

The effects of motivation on response rate: A hidden semi-Markov model analysis of behavioral dynamics

Eran Eldar^a, Genela Morris^c, Yael Niv^{a,b,*}

^a Princeton Neuroscience Institute, Princeton University, Green Hall, Princeton, NJ 08540, USA

^b Department of Psychology, Princeton University, Green Hall, Princeton, NJ 08540, USA

^c Department of neurobiology, Haifa University, Israel

ARTICLE INFO

Article history:

Received 25 November 2010

Received in revised form 25 June 2011

Accepted 28 June 2011

Keywords:

Hidden semi-Markov model

Sequential data analysis

Motivation

Response rate

ABSTRACT

A central goal of neuroscience is to understand how neural dynamics bring about the dynamics of behavior. However, neural and behavioral measures are noisy, requiring averaging over trials and subjects. Unfortunately, averaging can obscure the very dynamics that we are interested in, masking abrupt changes and artificially creating gradual processes. We develop a hidden semi-Markov model for precisely characterizing dynamic processes and their alteration due to experimental manipulations. This method takes advantage of multiple trials and subjects without compromising the information available in individual events within a trial. We apply our model to studying the effects of motivation on response rates, analyzing data from hungry and sated rats trained to press a lever to obtain food rewards on a free-operant schedule. Our method can accurately account for punctate changes in the rate of responding and for sequential dependencies between responses. It is ideal for inferring the statistics of underlying response rates and the probability of switching from one response rate to another. Using the model, we show that hungry rats have more distinct behavioral states that are characterized by high rates of responding and they spend more time in these high-press-rate states. Moreover, hungry rats spend less time in, and have fewer distinct states that are characterized by a lack of responding (Waiting/Eating states). These results demonstrate the utility of our analysis method, and provide a precise quantification of the effects of motivation on response rates.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Capturing the dynamics of animal behavior can be crucial to understanding the underlying patterns of neural activity and how these relate to overt behavior. In an experimental setting, animals often rapidly switch between different modes of action. However, such dynamics are rarely taken into account in commonly used methods of analysis. Typically, behavior is averaged across a group of subjects, resulting in a smooth curve that may not reflect the time-course of behavior of individual subjects in individual trials. For example, in a learning experiment, Papachristos and Gallistel (2006) showed that although individual mice exhibited abrupt stepwise changes in anticipatory head poking rates, the average group response was smooth, wrongly suggesting that learning proceeded gradually. In another study, the same phenomenon was observed when contrasting individual and group-average pigeon pecking rates (Gallistel et al., 2004). The authors concluded that the learning rate that seemingly characterizes the group-average

curve is meaningless with regard to the actual individual learning process (see also Daw and Courville, 2008). Similarly, averaging the behavior of an individual animal over multiple trials may also hide the true underlying characteristics of that behavior. For instance, if an animal transitions abruptly from a low rate to a high rate of responding within each trial, an average curve will inaccurately depict a smooth increase in response rates, as long as the abrupt change was timed differently on each trial (Matell et al., 2006). Such gradual curves are useful for some purposes (e.g., Church, 1984; Roberts and Boissvert, 1998; Drew et al., 2003; Balci et al., 2010), however, to accurately capture the dynamics of the behavior and neural activity of individual animals, a more precise characterization of single trial behavior is required. Here we describe a method of analysis, based on the statistical framework of hidden Markov models (HMMs; Rabiner, 1989) that can achieve such a precise quantification of non-continuous single trial dynamics, while still taking advantage of multiple instances of behavior to extract meaningful information from noisy data. In this paper we apply this method to analyzing behavioral response rates.

Response rate is one of the most commonly used dependent measures of behavior, utilized to examine the effect of a wide variety of experimental manipulations such as reinforcement schedule

* Corresponding author. Tel.: +1 609 258 1291; fax: +1 609 258 1113.

E-mail address: yael@princeton.edu (Y. Niv).

modifications, drug administration and neural lesions. Molar characteristics of animal response rates as a function of the frequency and magnitude of reinforcement have been described as adhering to the Matching Law (Herrnstein, 1961, 1997), and the effects of reinforcement on the dynamics of choice have been used to explain overall response rates across trials (Sugrue et al., 2004; Corrado et al., 2005). However, a basic understanding of how response rates change within a trial, and how these changes are influenced by experimental manipulations, is still lacking. In the current study, we were interested in precisely characterizing how the motivational state of individual animals affected their response rates. It is well accepted that the motivational state of an animal affects its behavior. For example, a hungry animal would be more likely to choose actions that are rewarded with food, and to perform these actions at a higher rate than a sated one (Batten and Shoemaker, 1961). While previous studies have established that response rate is affected by motivational factors such as palatability of food reward (Carper, 1953; Guttman, 1953; Smith and Kinney, 1956) and time of food or water deprivation (Batten and Shoemaker, 1961; Koch and Daniel, 1945; Skinner, 1950; Smith and Capretta, 1956), we set out to precisely quantify these effects of motivation on response rates of hungry or sated rats lever-pressing for food on a free-operant variable-interval schedule.¹

2. Method

2.1. Subjects

Thirty-nine male Sprague Dawley rats (Harlan Laboratories, Jerusalem, Israel) approximately three months old, weighing 329–473 g (mean 362 g) were housed 3–4 to a cage, in a vivarium maintained on a 12-h light–dark cycle (lights on 15:00–3:00). All behavioral training and testing occurred during the dark portion of the cycle. Animals were allowed one month familiarization with the vivarium before training began. Rats were then divided into two groups. Group HUNGRY was maintained on a 22-h food restriction schedule, with tap water available *ad libitum* in the home cage. Standard lab chow was provided in the home cage for 2 h each day, always after the daily session. Group SATED had continuous access to standard lab chow in their home cage. All animal research was carried out according to the guidelines of the Institutional Animal Care and Use Committee of Tel Aviv University, and efforts were made to minimize the number of animals used and their suffering.

2.2. Apparatus

Behavioral training and testing were conducted in four operant chambers (Campden Instruments, Loughborough, UK) fitted with a recessed food magazine and two retractable levers. Only the left lever was used in this experiment. The right lever remained retracted at all times. Access to the food magazine was through a hinged Perspex panel which activated a micro-switch when opened. All rats in the HUNGRY group and 9 rats in the SATED group received sucrose solution as reward. A peristaltic pump (RS Components, Northants, UK) attached to a silicon tube inaccessible to the rats, delivered approximately 0.25 ml of 20% sucrose solution (hand mixed) into the food magazine, over a period of 1 s. The 10 remaining rats in the SATED group received a single sucrose pellet as reward. The operant chambers were housed in sound attenuating boxes, and ventilating fans were mounted on the side of each box. The chambers could be illuminated by a house-light located at the ceiling. Equipment programming and data recording were

computer controlled by ABET software (Lafayette Instrument Co., Indiana, USA).

2.3. Procedure

2.3.1. Handling

On days 1–3, rats were individually handled for about 2 min daily. For the HUNGRY group, a 22-h food restriction schedule began one day prior to handling, and continued throughout subsequent training and testing.

2.3.2. Magazine training

On days 4–5, rats were trained to consume sucrose solution or pellets from the food magazine in the operant chamber, with the lever retracted. The session began with the onset of the house-light, which remained on for the entire session. Sucrose solution or pellets were delivered into the food magazine on a random time schedule, with a uniformly variable delay of 30–90 s (mean 60 s). On day 4, the magazine flap was taped back so that the magazine was constantly open and training continued until twenty-five rewards were delivered. On day 5, the session ended after twenty outcomes had been collected (as measured by the insertion of the rats' head into the food magazine), or until twenty-five outcomes had been delivered.

2.3.3. Lever-press training

On days 6–14, rats were trained to lever-press in order to obtain sucrose solution or pellets in a free operant variable interval procedure. We chose a variable interval schedule as in this type of schedule the rate of reward is relatively independent of the rate of responding. Moreover, in the intervals between baiting times, responses are entirely without consequence. Thus the schedule exerts minimal constraints on response rate dynamics, providing a suitable vehicle for measuring the effects of motivational state on response rates. Due to technical limitations we could only use a uniform distribution of interval durations, rather than the memory-less exponential distribution ("random-interval" schedule).

The beginning of each session was signaled by the onset of the house-light and the insertion of the left lever. On day 6, every press on the lever delivered an outcome into the food magazine. Individual shaping was used in this session to assist in acquisition of the lever-press response. On days 7–14, the variable interval schedule of reinforcement was introduced, with an outcome delivered only for the first lever-press after the programmed interval had elapsed. The mean interval was 2 s on day 7, 15 s on day 8, and 30 s for the following seven training sessions. Reinforcement baiting was equally likely at any time starting from half the nominal value of the schedule and ending at 150% of the nominal value of the schedule.

All sessions terminated when thirty outcomes had been delivered, except the session on day 6 which ended when thirty outcomes had been collected (as measured by insertion of the rat's head into the food magazine prior to the delivery of the next outcome), or when forty outcomes had been delivered. The lever was then withdrawn, and the house-light was turned off. Day 12 included two training sessions, separated by at least 1 h, such that overall, a total of ten lever-press training sessions were given over nine days. Both lever presses and head insertions into the food magazine were recorded.

Two rats required extra days for magazine training and/or lever press training, and thus received only two sessions of training on the 30 s variable interval schedule, instead of seven. One rat in group HUNGRY refused to consume the sucrose solution from the food magazine and was dropped from the study. This resulted in

¹ In the particular variable-interval schedule used in the experiment, intervals were distributed uniformly within a range of 15–45 s.

final Ns of 19 rats in the HUNGRY group and 19 rats in the SATED group.

2.3.4. Lever-press testing

On days 15–18, rats underwent 5 testing sessions which were identical to the final 7 training sessions. Day 15 included two training sessions, separated by at least 1 h. In all testing sessions the scheduled intervals between rewards were at least 15 s long, thus rats were not forced to choose between consuming food and maximizing the amount of food received.

2.4. Computational modeling of lever-press rates

To characterize the statistics of response rates across multiple trials in a free operant experiment and their malleability due to experimental manipulations such as motivation while still preserving an accurate description of the abruptness of within-trial behavioral changes, we propose to model behavior using a hidden Markov model (HMM; Rabiner, 1989). In an HMM, observed behavior is assumed to be probabilistically generated by a set of hidden states, for instance in our case, each corresponding to a certain response rate. Transitions between these states are abrupt, and occur with some fixed probability that depends on the current state and the new state. The number of distinct states that best describes the empirical observations, the response rates associated with each state and the probability of transitioning from one state to another can all be inferred by fitting the model to behavior using a maximum-likelihood procedure (details below). The result is a description of behavior that effectively pools behavior across trials, sessions and/or subjects to extract its statistical properties, while still allowing for (and correctly modeling) behavioral transitions that do not conform to smooth curves.

2.4.1. An HMM model of response rates

To examine how rat behavior varied overtime, both across and within trials, we developed an HMM in which hidden states reflected underlying response rates, stochastically emitting observations in the form of inter-press-intervals (IPIs; Fig. 1). Each state s was thus associated with a gamma distribution over intervals, defined by two parameters: its mean μ_s and standard deviation σ_s . The gamma distribution is suitable for modeling an initial low-response refractory period followed by a constant probability of response per unit time. In addition, since it can reflect the sum of multiple exponentially distributed random variables, it is capable of accounting for time intervals that include multiple successive actions, as might be the case in our experiment (e.g., a rat may consume food, groom or rest before preparing to press the lever again).²

Each IPI constituted an observed data point presumed to be emitted by the HMM's hidden state. Following each IPI the model could transition to a different state or stay in the same state. State transitions followed the fixed transition probability matrix T , where $T_{ss'}$ is the probability of transitioning to state s' given that the current state is s . In addition, the delivery of reinforcement was input to the model (formally, an input–output HMM; Bengio and Frasconi, 1995), with reward receipt inducing transition probabilities $T_{Rs'}$ to state s' , irrespective of the current state. Technically speaking, the model is a hidden semi-Markov model (HSMM), due to the non-equal state dwell times between transitions.

² An alternative analysis using a closely related interval distribution that approximates a normally distributed lag-time followed by a Poisson process (Haccou and Meelis, 1992; Conover et al., 2001) allowed a smaller number of free parameters but did not fit our data as well as the gamma distribution, and thus is not reported.

2.4.2. Model fitting

To fit the model to the observed behavior, we used an expectation-maximization (EM) algorithm (Dempster et al., 1977; also called the Baum–Welsh algorithm, Baum et al., 1970). This allowed us to find the maximum likelihood setting of the IPI mean and standard deviation for each state (μ_s and σ_s , respectively), the probabilities of occupying each state at the beginning of a testing session (π_s , henceforth ‘initial state probabilities’), the full matrix of transition probabilities between states T , and the vector of transitions predicated on reinforcement T_R . These parameters were fit to the sequence of rewards and IPIs from the five test sessions, separately for each rat.

Model-fitting proceeded as follows. *Initialization*: Initial state probabilities and transition probabilities were initialized to uniform probability. The mean and standard deviation parameters controlling the emission distributions were initialized at random values. Following initialization, an expectation step and a maximization step (described below) were repeated in alternation. *Expectation step*: In this step we inferred the state and transition probabilities given the observed data and the current model parameters. We first recursively computed α_s^i , the joint probability of occupying state s at the i th IPI and of the observed IPIs up to and including that IPI, given the model parameters:

$$\alpha_s^1 = \pi_s P(\text{IPI}^1 | \mu_s, \sigma_s) \quad (1)$$

$$\alpha_{s'}^{i+1} = \sum_s \alpha_s^i T_{ss'} P(\text{IPI}^{i+1} | \mu_{s'}, \sigma_{s'}) \quad (2)$$

where $P(\text{IPI}^i | \mu_s, \sigma_s)$ is the gamma probability density function of IPI^i given mean μ_s and standard deviation σ_s .

Then, we recursively computed β_s^i , the probability of all IPIs observed after IPI^i , given state s at the i th IPI and the model parameters:

$$\beta_s^I = 1 \quad (3)$$

$$\beta_s^i = \sum_{s'} T_{ss'} \beta_{s'}^{i+1} P(\text{IPI}^{i+1} | \mu_{s'}, \sigma_{s'}) \quad (4)$$

In computing β_s^i , the recursion starts with a value of 1 at the last IPI (IPI^I) and progresses backwards (hence the name ‘forward–backward algorithm’).

Finally, we used α_s^i and β_s^{i+1} to compute $\xi_{ss'}^i$, the probability of a transition from state s to state s' between the i th IPI and the following IPI given the observed data and the model parameters, and γ_s^i , the probability of occupying state s at the i th IPI given the observed data and the model parameters:

$$\xi_{ss'}^i = \frac{\alpha_s^i T_{ss'} P(\text{IPI}^{i+1} | \mu_{s'}, \sigma_{s'}) \beta_{s'}^{i+1}}{\sum_s \sum_{s'} \alpha_s^i T_{ss'} P(\text{IPI}^{i+1} | \mu_{s'}, \sigma_{s'}) \beta_{s'}^{i+1}} \quad (5)$$

$$\gamma_s^i = \sum_{s'} \xi_{ss'}^i \quad (6)$$

Note that all equations were used similarly for time intervals that followed reward delivery, except that transition probabilities there were T_{Rs} .

Maximization step: This step involved setting new values for the free parameters of the model so as to maximize the likelihood of the observations given the model, based on the results of the expectation step. Initial state probability of state s was set as the probability of that state at the first IPI as inferred in the expectation step (Eq. (7)). The probability of transition to state s' given state s was set as the sum of the corresponding time-dependent transition probabilities that were inferred in the expectation step, divided by the sum of probabilities of occupying state s :

$$\pi_s = \gamma_s^1 \quad (7)$$

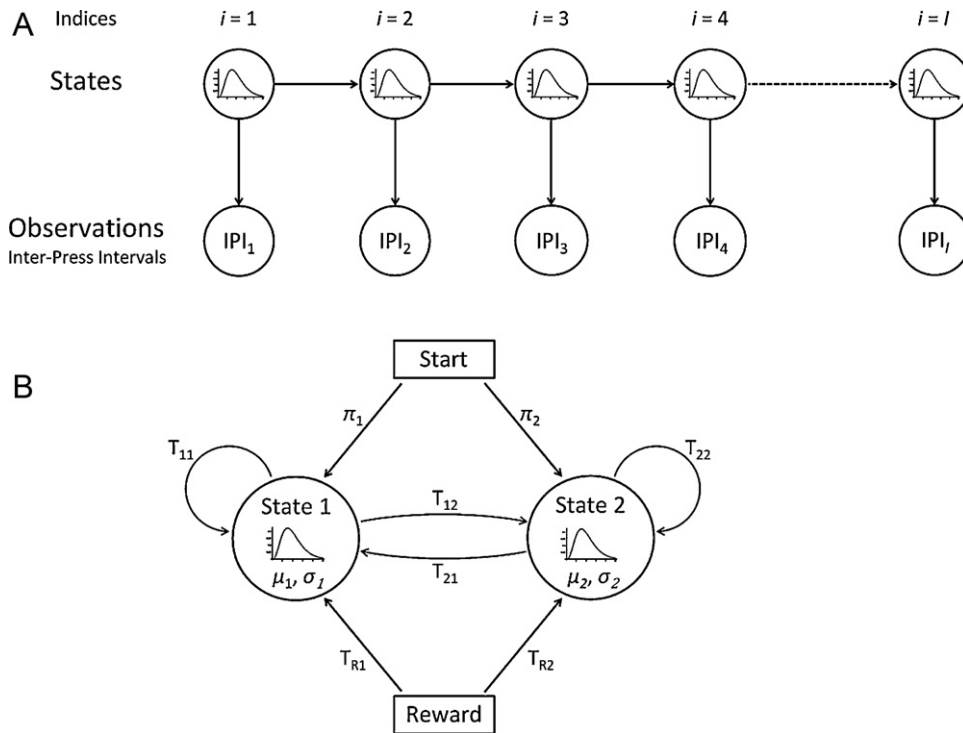


Fig. 1. Hidden semi-Markov model with two states. (A) Graphical model depiction of a sequence of states and observations. (B) State diagram showing all parameters fitted to the sequence of inter-press intervals of each rat. μ_s – mean of gamma emission distribution of state s , σ_s – standard deviation of gamma emission distribution of state s , π_s – probability of starting at state s , $T_{ss'}$ – probability of transitioning to state s' following state s , T_{Rs} – probability of transitioning to state s following reward receipt.

$$T_{ss'} = \frac{\sum_{i=1}^{l-1} \xi_{ss'}^i}{\sum_{i=1}^{l-1} \gamma_s^i} \quad (8)$$

Since no analytically closed expressions exist for maximum likelihood estimators of gamma distribution parameters, we estimated these numerically (Choi and Wette, 1969). For each state, every IPI was weighted by the probability of occupying that state at that time interval (γ_s , inferred in the preceding expectation state). The parameters of the gamma distribution of that state were then fitted to this weighted set of IPIs:

$$\mu_s = \arg \max_{\mu_s} P(IPI^{1:l} | \mu_s, \sigma_s, \gamma_s^{1:l}) \quad (9)$$

$$\sigma_s = \arg \max_{\sigma_s} P(IPI^{1:l} | \mu_s, \sigma_s, \gamma_s^{1:l}) \quad (10)$$

200 iterations of the Expectation and Maximization steps were performed for each model, which was sufficient for the estimates of the model parameters given the data to converge. To avoid local maxima, the fitting process was repeated 15 times for each model and each rat, with different randomly initialized parameters. Finally, we produced the most likely sequence of states for describing the rat's behavior using the Viterbi algorithm (Viterbi, 1967) given the model and its maximum likelihood parameters.

2.4.3. Model comparison

A major question of interest was the number of distinct states (response rates) that characterize the rats' behavior, and how this is affected by motivational state. However, when fitting a hidden (semi-) Markov model, the number of states must be set in advance. We thus fitted separate models with 1, 2, 3 and 4 states to the observations obtained from each rat, and selected the model that best accounted for the data according to the Bayesian Information Criterion (BIC; Schwarz, 1978). Adding states to the model means adding free parameters: 8 free parameters are required in a model with two states (Fig. 2), and that number grows exponen-

tially as a function of the number of states. In particular, a model with n states involves $2n$ gamma distribution parameters, $(n-1)$ initial state occupancy probabilities, $n(n-1)$ transition probabilities and $(n-1)$ reinforcement-induced transition probabilities. The BIC criterion adjusts for the inherently better explanatory power of models with more parameters, penalizing for the number of free parameters such that simpler models are preferred and overfitting of the data is minimized. In effect, the BIC score can be used to determine whether the increase in likelihood of the data due to an additional parameter is sufficiently large to justify the additional parameter.

We tested the differences in the numbers of distinct states between the SATED and HUNGRY groups for statistical significance using a chi-squared randomization test (McDonald, 2009). Results of both study groups were pooled together and 100,000 alternative result sets were randomly drawn from the pool with replacement. The chi-squared statistic was computed for each result set, with the pooled results serving as the expected distribution. Statistical significance level was then computed by comparing the chi-squared statistic of the actual result to this distribution.

3. Results

We trained hungry and sated rats to press a lever for food reward on a variable-interval schedule, in order to examine how motivation affects their patterns of responding. To avoid averaging over data between and within subjects and potentially obscuring the very effects we would like to quantify, we used an HSMM model of lever-pressing. HSMMs are particularly suitable for analyzing sequential stochastic behavioral responses. Furthermore, using input variables, it is possible to incorporate into the models the effect of events such as reward receipt, thereby taking into account the structure of the experiment without having to average over multiple trials. In our model, the rat presses a lever according to several discrete response rates, with each hidden state corre-

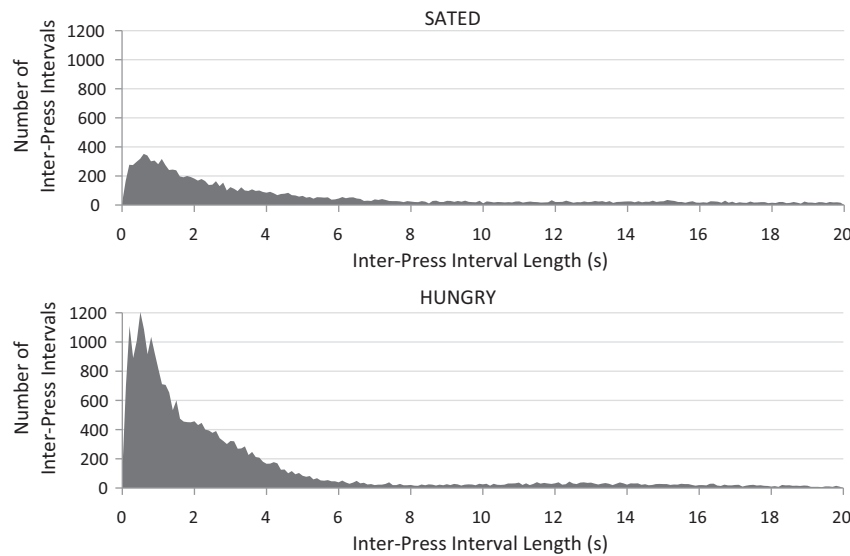


Fig. 2. Inter-press interval histograms of the SATED and HUNGRY groups. IPIs were binned into 100 ms bins.

sponding to a distribution over inter-press-intervals (IPIs). At each step of the model, the current state emits one IPI and then transitions to the next state (including, possibly, to itself) according to fixed transition probabilities. The occurrence of reinforcement perturbs behavior by effecting a transition to any of the states according to different, reinforcement-dependent transition probabilities. The model was fit separately to the data from each rat, allowing us to infer from each rat's behavior the underlying hidden states (i.e., the response rates characterizing behavior), the number of states sufficient to accurately describe the behavior, and how often each state was occupied during the behavioral sessions.

How can motivation affect responding in our model? One option is that hungry (motivated) rats press the lever at a higher rate such that each of (or one of) the hidden states corresponding to their behavior is associated with significantly shorter IPIs. A second option is that motivation does not change the response rates associated with each state, but rather, it affects the probability of occupying each of the states such that hungry rats spend more time in short IPI (fast lever-press) states. Finally, a third possibility is that motivated rats show more distinct rates of responding, that is, that higher motivation adds new states with even shorter IPIs to the rat's behavioral repertoire in the experiment.

3.1. Overall rate of responding

During testing, rats in the HUNGRY group showed a significantly higher ($p < 10^{-8}$, one-way ANOVA) average rate of lever pressing (mean 17.3 presses/min, SD 4.5) compared to the SATED group (mean 7.6 presses/min, SD 2.8). Accordingly, the HUNGRY group rats generally obtained rewards more quickly than the SATED group did and since each testing session ended when the thirtieth reward was earned, HUNGRY group testing sessions were significantly shorter ($p < 10^{-4}$, one-way ANOVA) than those of the SATED group. The mean HUNGRY group session length was 976 s (SD 30) while the SATED group sessions were on average 1164 s long (SD 177).

Fig. 2 shows IPI histograms of each of the study groups. The histograms show a skewed distribution in which very short and very long IPIs are rare. Thus the histograms suggest a gamma distribution (or a composition of several such distributions) over IPIs, prompting our choice to characterize response rates in the HSMM model using a gamma distribution.

3.2. Model-based analysis of response rates

Fig. 3 illustrates the results of the model-fitting process for one rat from the SATED group. The analysis indicated 2 distinct behavioral states: one state with a mean of 4.15 presses/min and another state with a mean of 24.28 presses/min (Fig. 3A). As shown in the state diagram in Fig. 3B, reward receipt caused the model to transition to the low press rate state ($P_{R1} = 1$). Then, following a single lever press the model was highly likely to transition to the high press rate state ($P_{12} = 0.89$), and this rate was typically maintained for multiple lever presses ($P_{22} = 0.95$) until the next reward. Plotting the behavior of the rat as a function of the time since the last reward (Fig. 3C) shows a correspondence between the rat's average press rate and the probability that the model is in the high press rate state.

Fig. 4 illustrates the model-fitting results for a different rat, from the HUNGRY group. Here, the analysis detected three distinct behavioral states with mean lever press rates of 4.73, 17.63 and 118.04 presses/min. Similar to the results for the previous rat, this rat's model also transitioned to the low press rate state following every reward, and was unlikely to stay in this state for more than a single inter-press interval. However, following the low press rate state the model was equally likely to transition to either the intermediate or the high press rate state, with frequent transitions between these two states until the next reward was obtained. Average lever press rate and state probabilities followed the same pattern as a function of the time that elapsed since the last reward (Fig. 4D). To further validate the inferred HSMMs, we used each model to generate a sequence of lever presses in response to the same reward schedule that was used in the experiment. The simulated data shown in Fig. 4C was generated using the model in Fig. 4B. As expected, the results of the simulation are qualitatively similar to the observed data that the model was inferred from (Fig. 4A).

Overall, the analysis indicated the presence of two distinct behavioral states in 9 out of 19 rats in the SATED group and in 5 out of 19 rats in the HUNGRY group. In contrast, three distinct behavioral states were indicated in more rats in the HUNGRY group (13 rats) than rats in the SATED group (8 rats). Four distinct states were indicated in only two rats in the SATED group and one rat in the HUNGRY group. These differences between the study groups were statistically significant ($p < 0.05$, chi-squared randomization test).

We next examined the mean lever-press rates characterizing each of the states indicated for each rat (Fig. 5). In general, mean

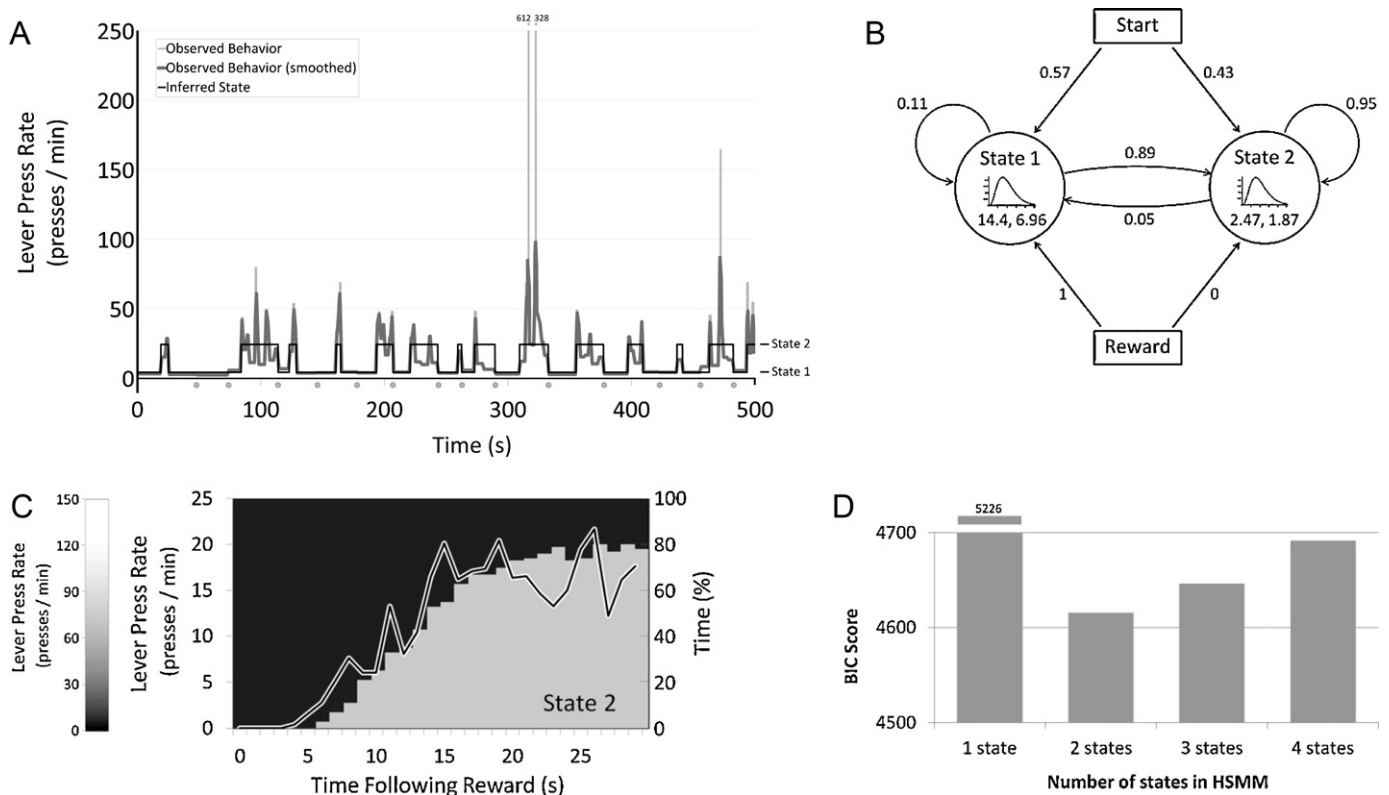


Fig. 3. Model-fitting results for one rat from the SATD group for which 2 distinct states were detected. (A) Observed press-rates and inferred states during the first half of one testing session. Circles denote reward receipt. The smoothed observed behavior curve, produced using a 15-s moving average, illustrates the correspondence between the data and the inferred behavioral states. (B) State diagram of the best fitting model (see Fig. 1B for corresponding parameters). (C) Average lever-press rate (line) and proportion of time attributed to each of the two states (background shading) as a function of time following reward receipt. Shades correspond to the average pressing rate in each state, according to scale shown on the left. (D) BIC scores for HSMMs with 1, 2, 3 and 4 states. Lower scores indicate a better fit of the model.

lever-press rates were higher for the HUNGRY group as compared to the SATD group. Interestingly, as no states were characterized by mean lever press rates in the range of 5–10 presses/min, a dichotomy emerged between very low and medium to high lever-press-rate states. Low press-rate states were characterized by mean lever-press rates lower than 5 presses/min. The probability of transitioning to one of these states following reward receipt was above 0.99 for all rats, while the probability of maintaining one of these states following a single lever press was on average 0.08 (SD 0.14). We thus interpret these states as ‘Eating/Waiting’ states in which rats consumed the available reward, and possibly engaged in other activities such as grooming, while waiting prior to resuming lever-pressing for the next reward. In contrast, we categorized ‘High-Rate Pressing’ states as those states that were characterized by mean lever press rates higher than 30 presses/min.³ The probability of transitioning to one of these states following reward receipt was negligible in all cases (<0.01), while the probability of maintaining one of these states following a single lever press was, on average, 0.56 (SD 0.32), which is significantly higher than the corresponding probability computed for the ‘Eating/Waiting’ states ($p < 10^{-10}$, one-way ANOVA).

Using this dichotomy, we classified the behavioral states indicated by the HSMM analysis as either ‘Eating/Waiting’ or ‘High-Rate Pressing’ states, and counted these separately for each of the study groups. Notably, fewer ‘Eating/Waiting’ states and more ‘High-Rate Pressing’ states were indicated for the HUNGRY group rats compared to the SATD group ($p < 0.05$, chi-squared randomization

test). Only a single ‘Eating/Waiting’ state was indicated for each rat of the HUNGRY group, compared to an average of 1.32 ‘Eating/Waiting’ states for rats of the SATD group. In contrast, an average of 0.89 ‘High-Rate Pressing’ states per rat were indicated for the HUNGRY group rats, compared to 0.32 ‘High-Rate Pressing’ states for rats of the SATD group.

In addition to exhibiting a higher number of distinct high press rates, the models indicated that rats from the HUNGRY group also spent more time pressing the lever at these high rates (Fig. 6). On average, the HUNGRY group rats spent 52% (SD 12) of the testing sessions in states with a mean press rate higher than 10 presses/min. This was significantly higher than the 32% (SD 17) the SATD group rats spent in such states ($p < 0.001$, one-way ANOVA). The difference in time spent at states with press rate higher than 30 presses/min is even more pronounced (19% SD 20 vs. 1% SD 3; $p < 0.001$, one-way ANOVA).

3.3. Model validation

Analysis using HSMMs is particularly useful for detecting abrupt changes between different behavioral states. However, one limitation of this method of analysis is its underlying assumption that subjects’ behavior does actually change abruptly. It is therefore important to examine the results that this method would produce if given a set of observations that are generated from a gradually changing process. The goal of this analysis was to show that if the true generative process underlying the rats’ lever-pressing behavior was a gradually changing one, our analysis method would have given different results.

We used the empirical mean lever-press-rate curves following reward receipt to stochastically generate sequences of inter-

³ This threshold was chosen to avoid the mass of medium press-rate states that characterized both study groups similarly.

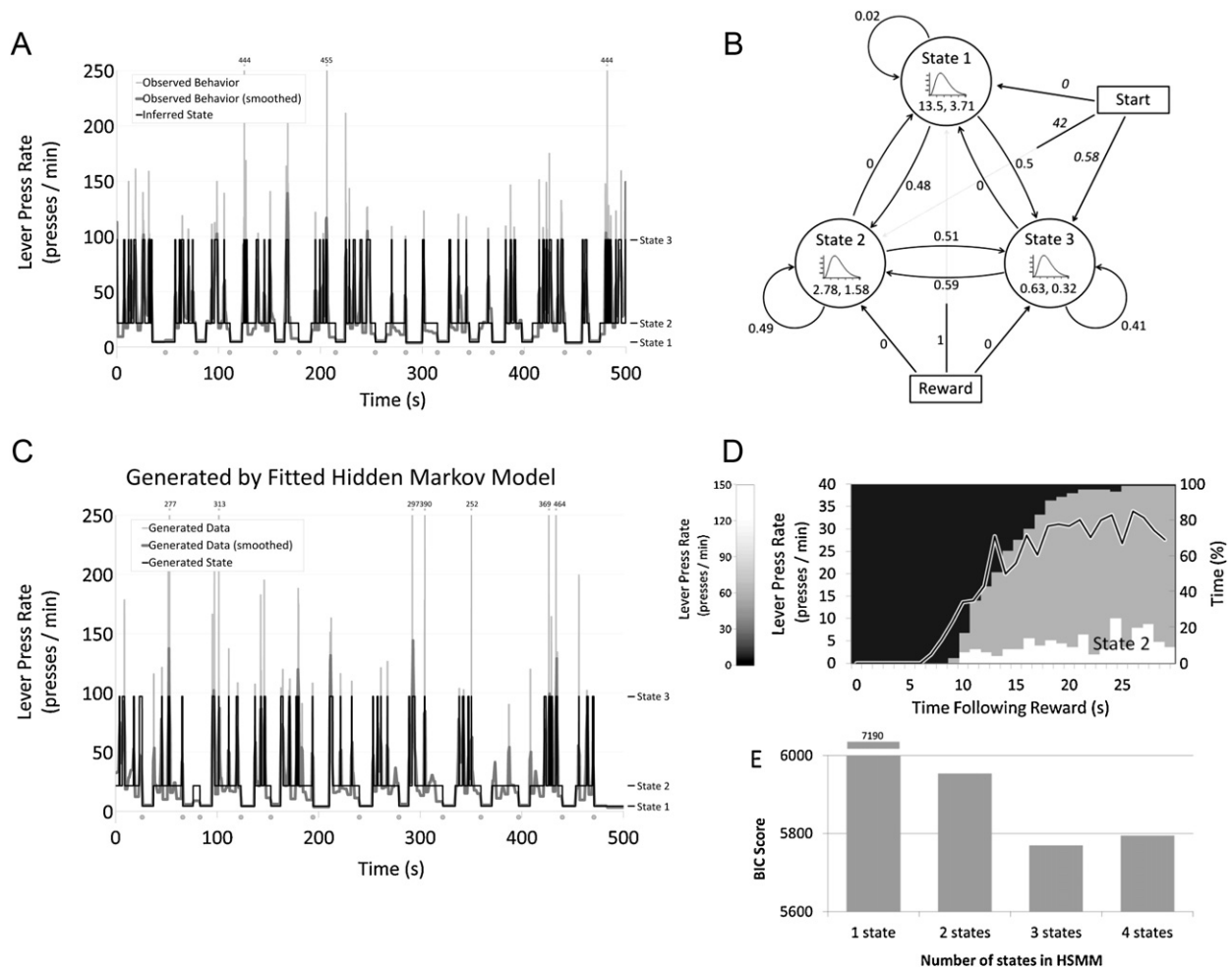


Fig. 4. Results for one rat from the HUNGRY group for which 3 distinct states were indicated. (A) Observed lever-press rates and inferred states during first half of one testing session. Circles mark rewards. The smoothed observed behavior curve, produced using a 15-s moving average, illustrates the correspondence between the data and the inferred behavioral states. (B) State diagram of the best fitting model. (C) Data generated using the fitted HSMM in a generative mode. (D) Average lever-press rate (line) and proportion of time in each inferred state (background shading) as a function of time following reward receipt. Shades correspond to average pressing rate in each state according to scale shown on the left. (E) BIC scores for HSMMs with 1, 2, 3 and 4 states. Lower scores indicate a better fit of the model.

press intervals that reflect gradually changing behavior. Empirical response-rate curves were separately computed for rats that showed two and three distinct behavioral states (Fig. 7), to examine the possibility that a difference in the shape or scale of these curves was the actual reason for the different number of states detected by our HSMM analysis. Ten data sets were generated from each of these curves by simulating lever-presses as generated by a Poisson process whose rate changes as a function of time following reward according to the empirical average curve. Similarly to the actual experiment, rewards were made available following a variable interval uniformly distributed in the range of 15–45 s, each session ended when the thirtieth reward was administered, and each data set consisted of five sessions.

We then applied the same HSMM analysis to the simulated data sets. All ten data sets based on the 2-state rats curve and nine out of the ten data sets based on the 3-state rats curve were found to have only two distinct behavioral states. Three behavioral states were detected in the single remaining data set. Thus, although the proposed HSMM analysis method does not reliably differentiate between rats changing their behavior gradually and rats alternating between two distinct behavioral states (for instance, by suggesting a large number of intermediate states for rats that change their behavior gradually), the results suggest that the finding of three or more distinct behavioral states cannot be simply attributed to

the gradually changing behavior described by either one of the averaged curves. Rather, the fact that our analysis detected more than two states suggests that multiple distinct clusters of IPIs were present in the data, at least to some extent.

We further tested our method of analysis with data generated explicitly from two distinct behavioral states. The probability of occupying each state, as a function of time after reward, was determined according to the empirical mean lever-press-rate curve. However, instead of using the two-state curve from Fig. 7 to determine the lever-press rates for the two states, the three-state curve was used. Nevertheless, the analysis still correctly detected only two distinct states in all of the ten data sets that were generated. This shows that the higher press rates that characterized the three-state rats were not sufficient to bias the analysis in favor of detecting more than two states.

Another potential confound for our analysis is the type of food reward that rats received. Some of the rats in the SATD group received sucrose solution as reward while others received sucrose pellets. In contrast, all rats in the HUNGRY group received sucrose solution. To test this, we divided the SATD group into two sub-groups on the basis of the type of reward they received. Fig. 8 shows that both rat performance, as measured by mean session length, and the number of distinct states detected, were not significantly affected by type of reward. In fact, if anything,

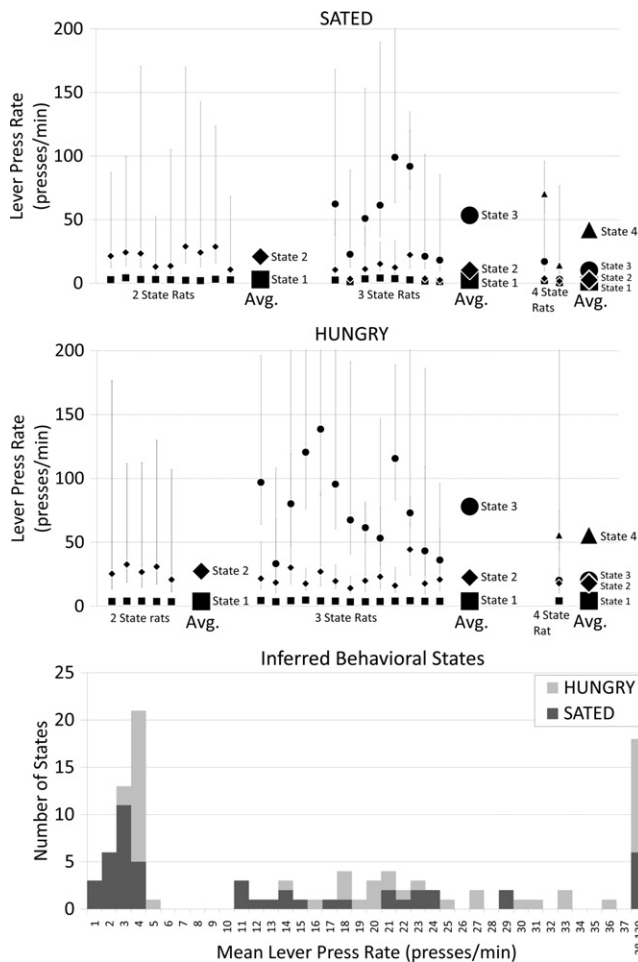


Fig. 5. Top: Mean lever press rates characterizing inferred behavioral states. Results are presented for individual rats arranged by number of states and study group. Error bars represent the standard deviation of the inferred gamma distribution. Error bars above 200 were truncated. Bottom: Histogram of inferred behavioral states as a function of mean lever press rate.

results for the HUNGRY group who received sucrose solution were nominally more similar to those of the pellets sub-group rather than the sucrose solution sub-group of the SATED rats.

In addition, the total number of lever presses and the mean testing session time differed between the SATED and the HUNGRY group. The number of lever presses is particularly important because it determined the number of IPIs, and consequently, the number of data points the HSMs were fitted to. To examine whether either of these factors had a significant effect on the number of states detected by the analysis, number of lever presses and mean session length were compared within each of the study groups between rats that showed two distinct behavioral states and rats that showed three distinct states. Fig. 9 shows that no significant differences were found in either variable, suggesting that neither of these variables substantially affected the results. To further ensure that our main findings were not a result of differences in session length and that they do not reflect differences that are specific to the later part of each session, we fitted HSMs to the behavioral data obtained from the first 10 min of each session. The main findings of this analysis were similar to those of the full data set analysis: HUNGRY group rats showed more distinct behavioral states than the SATED group rats ($p < 0.05$, chi-squared randomization test). In addition, we again found more 'High-Rate Pressing' states and fewer 'Waiting/Eating' states for the HUNGRY group

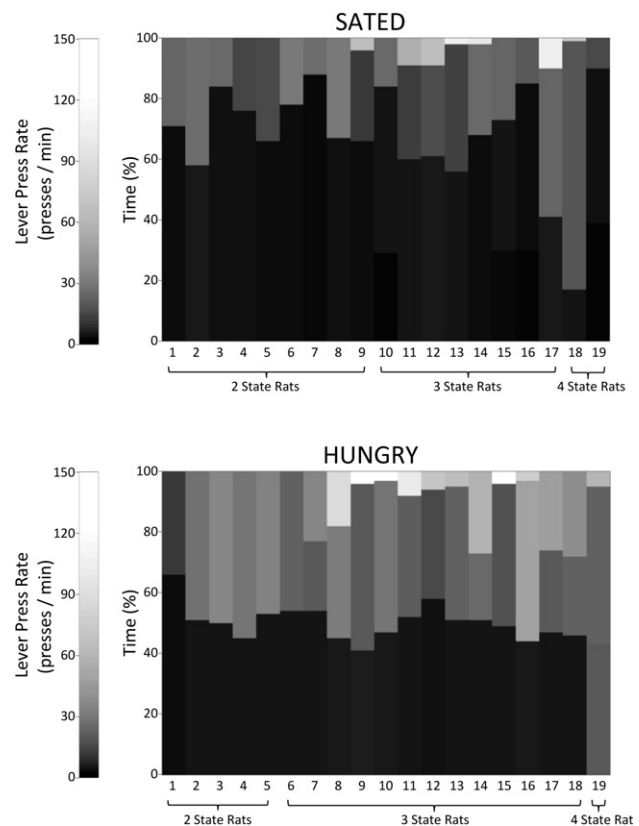


Fig. 6. Mean lever press rates (shade) and proportion of time spent (extent) in each behavioral state for each individual rat. Rats in the SATED group spend more time in 'Eating/Waiting' states and less time in 'High-Rate Pressing' states than rats in the HUNGRY group.

compared to the SATED group ($p < 0.05$, chi-squared randomization test).

Finally, to ensure that additional behavioral states were not detected simply due to systematic fluctuations in behavior across or within sessions, such as may be caused by less stable behavior in the first test sessions or by satiation developing over the course of a session, we compared the time spent by rats in the different states between the first and second halves of each session and between sessions (two-way ANOVA). No significant differences were found between halves or sessions, regardless of the number of states detected ($p > 0.05$). Thus, we conclude that the main results of our analysis do not reflect such systematic fluctuations.

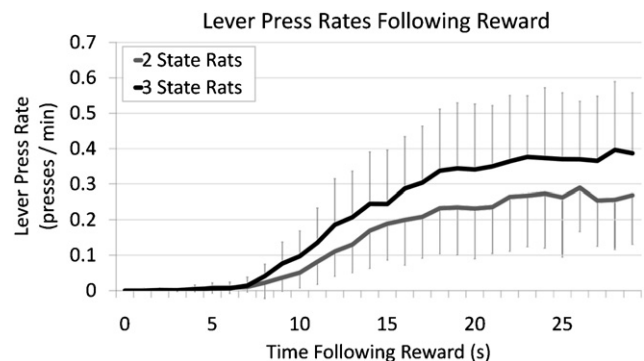


Fig. 7. Mean lever press rate following reward receipt, computed separately for rats with 2 and 3 inferred states. Separate analyses of response rates between the 30 and 35, the 35 and 40 and 40, and the 45 s marks (not shown) indicated that press-rates did not significantly change after the 30 s mark (one-way ANOVA, $p > 0.05$). Error bars represent standard deviation.

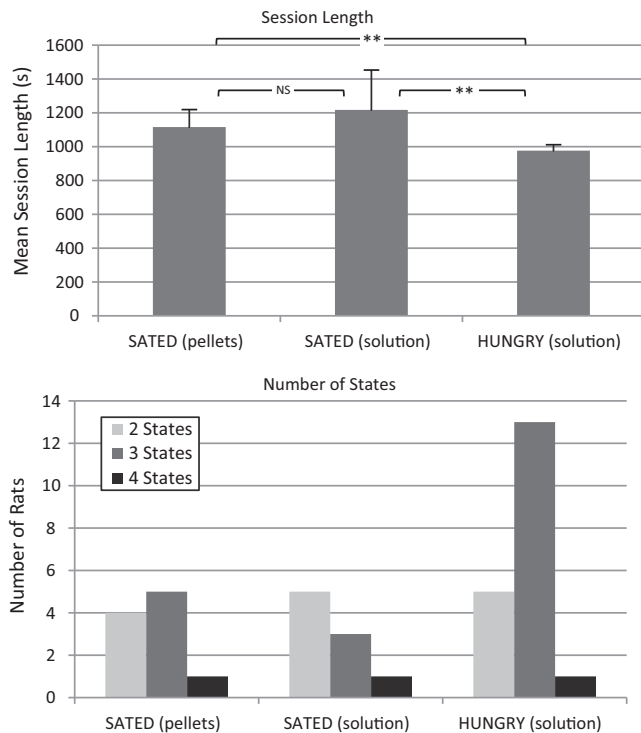


Fig. 8. Mean testing session length and number of detected behavioral states as a function of the type of food reward, shown separately for the SATED and HUNGRY groups. ** denotes significance at the level of $p < 0.001$, one-way ANOVA. No significant difference was found in the distribution of number of states between the SATED (pellets) and the SATED (solution) groups (chi-squared randomization test, $p > 0.05$).

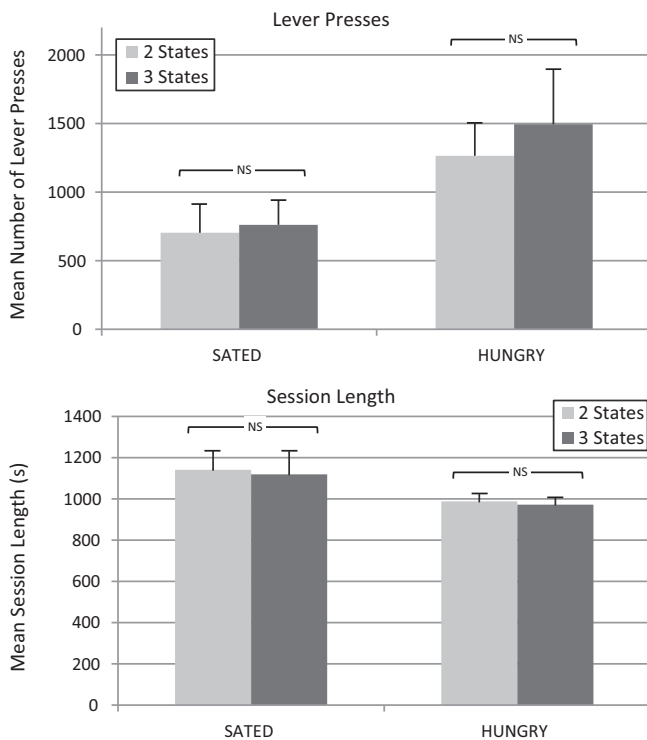


Fig. 9. Mean number of lever presses and mean testing session length as a function of the number of inferred behavioral states, shown separately for the SATED and HUNGRY groups.

4. Discussion

We introduced a new method for analyzing response rate dynamics, which is especially suited to characterizing processes that have sequential structure and involve abrupt changes. To illustrate the strengths of this approach, we examined the effect of food deprivation on rats' rate of lever-pressing for food reward in a uniformly-distributed variable-interval schedule. Generally, food-restricted rats pressed the lever more often than rats that had continuous access to food. Analysis using HSMs revealed that behavior was largely comprised of two or three distinct behavioral states, associated with different rates of lever-pressing. Food-restricted rats showed more distinct states with high lever press rates and fewer distinct states with low lever press rates. Additionally, food-restricted rats also tended to maintain high press-rate states for longer periods of time as compared to rats that had access to *ad libitum* food chow in their home cage.

Based on these results, we can conclude that hungry rats had more distinct modes of high-rate lever-pressing and fewer distinct modes of waiting than sated rats. Notably, rats in the food-restricted group had only one type of 'Waiting' state. This state was characterized by a lever press rate of around 4 presses/min, corresponding to a 15 s IPI which, perhaps coincidentally, is exactly the period of time that rats had to wait for the next reward to be possibly available. In contrast, some rats that had continuous access to food showed more than a single distinct 'Waiting' state, and some of these states were characterized by mean IPIs that were significantly longer than 15 s, indicating that despite the consequence of a lower reward rate, these rats opted for longer periods of waiting and shorter periods of lever-pressing compared to food-restricted rats.

As demonstrated in the present study, the main advantage of using HMMs and HSMs in data analysis lies in their ability to capture abrupt switches between distinct behavioral states. In addition, behavioral patterns that might be lost due to averaging can be revealed using our proposed method. Avoiding averaging over multiple trials does not mean that the structure of the experiment must be ignored: events such as trial start or reward receipt can easily be incorporated as input variables in the causal structure of the model. In addition, the sequential structure of HMMs allows them to detect distinct clusters of observations not only according to the value of the observations but also according to the timing of each observation in relation to other observations. This property allows a more sensitive analysis of the data that takes into account sequential aspects of the observed behavior. For instance, one long IPI per inter-reward-interval may not carry much weight in a traditional analysis, but if this IPI is consistently observed immediately following reward, it is sufficient to justify inference of a distinct behavioral state in our analysis. Similarly, a few short IPIs will have much more effect on the HSM analysis if they follow each other in succession as compared to randomly interspersed within or among trials, a structural difference that would be lost in averaging.

Several approaches have previously been proposed for avoiding averaging over multiple trials and detecting abrupt changes in behavior (Church et al., 1994; Cheng and Westwood, 1993; Schneider, 1969; Gallistel et al., 2001, 2004). The most powerful and generally applicable of these approaches is Gallistel et al.'s change point-algorithm, which identifies behavioral change-points by analyzing deviations from linearity in cumulative behavior plots and weighing the probabilistic evidence in favor of each potential change point. Our approach improves on the change-point algorithm in several respects. First, the change-point algorithm detects changes in behavior rather than behavioral states. Thus, in contrast to our approach it is agnostic as to whether two relatively similar behavioral states that are detected at different parts of the data set should be regarded as different states or as two instances of

the same state. Second, the change-point algorithm truncates the data at each behavioral change point before determining the location of the next one, so that data that comes before the previous change point is ignored when the next one is inferred. In contrast, in our approach the whole data set is used to infer the existence and timing of each behavioral state, as well as the change-points between them. Therefore, patterns of responses that repeat themselves at different parts of the data set but are not distinct enough for a local algorithm to detect, may nevertheless be detected by our algorithm, thereby revealing changes in behavior that would otherwise go unnoticed. Finally, rather than using a user-specified decision criterion for the detection of behavioral changes that is varied so as to produce the desired number of change-points, here we use the Bayesian Information Criterion which provides us with a principled way of determining the number of behavioral states, and enables us to infer a mathematical model of the process that generated the data.

More algorithmically similar to our approach is Smith et al.'s (2004) use of a state-space algorithm to infer the trial in which learning was achieved. There, responses are modeled as observations that are stochastically emitted by a hidden state that depends on a continuous learning-state parameter. Similarly to our study, Smith et al. demonstrate how a state-space algorithm is useful for analyzing the behavior of individual animals. However, a significant limitation of their method is that the rate of change of the learning-state parameter at any single time-point is probabilistically constrained by the average rate of change over the whole time course. For that reason, the algorithm is biased towards detection of gradual learning curves and is limited in its ability to detect abrupt changes that occur on single trials. Modeling behavior using an HMM with discrete hidden states, as we have done here, is more sensitive to such abrupt changes and thus suitable for dynamic processes that are not well-characterized by gradual dynamics.

The advantages that HMM-based analysis carries for behavioral data are also applicable to the analysis of neural data. For example, when analyzing spike trains, individual spikes or inter-spike-intervals can be modeled as observations stochastically emitted by a hidden state. This allows direct inference of the properties of different states of neural activity and characterization of abrupt transitions between them. Indeed, similar state-space methods were successfully used in the past for analyzing neural data (Gat et al., 1997; Brown and Barbieri, 2006; Chen et al., 2009). Augmenting these methods with a model selection procedure as we have done here can be useful for inferring the appropriate number of discrete states of neural activity.

On the other hand, a significant limitation of analysis using HMMs is that the reliability of detection of a hidden state depends on the number of observations that correspond to that state. For this reason, if behavior changes gradually (as in our simulated dataset) it is difficult to predict how many states will eventually be detected, as more data will allow more intermediate-rate behavioral states to surface. It is therefore important to establish the sensitivity of the analysis using artificially generated data sets that simulate gradually changing behavior that corresponds in scale and length to the actual data observed in the experiment.

Indeed, tests on simulated data sets showed that given the size of our data set, our method of analysis is not sensitive enough to distinguish gradually changing behavior from abrupt switching between two distinct behavioral states. However, our analyses detected at least *three* distinct behavioral states for most rats. In addition, peak lever-press rates as well as number of lever presses and experiment length could not, on their own, account for the differences in the number of states that our analysis detected for different rats. Tests on simulated data showed that three states were never detected for data generated from gradually increasing press rates, or from two distinct states, but rather only if

the data were indeed generated from three (or more) distinct states.

Considering these results, our findings indicate that the observed IPLs of at least some of the rats must have formed distinct clusters, which could reasonably be interpreted as representing distinct behavioral states with abrupt changes between them. The hidden behavioral states that we detect, and the punctate changes between them, correspond nicely to the predictions of a recently proposed reinforcement learning model of free-operant behavior in which a distinct response rate policy is learned for each state of a Markov decision process (Niv et al., 2007). In that model, however, motivation is expected to affect more strongly the rate of pressing in discrete states of the environment, rather than the number of discrete rates of responding (Niv et al., 2006). It should be noted though that the HMMs that were revealed by our method of analysis characterize rat behavior only in the specific task environment that our experiment involved. For example, state parameters and state transition probabilities would likely be different had the reward schedule been different.

Our study holds some resemblance to studies of timing behavior. Timing studies typically use fixed-interval reinforcement schedules so that the timing of reward is fully predictable and subjects may adjust their behavior accordingly. Since we were not interested in timing, we used a variable-interval reinforcement schedule. Nevertheless, since intervals between rewards were uniformly and not exponentially distributed, rats could have deduced that a reward was more likely when more time had passed since the last reward. This would predict response rates that increase with time, until reward receipt. We saw no evidence for such a response pattern. Instead, after an initial increase, response rates tended to plateau, suggesting that rats did not time their behavior optimally. That said, some timing behavior might be evidenced by the fact that response rates were lower in the first 15 s following reward, when the next reward was not yet available.

In sum, we have characterized the effects of a rat's motivational state on its response rate: hungrier rats demonstrated more distinct high-press-rate behavioral states, and spent more time in these states compared to sated rats. We have also demonstrated the advantages of using an HMM to analyze behavioral data that has sequential structure. Future directions for research include quantifying the effect of motivation on response rate in a generalizable way, as was previously done for the effects of reinforcement quantity and frequency. The main challenge here is to recognize and quantify the multiplicity of factors that may affect the motivational state of an animal, such as age, type of reward and treatment schedule. A different line of research involves pinpointing the conditions under which the effects of motivation on response rate are outcome-specific *versus* outcome-independent (Niv et al., 2006). Finally, investigating the patterns of neural activity that correspond to the inferred behavioral states may help clarify the neuronal mechanisms that link motivation and response rates. In any case, hidden Markov models provide a useful tool for understanding animal behavior, and can be used to augment current methods of neural data analysis.

Acknowledgements

This work was funded by a start-up grant from the United States-Israel Binational Science Foundation. We thank Peter Dayan for helpful discussions of this work and Daphna Joel for assistance with execution of the rat experiment.

References

- Balci F, Ludvig EA, Brunner D. Within-session modulation of timed anticipatory responding: when to start responding. *Behavioural Processes* 2010;85:204–6.

- Batten DE, Shoemaker HA. Effects of deprivation and incentive palatability on a conditioned operant response. *Journal of Comparative and Physiological Psychology* 1961;54:577–9.
- Baum LE, Petrie T, Soules G, Weiss N. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Annals of Mathematical Statistics* 1970;41:164–71.
- Bengio Y, Frasconi P. An input/output HMM architecture. In: Tesauro G, Touretzky D, Leen T, editors. *Advances in neural information processing systems*. Cambridge, MA: MIT Press; 1995. p. 427–34.
- Brown EN, Barbieri R. Dynamic analyses of neural representations using the state-space modeling paradigm. Cold Spring Harbor Laboratory Press; 2006.
- Carper JWA. comparison of the reinforcing value of a nutritive and a non-nutritive substance under conditions of specific and general hunger. *American Journal of Psychology* 1953;66:270–7.
- Chen Z, Vijayan S, Barbieri R, Wilson MA, Brown EN. Discrete- and continuous-time probabilistic models and algorithms for inferring neuronal UP and DOWN states. *Neural Computation* 2009;21:1797–862.
- Cheng K, Westwood R. Analysis of single trials in pigeons timing performance. *Journal of Experimental Psychology: Animal Behavior Processes* 1993;19:56–67.
- Choi SC, Wette R. Maximum likelihood estimation of parameters of gamma distribution and their bias. *Technometrics* 1969;11:683–90.
- Church RM. Properties of the internal clock. *Annals of the New York Academy of Sciences* 1984;423:566–82.
- Church RM, Meck WH, Gibbon J. Application of scalar timing theory to individual trials. *Journal of Experimental Psychology: Animal Behavior Processes* 1994;20:135–55.
- Conover KL, Fulton S, Shizgal P. Operant tempo varies with reinforcement rate: implications for measurement of reward efficacy. *Behavioural Processes* 2001;56:85–101.
- Corrado GS, Sugrue LP, Seung HS, Newsome WT. Linear-nonlinear-Poisson models of primate choice dynamics. *Journal of the Experimental Analysis of Behavior* 2005;84:581–617.
- Daw N, Courville A. The pigeon as particle filter. In: Platt J, Koller D, Singer Y, Roweis S, editors. *Advances in neural information processing systems*, 20. Cambridge, MA: MIT Press; 2008.
- Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via EM algorithm. *Journal of the Royal Statistical Society Series B: Methodological* 1977;39:1–38.
- Drew MR, Fairhurst S, Malapani C, Horvitz JC, Balsam PD. Effects of dopamine antagonists on the timing of two intervals. *Pharmacology Biochemistry and Behavior* 2003;75:9–15.
- Gallistel CR, Fairhurst S, Balsam P. The learning curve: implications of a quantitative analysis. *Proceedings of the National Academy of Sciences of the United States of America* 2004;101:13124–31.
- Gallistel CR, Mark TA, King AP, Latham PE. 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* 2001;27:354–72.
- Gat I, Tishby N, Abeles M. Hidden Markov modeling of simultaneously recorded cells in the associative cortex of behaving monkeys. *Network: Computation in Neural Systems* 1997;8:297–322.
- Guttman N. Operant conditioning extinction, and periodic reinforcement in relation to concentration of sucrose used as reinforcing agent. *Journal of Experimental Psychology* 1953;46:213–24.
- Haccou P, Meelis E. Statistical analysis of behavioural data: an approach based on time-structured models. Oxford, UK: Oxford University Press; 1992.
- Herrnstein RJ. Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior* 1961;4:267–72.
- Herrnstein RJ. The matching law: papers in psychology and economics. Harvard University Press; 1997.
- Koch S, Daniel WJ. The effect of satiation on the behavior mediated by a habit of maximum strength. *Journal of Experimental Psychology* 1945;35:167–87.
- Matell MS, Bateson M, Meck WH. Single-trials analyses demonstrate that increases in clock speed contribute to the methamphetamine-induced horizontal shifts in peak-interval timing functions. *Psychopharmacology* 2006;188:201–12.
- McDonald J. Handbook of biological statistics. 2nd ed. Baltimore, Maryland: Sparky House Publishing; 2009.
- Niv Y, Daw ND, Joel D, Dayan P. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology* 2007;191:507–20.
- Niv Y, Joel D, Dayan P. A normative perspective on motivation. *Trends in Cognitive Sciences* 2006;10:375–81.
- Papachristos EB, Gallistel CR. Autosshaped head poking in the mouse: A quantitative analysis of the learning curve. *Journal of the Experimental Analysis of Behavior* 2006;85:293–308.
- Rabiner LR. A tutorial on hidden Markov-models selected applications in speech recognition. *Proceedings of the IEEE* 1989;77:257–86.
- Roberts WA, Boissvert MJ. Using the peak procedure to measure timing and counting processes in pigeons. *Journal of Experimental Psychology-Animal Behavior Processes* 1998;24:416–30.
- Schneider BA. A 2-state analysis of fixed-interval responding in pigeon. *Journal of the Experimental Analysis of Behavior* 1969;12:677–87.
- Schwarz G. Estimating the dimension of a model. *Annals of Statistics* 1978;6:461–4.
- Skinner BF. Are theories of learning necessary? *Psychological Review* 1950;57:193–216.
- Smith AC, Frank LM, Wirth S, Yanike M, Hu D, Kubota Y, et al. Dynamic analysis of learning in behavioral experiments. *Journal of Neuroscience* 2004;24:447–61.
- Smith M, Kinney GC. Sugar as a reward for hungry and nonhungry rats. *Journal of Experimental Psychology* 1956;51:348–52.
- Smith MP, Capretta PJ. Effects of drive level and experience on the reward value of saccharine solutions. *Journal of Comparative and Physiological Psychology* 1956;49:553–7.
- Sugrue LP, Corrado GS, Newsome WT. Matching behavior and the representation of value in the parietal cortex. *Science* 2004;304:1782–7.
- Viterbi AJ. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory* 1967;13:260–9.