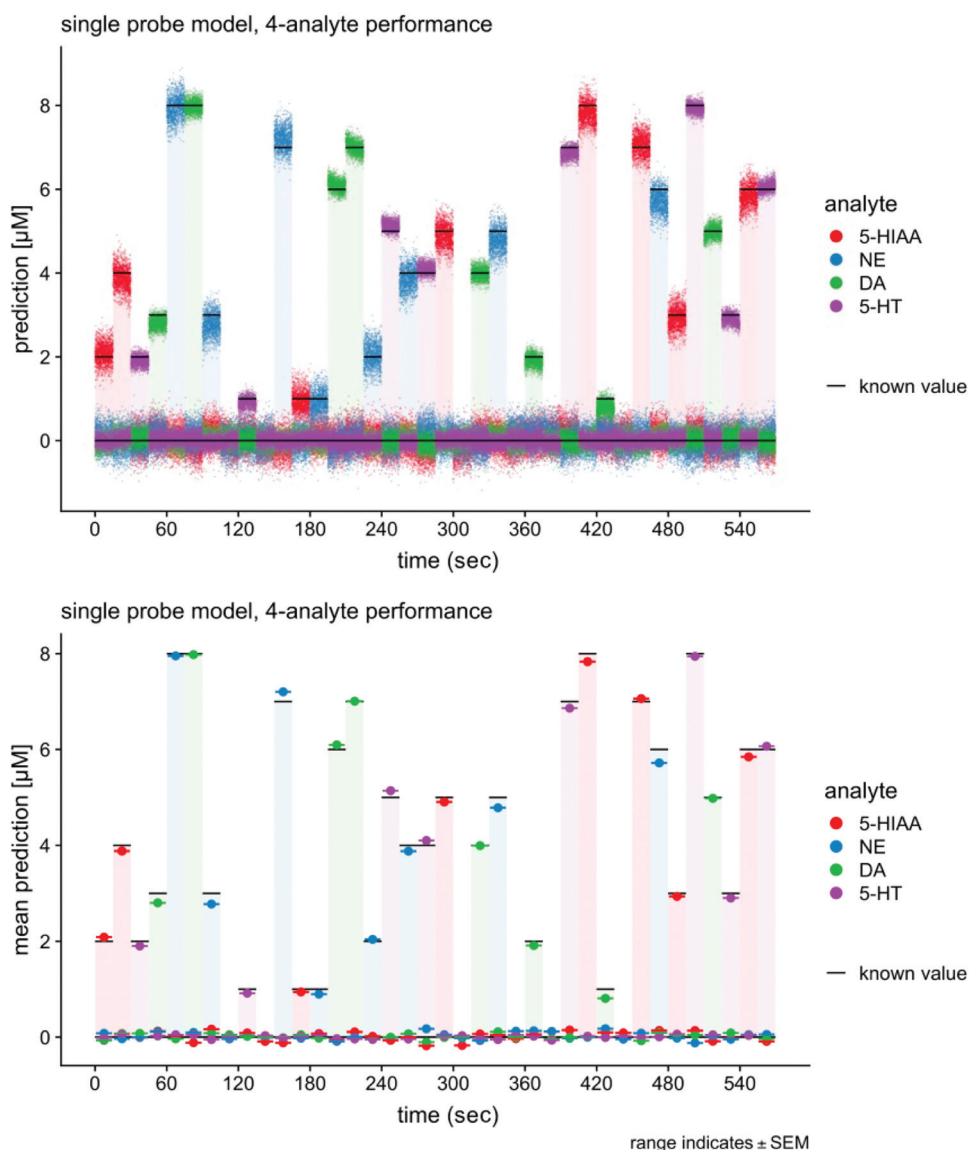
Figure 3. Four analyte model performance at ~10 msec per estimate. (*Top*) Moment-to-moment predictions (10.3 msec per estimate [97 Hz]) of four analyte model extracted by elastic net from a four-analyte mixture in a flow cell. (*Bottom*) Average performance of the four-analyte model. Notice the model can also separate serotonin from 5-hydroxy-indole-acetic acid. Carbon fiber electrode.

400,000 sweeps and hundreds of concentrations) and across many electrodes (Fig. 2 shows model extracted across recordings from 20 electrodes). Further, no experimenter judgement is necessary regarding the shape of the voltammogram—with large amounts of data the nonspecific variability in these responses is regressed out of the resulting model. In building models this way, we believe we have provided a path to standardization of these approaches in that we have removed the experimenter judgement bias that is implicit in standard voltammetric inference methods. What bias remains, importantly, is reportable in that all of the bias resides in the calibration data sets used to train a given model. This allows reinterrogation of existing data and a scientific approach to identifying features that improve or degrade the precision of new models.

Elastic Net Electrochemistry on MultiAnalyte Mixtures and Common Electrodes

A major challenge for recording dopamine, serotonin, and norepinephrine concurrently is the selectivity of the extracted models. That is, the goal of any inference method is to determine models that predict out-of-sample measurement of analyte mixtures (e.g., dopamine and serotonin), distinguish the analytes from one another, and distinguish them from other compounds that could confuse the measurements (e.g., norepinephrine, pH, and 5-hydroxy-indole-acetic acid). Figure 3 shows just such a separation (at 10.3 msec per estimate in which the 90-msec "waiting period" from Figure 1A has been dropped). These measurements were made in a flow cell where the concentrations of dopamine, norepinephrine, and serotonin, 5HIAA
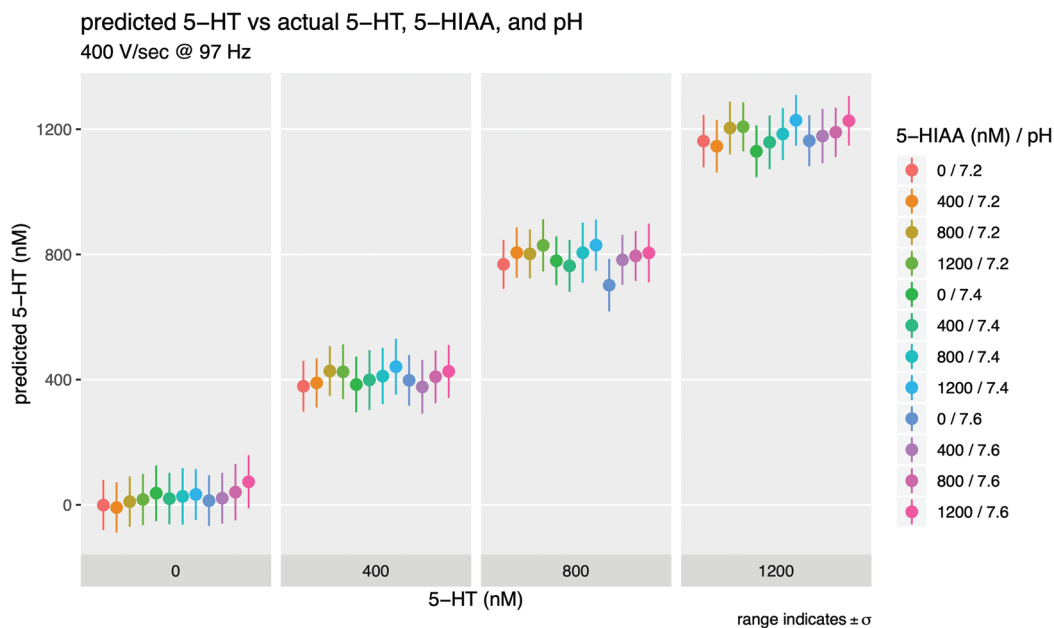
Figure 4. Serotonin model performance versus pH and 5HIAA (10.3 msec per estimate). A three-analyte model was trained to predict serotonin concentration in the context of varying pH and varying concentration of 5-HIAA, a known metabolite of serotonin that confuses prior voltammetric inference methods. From *left* to *right*, serotonin concentration predictions are stable against a background of increasing pH or increasing 5-HIAA concentration.

were controlled exactly. The top panel in Figure 3 shows moment-by-moment predictions of the model (at 10-msec resolution) and the bottom panel shows the average value of these predictions for each known concentration. Figure 4 shows the elastic net extracted 5HT model prediction as a function of contaminating pH and 5HIAA concentrations (again, 10.3 msec per estimate as in Fig. 3).

Figure 5 shows predictions of a dopamine, serotonin, and 5HIAA model extracted using an AdTech electrode used for stereo-EEG recording in humans. As noted in the figure legend, the models were extracted for the micro-contacts as indicated. These estimates are at 10.3 msec per estimate (~97 Hz) similar to the results reported in Figures 2 and 3. These electrodes are just one example of platinum–iridium electrodes in common use in humans and similar electrodes in common use in model organisms. For model organisms, these methods open up many possibilities for neuromodulator recordings, but for humans they could transform any platinum–iridium contact (within an appropriate impedance range) into a source of fast neurochemical information about dopamine, serotonin, norepinephrine, and even the oxidative metabolite of serotonin, 5HIAA. This kind of information could be used in a host of cognitive paradigms, sleep, or changes in consciousness to extract their dependence on these important neuromodulators.

Figure 6 suggests that the fast information at 10.3 msec per sample (Figs. 2–4) could be reduced to the order of a millisecond or even faster. This latter possibility would put these measurements on the same order of magnitude as action potentials and modulations in their rate, which would make these neuromodulator measurements capable of tracking changes in spiking rates thought to carry pre-

diction error signals to target neural structures. Figure 6A illustrates the impetus supporting the idea that order-millisecond estimates are possible. Here, the 10-msec triangular voltage waveform is subsampled at random, using only 50% of the points in the measured time series; a finite time difference is computed across the remaining down-sampled data (here it is actually a finite index difference because consecutive points are not necessarily contiguous in time); and this "time" difference is used to fit an elastic net regression–based concentration-prediction model. As shown, this procedure works well at 50% and 10% down-sampling. The predictions shown are out-of-sample concentration estimates, but generalization to naive probes for such down-sampling awaits future experiments.

Applications of Elastic Net Electrochemistry in Human Striatum during Active Investment Game

We have presented a summary of an approach to electrochemical detection that makes possible the colocalized detection of dopamine, norepinephrine, and serotonin from both carbon fibers and high-impedance platinum–iridium electrodes that are in routine clinical use around the world. These technical developments open up the possibility of using human subjects in whom neuromodulator recordings at 10 msec or better could be paired with quantitative behavioral estimates and using electrodes put in place for other reasons (currently clinical reasons). These same methodologies also offer promise for model organism research and should be very useful in calibrating new expressible optical sensors for neuromodulators—for example, the new and exciting DLIGHT reporter for dopamine (Patriarchi et al. 2018). Notably, simultaneous
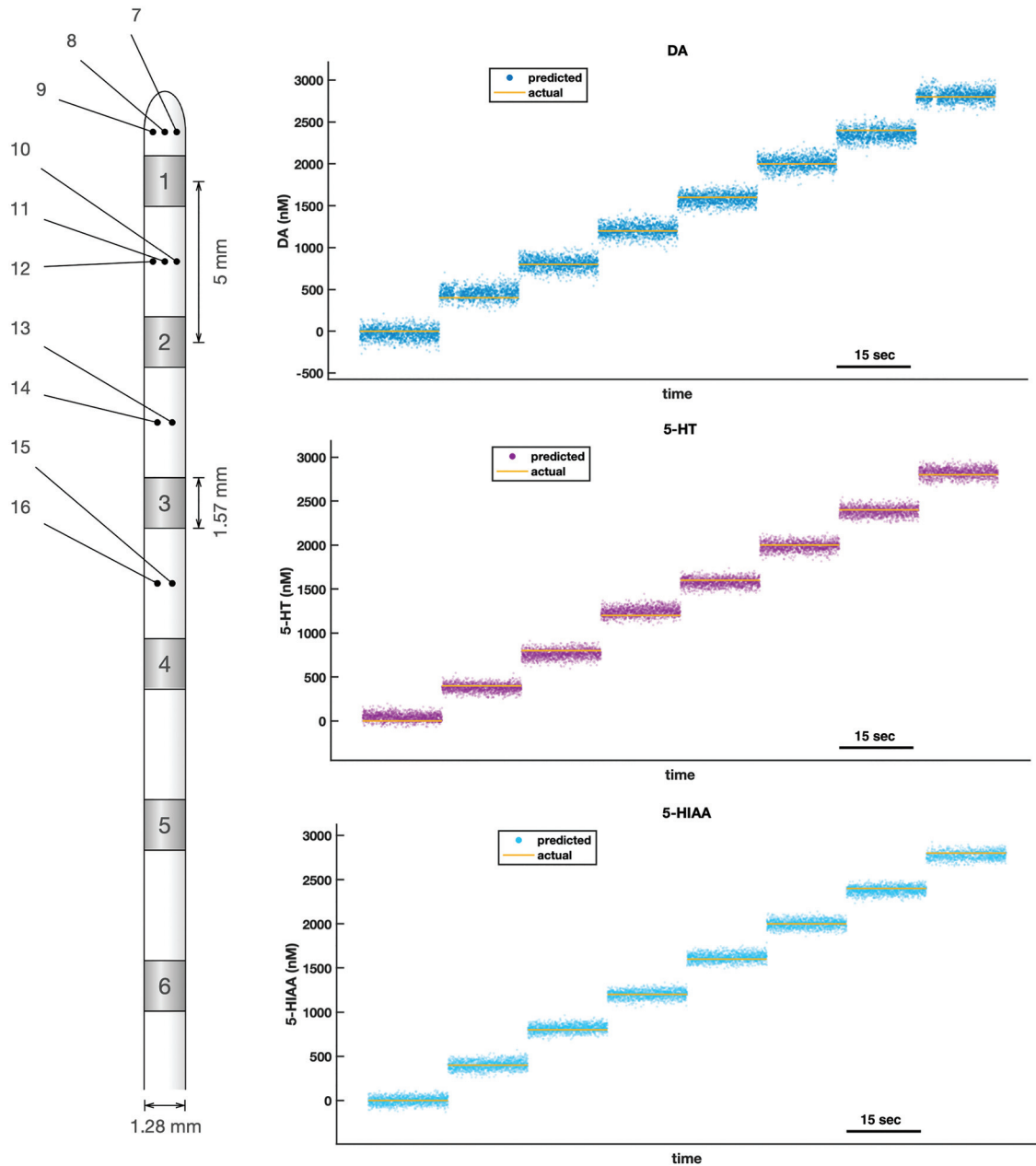
Figure 5. Elastic net electrochemistry models on human electrophysiology electrodes. Three-analyte model on an AdTech human depth electrode with low-impedance macrocontacts and high-impedance microcontacts. Cartoon at *left* shows electrode configuration. Microcontacts are distributed radially (and uniformly), but shown here as small dots indicating number of contacts at each location. These data (shown to the *right*) are from a model extracted in a flow cell using contacts 7 and 10. Each colored dot (labeled prediction in each panel) is a prediction for each 10.3-msec bin. Here a triplet mixture of dopamine, serotonin, and 5HIAA is used. All predictions shown here are made from measurements not used in training the model.

multianalyte detection does not currently seem to be feasible using receptor-based optogenetic methods suggesting complementary, but distinct, roles for high-speed electrochemical detection methods alongside rapidly developing optogenetic approaches.

Using carbon fiber microelectrodes and the opportunity afforded by deep brain stimulating (DBS) electrode implantation in humans (for Parkinson's disease or essential tremors), elastic net electrochemistry has been performed in conscious human subjects during the execution of a simple investment game (Kishida et al. 2016; Moran et al. 2018; also see Platt and Pearson 2016 for commentary). This game is cartooned in Figure 7A. Subjects are endowed with $100 and presented with a market trace, they invest between 0% and 100% of their holdings, the market fluctuates, and they experience a gain or loss. This repeats for 20 rounds for each market. The surgical patients practiced this game before surgery. In the surgical
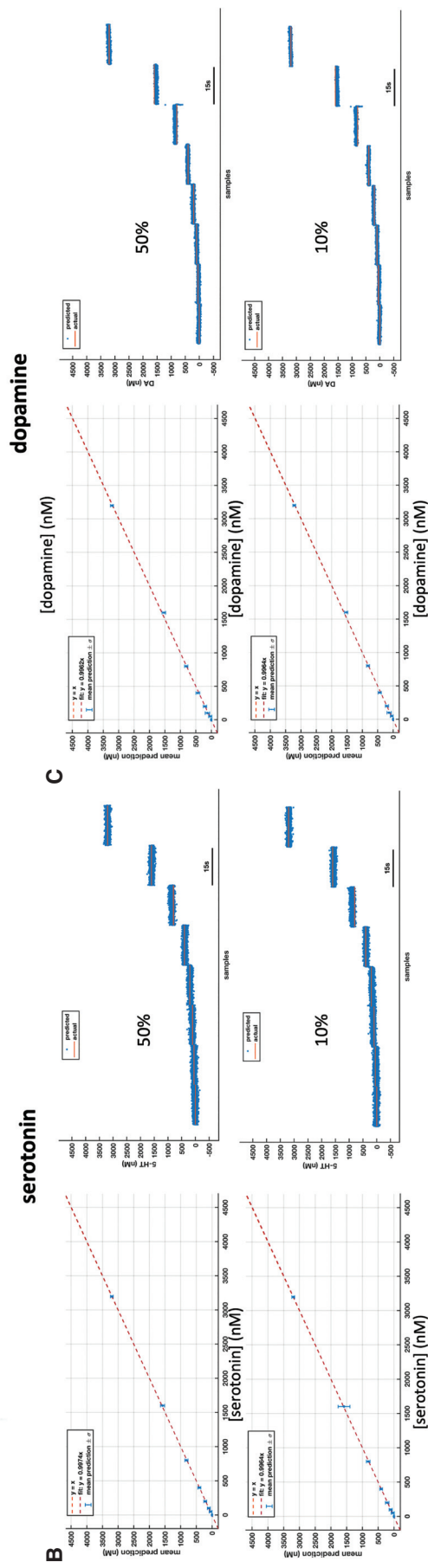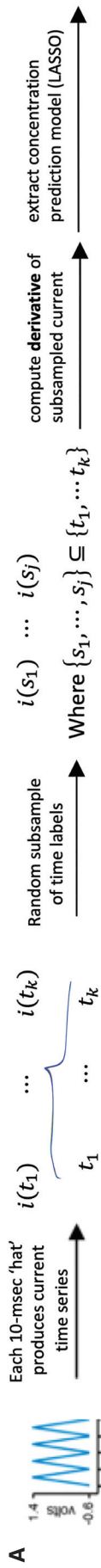
Figure 6. Evidence that 1 msec of data is sufficient for elastic net electrochemistry to extract a concentration-prediction model. (*A*) Triangular voltage wave repeating at ~10 msec is randomly down-sampled, respectively at 50% and 10%, the derivative of the subsampled current time series is computed, and, as above, this "derivative" is entered into an elastic net regression to extract dopamine and serotonin prediction models. Note that the randomly chosen points, although ordered in time, will not be contiguous at the base sampling rate; nevertheless, the prediction models remain excellent. The plots in *B* and *C* show out-of-sample predictions for the mixtures averaged for two out-of-sample probes.
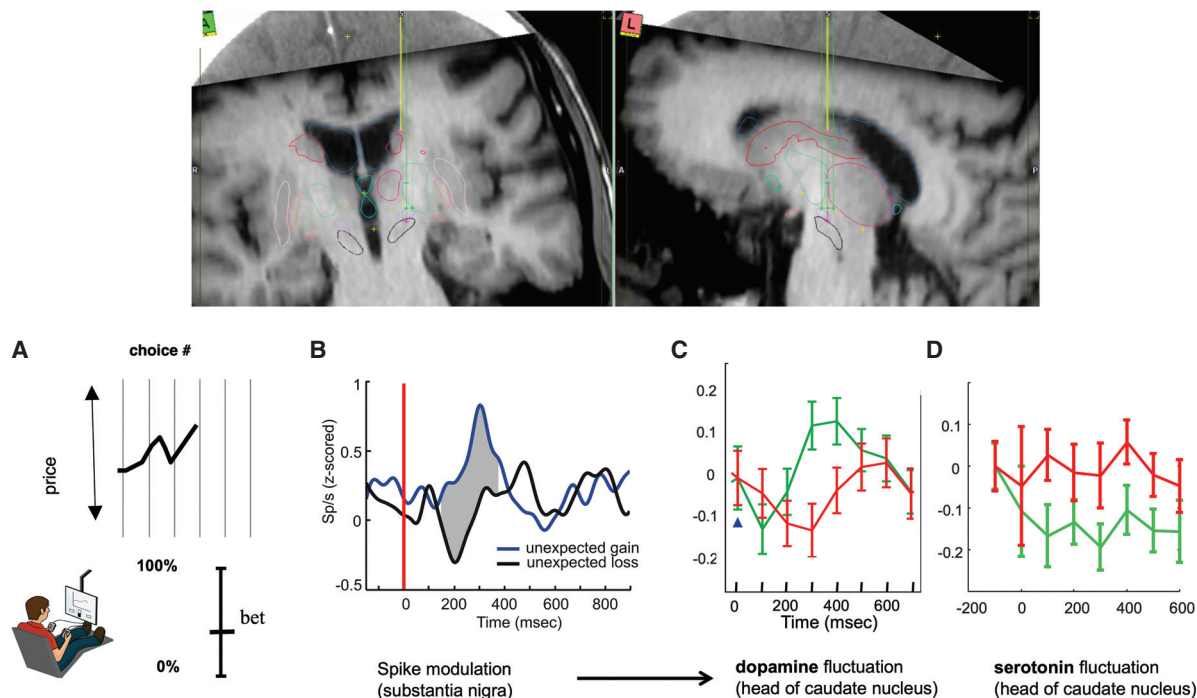
Figure 7. Application of elastic net electrochemistry in human striatum during investment game. *Top inset* shows placement of an electrochemical sensor in the caudate during a DBS electrode implantation procedure for a patient with Parkinson's disease. The electrochemical sensor (see Fig. 1, *top*) follows the yellow path before functional mapping of the eventual DBS electrode path (green path terminating in purple cross in *left* panel). (*A*) Market investment task (Lohrenz et al. 2007; Kishida et al. 2016; Moran et al. 2018); (*B*) Reward prediction errors during a simple card game encoded in spike modulation in human substantia nigra (Zaghloul et al. 2009). (*C*) Subsecond dopamine release in the caudate encodes reward prediction errors during investment task when investments are 100% of participant's portfolio Kishida et al. (2016). (*D*) Subsecond serotonin release encodes an opponent signal to dopamine release for reward prediction errors (Moran et al. 2018) in the same task events as in *C*.

suite, subjects played six markets (see for BOLD imaging on this task, Lohrenz et al. 2007; for dopamine recordings in caudate on this task, Kishida et al. 2016; for serotonin recordings in caudate during this task, Moran et al. 2018).

For Parkinson's patients, this is a very engaging task, which is important, because subjects off their dopamine precursor medication fatigue quickly. One basic finding that comports with extant data from human single unit recordings in substantia nigra is that for high bets or bets "all in," changes in dopamine delivery to the caudate encode positive prediction errors as positive-going transients (Fig. 7C) and negative reward prediction errors as negative-going transients (Fig. 7C). Using a slightly modified version of elastic net electrochemistry, Moran et al. (2018) showed an opponent pattern to serotonin fluctuations in human caudate nucleus (from the same carbon fiber that recorded the dopamine transients)—positive-going serotonin transients encoded negative reward prediction errors and negative-going transients encoded positive reward prediction errors. This task is designed to ask how on a fixed budget subjects allocate their money with the amount "not risked in the market" remaining in their pocket. Together, these data are the first subsecond recordings of either dopamine or serotonin in human subjects and the first clear subsecond report of such opponent encodings.

DISCUSSION

We have presented a new approach to the selective detection of biogenic amines that springboards off work in fast-scan cyclic voltammetry on carbon fiber electrodes but apparently exploits a feature of the current time series data that had not been targeted by previous approaches. We highlight this feature—that information encoding specific concentrations of dopamine and serotonin is distributed coherently throughout the electrochemical current time series. We showed how this information can be easily exploited by a modern machine learning method—the elastic net—to extract a concentration-prediction model for multiple analytes that include dopamine, serotonin, and norepinephrine. This was accomplished with off-the-shelf hardware and software, which we believe warrants the coupling of other, perhaps more sophisticated inference approaches, to similar electrochemical approaches. These methodological steps forward and their standardization open up the possibility of testing important hypotheses about dopamine, serotonin, and norepinephrine function at order millisecond timescales and in human brains (our focus in this paper).

We presented an array of experiments supporting the separation of dopamine, serotonin, and norepinephrine from one another and from pH and at least one oxidative metabolite of serotonin, 5-hydroxy-indole-acetic acid.

These models were extracted in a flow cell environment by comparison to a stainless-steel ground capable of being used in human beings, and we reviewed one application of the approach to dopamine and serotonin detection in the striatum of conscious humans (Kishida et al. 2016; Moran et al. 2018). We also showed a preliminary model extracted similarly from a commercially available epilepsy electrode in routine use in human depth electrode monitoring suggesting that such depth electrodes could become sources of (potentially) ultrafast neurochemical information about dopamine, serotonin, norepinephrine, and perhaps other oxidizable species. This exciting use of the approach has the potential to provide new information about neuromodulatory function in humans and invites similar work in model organisms where a flexible method for their colocalized recording has been lacking.

Last, we showed through a randomized down-sampling procedure that order-millisecond or better estimates were possible using elastic net electrochemistry (Fig. 6). This demonstration suggests that a random 10% of the points collected during the ∼10-msec triangular sweep contains sufficient information to estimate a reasonably accurate out-of-sample predictive model for both dopamine and serotonin. In Figure 6, there are 1000 time points defining the triangular waveform (the voltage forcing function), which means the 10% case is only 100 random points spread throughout the ∼10-msec duty cycle. These data are not yet definitive because one needs to test these kinds of manipulations out-of-probe on naive electrodes after training on a group of electrodes; however, it suggests that there is only a loose dependence of the models on the time ordering of the points and even on the voltage. These two observations together suggest radically different approaches might also be possible, but those await future experiments.

Testing the Reward Prediction Error Hypothesis throughout the Human Brain

Dopamine signaling in the human brain represents a crucial physical substrate that supports motivated learning (Wise 2004; Bromberg-Martin et al. 2010), value-dependent action choice (Montague et al. 2004), working memory (Cools and Esposito 2011), motor learning (Graybiel 1995), and a variety of other cognitive functions. Consequently, perturbed dopamine signaling plays a major, but complicated role in a range of conditions including drug addiction and Parkinson's disease. Despite the importance of dopamine signaling in human mental function, there has previously been no method to gain access to ongoing fast changes (subsecond) in dopamine delivery in the human brain.

As presented here, elastic net electrochemistry addresses this gap in our understanding of dopamine signaling by implementing a new methodology for recording ultrafast (potentially order-millisecond) dopamine fluctuations on standard electrodes used in human electrophysiological recordings and using this methodology to test an influential computational model of reward learning: the temporal difference (TD) reward prediction error hypothesis for dopamine (Glimcher 2011; Dayan 2012; Platt and Pearson 2016; also see Montague et al. 1993, 1994, 1995, 1996, 2004, 2006; Schultz et al. 1997; McClure et al. 2003; O'Doherty et al. 2003, 2004; Tobler et al. 2005).

What is the reward prediction error hypothesis for dopamine? To quote Michael Platt and John Pearson (2016),

> *Dopamine encodes a key variable posited by theories of reinforcement learning. These theories posit that animals select behaviors on the basis of which ones they expect to result in reward, updating their beliefs on the difference between expectations and observed outcomes, good or bad (3). This reward prediction error is large when rewards are unexpected and small when rewards are fully predicted, and its magnitude and sign drive the speed and direction of learning, respectively.*

(Also see Montague et al. 1993, 1995, 1996, 2004, 2006; Montague and Sejnowski 1994; Schultz et al. 1997; Bayer and Glimcher 2005; Dayan and Daw 2008; Dayan and Niv 2008; Glimcher 2011; Dayan 2012.) This model has been tested in nonhuman primates at the level of spiking activity in midbrain dopamine neurons (e.g., Hollerman and Schultz 1998; Tobler et al. 2005; also see Schultz et al. 2015 for review). However, in human and nonhuman primates, the hypothesis has never been tested in terms of fast changes in dopamine delivery—this is a central gap in our understanding of reward learning in humans and its contribution to important features of human health.

A number of health disorders in humans involve changes in reward processing, value-based decision-making, and reward learning. These include (but are not limited to) substance use disorders (Bickel and Marsch 2001; Chiu et al. 2008; Gu et al. 2015), psychosis (Sevy et al. 2007), anxiety disorders (Tolin et al. 2003; Casada and Roache 2005; Jovanovic et al. 2010), and mood disorders (Pizzagalli 2014; also see Beevers et al. 2013; Chiu and Deldin 2007; Dayan and Huys 2008; Kumar et al. 2008; Chase et al. 2010; Gradin et al. 2011; Kunisato et al. 2012; Huys et al. 2013; Greenberg et al. 2015; Rothkirch et al. 2017). Consequently, a central question arises for human health: How does subsecond dopamine delivery conform to the reward prediction error hypothesis? In humans, a clear answer to this question and its dependence on the target neural region would represent a major step forward in our understanding of dopamine's role in neural processing and hence its potential contributions to diminished human health.

In humans, there is one experiment (a card game) showing clearly that spiking activity in neurons in substantia nigra changes according to a simple prediction error signal (Fig. 7B; Zaghloul et al. 2009); however, there has been no direct assessment of the "other end of the problem" in neural structures where dopaminergic neurons project (with the exception of Kishida et al. 2016 and Moran et al. 2018 as discussed). So, the question arises, "Do measured fluctuations in dopamine delivery actually encode reward prediction error signals throughout the human brain?" The new methodological approaches presented here will allow this question to be asked with selectivity, that is, separating dopamine from serotonin and norepi-

nephrine and real-time temporal resolution. As for specific behavioral experiments designed around serotonin and norepinephrine delivery, those are far too numerous to list here; however, we anticipate that having the technology to ask such questions will inspire the next generation of human neuroscience research.

ACKNOWLEDGMENTS

REFERENCES

Bayer HM, Glimcher PW. 2005. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47:** 129–141. doi:10.1016/j.neuron.2005.05.020

Beevers CG, Worthy DA, Gorlick MA, Nix B, Chotibut T, Todd Maddox W. 2013. Influence of depression symptoms on history-independent reward and punishment processing. *Psychiatry Res* **207:** 53–60. doi:10.1016/j.psychres.2012.09.054

Bickel WK, Marsch LA. 2001. Toward a behavioral economic understanding of drug dependence: delay discounting processes. *Addiction* **96:** 73–86. doi:10.1046/j.1360-0443.2001.961736.x

Bromberg-Martin ES, Matsumoto M, Hikosaka O. 2010. Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* **68:** 815–834. doi:10.1016/j.neuron.2010.11.022

Casada JH, Roache JD. 2005. Behavioral inhibition and activation in posttraumatic stress disorder. *J Nerv Ment Dis* **193:** 102–109. doi:10.1097/01.nmd.0000152809.20938.37

Chase HW, Frank MJ, Michael A, Bullmore ET, Sahakian BJ, Robbins TW. 2010. Approach and avoidance learning in patients with major depression and healthy controls: relation to anhedonia. *Psychol Med* **40:** 433–440. doi:10.1017/S0033291709990468

Chiu PH, Deldin PJ. 2007. Neural evidence for enhanced error detection in major depressive disorder. *Am J Psychiatry* **164:** 608–616. doi:10.1176/ajp.2007.164.4.608

Chiu PH, Lohrenz TM, Montague PR. 2008. Smokers' brains compute, but ignore, a fictive error signal in a sequential investment task. *Nat Neurosci* **11:** 514–520. doi:10.1038/nn2067

Clark JJ, Sandberg SG, Wanat MJ, Gan JO, Horne EA, Hart AS, Akers CA, Parker JG, Willuhn I, Martinez V, et al. 2010. Chronic microsensors for longitudinal, subsecond dopamine detection in behaving animals. *Nat Methods* **7:** 126–129. doi:10.1038/nmeth.1412

Cools R, D'Esposito M. 2011. Inverted-U-shaped dopamine actions on human working memory and cognitive control. *Biol Psychiatry* **69:** e113–e125. doi:10.1016/j.biopsych.2011.03.028

Dayan P. 2012. Twenty-five lessons from computational neuromodulation. *Neuron* **76:** 240–256. doi:10.1016/j.neuron.2012.09.027

Dayan P, Daw ND. 2008. Decision theory, reinforcement learning, and the brain. *Cogn Affect Behav Neurosci* **8:** 429–453. doi:10.3758/CABN.8.4.429

Dayan P, Huys QJ. 2008. Serotonin, inhibition, and negative mood. *PLoS Comput Biol* **4:** e4. doi:10/1371/journal.pcbi.0040004

Dayan P, Niv Y. 2008. Reinforcement learning: the good, the bad and the ugly. *Curr Opin Neurobiol* **18:** 185–196. doi:10.1016/j.conb.2008.08.003

Glimcher PW. 2011. Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proc Natl Acad Sci* **108:** 15647–15654. doi:10.1073/pnas.1014269108

Gradin VB, Kumar P, Waiter G, Ahearn T, Stickle C, Milders M, Reid I, Hall J, Steele JD. 2011. Expected value and prediction error abnormalities in depression and schizophrenia. *Brain* **134:** 1751–1764. doi:10.1093/brain/awr059

Graybiel AM. 1995. Building action repertoires: memory and learning functions of the basal ganglia. *Curr Opin Neurobiol* **5:** 733–741. doi:10.1016/0959-4388(95)80100-6

Greenberg T, Chase HW, Almeida JR, Stiffler R, Zevallos CR, Aslam HA, Deckersbach T, Weyandt S, Cooper C, Toups M, et al. 2015. Moderation of the relationship between reward expectancy and prediction error-related ventral striatal reactivity by anhedonia in unmedicated major depressive disorder: findings from the EMBARC study. *Am J Psychiatry* **172:** 881–891. doi:10.1176/appi.ajp.2015.14050594

Gu X, Lohrenz T, Salas R, Baldwin PR, Soltani A, Kirk U, Cinciripini PM, Montague PR. 2015. Belief about nicotine selectively modulates value and reward prediction error signals in smokers. *Proc Natl Acad Sci* **112:** 2539–2544. doi:10.1073/pnas.1416639112

Hoerl AE, Kennard RW. 1970. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics* **12:** 55–67. doi:10.1080/00401706.1970.10488634

Hollerman JR, Schultz W. 1998. Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat Neurosci* **1:** 304–309. doi:10.1038/1124

Huffman ML, Venton BJ. 2009. Carbon-fiber microelectrodes for in vivo applications. *Analyst* **134:** 18–24. doi:10.1039/B807563H

Huys QJ, Pizzagalli DA, Bogdan R, Dayan P. 2013. Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biol Mood Anxiety Disord* **3:** 12. doi:10.1186/2045-5380-3-12

Jovanovic T, Norrholm SD, Blanding NQ, Davis M, Duncan E, Bradley B, Ressler KJ. 2010. Impaired fear inhibition is a biomarker of PTSD but not depression. *Depress Anxiety* **27:** 244–251. doi:10.1002/da.20663

Kishida KT, Sandberg SG, Lohrenz T, Comair YG, Sáez I, Phillips PE, Montague PR. 2011. Sub-second dopamine detection in human striatum. *PLoS One* **6:** e23291. doi:10.1371/journal.pone.0023291

Kishida KT, Saez I, Lohrenz T, Witcher MR, Laxton AW, Tatter SB, White JP, Ellis TL, Phillips PE, Montague PR. 2016. Sub-second dopamine fluctuations in human striatum encode superposed error signals about actual and counterfactual reward. *Proc Natl Acad Sci* **113:** 200–205. doi:10.1073/pnas.1513619112

Kuhr WG, Wightman RM. 1986. Real-time measurement of dopamine release in rat brain. *Brain Res* **381:** 168–171. doi:10.1016/0006-8993(86)90707-9

Kumar P, Waiter G, Ahearn T, Milders M, Reid I, Steele JD. 2008. Abnormal temporal difference reward-learning signals in major depression. *Brain* **131:** 2084–2093. doi:10.1093/brain/awn136

Kunisato Y, Okamoto Y, Ueda K, Onoda K, Okada G, Yoshimura S, Suzuki S, Samejima K, Yamawaki S. 2012. Effects of depression on reward-based decision making and variability of action in probabilistic learning. *J Behav Ther Exp Psychiatry* **43:** 1088–1094. doi:10.1016/j.jbtep.2012.05.007

Lohrenz T, McCabe K, Camerer CF, Montague PR. 2007. Neural signature of fictive learning signals in a sequential investment task. *Proc Natl Acad Sci* **104:** 9493–9498. doi:10.1073/pnas.0608842104

McClure SM, Berns GS, Montague PR. 2003. Temporal prediction errors in a passive learning task activate human striatum. *Neuron* **38:** 339–346. doi:10.1016/S0896-6273(03)00154-5

Mermet C, Gonon F. 1986. In vivo voltammetric monitoring of noradrenaline release and catecholamine metabolism in the hypothalamic paraventricular nucleus. *Neuroscience* **19:** 829–838. doi:10.1016/0306-4522(86)90301-5

Montague PR, Sejnowski TJ. 1994. The predictive brain: temporal coincidence and temporal order in synaptic learning mechanisms. *Learn Mem* **1:** 1–33.

Montague PR, Dayan P, Nowlan SJ, Pouget A, Sejnowski TJ. 1993. Using aperiodic reinforcement for direct self-organization. In *Advances in neural information processing systems 5* (ed. Hanson SJ, Cowan JD, Giles CL), pp. 969–976. Morgan Kauffman, San Mateo, CA.

Montague PR, Dayan P, Person C, Sejnowski TJ. 1995. Bee foraging in uncertain environments using predictive Hebbian learning. *Nature* **377:** 725–728. doi:10.1038/377725a0

Montague PR, Dayan P, Sejnowski TJ. 1996. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* **16:** 1936–1947. doi:10.1523/JNEUROSCI.16-05-01936.1996

Montague PR, Hyman SE, Cohen JD. 2004. Computational roles for dopamine in behavioural control. *Nature* **431:** 760–767. doi:10.1038/nature03015

Montague PR, King-Casas B, Cohen JD. 2006. Imaging valuation models in human choice. *Annu Rev Neurosci* **29:** 417–448. doi:10.1146/annurev.neuro.29.051605.112903

Moran RJ, Kishida KT, Lohrenz T, Saez I, Laxton AW, Witcher MR, Tatter SB, Ellis TL, Phillips PE, Dayan P, et al. 2018. The protective action encoding of serotonin transients in the human brain. *Neuropsychopharmacology* **43:** 1425–1435. doi:10.1038/npp.2017.304

O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ. 2003. Temporal difference models and reward-related learning in the human brain. *Neuron* **38:** 329–337. doi:10.1016/S0896-6273(03)00169-7

O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. 2004. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* **304:** 452–454. doi:10.1126/science.1094285

Patriarchi T, Cho JR, Merten K, Howe MW, Marley A, Xiong WH, Folk RW, Broussard GJ, Liang R, Jang MJ, et al. 2018. Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. *Science* **360:** eaat4422. doi:10.1126/science.aat4422

Phillips PE, Stuber GD, Heien ML, Wightman RM, Carelli RM. 2003. Subsecond dopamine release promotes cocaine seeking. *Nature* **422:** 614–618. doi:10.1038/nature01476

Pizzagalli DA. 2014. Depression, stress, and anhedonia: toward a synthesis and integrated model. *Annu Rev Clin Psychol* **10:** 393–423. doi:10.1146/annurev-clinpsy-050212-185606

Platt ML, Pearson JM. 2016. Dopamine: context and counterfactuals. *Proc Natl Acad Sci* **113:** 22–23. doi:10.1073/pnas.1522315113

Qian J, Hastie T, Friedman J, Tibshirani R, Simon N. 2013. Glmnet for Matlab. http://www.stanford.edu/~hastie/glmnet_matlab/.

Robinson DL, Hermans A, Seipel AT, Wightman RM. 2008. Monitoring rapid chemical communication in the brain. *Chem Rev* **108:** 2554–2584. doi:10.1021/cr068081q

Rothkirch M, Tonn J, Köhler S, Sterzer P. 2017. Neural mechanisms of reinforcement learning in unmedicated patients with major depressive disorder. *Brain* **140:** 1147–1157. doi:10.1093/brain/awx025

Schultz W, Dayan P, Montague PR. 1997. A neural substrate of prediction and reward. *Science* **275:** 1593–1599. doi:10.1126/science.275.5306.1593

Schultz W, Carelli RM, Wightman RM. 2015. Phasic dopamine signals: from subjective reward value to formal economic utility. *Curr Opin Behav Sci* **5:** 147–154. doi:10.1016/j.cobeha.2015.09.006

Sevy S, Burdick KE, Visweswaraiah H, Abdelmessih S, Lukin M, Yechiam E, Bechara A. 2007. Iowa gambling task in schizophrenia: a review and new data in patients with schizophrenia and co-occurring cannabis use disorders. *Schizophr Res* **92:** 74–84. doi:10.1016/j.schres.2007.01.005

Stamford JA. 1990. Fast cyclic voltammetry: measuring transmitter release in 'real time'. *J Neurosci Methods* **34:** 67–72. doi:10.1016/0165-0270(90)90043-F

Stamford JA, Kruk ZL, Millar J, Wightman RM. 1984. Striatal dopamine uptake in the rat: in vivo analysis by fast cyclic voltammetry. *Neurosci Lett* **51:** 133–138. doi:10.1016/0304-3940(84)90274-X

Sutton RS, Barto AG. 2018. *Reinforcement learning: an introduction*. MIT Press, Cambridge, MA.

Tibshirani R. 1996. Regression shrinkage and selection via the lasso. *J R Stat Soc Series B Methodol* **58:** 267–288. doi:10.1111/j.2517-6161.1996.tb02080.x

Tobler PN, Fiorillo CD, Schultz W. 2005. Adaptive coding of reward value by dopamine neurons. *Science* **307:** 1642–1645. doi:10.1126/science.1105370

Tolin DF, Abramowitz JS, Brigidi BD, Foa EB. 2003. Intolerance of uncertainty in obsessive–compulsive disorder. *J Anxiety Disord* **17:** 233–242. doi:10.1016/S0887-6185(02)00182-2

Wise RA. 2004. Dopamine, learning and motivation. *Nat Rev Neurosci* **5:** 483–494. doi:10.1038/nrn1406

Zaghloul KA, Blanco JA, Weidemann CT, McGill K, Jaggi JL, Baltuch GH, Kahana MJ. 2009. Human substantia nigra neurons encode unexpected financial rewards. *Science* **323:** 1496–1499. doi:10.1126/science.1167342

Zou H, Hastie T. 2005. Regularization and variable selection via the elastic net. *J R Stat Soc Series B Stat Methodol* **67:** 301–320. doi:10.1111/j.1467-9868.2005.00503.x