



Mice make temporal inferences about novel locations based on previously learned spatiotemporal contingencies

Ezgi Gür^{1,2} · Yalçın A. Duyan^{2,3} · Fuat Balcı^{1,2}

Received: 1 August 2022 / Revised: 7 November 2022 / Accepted: 9 November 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Animals learn multiple spatiotemporal contingencies and organize their anticipatory responses accordingly. The representational/computational capacity that underlies such spatiotemporally guided behaviors is not fully understood. To this end, we investigated whether mice make temporal inferences of novel locations based on previously learned spatiotemporal contingencies. We trained 18 C57BL/6J mice to anticipate reward after three different intervals at three different locations and tested their temporal expectations of a reward at five locations simultaneously, including two locations that were not previously associated with reward delivery but adjacent to the previously trained locations. If mice made spatiotemporal inferences, they were expected to interpolate between duration pairs associated with previously reinforced hoppers surrounding the novel hopper. We found that the maximal response rate at the novel locations indeed fell between the two intervals reinforced at the surrounding hoppers. We argue that this pattern of responding might be underlain by spatially constrained Bayesian computations.

Keywords Bayesian averaging · Interval timing · Mice · Conditioning · Peak interval procedure

Introduction

Interval timing is typically investigated using well-defined spatiotemporal contingencies as a result of which subjects learn when an outcome will occur at different locations. For instance, in the dual-interval peak procedure, two different levers are associated with two intervals, and subjects are tested separately on one lever at a time (Buhusi et al. 2018). The peaks of response rates at two levers are typically located around the outcome delays associated with the corresponding lever. When two food hoppers associated with short and long intervals are illuminated simultaneously in a single-phase experiment (Balcı et al. 2008; see also test phase of Tosun et al. 2016), subjects start responding on the short-latency hopper and then switch to the long-latency hopper when the short duration elapses with no reward

delivery. In both these procedures, timing behaviors are robustly guided by associations between location and time intervals. This is because even when there is uncertainty about which location would pay off in a given trial, there is no uncertainty about the association between locations and delays. Whether spatial relationships among different durations can determine the timing behavior of subjects at novel locations with no previously learned temporal relations is not known (but see Malet-Karas et al. 2019). For instance, a common metric system (e.g., Walsh 2003; A Theory of Magnitude, ATOM) that processes different magnitudes can foster cross-domain (e.g., spatiotemporal) transfer of judgments based on an isomorphic mapping between those domains.

This is an important question as it may shed light on the kind of model-based inferences subjects can spontaneously make based on the representations lent by the spatiotemporal contingencies experienced in the same context. For instance, Tosun et al. (2016) first trained mice to expect reward delivery independently at two different locations associated with two different delays and probabilities. In the training trials, the active location (hopper that would pay off) in that trial was signaled with a visual stimulus. Critically, when both locations were signaled during testing (resulting in

✉ Fuat Balcı
fuat.balci@umanitoba.ca

¹ Department of Biological Sciences, University of Manitoba, 50 Sifton Road, Room 222, Winnipeg R3T 2M5, Canada
² Department of Psychology, Koç University, Istanbul, Turkey
³ Department of Psychology, MEF University, Istanbul, Turkey

ambiguity), mice spontaneously adopted an optimal strategy by starting to respond at the location associated with the short delay to reward and then switching to the location associated with the long delay to reward. Critically, the switch times were sensitive to the previously experienced relative frequencies of reward delivery at the two different locations. In the context of the current study that focuses on the interaction between temporal and spatial representations, it is possible that animals make inferences about the temporal correlates of novel locations based on the neighboring spatiotemporal contingencies (spatiotemporal context), just temporal context independent of spatial context, or simply treat these novel locations as auxiliary visiting spots to be explored en route as they move between two previously reinforced locations. The current study aimed to answer this question by training mice to associate increasingly longer delays with three different serially ordered hoppers (e.g., Hoppers 1, 3, and 5) and testing them at all hoppers, including the interleaved hoppers (e.g., Hoppers 2 and 4).

There are at least two theoretical accounts regarding the kind of inferences that can be made for the temporal correlates of the novel hoppers. One of these possibilities stems from earlier research that showed an interaction between magnitude representations. For instance, human research has shown an interaction between time and space/size (Henik and Tzelgov 1982) as well as number and space (Dehaene et al. 1993). Some of these relations were also exhibited by non-human animals (Rugani et al. 2020). For instance, Rugani et al. (2015) showed that baby chicks spontaneously associate a smaller number with the left and a larger number with the right side of space (akin to the SNARC effect in humans, Dehaene et al. 1993). Merritt et al. (2010) trained two rhesus monkeys to categorize either duration or length of a line, and varied the line lengths in duration judgements and presentation duration in length judgements. They found that varying the dimension irrelevant to choice shifted the choice proportions such that when the same-length line is presented for a longer duration, monkeys were more likely to judge lines as larger. Moreover, this observed effect was bidirectional between time and space. The same pattern of results was also observed in pigeons in the bisection task (De Corte et al. 2017). Regarding the processing of spatiotemporal information, Cheng et al. (1996) trained pigeons to respond to a target moving at a constant speed. A reward was delivered upon the first response after a fixed amount of time had passed since the target started moving and/or a fixed amount of distance was covered. Hence, pigeons could rely on spatial or temporal information or both. When the speed of the target was varied, the pigeons' pattern of peak responding (duration or position) at different speeds changed systematically, best explained by a linear fit. The regression slope was somewhere between where it would be expected if pigeons relied on only temporal or only spatial information,

suggesting that the animals averaged time and spatial position, with different weights given to each type of information (see also Light et al. 2019 for a similar observation for temporal and numerosity judgments).

Based on these earlier findings, it is possible that subjects generate an isomorphic mapping between space and time across the experienced range of values, and that interpolation between learned temporal relations based on spatial reference points is built into these representations (Fig. 1, top panel). For instance, this can be underlain by the rescalability of the representations of spatial and temporal relations by CA1 pyramidal neurons (O'Keefe and Burgess 1996). Another way of achieving this would be averaging intervals associated with the two hoppers that are immediately neighboring the novel location. One good example of these computations is the temporal averaging phenomenon that emerges when stimuli that are individually associated with different intervals are presented as a compound stimulus (De Corte and Matell 2016). Different from experimental preparations of these earlier studies, in our example, averaging-like behavior would be induced by the ambiguity regarding the temporal correlates of the novel hoppers. Crucially, the same computation can also take place irrespective of the relative location of the novel hoppers, in which case the second possibility would be manifested (Fig. 1, bottom panel). Specifically, subjects would target the middle of the range of experienced intervals as a form of regression to the mean (Bayesian inference that is disassociated from spatial references) when faced with uncertainty regarding temporal correlates of novel hoppers.

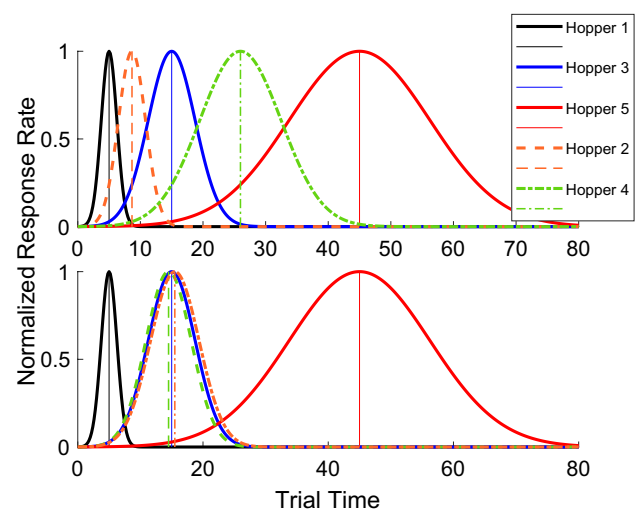


Fig. 1 Graphical illustration of the predictions of metric spatiotemporal mapping/local temporal averaging (top panel) and global temporal averaging (bottom panel). Solid curves represent the expected response pattern for the trained locations while dashed curves represent the expected response pattern for untrained locations during testing

Finally, it is possible that subjects choreograph a time-based response strategy (e.g., timed switching) between previously learned spatiotemporal contingencies (e.g., short → medium → long) while exploring the novel locations en route without attributing any temporal contingencies to them or generalize the previously learned temporal rule across space. In this case, we might observe no temporal control at all or response peaks at trained intervals in novel hoppers. Regarding the spatial generalization under uncertainty account, Malet-Karas et al. (2019) trained rats to different spatiotemporal contingencies (Hopper 2–15 s and Hopper 4–60 s) in different contexts marked by different house light color and tone frequency (A and B, respectively) in a setting with a five-hole nose-poke wall, and later they changed the spatiotemporal rule (Hopper 1–30 s) in context B with no discriminative cue. On the first day of testing, all nose-poke holes were illuminated in context B. Results showed that response rates were high not only in the new location (Hopper 1) but also in the old location (Hopper 4), its neighboring location (Hopper 3 which has never been reinforced) as well as the location associated with context A (Hopper 2). Importantly, they observed that the timing of responses in the non-reinforced locations was controlled by the context suggesting less generalization for time than space.

These possible accounts have different predictions regarding the timing of behavior at novel locations. The first account, which relies on spatiotemporal inference/averaging, predicts that the timing behavior of subjects would cluster around the average of the neighboring intervals, respectively. The second account, which relies on temporal averaging/regression to the mean, predicts that the timing behavior at the novel locations would cluster around the mean of all intervals (i.e., middle interval). What type of averaging would be observed depends on whether response weights are based on (inverse) variance or CV (SD/mean); therefore, peaks in the novel locations can be anywhere in between geometric and arithmetic means. Finally, the third account predicts that the responses on the novel hoppers would lack temporal control or might present bimodality (targeting intervals associated with the neighboring hoppers; see Spence (1937) for a similar reasoning based on excitation gradients). The current study tested these predictions in mice.

Methods

Subjects

Eighteen naive male C57BL/6J mice, bred in Koç University Animal Research Facility, were approximately 10 weeks old at the beginning of the experiment. C57BL/6J mice were tested

since this strain has been shown to robustly exhibit scalar property of interval timing (Buhusi et al. 2009) and male mice were used as previous research does not point to a large sex difference in this behavioral domain when measured with standard tasks (e.g., Buhusi et al. 2017). The initial sample size was determined using the resource equation approach (e.g., Arifin and Zahiruddin 2017), which suggested that 15 mice were sufficient, and three subjects were added considering any possible attrition/exclusion. Mice in groups of three to five were housed in individually ventilated cages (Allentown type I long) in a room on a 12:12 h light/dark cycle. Daily experimental sessions were 1-h long and were run during the light cycle on consecutive days, such that 1-h long sessions were run at the same time of the day. To maintain mice at 85% of their free-feeding weights, a food-restriction regimen was started 3 days prior to the experiment. Their weights were monitored on a daily basis, and the amount of food pellet they received was adjusted accordingly. Access to water was ad libitum in the home cages, and additional food pellets were provided only after the experimental sessions. All the procedures were in compliance with national ethical guidelines for animal research; Koç University Animal Research Local Ethics Committee approved all the procedures (Protocol #: 2016-34).

Apparatus

Five-choice serial reaction-time task boxes (5CSRTT; SOMED-NP5M-B1; Med Associates Inc.) placed in sound-attenuating cubicles (ENV022MD; Med Associates Inc.) were used to conduct the experiment. One of the side walls of the box had five illuminable nose-poke holes, set side by side. Three of these holes (H1, H3, and H5) were associated with different time intervals while the other two holes (H2 and H4) were used only during testing. None of these holes were blocked during training and test phases but illuminated depending on the type of active trial and experimental phase. The opposite wall had an illuminable pellet receptacle (ENV-303W and ENV-303RL, Med Associates Inc.) to deliver sucrose pellets (TestDiet Sucrose Tablets 20 mg) via a pellet dispenser (ENV-203-20, Med Associates Inc.), a houselight (ENV-315W, Med Associates Inc.) above the pellet receptacle, and a speaker (ENV-324W, Med Associates Inc.) for auditory stimulus delivery. Head entries to the nose-poke holes and the pellet receptacle were detected via IR-beam-break detectors. MED-PC IV software was used to control boxes and record event logs and their timestamps with a resolution of 10 ms.

Procedure

Magazine and nose-poke training

First, mice had three sessions of Fixed Time 60 s and Fixed Ratio 1 (FT60s and FR1) training on three consecutive days.

All sessions were started by illumination of the box, and one of the three training holes (H1, H3, or H5) was illuminated to signal food availability upon a nose-poke response at that location (FR1). When there was no response to the illuminated nose-poke hole, food was delivered at the end of the 60 s (FT 60 s). The trials were terminated by the food delivery, and a new trial began. Illumination of the holes was random with equal probability. After three sessions, FT60s trials were dropped, and mice continued FR1 training until they received at least 15 tablets from each hole for two consecutive days.

Phase 1: peak interval procedure (isolated interval–location trainings)

At this stage, mice were trained to learn the relationship between three intervals (5 s, 15 s, and 45 s) and the corresponding locations (H1, H3, and H5) in different training phases (i.e., different sessions/days). The order for interval–location training was random among mice. Durations were chosen based on a pilot study and considering Weber's law to keep the discriminability of the consecutive durations constant. While 15 s interval was always paired with H3, the pairing of 5 s and 45 s with H1 and H5 was counterbalanced between subjects. Within an experimental session, individual trial types were of two kinds: fixed interval (FI) and peak interval (PI) trials. PI trials constituted 1/5 of the total number of trials. All trials started with the illumination of the box (i.e., houselight) and the hole associated with the target interval, and the presentation of white noise. In FI trials, a reward was delivered for a nose-poke made into the illuminated hole at or after its associated time; the reward was delivered by illumination of the food receptacle and all other stimuli were turned off upon reward delivery. In PI trials, illumination of the associated hole was accompanied by houselight and white noise for a duration that was three times the target time plus a random number between 1 and 5 s (uniformly distributed with increments of 1 s); no reward was delivered upon any response. If there were no responses in an FI trial, trial termination was set to a time calculated as PI trial time. Average intertrial interval (ITI) was 30 s. Inter-trial interval was a right-shifted (by 15 s) exponentially distributed random variable with a mean of 15 s. This phase lasted 15 sessions for 5 s and 15 s, and 20 sessions for 45 s.

Phase 2: intermixed FI sessions

After the completion of the PI procedure separately for each interval–location pairing, FI trials for each interval–location pairing were presented within a session in random order and with equal probability. ITI was set as described in the previous phase. This phase lasted for 10 sessions.

Phase 3: intermixed peak interval procedure

This phase was an intermixed version of the PI procedure for three interval–location pairings. First, one of three nose-poke locations was chosen randomly, and then, an FI or PI trial for the corresponding interval–location pairing was initiated as described in the PI procedure section. The ratio of FI trials to PI trials per pairing was 2:1. ITI was the same as in the previous phases. This phase lasted for 39 sessions on average (38–41 sessions).

Phase 4: intermixed peak interval procedure with test trials

In the test phase, there were FI trials of three interval–location pairings and test trials similar to PI trials. Different from the previous phase(s), all trials, irrespective of the type of trial, started with the illumination of all five nose-poke holes accompanied by white noise and houselight such that there was no discriminative cue signaling the type of trial. In each 1-h-long daily session, 6 trials out of 9 were FI trials for 5, 15, or 45 s (2 of each). Therefore, although not signaled, mice collected a reward if they responded to the active nose-poke hole at or after the time associated with that nose-poke hole in FI trials. In test trials, trials were terminated at the end of 135 s (i.e., longest target interval $\times 3$) without a reward irrespective of the time or location of the responses. ITI was the same as the previous phase. This phase lasted for 15 sessions.

Data analysis

Response times and their frequency in PI or test trials were extracted for each phase (excluding Phase 2, which had no PI trials). Our primary measurement for the statistical comparisons was the peak time. Peak times were calculated separately per schedule (interval–location pairings) or test locations and determined using the average response curves. In each curve, the time that corresponded to the highest response frequency was marked as the peak time (e.g., Balci et al. 2009). Three subjects were excluded prior to the data analysis because they had peak times longer than 90 s ($2 \times$ longest FI schedule). For the first and third phases, we analyzed the data from the last 5 sessions to show that mice learned/differentiated interval–location pairings prior to the testing of novel locations. For the test phase, we analyzed the first five sessions (unless indicated otherwise) to examine when the time of the reward anticipation was highest in the test locations that had never been paired with an interval relative to the trained locations during test trials. For the training phases, we compared the peak times of PI trials among three schedules (associated with 5, 15, and 45 s). For the test phase, we compared the peak times in test trials among five locations. Friedman test was used for the comparisons

due to the violation of normality assumption (tested using Shapiro–Wilk test of normality), Kendall’s W test was used to provide an estimate of effect size, and Holm–Bonferroni corrected p values are reported for the pairwise comparisons (Durbin–Conover).

In order to characterize whether mice exhibited a specific pattern of responding across space as a function of time during the test trials, we computed the transition matrix for directional switches between different hoppers (not considering the stays at a hopper). If the peak times were artifacts of mice visiting the novel hoppers as they were transitioning between the previously reinforced hoppers during the test trials (part of the third account), then one would expect a higher transitioning probability between the neighboring hoppers in a direction that is dictated by the spatiotemporal associations (i.e., moving from the hoppers associated with shorter interval toward neighboring hoppers in the direction of the longer interval).

Finally, to elaborate on what kind of averaging (arithmetic vs. geometric) was in effect, peak times at novel locations were regressed on arithmetic and geometric averages of the peak times of the neighboring locations. In addition, normalized variability (as a proxy of CV) estimates were calculated using the spreads and peak times derived from the normalized average curves (as in Balci et al. 2009), and comparison of normalized variability estimates among five locations was made using Repeated-Measures ANOVA. Data processing and statistical analysis were done using Matlab (R2020a) and Jamovi (1.8.1).

Results

The comparison of peak times among three schedules for the initial training phase in which mice were trained for each schedule in different sessions revealed a significant result, $\chi^2(2) = 30.00$, $p < 0.001$, $W = 1.00$. Median peak times were 7 s (IQR = 0.50), 16 s (IQR = 8), and 46 s (IQR = 18.50) for PI5, PI15, and PI45 trials, respectively (p values for pairwise comparisons < 0.001). The significant differences observed among peak times were still present by the end of intermixed training (phase 3, last five sessions, $\chi^2(2) = 29.50$, $p < 0.001$, $W = 0.98$). Median peak times were 7 s (IQR = 3.50), 16 s (IQR = 8), 37 s (IQR = 20), respectively (p values for pairwise comparisons < 0.001). Figure 2 shows normalized average response curves for PI5, PI15, and PI45 trials for isolated and intermixed interval training phases. These results show that subjects learned the interval–location associations during the first phase, and mixing schedules within a session did not vastly disturb their temporal expectations per response location.

For the test phase, we compared the peak times among five adjacent response locations in the test trials to see how

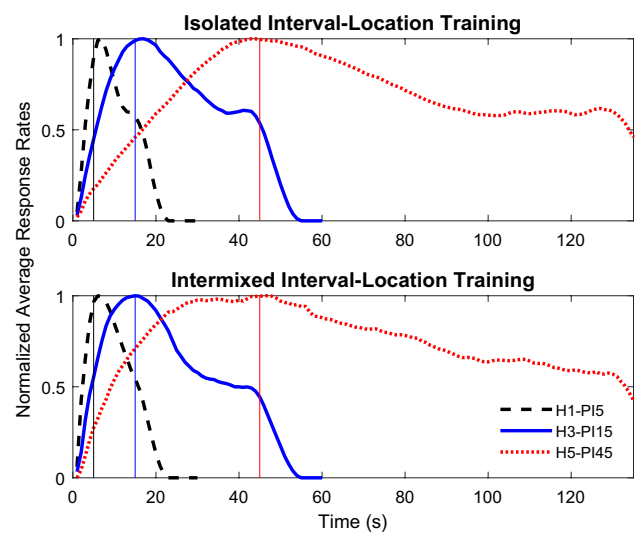


Fig. 2 Normalized average response curves for PI5 (H1), PI15 (H3), and PI45 (H5) trials in isolated interval–location training (phase 1, top panel) and intermixed interval–location training (in phase 3, bottom panel). Vertical lines show trained schedules

temporal expectations were shaped in the novel locations introduced during the test (H2 and H4) in comparison to the locations previously associated with different intervals (H1–PI5, H3–PI15, H5–PI45). Figure 3 shows normalized average response curves for each location during test trials. Friedman test revealed a significant result, $\chi^2(4) = 50.20$, $p < 0.001$, $W = 0.84$. Median peak times were 6 s (IQR = 2), 13 s (IQR = 10), 18 s (IQR = 8), 34 s (IQR = 18.50), and 45 s (IQR = 21.50) for H1, H2, H3, H4, and H5, respectively. All pairwise comparisons were significant after correction ($p_{H2vsH3} = 0.04$, $p_{H4vsH5} = 0.008$, for all other comparisons $p < 0.001$) indicating that mice had distinct temporal expectations for the novel locations. Further analyses showed that these differences disappeared with further testing (see Supplementary Online Material, SOM Fig. 1 and SOM Table 1).

We computed a directional transition index by getting the sum of transition probabilities for transitions at short-to-long direction and compared this index to the chance level using a one-sample t -test. All transition probabilities are shown in Fig. 4. Proportion of transitions from short-to-long locations ($M = 0.54$, $SEM = 0.02$) was marginally higher than the chance level ($= 0.50$), $t(14) = 1.76$, $p = 0.05$ (one-tailed), $d = 0.45$. Finally, in order to directly test whether the temporal characteristics of anticipatory responses at the novel hoppers are an artifact of the animals stopping by these locations while switching between the previously reinforced hoppers, we compared the transition probability for $H1 \Rightarrow H2$ vs $H3 \Rightarrow H2$ and $H3 \Rightarrow H4$ vs $H5 \Rightarrow H4$. Results showed that there was no significant difference between the transition probabilities for either pair; $t(14) = 0.08$, $p = 0.94$ and $t(14) = -0.14$, $p = 0.89$, respectively (Wilcoxon signed ranks

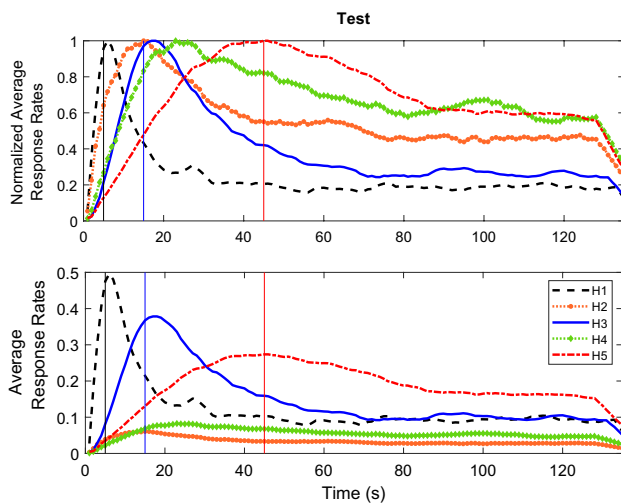


Fig. 3 Normalized and absolute average response curves for each location in test trials are shown in top and bottom panels, respectively. Vertical lines for the locations active during training mark the trained schedules. H1 is the location paired with 5 s and H5 is the location paired with 45 s

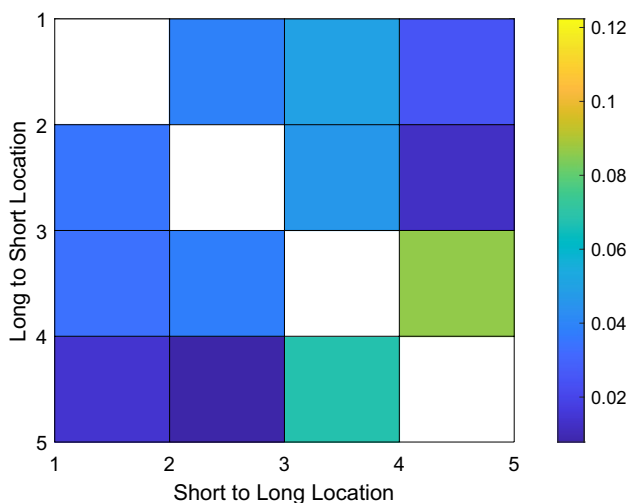


Fig. 4 Probability of possible transitions between hoppers in both directions during test trials

test led to the same results). According to these results, any temporal artifact due to animals switching between locations should have appeared at two points during the trial: for H2 once between 5 and 15 s and once between 45 and 15 s, and for H4 once between 15 and 45 s and once after 45 s. This was not observed in the data despite the heavier tail of the response curves for the novel locations.

Since our observation pointed out a local averaging strategy, regression analyses were run to see what averaging strategy was in effect. Arithmetic and geometric averages of the peak times of H1 and H3 pair and H3 and H5 pair

were calculated for the prediction of the peak times of H2 and H4, respectively. The arithmetic mean of the durations associated with neighboring hoppers was not a significant predictor of peak time for H2 [$\beta = 0.46$ (0.28), $p = 0.12$] or for H4 [$\beta = 0.49$ (0.27), $p = 0.089$]. When the analyses were run with the geometric means we found $\beta = 0.62$ (0.32), $p = 0.077$ for H2 and $\beta = 0.63$ (0.26), $p = 0.03$ for H4. These results suggest that geometric mean was a better predictor of the peak times at novel locations (only significant for H4, and a trend for H2 with slopes closer to unity line compared to arithmetic mean). SOM Fig. 2 shows the data and regression slopes gathered from these analyses. Finally, the normalized variability estimates among five locations were not significantly different from each other $F(4, 56) = 1.46$, $p = 0.23$.

Discussion

Different quantities have been claimed to be underlain by a common magnitude mechanism (Walsh 2003). To this end, for instance, researchers found that numerical magnitudes affect time estimates (Karşilar and Balci 2019), judgements can be transferred from duration to numerosity (Balci and Gallistel 2006; Meck and Church 1983 in rats), numerical magnitudes are associated with different sides in space (Dehaene et al. 1993; see Rugani et al. 2015 for similar results in baby chicks), and temporal and spatial information interact when a task demands it (Cheng et al. 1996). But related theoretical approaches typically do not specify how the metric properties of one domain can be spontaneously translated into metric estimates about another domain. To this end, Meck and Church (1983) trained rats to categorize durations and then induced them to transfer these judgments to numerosities. They concluded that each counted number corresponded to 200 ms for their task parameters. Balci and Gallistel (2006) tested the same question in humans by characterizing the nature of mapping from duration discrimination to numerosity discrimination. Critically, they varied the range of numerosities as well as the proportions between the low- and high-numerosity references and showed that duration-to-numerosity judgments were based on the cross-domain alignment/comparability of within-domain proportions. The current study investigated whether mice can make novel inferences about time intervals using spatial locations based on previously learned spatiotemporal associations. To this end, mice first learned the association between three locations (far left, middle, and far right) on the operant wall with three different intervals in ascending order, and their timing behavior was tested at five locations including the intermediate two locations between far left and middle and between middle and far right.

A metric mapping between spatial and temporal domains predicted that the peak of timed responses at novel locations would be located between the peak times observed at the previously reinforced neighboring locations. This can be manifested simply by local temporal averaging (Fig. 1, top panel). Alternatively, if temporal averaging would take place without taking account of locations but only previously experienced time intervals, the central tendency effect would predict peak times at novel locations to be similar to the peak location previously observed in the middle location (Fig. 1, bottom panel). The last alternative is that mice responded at novel locations as they were transitioning between previously reinforced locations in one direction (i.e., from short to long) such that peak times at the novel locations would exhibit bimodality or their responses in novel locations would lack temporal control. For example, Malet-Karas et al. (2019) showed that when different contexts are associated with different spatiotemporal contingencies and tested in a single context (i.e., context A in Malet-Karas et al. 2019), the temporal pattern of responses in non-reinforced locations was controlled by the context, and there was spatial generalization. In their experimental protocol, it would be more difficult to use isomorphic relationships between time and space compared to an experimental scenario with multiple anchors (as in the current task); thus, spatial generalization would be more likely compared to our case.

Our results suggest that when faced with uncertainty, mice parameterized their timing behavior at novel locations such that they targeted in between the intervals associated with the neighboring hoppers. As mentioned above, mice might have made temporal inferences about novel locations based on learned spatiotemporal associations that immediately surround a novel location. This could result from the cross-domain transfer of quantitative spatial judgments to temporal judgments, which has been shown earlier between different quantitative domains in humans (e.g., Balci and Gallistel 2006) and non-human animals (e.g., Meck and Church 1983; Rugani et al. 2015). Alternatively, mice might have simply targeted the average of the time intervals that are associated with the immediately neighboring locations (local averaging). To this end, our data might be pointing at a local manifestation of the central tendency effect observed when mice have no prior experience regarding the temporal correlates of the test locations (e.g., under uncertainty regarding spatiotemporal contingency).

In case of local averaging, we predicted that if response weights are based on inverse variances, one would expect to observe responding closer to the geometric mean since the relative weight of the longer interval would be lower under scalar property. If the response weights are based on inverse CV (assuming scalar property), then responding would be closer to the arithmetic mean. Prediction of the peak times

for the novel locations using the arithmetic vs. geometric averages of the peak times of the neighboring location revealed that geometric averages were better at predicting the peak times for the novel locations.

A similar pattern of responding (i.e., temporal averaging) was observed in previous work that trained animals to associate two different intervals with two different stimuli and tested them under a compound stimulus (Swanton et al. 2009; see Gür et al. 2021 for similar results gathered in a counting task). In temporal/numerical averaging studies, uncertainty is caused by ambiguity in the temporal/numerical associations of the compound stimulus. These studies showed that animals' timing behavior is characterized by the Bayesian average of the previously experienced intervals. Importantly, these studies showed that response distributions were scalar in compound-cue trials (under specific conditions, see Matell and Kurti 2014; Swanton and Matell 2011), as would be predicted by an averaging explanation, supporting the conclusion that temporally controlled responses in compound trials were governed by a single memory computed as the average of distinct temporal memories associated with each cue. In our study, the fact that we did not find a significant difference among the normalized variability estimates (as a proxy of CV) of five response locations also supports the conclusion that mice relied on averages from temporal memory when responding at novel locations.

Under this rationale, these results point to the ubiquity of Bayesian inference in the magnitude domains that are subject to representational precision limits (e.g., Jazayeri and Shadlen 2010; Petzschner et al. 2015; Shi et al. 2013). Our results also show that Bayesian inference can operate based on local contingencies. This is consistent with earlier data that showed that mice (but not humans) can do local computations of probabilities (Çoşkun et al. 2015; see also Wu et al. 2009 for deviations from local computations in humans). Çoşkun et al. (2015) tested humans and mice in a task that had three locations each of which was associated with a different interval (ascending order from left to right or vice versa, counterbalanced). All three options were always available but in a given trial only one of these options would result in reward delivery after the associated time interval. The probability of each option being active in a trial was manipulated. Çoşkun et al. (2015) found that the timing of switching from the first and second options was adaptively modulated by the relative frequency by which those options paid off whereas the timing of switching from the second option to the third option was adaptively modulated by the relative frequency only for mice.

Finally, mice might have simply stopped and poked at the novel location as they were traveling between locations associated with rewards at different delays. This could result in a number of possible patterns of responding such as loss of temporal control, bimodal distribution with peaks at

previously trained targets at novel locations or even local averaging under certain conditions. The analysis of the transition probabilities between the hoppers of interest refutes the latter possibility: the probabilities of mice transitioning to novel hoppers were equal in both directions. If the temporal modulation of responding was an artifact of such bidirectional movement patterns, one would expect bimodality in response curves, which is not seen in the data. Thus, our results support the temporal inference based on spatiotemporal relations/local averaging.

Although we observed clear peaks for each trained interval–location pair prior to and during the final test, some might argue that the bumps with lower amplitudes on the tails of the response curve during initial training sessions (but not the test session) were because the learning of the location–interval pairs was not fully achieved, and mice could have benefited from longer training. We trained mice overall for 94 sessions for 3 interval–location pairs prior to testing, which can be considered long enough for peak interval procedure with three intervals (e.g., Buhusi et al. 2009). Moreover, previous studies using peak interval procedure (or similar tasks) suggested that animals learn temporal characteristics of the task in an abrupt fashion early in training while increase in response magnitude (i.e., amplitude of response curves) and emergence of controlled stops are observed with further training (Balci et al. 2009; Drew et al. 2005; Gür et al. 2018). In our experiment, even though intermediate (novel) locations have never been paired with a reward, temporal properties attributed by mice to the novel locations show that despite the low response rates at these locations, temporal expectations were spontaneously formed without any differential training for these locations. Although the intermediate-hopper nose pokes had temporal control (derived from averaging) over them, the amplitudes of response curves suggest a low expectancy of reward at these hoppers. This could arise partially because these nose pokes could be exploratory and/or due to extinction effects.

A common magnitude system is now a well-accepted approach to the processing of time, numerosities, distances, etc. in the brain (Walsh 2003). One of the assumptions of these approaches is that information processing in one domain can influence magnitude estimates in another domain and/or different magnitudes can be expressed on a common metric allowing cross-domain transfer of judgements. From different magnitudes, time and space are inevitably processed together in nature. However, most of these accounts fall short of specifying the exact nature of cross-domain information exchange. Our results suggest that this information exchange might rely on a flexible and inferred isomorphic relationship between time and space. This flexibility might be powered by the unit-free nature of information exchange (e.g., mapping different domains

based on within-domain proportions as suggested by Balci and Gallistel 2006). Briefly, our results contribute to the nature of cross-domain interactions assumed by the common magnitude system.

The current study shows that mice make temporal and/or Bayesian inferences locally based on their previous experiences with time intervals when they respond at a location with no prior contingency with a time interval. It is important to note that our analysis approach was based only on average response curves. Analysis incorporating a single-trial approach could have provided more detailed information for the interpretation of our current observations. However, the relatively lower levels of responses in the novel locations (Fig. 3, bottom panel) did not render this approach possible. Future studies should test the timing behavior at the novel locations by only signaling those locations individually to minimize the effect of timing behaviors at previously reinforced locations (e.g., better controlling for the third account).

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10071-022-01715-4>.

Author contributions EG: conceptualization, methodology, software, formal analysis, investigation, writing—original draft and review and editing, and visualization. YAD: conceptualization, methodology, software, investigation, and writing—review and editing. FB: conceptualization, methodology, formal analysis, writing—original draft and review and editing, visualization, project administration, and funding acquisition.

Funding This research was supported by a grant from the Scientific and Technological Research Council of Turkey (TÜBİTAK) to FB [Grant number: 117K370]. EG was supported by TÜBİTAK through the National Scholarship Program for Ph.D. students (BİDEB 2211E).

Availability of data and materials Available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethics approval All the procedures performed in this study were in accordance with the ethical standards of the institutional and/or national research committee. The study was approved by the Koç University Animal Research Local Ethics Committee (Protocol #: 2016-34).

References

- Arifin WN, Zahiruddin WM (2017) Sample size calculation in animal studies using resource equation approach. *Malays J Med Sci*. <https://doi.org/10.21315/mjms2017.24.5.11>

- Balcı F, Gallistel CR (2006) Cross-domain transfer of quantitative discriminations: Is it all a matter of proportion? *Psychon Bull Rev*. <https://doi.org/10.3758/BF03193974>
- Balcı F, Papachristos EB, Gallistel CR, Brunner D, Gibson J, Shumyatsky GP (2008) Interval timing in genetically modified mice: a simple paradigm. *Genes Brain Behav*. <https://doi.org/10.1111/j.1601-183X.2007.00348.x>
- Balcı F, Gallistel CR, Allen BD, Frank KM, Gibson JM, Brunner D (2009) Acquisition of peak responding: What is learned? *Behav Processes*. <https://doi.org/10.1016/j.beproc.2008.09.010>
- Buhusi CV, Aziz D, Winslow D, Carter RE, Swearingen JE, Buhusi MC (2009) Interval timing accuracy and scalar timing in C57BL/6 mice. *Behav Neurosci*. <https://doi.org/10.1037/a0017106>
- Buhusi M, Bartlett MJ, Buhusi C (2017) Sex differences in interval timing and attention to time in C57Bl/6J mice. *Behav Brain Res*. <https://doi.org/10.1016/j.bbr.2017.02.023>
- Buhusi CV, Reyes MB, Gathers CA, Oprisan SA, Buhusi M (2018) Inactivation of the medial-prefrontal cortex impairs interval timing precision, but not timing accuracy or scalar timing in a peak-interval procedure in rats. *Front Integr Neurosci*. <https://doi.org/10.3389/fnint.2018.00020>
- Cheng K, Spetch ML, Miceli P (1996) Averaging temporal duration and spatial position. *J Exp Psychol Anim Behav Process*. <https://doi.org/10.1037/0097-7403.22.2.175>
- Çoşkun F, Berkay D, Sayalı ZC, Balcı F (2015) Sequential temporal discrimination in humans and mice. *Int J Comp Psychol*. <https://doi.org/10.46867/ijcp.2015.28.02.08>
- De Corte BJ, Matell MS (2016) Interval timing, temporal averaging, and cue integration. *Curr Opin Behav Sci*. <https://doi.org/10.1016/j.cobeha.2016.02.004>
- De Corte BJ, Navarro VM, Wasserman EA (2017) Non-cortical magnitude coding of space and time by pigeons. *Curr Biol*. <https://doi.org/10.1016/j.cub.2017.10.029>
- Dehaene S, Bossini S, Giraux P (1993) The mental representation of parity and number magnitude. *J Exp Psychol Gen*. <https://doi.org/10.1037/0096-3445.122.3.371>
- Drew MR, Zupan B, Cooke A, Couvillon PA, Balsam PD (2005) Temporal control of conditioned responding in goldfish. *J Exp Psychol Anim Behav Process*. <https://doi.org/10.1037/0097-7403.31.1.31>
- Gür E, Duyan YA, Balcı F (2018) Spontaneous integration of temporal information: implications for representational/computational capacity of animals. *Anim Cogn*. <https://doi.org/10.1007/s10071-017-1137-z>
- Gür E, Duyan YA, Balcı F (2021) Numerical averaging in mice. *Anim Cogn*. <https://doi.org/10.1007/s10071-020-01444-6>
- Henik A, Tzelgov J (1982) Is three greater than five: The relation between physical and semantic size in comparison tasks. *Mem Cognit*. <https://doi.org/10.3758/BF03202431>
- Jazayeri M, Shadlen MN (2010) Temporal context calibrates interval timing. *Nat Neurosci*. <https://doi.org/10.1038/nn.2590>
- Karşılar H, Balcı F (2019) Symbolism overshadows the effect of physical size in supra-second temporal illusions. *Atten Percept Psychophys*. <https://doi.org/10.3758/s13414-019-01748-x>
- Light KR, Cotten B, Malekan T, Dewil S, Bailey MR, Gallistel CR, Balsam PD (2019) Evidence for a mixed timing and counting strategy in mice performing a Mechner counting task. *Front Behav Neurosci*. <https://doi.org/10.3389/fnbeh.2019.00109>
- Malet-Karas A, Noulhiane M, Doyère V (2019) Dynamics of spatio-temporal binding in rats. *Timing Time Percept*. <https://doi.org/10.1163/22134468-20181124>
- Matell MS, Kurti AN (2014) Reinforcement probability modulates temporal memory selection and integration processes. *Acta Psychol*. <https://doi.org/10.1016/j.actpsy.2013.06.006>
- Meck WH, Church RM (1983) A mode control model of counting and timing processes. *J Exp Psychol Anim Behav Process*. <https://doi.org/10.1037/0097-7403.9.3.320>
- Merritt DJ, Casasanto D, Brannon EM (2010) Do monkeys think in metaphors? Representations of space and time in monkeys and humans. *Cognition*. <https://doi.org/10.1016/j.cognition.2010.08.011>
- O'Keefe J, Burgess N (1996) Geometric determinants of the place fields of hippocampal neurons. *Nature*. <https://doi.org/10.1038/381425a0>
- Petzschner FH, Glasauer S, Stephan KE (2015) A Bayesian perspective on magnitude estimation. *Trends Cogn Sci*. <https://doi.org/10.1016/j.tics.2015.03.002>
- Rugani R, Vallortigara G, Priftis K, Regolin L (2015) Number-space mapping in the newborn chick resembles humans' mental number line. *Science* 347(6221):534–536. <https://doi.org/10.1126/science.aal1379>
- Rugani R, Vallortigara G, Priftis K, Regolin L (2020) Numerical magnitude, rather than individual bias, explains spatial numerical association in newborn chicks. *Elife*. <https://doi.org/10.7554/eLife.54662>
- Shi Z, Church RM, Meck WH (2013) Bayesian optimization of time perception. *Trends Cogn Sci*. <https://doi.org/10.1016/j.tics.2013.09.009>
- Spence KW (1937) The differential response in animals to stimuli varying within a single dimension. *Psychol Rev*. <https://doi.org/10.1037/h0062885>
- Swanton DN, Matell MS (2011) Stimulus compounding in interval timing: The modality–duration relationship of the anchor durations results in qualitatively different response patterns to the compound cue. *J Exp Psychol Anim Behav Process*. <https://doi.org/10.1037/a0020200>
- Swanton DN, Gooch CM, Matell MS (2009) Averaging of temporal memories by rats. *J Exp Psychol Anim Behav Process*. <https://doi.org/10.1037/a0014021>
- Tosun T, Gür E, Balcı F (2016) Mice plan decision strategies based on previously learned time intervals, locations, and probabilities. *Proc Natl Acad Sci U S A*. <https://doi.org/10.1073/pnas.1518316113>
- Walsh V (2003) A theory of magnitude: common cortical metrics of time, space and quantity. *Trends Cogn Sci*. <https://doi.org/10.1016/j.tics.2003.09.002>
- Wu SW, Dal Martello MF, Maloney LT (2009) Sub-optimal allocation of time in sequential movements. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0008228>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.