Mice plan decision strategies based on previously learned time intervals, locations, and probabilities

Tuğçe Tosun, Ezgi Gür, and Fuat Balcı[1]

Department of Psychology, Koç University, Rumelifeneri Yolu, Sarıyer, Istanbul 34450, Turkey

Edited by Warren H. Meck, Duke University, Durham, NC, and accepted by the Editorial Board December 10, 2015 (received for review September 15, 2015)

Animals can shape their timed behaviors based on experienced probabilistic relations in a nearly optimal fashion. On the other hand, it is not clear if they adopt these timed decisions by making computations based on previously learnt task parameters (time intervals, locations, and probabilities) or if they gradually develop their decisions based on trial and error. To address this question, we tested mice in the timed-switching task, which required them to anticipate when (after a short or long delay) and at which of the two delay locations a reward would be presented. The probability of short trials differed between test groups in two experiments. Critically, we first trained mice on relevant task parameters by signaling the active trial with a discriminative stimulus and delivered the corresponding reward after the associated delay without any response requirement (without inducing switching behavior). During the test phase, both options were presented simultaneously to characterize the emergence and temporal characteristics of the switching behavior. Mice exhibited timed-switching behavior starting from the first few test trials, and their performance remained stable throughout testing in the majority of the conditions. Furthermore, as the probability of the short trial increased, mice waited longer before switching from the short to long location (experiment 1). These behavioral adjustments were in directions predicted by reward maximization. These results suggest that rather than gradually adjusting their time-dependent choice behavior, mice abruptly adopted temporal decision strategies by directly integrating their previous knowledge of task parameters into their timed behavior, supporting the model-based representational account of temporal risk assessment.

decision making | interval timing | temporal risk assessment | probabilities | mice

Many vertebrate species can build temporal expectancies and cluster their anticipatory behaviors around intervals that lead to critical outcomes (1). The resultant timed behaviors have been shown to be sensitive to other crucial elements of environmental statistics, such as the probabilities of different outcomes (2–4). However, how steady-state choice behavior emerges in temporal decision-making tasks that contain probabilistic contingencies remains to be answered. Do animals directly manifest their knowledge of quantities (e.g., time interval, probability) and locations in their timed behavior or do they gradually acquire differential timed response patterns based on reinforcement learning in a fashion stripped of representations? This study aimed to address this fundamental question using a simple temporal decision-making task.

A class of interval timing paradigms requires the animals to distribute their responses between two (short vs. long latency) options, each of which predicts reward after the corresponding fixed delay. In these cases, the emergent response pattern is first behaviorally investing in the option with a short delay to the reward, and if responding after the short interval does not result in reward delivery, then switching to the option with a longer delay to the reward (5–7). Using this paradigm, a number of studies have consistently shown that mice and humans can integrate the probabilities of different options into their time-based decisions. For instance, if the probability of receiving reward at the short-latency location is higher than it is at the long-latency location, animals switch from the short-latency option to the long-latency option later in the trial, and they make this decision in a nearly optimal fashion (3, 8, 9). To investigate how mice attained the optimal timed switching behavior, Kheifets and Gallistel (8) manipulated the probabilities of different options over the course of testing and characterized the performance of mice after these manipulations of the task parameters. Their results suggested that mice adjusted the timing of switching behavior in an optimal fashion very rapidly and abruptly. These results favored the model-based representational account of timed switching behavior.

A critical test of this claim in the context of a timed switching paradigm is initially training the animals with different options individually (with predetermined probabilities) without inducing the switching behavior and then presenting both short- and long-latency options simultaneously to characterize the emergence and parameterization of timed switching behavior. If animals represent the task parameters (i.e., time intervals, locations, probability of different trial types) and directly apply them to their timed choice behavior, then the timed switching behavior should emerge starting from the very first test trials and its temporal parameterization should reflect the estimates of delays to reinforcement as well as the probabilistic relations acquired during the training phase. In other words, subjects would approximately make Bayesian inference by combining their prior knowledge (e.g., prior probabilities of different options) with their real-time judgments of elapsed duration and incorporate it into their timed decisions.

In the first experiment, we conducted this critical test with mice in three different probability conditions. The results of this experiment showed that mice can integrate previously learned time intervals, their relation to different locations, and probabilities of different trial types into their decisions when first confronted with a choice in which all three are factors and adjust their decisions in

Significance

A fundamental question is whether experience shapes appropriate behavior by selective reinforcement of that behavior or whether it instills knowledge of environmental parameters, which knowledge may then be deployed to direct behavior in novel situations. We initially trained mice on three different environmental parameters (time intervals, locations, and probabilities) and then confronted these mice for the first time with the situation that required them to combine and use these parameters to predict reward. Mice deployed this knowledge to inform an appropriate behavioral strategy. Their behavioral strategies appeared immediately and did not improve further. These results suggest that mice can directly apply their prior knowledge about time intervals, locations, and probabilities to their decisions without any need for trial and error.

Author contributions: F.B. designed research; T.T., E.G., and F.B. performed research; T.T., E.G., and F.B. analyzed data; and T.T., E.G., and F.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. W.H.M. is a guest editor invited by the Editorial Board.

[1]To whom correspondence should be addressed. Email: fbalci@ku.edu.tr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1518316113/-/DCSupplemental.

directions predicted by optimality. In the second experiment, the same test was conducted with a new set of mice in two probability conditions. The second experiment had a shorter training period, with additional probe trials for attaining a more sensitive test of timed switching tendencies that might be present during training. The key findings of the first experiment regarding the immediate and abrupt emergence of the timed switching behavior in the test phase were confirmed by the results of the second experiment.

Results

Experiment 1. Fig. 1 shows the raster plots of the responses of each mouse in the last training session (phase 1) and the first test session (phase 2) separately for different probability conditions. Mice responded almost exclusively at the correct location when it was signaled by a discriminative stimulus. However, when both options were signaled in the test phase, mice went first to the short-latency hopper, and if the reward was not received there after the short delay, they switched to the long-latency hopper (averaged response curves are shown in Fig. S1). Importantly, this pattern of behavior was immediately evident. The average rates of switching in the last training session were 6%, 3%, and 6% (SDs 5%, 5%, and 6%) for 0.25, 0.50, and 0.75 short-trial probability conditions, respectively. These values increased to a significantly higher level [76%, 71%, and 65% (SDs 13%, 28%, and 21%), respectively] in the first test session ($Z = -2.02$, $P < 0.05$ for all groups). These results indicated that mice could immediately build a temporally controlled motor plan (timed-switching behavior) when the task demands require it

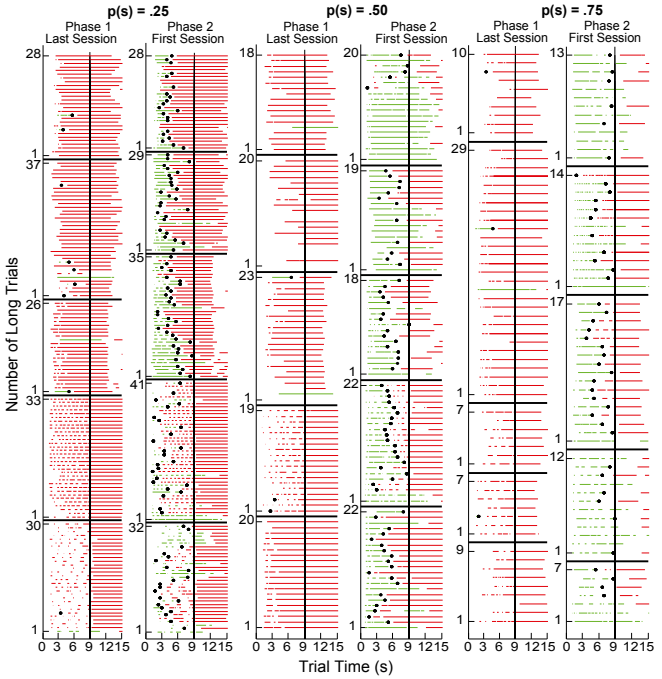


Fig. 1. Raster plots of short (green line) and long (red line) hopper responses in the long trials in the last session of phase 1 (training phase) and in the first session of phase 2 (test phase) of experiment 1. Black dots mark the switches from the short-latency to long-latency hopper. Horizontal thick black lines separate the data collected from different mice. The short reward latency was 3 s, whereas the long reward latency was 9 s (vertical black lines). Notice that switches were rare at the end of phase 1, but they appeared immediately in phase 2. Notice, second, that the switches occur after the expected time for the short-latency feed and before the expected time for the long-latency feed. Notice, third, that the higher the probability of a short trial [p(s)], the later the mice tended to switch, and, finally, that this between-group difference is apparent from the outset.

(between-group comparisons of the switch rates are presented in *SI Text*, section 1.1).

On the other hand, it is possible that mice directly transferred their independent short- and long-latency–related response patterns developed during training to the test trials, resulting in the impression that the timed switching pattern emerged during testing. This possibility can be tested by the between-phase comparisons of the trial times at which mice terminated their short-latency location response ($T_{Short,Stop}$) and started their first long-latency location response ($T_{Long,Start}$). If mice directly transferred their training response patterns to the test phase, one would not expect a between-phase difference in either of these parameters. Only $T_{Long,Start}$ values were analyzed in experiment 1, because short-latency location responses were rare in long trials and it was not possible to estimate $T_{Short,Stop}$ values reliably in short trials due to the strict procedural censor.

The $T_{Long,Start}$ values in long trials were significantly earlier during training compared with testing in the 0.25 (median: 2.48 s vs. 5.34 s) and 0.50 (median: 2.67 s vs. 6.12 s) short trial probability groups ($Z = -2.02$, $P < 0.05$, in both groups). Although this difference was in the same direction (median: 3.40 s vs. 6.98 s), it was marginally significant in the 0.75 short trial probability group ($Z = -1.83$, $P = 0.07$). These between-phase differences suggest that mice did not directly manifest their training response patterns in the test phase (Fig. S2). This conclusion was further supported by the results of the between-group comparisons of $T_{Long,Start}$ values separately during training and testing. Although $T_{Long,Start}$ values did not differ between probability groups during training [$\chi^2(2) = 2.42$, $P = 0.30$], there was a significant between-group difference during testing [$\chi^2(2) = 6.09$, $P < 0.05$]. Pairwise comparisons revealed a significant difference between the 0.25 and 0.75 short trial probability groups (difference held after Holm–Bonferroni correction). The results of the same comparisons conducted for the latency to first short-latency location response (gathered from the short trials) are presented in *SI Text*, section 1.2. Finally, our analyses showed that the differences between $T_{Long,Start}$ and $T_{Short,Start}$ (latency to the first short-latency location response) values were overall substantially larger during testing than during training, suggesting a stronger temporal control over responding during testing (*SI Text*, section 1.3 and Fig. S1). The summary statistics regarding these parameters are presented in Table S1.

To examine whether mice adjusted their switch latencies over the trials within the first test session, and because conventional null-hypothesis significance tests cannot provide evidence in favor of the null hypothesis, we conducted a Bayesian analysis (10). A simple Bayesian analysis compares the marginal likelihoods of two competing statistical models for the data. To this end, the Bayes factor (BF) (11) is a commonly used measure that indicates the probability of data under one model in comparison to another model (i.e., odds ratio). In the present case, the first model was that the slope of the regression was exactly 0. The alternative model was that the slope of the regression was somewhere within the range of the maximum-likelihood slopes obtained by the conventional regression. Fig. 2*A* gives the cumulative distribution of the weight of the evidence (the common log of the BF) for the 14 analyses (excluding one subject with only five unusual data points yielding both slope and intercept estimates that were gross outliers). A BF of 1 (weight of 0) provides no evidence, BF > 1 (positive weights) provides evidence in favor of the alternative hypothesis, and BF < 1 (negative weights) provides evidence in favor of the null hypothesis. The weight of the evidence favored the null model in the majority of cases [10 of 14 subjects (71%)], often by a substantial value. The median weight was −0.36, which gives odds of 2.31:1 favoring the null model (further details are provided in *SI Text*, section 1.4). The results of the corresponding conventional regression analysis are presented in *SI Text*, section 1.5.

The means, SDs, and coefficients of variation (CVs) of the switch latencies were estimated for each subject from the first test session data, which were then compared between the three probability groups. The switch latencies (Fig. 3*A*) differed significantly between probability groups [$\chi^2(2) = 10.14$, $P < 0.01$] whereas

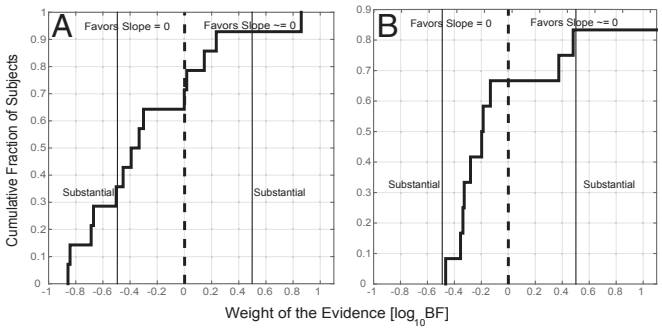


Fig. 2. Cumulative distribution of the weight of the evidence in favor of the null model (negative weights) or in favor of the model that assumes that slopes may, in fact, cover the range obtained from conventional linear regressions on the data from each of 14 subjects in experiment 1 (*A*) and 12 subjects in experiment 2 (*B*). (*A*) In experiment 1, in 10 of the 14 datasets, the weight of the evidence favors the null model. In six of these cases, the weight approaches or exceeds the conventional level considered to betoken substantial evidence (BF = 3:1, vertical solid line at −0.48). In only one case does the weight in favor of the alternative approach or exceed this criterion (vertical solid line at 0.48). (*B*) In experiment 2, in 8 of the 12 datasets, the weight of the evidence favors the null model. In one of these cases, the weight approached the conventional level considered to betoken substantial evidence. In three cases, the weight in favor of the alternative approaches or exceeds this criterion.

there was no significant difference between the groups in terms of SD [$\chi^2(2) = 1.94$, $P = 0.38$] or CV [$\chi^2(2) = 3.92$, $P = 0.14$] estimates. Pairwise comparisons of the switch latencies revealed significant differences for each pair (all P values < 0.05). After applying Holm–Bonferroni correction, the difference between the 0.25 [median = 4.30 s, interquartile range (IQR) = 0.80] and 0.75 (median = 6.40 s, IQR = 1.70) short trial probability groups remained significant (for the 0.50 short trial probability group, median = 5.43 s, IQR = 1.04).

Briefly, mice tested in the highest short probability condition had longer switch latencies compared with the ones tested in lowest short trial probability condition. The follow-up Bayesian t test (12) provided evidence in favor of the hypothesis that the switch latencies were different between each pair of the three probability groups. The BFs for the short trial probability pairs of 0.25–0.50, 0.50–0.75, and 0.25–0.75 were 3.45, 2.82, and 40.21, respectively.

The optimality analysis of the empirical switch latencies was conducted using the expected gains associated with the estimated target switch times, endogenous timing uncertainty [coefficient of variation (CV)], probabilities of different trial types, and payoffs. The gain index was computed by taking the ratio of the expected gain to the maximum possible expected gain (MPEG) of each subject, given the level of timing uncertainty (6). In the first session of testing, on average, subjects earned 99.23%, 98.98%, and 99.56% (all SDs = 1%) of their MPEG in the 0.25, 0.50, and 0.75 short trial probability groups, respectively. The more stringent definition of the gain index, which considered the minimum gain as the gain predicted by random choice between the two options, revealed very similar results (*SI Text*, section 1.6). Thus, even in the first test session, the performance of the subjects in each probability condition was nearly optimal in terms of the gains attained.

There was a significant relation between the empirical and optimal switch latencies ($\beta = 2.43$, $P < 0.01$), although the slope was higher than 1 (Fig. 3*B*). The observed slope suggests that mice overparameterized their timed switching behavior in response to different exogenous probabilities. In other words, mice tended to switch earlier than the optimal latency for the lowest short trial probability condition, although later than the optimal latency for the highest short trial probability condition. However, this overparameterization resulted in only negligible loss in the reward attained in this task. Furthermore, the Bayesian t test comparison of empirical and optimal switch latencies across all conditions provided substantial evidence favoring the theoretically important null hypothesis that empirical and optimal switch latencies did not differ (BF = 4.51 in favor of the null hypothesis). In no analysis of the individual probability conditions did the BF favor the alternative model.

Experiment 2. It is conceivable that the design of the first experiment masked the behavioral manifestation of the possible timed switching tendencies that could be present during training. Consequently, it is possible that the immediate and abrupt emergence of timed switching behavior in the test phase of experiment 1 was an artifact. Properly designed probe trials might indeed be necessary for the induction of switching behavior. Experiment 2 was designed to address these potential issues. In addition to the signaled short and long trials, the training phase of the second experiment contained signaled short and long probe trials that lasted much longer than the corresponding reinforcement trials. Importantly, the reinforcement was omitted in these probe trials.

Furthermore, the training phase of experiment 1 lasted until each mouse met preset criteria that could have indirectly suppressed the behavioral manifestation of the switching tendency. These success criteria were eliminated, and the training period was kept fixed at a length that was shorter than it was in experiment 1. The lowest short trial probability condition was not tested in this experiment (given its primary aim) due to the very low number of the critical short probe trials that would result in this condition.

As in experiment 1, mice immediately started switching from the short-latency to long-latency hopper if the reward was not presented in the short-latency hopper after the short delay in the first test session of experiment 2. Raster plots of each mouse are presented in Fig. 4 for the last training session (first and third columns) and the first test session (second and fourth columns; averaged response curves are shown in Fig. S3). The average rate of switching in the long trials of the last training session were 24% and 17% (SDs 20% and 17%) for the 0.50 and 0.75 short trial probability groups, respectively. These values increased to 67% and 77% (SDs 25% and 21%) in the first test session. Between-phase differences were significant for both groups (both $Z = -2.20$, $P < 0.05$). The between-group comparisons of the switch rates are presented in *SI Text*, section 2.1).

Additionally, we compared the average rate of switching in the short probe trials (from the last five sessions of training; fifth and sixth columns of Fig. 4) with the switch rates in the long trials of first test session for both groups. The average switch rate in short probe trials were 13% and 10% (SDs 8% and 13%) for the 0.50 and 0.75 short trial probability groups, respectively. Comparison of the switch rates between training (short probe trials) and the testing revealed significant differences for both probability groups (both $Z = -2.20$, $P < 0.05$). The between-group comparisons of the switch rates in short probe trials (*SI Text*, section

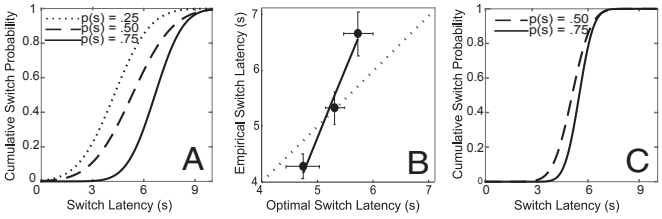


Fig. 3. Cumulative normal distribution functions were built based on the group-averaged switch latencies and SDs estimated from each probability condition in experiment 1 (*A*) and experiment 2 (*C*). Dotted, dashed, and solid curves correspond to the 0.25, 0.50, and 0.75 short trial probability conditions, respectively. (*B*) Average empirical switch latencies as a function of average optimal switch latencies for different probability conditions in experiment 1. The solid line is the best-fit orthogonal regression line, and the dotted line is the identity line. Error bars show SEM.

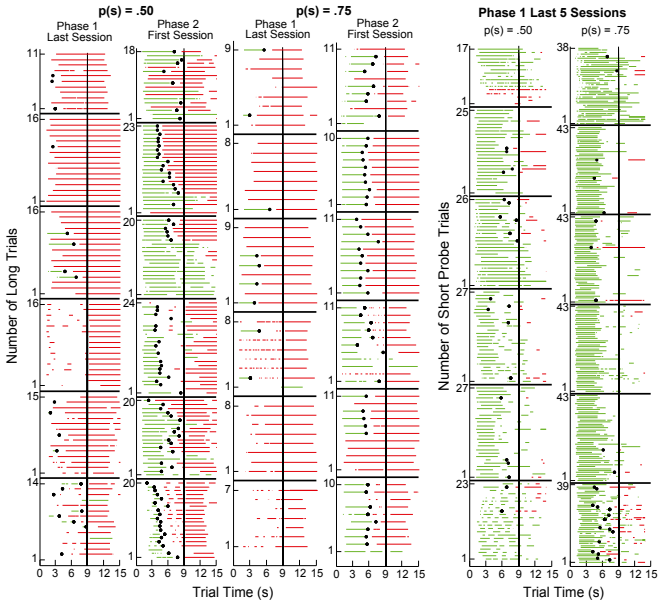


Fig. 4. Raster plots of short and long hopper responses in the long trials in the last session of phase 1, in the long trials in the first session of the phase 2, and in the short probe trials of the last five sessions of phase 1 of experiment 2. Black dots mark the switches from the short-latency to long-latency hopper. Notice that switches were rare at the end of phase 1 (in both the long trials and short probe trials), but they appeared immediately in phase 2. Notice, second, that the switches occur after the expected time for the short-latency feed and before the expected time for the long-latency feed.

2.1), the analysis of average switch rates in the long probe trials (*SI Text*, section 2.2), and the comparison of switch rates between short probe and long trials of training are presented in *SI Text*, section 2.3. Overall, these findings suggest that the timed switching tendency was not present during training.

Despite these clear findings that pointed to the novelty of the timed switching response pattern during testing, we investigated if it was possible that mice directly transferred their timed behaviors developed during training, which could have been independently parameterized for short-latency and long-latency locations, to the test phase. We tested this possibility by focusing primarily on $T_{Long,Start}$ and $T_{Short,Stop}$ values. The summary statistics regarding these parameters are presented in Table S1 (also Fig. S4).

We first compared $T_{Long,Start}$ values between phases and probability groups. The $T_{Long,Start}$ values in long trials of the training (medians: 3.69 s and 3.63 s for 0.50 and 0.75 short trial probability conditions) and test phase (corresponding medians: 6.46 s and 6.21 s) differed significantly in both probability groups ($Z = -2.20$, $P < 0.05$ for both groups). There were no differences between the two probability groups in terms of $T_{Long,Start}$ values in either phase ($Z = -0.16$, $P = 0.87$ for the training phase and $Z = -0.48$, $P = 0.63$ for the test phase). Similar comparisons of $T_{Short,Start}$ values are presented in *SI Text*, section 2.4 and Table S1. Furthermore, $T_{Short,Stop}$ values between the long trials of the test phase and short probe trials of training (last five sessions) were compared for each group. The median $T_{Short,Stop}$ values in short probe trials of training were 7.78 s and 7.10 s in the 0.50 and 0.75 short trial probability conditions, whereas these values decreased significantly to 6.21 s ($Z = -1.99$, $P < 0.05$) and 6.00 s ($Z = -2.20$, $P < 0.05$) in the long trials of the test phase. The two probability groups did not differ from each other in terms of $T_{Short,Stop}$ values during training ($Z = -1.76$, $P = 0.08$) or the test phase ($Z = -0.64$, $P = 0.52$). Finally, we observed that the differences between $T_{Long,Start}$ and $T_{Short,Start}$ values were overall substantially larger in the test phase compared with the training phase, again suggesting a stronger temporal control over responding during testing (*SI Text*, section 2.5, Fig. S3, and Table S1). Consistent with the conclusion of experiment 1, these results overall suggest that mice did not simply manifest their independent short-latency and long-latency location response patterns developed during training in the test phase.

As in the case of experiment 1, we conducted the Bayesian analysis of the change in switch latencies over the course of the test session (Fig. 2*B*). In eight of the 12 datasets (67% of the subjects), the weight of the evidence favored the null model. The median weight was −0.19, which gives odds of 1.56:1 (weak evidence) in favor of the null model (additional details are provided in *SI Text*, section 2.6). The results of the corresponding conventional regression analysis are presented in *SI Text*, section 2.7.

Median switch latencies were 5.49 s (IQR = 2.02) and 5.56 s (IQR = 0.79) for the 0.50 and 0.75 short trial probability groups, respectively (Fig. 3*C*). The mean switch latencies ($Z = -0.16$, $P = 0.87$), SDs ($Z = -0.64$, $P = 0.52$), and CVs ($Z = -0.96$, $P = 0.34$) estimated from the first test session were not different between the two probability groups. Note that the difference in the switch latencies between the 0.50 and 0.75 short trial probability groups also disappeared in experiment 1 after Holm–Bonferroni correction. However, different from experiment 1, the Bayesian *t* test provided evidence in favor of the null hypothesis that the switch latencies did not differ between the two probability groups (BF = 2.68 in favor of the null model) in the second experiment.

The optimality analysis of the empirical switch latencies was conducted as in experiment 1. On average, subjects earned 99.05% and 99.64% (SDs = 1%) of the MPEG in the 0.50 and 0.75 short trial probability groups, respectively. The results of the more conservative approach to the gain index were consistent with these results (*SI Text*, section 2.8). Consistent with the results of the first experiment, the Bayesian *t* test comparison of empirical and optimal switch latencies across all conditions provided substantial evidence in favor of the null hypothesis that empirical and optimal switch latencies did not differ (BF = 3.11 in favor of the null model). In the analysis of none of the individual conditions did the BF favor the alternative hypothesis. Thus, the results of these optimality analyses were overall consistent with the results gathered from experiment 1.

Discussion

The switching behavior observed during the test phase of the first experiment was based on knowledge of three different environmental parameters acquired in the preceding training phase, during which switch behavior was not appropriate, never reinforced, and rarely seen. A long-standing question of fundamental importance in psychology, cognitive science, and neuroscience is whether experience shapes appropriate behavior by selective reinforcement of that behavior or whether it instills knowledge of environmental parameters, which knowledge may then be deployed to direct behavior in novel situations. Our results imply that during the training phase, mice learned the relative frequencies of each location providing a reward after the latency associated with that location. When confronted for the first time with the situation in which it was ambiguous which location to go to, mice deployed this knowledge to inform an appropriate behavioral strategy. They went first to the short-latency location and then switched to the long-latency location if and when the short-latency location failed to pay off. This behavior appeared immediately and did not improve further.

The key findings of the first experiment were replicated in the second experiment. However, the training phase of the second experiment also contained the critical probe trials that lasted longer than the corresponding latencies to reward, and the reward was omitted in these trials. These probe trials specifically aimed to set the occasion for the occurrence of switching behavior in case the design of the first experiment masked the manifestation of the timed switching tendency. Consistent with the findings of the first experiment and data gathered from the reinforced trials of the second experiment, mice exhibited very low rates of switching in these probe trials. Finally, the comparisons of temporal characteristics of independent critical response

elements (i.e., latency to the first long-latency location response and end of the short-latency location response) showed consistent differences between the training and test phases, and these parameters were more sensitive to the manipulations of trial type probabilities during the test phase. These findings suggested that mice did not simply transfer their independent short-latency and long-latency location response patterns developed during training to the test phase. Based on this convergent evidence, we conclude that the rewards received in the training phase were important for the information they provided about parameters of the environment, rather than for a strengthening effect on the "reward-producing behavior" observed in the test phase.

The most striking manifestation of this previously acquired knowledge is that the distribution of switch latencies was appropriately and immediately determined not only by the previously experienced differences in feed latencies but also by the previously experienced differences in the relative frequencies of obtaining food at the two different locations. In experiment 1, the higher the previous probability of receiving food at a short latency at the short-latency location was, the longer mice tended to stay there before switching to the long-latency location. These behavioral adjustments made in the face of probabilistic manipulations were in directions predicted by reward maximization. However, the corresponding difference between the 0.50 and 0.75 short trial probability conditions was not observed in experiment 2. Note that this difference did not survive the correction for multiple comparisons in the first experiment either. These particular probability conditions were picked for experiment 2 due to relatively larger number of short probe trials (the critical test trials) that would result in these conditions.

Consistent with this asymmetry, Akdoğan and Balcı (13) recently found that in the temporal bisection task (retrospective variant of the timed switching task), the leftward shifts in the psychometric functions of mice in response to a reduction in the short reference duration probability were more pronounced than the rightward shifts in response to an increase in this probability [a similar effect with reward magnitude manipulations is described by Galtress and Kirkpatrick (14)]. These differential findings between the two experiments regarding the effect of probabilistic manipulations on switch latencies could also be due to the substantially shorter training in the second experiment, the presence of probe trials that might have disrupted the estimation of discrete probabilities, and/or the elimination of the global stimulus (white noise) in experiment 2 that accompanied the local stimulus (hopper illumination) in experiment 1.

Finally, the finding that animals can immediately and abruptly incorporate their previously acquired knowledge regarding time intervals into their choice behavior in an adaptive fashion is consistent with the findings that rats had temporally specific savings in cross-modal transfer of temporal discriminations (e.g., 15). However, different from these earlier important findings, in our study, this time-based knowledge is manifested in the form of a new temporally and spatially controlled response pattern that (at least for the probability conditions tested as part of experiment 1) was also sensitive to the experienced probabilistic contingencies.

Future studies are needed to test the neural basis of our findings, and previous studies provide insights to this end. Distributed brain networks that include the basal ganglia, parietal cortex, frontal cortex, and thalamus are implicated for interval timing (1, 16–18). The striatum is particularly thought to have a central role in timing behavior (19). Most relevant to our study, Portugal et al. (20) have shown that compared with the periods of responding that targeted individual options [i.e., fixed interval (FI) and variable interval (VI) schedules], the switching between these delayed options was more effective in leading to maximal change in the firing rate of the task-modulated striatal neurons. These results suggested that rather than event times, the striatum might encode the timing of adaptive behavioral transitions between options based on the environmental statistics (20), which included multiple elements in our task (e.g., durations, probabilities, locations). Parts of the frontal cortex, such as the anterior cingulate cortex, were previously shown to multiplex the representations of key decision variables (e.g., probability, payoffs), pointing to their role for optimal decision making (21, 22). The integration of this information into temporal control might, in turn, be realized through the repeated propagation of temporal information through the corticostriatal-thalamic loop and the resultant dynamic modulation of time-dependent striatal activity during this process (17, 23). Consistent with this assumption, the effects of prior probabilities in perceptual decision making have also been shown to be underlain by the activation of the corticostriatal circuit in humans (24). Importantly, recent evidence further suggests that the cortical modulation during decision making can take the form of discrete steps characterized by abrupt transitions within a trial (25), capturing the discrete nature of timed switching behavior (short vs. long) in our task.

The most critical aspect of our findings is the abrupt and immediate appearance of adaptive timed decisions. A recent electrophysiological study provides insights also regarding the neural underpinnings of these features of the timed behaviors. Mello et al. (26) showed that striatal neuronal populations can very quickly adapt to the immediate task demands with high flexibility by rescaling their temporal tuning to abrupt changes in delays to reinforcement. These neuronal populations were shown to multiplex information regarding time, sensorimotor state, and other task relevant information, predicting the timing behavior of rats as they adapted to new critical delays. Our experimental design also introduces an immediate change in task demands with the simultaneous presentation of two response options during testing. The resultant uncertainty might constitute a relevant signal for triggering quick adaptive alterations in striatal activity to meet the new task requirements better (27), leading to immediate changes in timed behavior. The effect of the new task demands could be incorporated into ongoing behavior, again through the corticostriatal-thalamic loop, and the modulation of temporal tuning of striatal neurons during this process might be weighted by the differential valuations of the two options [in our case, due to their different probabilities (24)].

Overall, the findings of the current study have important implications regarding the processes that underlie the emergence of timed decisions. Our results suggest that mice can directly apply their prior knowledge about time intervals, locations, and probabilities into their time-based decision making without any need for trial and error. These results are in contrast to the findings gathered from the matching paradigm that used a similar approach: training subjects separately on each individual option (each VI schedule) and testing them in the presence of two options presented simultaneously (concurrent VI schedules) to characterize the emergent choice behavior of the animals (28–30). In these studies, animals exhibited overmatching; they almost exclusively preferred the richer option without the frequent switching back and forth between the options, a characteristic of matching behavior (28–30). One possible reason underlying this difference between our study and matching studies could be the stronger temporal determinism in the switch task compared with the matching paradigm due to the use of fixed vs. variable delays to reward in these tasks, respectively. The manifestations of quantity representations might be more likely in temporally deterministic tasks compared with the nondeterministic ones.

Methods

Subjects. Twenty-seven experimentally naive C57BL/6J male mice obtained from the Koç University Animal Research Facility were used in experiment 1 (15 mice) and experiment 2 (12 mice). All mice were approximately 8 wk old at the beginning of the experiment. Animals were housed in groups of four in polycarbonate cages (Allentown type I long individually ventilated cages) in rooms, which were illuminated on a 12-h/12-h light/dark cycle period (lights on at 6:00 AM). Mice were tested in daily 1-h-long sessions during the light cycle. Three days before the beginning of the experiment, mice were subjected to a food deprivation procedure with ad libitum access to water. After each session, they were fed with additional food pellets to maintain them at 85% of their free-feeding weights. All procedures were approved by the Koç University Animal Research Local Ethics Committee.

Apparatus. Mice were tested in the operant chambers (Med Associates) placed in sound-attenuated boxes. Three illuminable feeding hoppers were mounted on one of the metal walls; only the ones at the extreme ends were used (Fig. S5*C*). The reinforcement was delivered from the active hoppers. A nose poke hole in the middle of the opposite metal wall was used to initiate the trials (Fig. S5*C*). A sound generator delivered white noise (75 dB) as an auditory stimulus for programmable durations. MED-PC IV software was used to control the experiment and record the data. Event times were logged and stamped with a resolution of 10 ms (further details are provided in *SI Text*, section 3.1).

Procedure.

General procedure. There were two types of trials: 3-s short trials and 9-s long trials. In experiment 1, on both the short and long trials, a visual signal and an auditory signal were presented for an interval associated with the corresponding trial type. In experiment 2, only the visual stimulus was presented to signal the trial type. Rewards were presented at different locations in the short and long trials (counterbalanced between mice). Subjects were divided into three short trial probability groups (0.25, 0.50, and 0.75) in experiment 1. In experiment 2, mice were tested only in the 0.50 and 0.75 short trial probability conditions.

Phase 1 (training phase). In experiment 1, when the hopper on the opposite wall of the feeding hoppers was illuminated, mice could initiate the trial by nose poking into this hopper. By means of this requirement, it was ensured that the location of each subject was fixed at the beginning of each trial. When mice made the trial-initiating nose poke, the light in the control hopper was turned off, the light in the feeding hopper associated with the active trial type was turned on to signal the correct location, and the white noise started. The white noise (used as the global timing stimulus) and the light in the active hopper terminated at the end of the interval associated with the corresponding trial type. The reward was delivered in the appropriate hopper irrespective of response. There was a 30-s fixed-delay intertrial interval and a VI of 60 s (exponentially distributed). To complete the first phase, subjects had go directly to the active feeding hopper (without going to the other hopper) in at least 85% of both short and long trials, had to initiate at least a total of 32 trials, and had to collect the reward (within 6 s after the delay associated with the active trial type) in at least 75% of both short and long trials in three consecutive experimental sessions. All subjects completed the first phase in 30 sessions on average.

A similar procedure was used in the first phase of experiment 2. However, different from experiment 1, only the feeder hopper lights were used for signaling the time interval (the auditory stimulus was eliminated), subjects were tested in this phase for a fixed number of sessions (i.e., 20 sessions), and the success criteria for completing the training phase were eliminated. Furthermore, we included the probe trials for both short and long trial types (constituting 25% of all trials). In the probe trials, the corresponding feeder hopper light stayed on for threefold longer than the corresponding trial duration (i.e., 9 s for short trial and 27s for long trial types) and no reward was delivered. Data from the short and long trials of the last session of phase 1 (experiments 1 and 2) and from the probe trials of the last five sessions of phase 1 (experiment 2) were used in the analysis.

Phase 2 (test phase). In each trial, both feeding hoppers were illuminated (for the duration of the active trial type) regardless of the trial type in both experiment 1 and experiment 2 as opposed to single hopper illumination during training. The auditory signal (white noise) accompanied illumination of the hoppers only in experiment 1. The probe trials were eliminated in phase 2 of experiment 2. Otherwise, all procedures in the second phase were identical to the procedures in phase 1 of the corresponding experiment. All subjects were tested with the short trial probabilities that were in effect in phase 1. Data from the first session of phase 2 were used in the analysis. Fig. S5 *A* and *B* present the schematic representation of the training and testing procedures.

Data Analysis. The primary unit of analysis was the switch latencies in the long trials. Switch latencies longer than the duration of the long trials (9 s) were excluded (these data can be observed in Figs. 1 and 4). The average switch latencies, CVs and SDs of switch latencies, and switch rates (e.g., proportion of long trials that contained a switch) were compared between three different experimental groups using the Kruskal–Wallis test in experiment 1 and between two different experimental groups using the Mann–Whitney *U* test in experiment 2. When appropriate, pairwise comparisons were conducted using the Mann–Whitney *U* test. Asymptotic and exact *P* values gathered from Mann–Whitney U tests led to the same conclusions. Holm–Bonferroni corrections were applied to correct for multiple pairwise post hoc comparisons. Where appropriate, similar comparisons were also made for the latencies to onset of the long-latency location responses as well as the latencies to the onset and offset of the short-latency location responses (short-latency location response offsets longer than 9 s in short probe trials were fixed to 9 s). The switch rates and the latency measures in phase 1 and phase 2 were compared separately for each experimental group using Wilcoxon signed rank tests. Note that one mouse [in p(s) = 0.75 condition in experiment 1] was excluded from the analyses of the first long-latency location responses during the test phase since it initiated all of its long-latency responses after the end of the long trial (9 s) in this phase. The relationship between empirical and optimal switch latencies was analyzed using orthogonal regression in experiment 1. Several key comparisons were also made using Bayesian methods (10–12). The Bayesian *t* tests (12) were conducted with the scale r on effect size = 1 and the Jeffrey-Zellner-Siow (JZS) BFs are reported. The optimality analysis is presented in *SI Text*, section 3.2.

ACKNOWLEDGMENTS. We thank Charles R. Gallistel for his comments on an earlier version of this manuscript and his help with the Bayesian analysis. This work was supported by Grant 111K402 of the Scientific and Technological Research Council of Turkey, TÜBİTAK (to F.B.).

1. Buhusi CV, Meck WH (2005) What makes us tick? Functional and neural mechanisms of interval timing. *Nat Rev Neurosci* 6(10):755–765.
2. Balci F, et al. (2011) Optimal temporal risk assessment. *Front Integr Neurosci* 5:56.
3. Coşkun F, Sayalı ZC, Gürbüz E, Balcı F (2015) Optimal time discrimination. *Q J Exp Psychol (Hove)* 68(2):381–401.
4. Jozefowiez J, Polack CW, Machado A, Miller RR (2014) Trial frequency effects in human temporal bisection: Implications for theories of timing. *Behav Processes* 101:81–88.
5. Balci F, et al. (2008) Interval timing in genetically modified mice: A simple paradigm. *Genes Brain Behav* 7(3):373–384.
6. Balci F, Freestone D, Gallistel CR (2009) Risk assessment in man and mouse. *Proc Natl Acad Sci USA* 106(7):2459–2463.
7. Stubbs DA (1976) Response bias and the discrimination of stimulus duration. *J Exp Anal Behav* 25(2):243–250.
8. Kheifets A, Gallistel CR (2012) Mice take calculated risks. *Proc Natl Acad Sci USA* 109(22):8776–8779.
9. Maggi S, et al. (2014) A cross-laboratory investigation of timing endophenotypes in mouse behavior. *Timing Time Percept* 2(1):35–50.
10. Gallistel CR (2009) The importance of proving the null. *Psychol Rev* 116(2):439–453.
11. Wetzels R, et al. (2011) Statistical evidence in experimental psychology: An empirical comparison using 855 t tests. *Perspect Psychol Sci* 6(3):291–298.
12. Rouder JN, Speckman PL, Sun D, Morey RD, Iverson G (2009) Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon Bull Rev* 16(2):225–237.
13. Akdoğan B, Balcı F (2015) Stimulus probability effects on temporal bisection performance of mice (Mus musculus). *Anim Cogn*, 10.1007/s10071-015-0909-6.
14. Galtress T, Kirkpatrick K (2010) Reward magnitude effects on temporal discrimination. *Learn Motiv* 41(2):108–124.
15. Church RM, Meck WH (1984) Acquisition and cross-modal transfer of classification rules for temporal intervals. *Quantitative Analyses of Behavior: Discrimination Processes* 4:75–97.
16. Matell MS, Meck WH, Nicolelis MAL (2003) Interval timing and the encoding of signal duration by ensembles of cortical and striatal neurons. *Behav Neurosci* 117(4):760–773.
17. Merchant H, Harrington DL, Meck WH (2013) Neural basis of the perception and estimation of time. *Annu Rev Neurosci* 36:313–336.
18. Allman MJ, Teki S, Griffiths TD, Meck WH (2014) Properties of the internal clock: First- and second-order principles of subjective time. *Annu Rev Psychol* 65:743–771.
19. Matell MS, Meck WH (2000) Neuropsychological mechanisms of interval timing behavior. *BioEssays* 22(1):94–103.
20. Portugal GS, Wilson AG, Matell MS (2011) Behavioral sensitivity of temporally modulated striatal neurons. *Front Integr Neurosci* 5:30.
21. Kennerley SW, Dahmubed AF, Lara AH, Wallis JD (2009) Neurons in the frontal lobe encode the value of multiple decision variables. *J Cogn Neurosci* 21(6):1162–1178.
22. Rushworth MFS, Behrens TEJ (2008) Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat Neurosci* 11(4):389–397.
23. Matell MS, Meck WH (2004) Cortico-striatal circuits and interval timing: Coincidence detection of oscillatory processes. *Brain Res Cogn Brain Res* 21(2):139–170.
24. Forstmann BU, Brown S, Dutilh G, Neumann J, Wagenmakers EJ (2010) The neural substrate of prior information in perceptual decision making: A model-based analysis. *Front Hum Neurosci* 4:40.
25. Latimer KW, Yates JL, Meister MLR, Huk AC, Pillow JW (2015) Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. *Science* 349(6244):184–187.
26. Mello GBM, Soares S, Paton JJ (2015) A scalable population code for time in the striatum. *Curr Biol* 25(9):1113–1122.
27. Courville AC, Daw ND, Touretzky DS (2006) Bayesian theories of conditioning in a changing world. *Trends Cogn Sci* 10(7):294–300.
28. Cerutti DT, Staddon JER (2004) Immediacy versus anticipated delay in the time-left experiment: A test of the cognitive hypothesis. *J Exp Psychol Anim Behav Process* 30(1):45–57.
29. Donahoe JW, Palmer DC (1994) Selection in the experienced learner. *Learning and Complex Behavior*, ed Dorsel VP (Allyn & Bacon, Boston), pp 112–113.
30. Gallistel CR, Gibbon J (2000) Time, rate, and conditioning. *Psychol Rev* 107(2): 289–344.