TITLE

Harmonization of the fastest and densest responses embodies human-like reaction time of mice

AUTHOR

Chan Hee Kim[1, *]

[1]Department of Physiology and Dental Research Institute, Seoul National University School of Dentistry, Seoul, 03080, Republic of Korea.

* chanypia@snu.ac.kr

ABSTRACT

Reaction time (RT) is a crucial component for evaluating delayed lstency in behavior. Animals' RT, in which the stimulus-response relationship is not one-to-one due to repeated responses per trial, unlike humans, may not be a single latency but may involve two peak latencies of the fastest and densest responses within distributed responses. We investigated whether the two peak latencies are aligned for a single RT by handling stimulus duration. In delay conditioning with C57BL/6 mice using sound cues of 10, 5, and 2 s, the 2-s group only showed the strongest positive correlations between the two peaks, alongside HE frequency and accuracy rate. This suggests coupling of fastest and densest responses, and a one-to-one relation of stimulus and response, respectively. We propose that the harmonization between the two peaks, elicited by stimuli prompting minimal responses, serves as a criterion for designing animal experiments to mimic human-like RT.

KEYWORDS

Response time; Head entry peak; Fastest head entry peak; Stimulus duration; Latency; Frequency; Mice; Behavior; Delay conditioning

INTRODUCTION

Timing is fundamental for exploring behavioral response[1]. Reaction time (RT), also known as response time[2-5], is a crucial component in behavior analysis, serving as a marker reflecting individual subjects' backgrounds[6], sensory and cognitive processes[7-9] and clinical perspective[10,11]. RT also is related to neural correlation involving diverse brain regions[12-14]. Terms like reaction speed[15,16] and response latency[17,18] are used interchangeably with RT in human and animal studies. However, there are subtle methodological differences between the two due to their distinct behavioral characteristics, even when directly comparing RTs between animals and humans[19-21]. Even if there are comparable aspects between human and animal behaviors under varying conditions, there may still be debatable regarding whether the results can be interpreted similarly. The present study proposes theoretical criteria for extracting human-like RT from animals' data.

RT is typically assessed within a specific time limit for each trial[22], accompanied by the anticipation of its time point[23,24]. Clear instructions provided by researchers to human subjects before experiments are crucial for RT sensitivity. Human subjects are occasionally asked to provide an as-soon-as-possible response[25,26]. However, instructing animals to respond quickly may be impractical due to communication limitations. Even with clear instruction for animal subject in experimental paradigm involving long training period, animals may delay their responses intentionally. Additionally, their behavior during a trial, with repeated actions such as licking, nose poking, and lever pressing, complicates the interpretation of RT-like measure. In animal studies, the relation between stimulus and response may not be one-to-one, unlike in human studies. It is uncertain whether the delayed latencies of repeated responses accurately reflect the animals' decision-making points in cognitive process. The averaged animals' responses may include error responses where the intention cannot be accurately determined. Comparing an averaged latency in animals with RT in human studies directly would be challenging. Therefore, the frequency of repeated responses, such as head entries (HEs) toward food cup, is essential for evaluating the behavioral performance of animals[27-29]. During the training period, repeated HEs for repeated stimulations gradually represent the densest point, forming the peak of HE in its distribution, which reflects subjects' behavioral habituation and sensitization[30-34]. However, beyond the HE peak latency, the fastest HE (FHE) and its peak also can exist within the distribution of HEs (Figure 1). Time points of RTs in animal data would fundamentally be not a single latency as a subset of both components. The coexistence of distinct two components may raise an issue that needs to be identified a real RT in the data. Considering one-to-one relationship like that in humans, the FHE peak of the first HE after stimulus onset is adequate. However, considering the reliability based on higher frequency, the HE peak is better. If FHE and HE peaks can be aligned in a single peak, it indicates a coupling of the fastest and densest responses. This is an ideal option closer to humans' RT.

Temporal aspects[35,36] involved in the experimental paradigm, such as the inter-stimulus interval, inter-trial interval, interval between conditioned stimulus (CS) and unconditioned stimulus (US), and CS duration, individually and collectively influence the response. In this study, particular emphasis is placed on investigating the impact of CS duration on determining the CS-US interval in delay conditioning. In classical conditioning[37], the interval between the stimulus and the reward determines the time within which animals are permitted to respond. The US of food following the CS onset of stimulus cue guides the HEs. Some animals can delay their actions after decision-making during long CS-US intervals. A short CS-US interval would be better for achieving more prompt response from animals, while, a long CS-US interval would lead to acquire more responses safely without a missing response. A higher level of frequency on delayed latencies may indicate animals' behavior resembling food addiction[38] as well as their deep awareness of task[39]. Latency and frequency have probably greater distribution variance in longer CS duration than in short CS duration, reflecting individual differences such as animals' preferences, hesitation, and laziness of subjects, deviating from human-like RT. Indeed, numerous studies have examined the effect and importance of the CS-US interval (or CS duration) based on the tests using diverse CS durations[27-29,35,40-44]. However, the specific durations of the stimulus depend on the apparatus and study paradigm, and an absolute value cannot be a criterion for determining the appropriate duration for designing an experiment.

RT can be influenced by various conditions within the experimental paradigm. Our first goal was to identify a temporal factor in experimental design that can consolidate the two different latencies of FHE and HE peaks into one cohesive unit of RT, rather than treating them as independent entities. We though that the factor was CS duration. To explore this, we conducted experiments using separate groups of C57BL/6 mice (8 male mice for each group) with sound cues of different CS durations (10 s, 5 s, and 2 s) (Figure 1). We hypothesized that the latency and frequency between FHE and HE peaks would show different patterns depending on various CS durations. If FHE and HE peak latencies were synchronized, their frequencies would also align into one peak, as the relationship between latency and frequency may be interdependent. Furthermore, we hypothesized that the optimal CS duration could be determined based on the relationship between these two peaks. For this, we recorded the HEs of freely moving mice in operant chamber based on delay conditioning[45], employing white noise and condensed milk. To eliminate additional factors except for training food-reward associations, and to observe mice' natural behaviors throughout the duration of sound cue, we did not incorporate no-reward or punishment for missed responses within the experimental paradigm. Delay conditioning, rather than trace conditioning related to memory-based process[46], was employed to induce the response depending on stimulus duration without an interval of no sound. In data analysis, to evaluate the impact of CS duration comprehensively, we conducted a detailed analysis of behavioral data, which can be examined at both temporal and frequency levels. Our analysis

proceeded in several steps: First, we assessed the difference and correlation between FHE and HE peaks to understand their relationship. Next, we examined group differences in FHE and HE peaks independently to understand their characteristics and implications. Third, we reviewed classical components such as HE frequency during the CS and Pre-CS periods, as well as HE accuracy, to confirm the effect of CS duration on subjects' performance. We also considered how shorter CS durations might affect mice's ability to reach the food cup compared to longer CS durations, and how CS duration could influence HE frequency. Based on these analyses, we investigated how subjects' performance was altered according to the relationship between FHE and HE peaks. Finally, we tested the correlation between the number of HEs for each trial and accuracy across each group to assess if the relationship between HE responses and correctness reflected human-like behavior. We ultimately proposed a criterion for determining an appropriate CS duration for eliciting a human-like RT based on the alignment of FHE and HE peaks.

RESULTS

Difference between FHE and HE peaks for each group

Nonparametric paired t-test results for all training days revealed differences between FHE and HE peaks solely in the 10-s group, both in latency and frequency (Figure 2). The significance of these results varied across training days (D). Specifically, for latency, HE peaks exceeded FHE peaks on D5, D7, D8, D9, and D11 (see Figure 2A; Wilcoxon Signed Ranks test, $P < 0.05$ in all cases; detailed statistical results in Table S1). Likewise, for frequency, HE peaks surpassed FHE peaks on D5 and D7 through D11 (see Figure 2B; Wilcoxon Signed Ranks test, $P < 0.05$ in all cases; Table S2). However, no differences were observed between FHE and HE peaks in both latency and frequency for the 5-s and 2-s groups ($P > 0.05$ in all cases).

Correlation between FHE and HE peaks for each group

We performed a Spearman test to assess the correlation between FHE and HE peaks, based on the nonparametric paired t-test results. As expected, latencies of FHE and HE peaks were strongly and positively correlated during the later training days (D4 – D6 and D8 – D12) for the 2-s group. Similarly, for the 5-s group, positive correlations were observed on D1, D4, D8, D10, D11, and D12. For the 10-s group, D1 and D2 were only significant (Figure 3B and Table S4). Regarding the frequencies of FHE and HE peaks, strong positive correlations were found in the 2-s group across all 12 days, except for D3 (Figure 3A and Table S3). Significant positive correlations were also observed in the 5-s on all days except for D1, D8, and D9, however, in the 10-s group, those were observed only in the early training period of D3 and D5, with a notable negative correlation noted on D1. The correlation results refer that FHE and HE peaks were similar in groups with short and shorter CS

durations, but not the group with the long CS duration. Particularly, the two peaks were almost overlapping in the 2-s group.

Group difference in FHE and HE peaks

In the results of one-way analysis of variances (ANOVAs) with the Group factor (10 s, 5 s, and 2 s) using Kruskal-Wallis's test, FHE peaks showed a significant Group effect across the training days for both latency (in cases of D1, D2, D5, D7, D10, D11, and D12) and frequency (in cases of D2, D3, D4, D5, D7 D9, D10, and D12) (Figure 4A). Post-hoc analysis revealed significant differences mostly between the 10-s and 2-s groups or between the 10-s and 5-s groups for both latency and frequency (Mann-Whitney test, Bonferroni corrected $P < 0.0167$; Table S5 and S6).

Meanwhile, one-way ANOVA results for HE peaks revealed a pattern where the significant Group effect for latency on D5 – D12 seemed to follow the significances observed in frequency on D1 – D4 (Figure 4B). In post-hoc analysis for frequency, group differences were significantly highest in the 2-s group compared to other groups in the early training periods of D1 to D4 (Mann-Whitney test, Bonferroni corrected $P < 0.0167$; Table S7). In post-hoc analysis for latency, group differences were mostly observed between the 10-s group and other groups during training periods from D5 to D12 (Mann-Whitney test, Bonferroni corrected $P < 0.0167$; Table S8). The 10-s group's latency was significantly longer compared to other groups across D5 to D12, and differences between the 5-s and 2-s groups were also observed in D5 and D11.

Given the results of FHE and HE peaks, group differences in FHE peaks seemed to appear sporadically and irregularly across training days, while group differences in HE peaks seemed to follow a more consistent pattern. We inferred that the results in HE peak reflected the adaptation to the stimulus after the early training days of D1 to D4, as differences between the longest and shortest CS groups in HE peak latency remained consistent from D5 to D12. Our data indicate that HE peaks explain the effect of CS duration better than FHE peaks.

HE accuracy

HE accuracy was considered as correct if the mice's head touched down in the food cup within the CS duration of 10 s, 5 s, and 2 s. In one-way ANOVAs' results using Kruskal-Wallis's test, Significant group differences in HE accuracy were observed on D7, D8, and D11 (Figure 5A). Post-hoc analysis revealed variations between the 5-s and 2-s groups on D7, and between the 10-s and 2-s groups on D8 and D11 (Mann-Whitney test, Bonferroni corrected $P < 0.0167$; Table S11). However, no significant difference was found on the final day of D12. Overall, subjects achieved high accuracy rates exceeding 80%, regardless of CS duration.

HE frequency

HE frequency was computed for both CS and Pre-CS periods, with HEs normalized per 1 s for the three groups. In the CS period, significant group differences were observed on seven days (D3 – D7, D11, and D12) (Figure 5B and Table S12). Post-hoc analysis revealed that the 2-s group had significantly higher frequencies than the 5-s group. Moreover, the 10-s group showed higher frequencies than the 2-s group only on D3 – D7. In the late training period, the significant difference was observed only between the 5-s and 2-s groups on D11 and D12 (Mann-Whitney test, Bonferroni corrected $P < 0.0166$). HE frequency decreased relatively in the 5-s group.

For the Pre-CS period, one-way ANOVAs revealed significant group effects across eight days except for D3, D4, D5, and D11 (Figure 5B and Table S14). Post-hoc analysis showed higher frequencies in the 2-s group compared to the 10-s group. Differences between the 10-s and 5-s groups were observed on D1 and D9, while differences between the 5-s and 2-s groups were observed on D2, D6, D7, D8, and D12 (Mann-Whitney test, Bonferroni corrected $P < 0.0166$). These results indicate the influence of CS duration on Pre-CS resting state. Different CS durations likely elicited frequent or infrequent HEs from the subjects.

Further analysis was conducted on baseline-corrected HE frequency in CS and HE peak. Mean values of HE frequency and HE peak in Pre-CS were set as the baseline, respectively, and each baseline effect was removed from HE frequency in CS and HE peak. However, there was no prominent difference compared to uncorrected results (Figure S6 and Tables S9, S10, and S13). We concluded that responses in the off-stimulus of Pre-CS did not directly affect HE peak and frequency in the on-stimulus of CS.

The number of HE and correlation with accuracy

For each group, we tallied the number of HEs in the CS period without normalizing per 1 s, as done in HE frequency. As expected, subjects responded more frequently to longer CS durations (Figure S7A). In the 10-s group, HE occurrences increased to about 3 after D8. In contrast, in the 2-s group, HE occurrences remained below 1, while in the 5-s group, HE slightly exceeded 1. On D12, the mean and SD values were 3.325 ± 1.676 in the 10-s group, 1.123 ± 0.285 in the 5-s group, and 0.888 ± 0.099 in the 2-s group. Significant group differences were observed between the 10 s and 2 s from D8 to D12, between 10 s and 5 s on D4, and between 5 s and 2 s on D11 and D12 (Mann-Whitney test, Bonferroni corrected $P < 0.0166$; Table S15).

Next, we examined the correlation between accuracy rate and the number of HEs per trial, normalized into a percentage value for better comparison with accuracy rate. Each occurrence of HE was scored as 10. The results were highly significant across all training days except for D7 in the 2-s group

(Spearman test, $P < 0.001$; Figure 6A and Figure S7B). The 2-s group showed a linear trend wherein accuracy rate improved with an increase in the number of HEs across all training days (Figure 6B and Table S16), with minimal HEs occurring less than once per trial (Figure 6C). However, the significances of the correlation results for the 5-s and 10-s groups varied across training days and weakened toward the final day. The lack of significance in the correlation results for the 10-s and 5-s groups might be influenced by the ceiling effect of accuracy rate, or alternatively, by an excess of HEs relative to the number of trials.

The repetition of HE

In the 2-s group, the one-to-one matched relation might be emerged by mice' restrained impulsivity for anticipated food reward, or it might simply reflect their natural tendency to momentarily pause after reaching the food cup. To explore these possibilities, we organized the HEs for each trial and calculated the intervals between them: from CS onset to the first HE, from the first to the second HE, and between subsequent HEs. We investigated the number of mice engaged in each HE order and the frequency of responses occurring in each HE order. If a mouse responded at least once out of 10 trials, this response was counted as 1 mouse engaged in the order of that HE. The number of HEs for each order was normalized to percentile, with 80 responses to 80 trials representing 100%. Due to incomplete successive observations of responses beyond the second HE for all subjects, statistical analysis could not be conducted in this analysis.

The repetitions of HEs were observed up to 10 across all trials for all training days (Figure 7B and Figure S8). In 10-s group, one mouse responded up to 10 in D9 and D10. The seventh HE in 5-s group and the third HE in 2-s group were the maximum observed. The number of subjects who produced the first HE increased by D12 in all groups. Additionally, up to 8 mice in both the 10-s and 5-s groups produced the second HE by D12, indicating continuity between first and second HEs. However, this pattern was not observed in the 2-s group (Figure 7). The second HE of the 2-s group from D3 to D7, similar with 5-s group in the number of participating subjects, was fast than 5-s group, as intervals mostly between 0 and 1 s. However, the number of subjects did not increase toward the D12 as observed in the 5-s group. Despite responding similarly within the 2-s CS period, mice in this group showed fewer repeated HEs compared to the other groups. These findings suggest that mice in the 2-s group learned the stimulus-response relation better. In contrast, mice in the 10-s and 5-s groups showed more variable responses across training days, as indicated by fluctuations in the correlation results between the number of HEs and accuracy rate.

DISCUSSION

In the present data, between the two peaks of FHE and HE, which could be considered as RT, it was challenging to determine which one is closer to RT. Positive correlation results between FHE and HE peaks offered a solution to the dilemma of choosing between the two peaks. As CS duration decreased, FHE and HE peak latencies became more aligned, resulting in fewer HE occurrences and more monotony, resembling human responses. Basic properties of HE accuracy and frequency showed that mice were able to reach and respond to the food cup regardless of different CS durations. HE accuracy score reached to all over 80 % for all groups of 10 s, 5 s, and 2 s. HE frequency in short and shorter CS durations did not differ from that in long CS. Group differences in pre-CS didn't affect frequency in CS or HE peaks.

The latency of RT provides an additional assessment of subjects' task perception alongside accuracy[2]. In human studies[10,12], accuracy and RT are extracted and evaluated from the same data point. However, in animal studies, subjects may respond repeatedly, resulting in multiple RTs within multiple data points, forming an accuracy. RT may be a more sensitive indicator compared to accuracy due to their differing mechanisms[47], which can fluctuate greatly based on researcher instructions. Subjects can be prompted to respond slower or faster with minimal impact on accuracy scores. In our study, all three groups responded well to each stimulus (Figure 5A), FHE and HE peak latencies varied (Figure 4), suggesting that different CS durations reflected researchers' instructions regarding response timing. Longer CS durations allowed subjects more freedom in choosing response timing, while shorter durations may have induced anxiety to respond within a limited time frame. Thus, the shorter CS duration in our study represented the instruction for a human-like RT, wherein FHE and HE peaks were nearly coincident in terms of both latency and frequency, and accuracy rate and the number of HE were positively correlated over the training periods.

FHE and HE peaks showed the different patterns between groups. HE peak clearly demonstrated the effect of different CS durations on latency and frequency across early to late training days, indicating a strong relation to animals' habituated behaviors for the stimulus[30-34]. Conversely, FHE peak showed rarely significant group differences on several days. The frequencies act as reliability indicators, reflecting the overlapping responses of individual subjects at specific time points[27-29]. Non-significant group difference of HE peak's frequency after D5 (Figure 4B) was consistent with saturation of HE frequency (Figure 5B). The results of HE peak indicate that all subjects of three groups adapted to the CS duration and responded within each CS period well. However, the FHE peak results indicate that the response of the longer CS group was dispersed even within subject's data due to the low density of responses depending on extended CS period, compared to the shorter CS groups. Group differences in FHE and HE revealed dissimilar characteristics in terms of both latency and frequency. Based on

our findings, we proposed that the 2-s CS duration showing synchronized FHE and HE peaks in the current experimental paradigm, represents the optimal CS duration for inducing the unbiased results in mice behavior.

The frequency during the Pre-CS period was notably highest in the 2-s group, indicating more frequent checking of the food cup under shorter CS durations. Moreover, a linear trend was observed between HE frequency and CS duration for the Pre-CS period. However, it is noteworthy that behavior during the Pre-CS period did not influence behavior during the on-stimulus period, as indicated by baseline-corrected results for HE peak frequency and latency during CS. Despite the highest number of HE occurrences during the CS period being observed in the 10-s group, there were no significant differences in HE frequency between the 10-s group and the 5-s or 2-s groups. A linear trend between CS duration and response[43,44] was only observed between the 5-s group and the 2-s group, with the significantly higher HE frequency for the 2-s group than for the 5-s group. In contrast to the linear trend in Pre-CS period, the non-linear trend in CS period indicates a differential role of CS duration between on-stimulus and off-stimulus.

HE frequency during the CS period for the 5-s group was significantly lower than for the 2-s group. This infers that the 5-s and 2-s group's response patterns after the onset of sound cue would be probably similar. Interestingly, the latency and frequency of FHE and HE did not significantly differ between 5-s and 2-s groups, except for HE peak latency in D10 (Figure 4). The 5-s and 2-s group both revealed a positive correlation between FHE and HE peaks. However, the 2-s group, showing a stronger correlation between FHE and HE peaks, also demonstrated a correlation between the number of HE and accuracy across training days. Meanwhile, the 5-s group's correlation results show that mice behaviors were subtle arbitrary in the late days like 10-s group. Moreover, 2-s group responded less than 1 per trial, while the 10-s and 5-s groups responded with over 1 HE. These findings refer that the 5-s and 2-s groups' performances seemed to be similar, however, the 2-s group's mice better learned the one-to-one matched relationship between stimulus and response. In addition, in the results of the second HE, mice that produced the first HE repeated it at least once for any trial, but this was observed only in the 2-s group among all groups (Figure 7). Our data indicates that mice in the 2-s group chose to stop responding during the time remaining until the US was presented after the first HE within the CS period. In contrast, the 5-s group showed a steep linear increase in the number of subjects involved in the second HE from D8, ultimately reaching all subjects producing second HEs by D12, similar to the 10-s group. The synchrony between FHE and HE peaks, and between the number of HEs and accuracy in the 2-s group, resulted from the absence of the second HE. In our data, the prediction for the US involved not only enhancement of response[48-50] but also both inhibition and promotion of response according to CS duration. We suggest that the second HE, affected by CS duration, can be interpreted as an indicator of impulsivity toward the US in our data.

We aimed to suggest that mice's RT could be comparable to humans based on the relationship of FHE and HE peaks. However, it's important to note that results obtained using freely moving C57BL/6 mice may vary depending on the experimental paradigm, apparatus, and specific characteristics of the animals. In our paradigm, the shorter 2-s CS revealed a one-to-one matched relation with response frequencies lower than 0.5 Hz. However, if it involves licking behavior with higher-frequency responses under a head-fixed model, the relation may change, despite observing synchronization between FHE and HE peaks. The one-to-one matched response would be specific to our paradigm. Nonetheless, our data underscore the diversity of personalities among animals and how this can be managed in terms of CS duration. Establishing a time limit for response can promote sincere participation without hesitation, enhancing the sincerity and reliability of behavioral data. Ultimately, our findings propose quantifiable criteria for evaluating the duration of stimulus in experimental paradigms related to RT, based on the harmonization between FHE and HE peaks, as well as between the number of HEs and accuracy. If animals are capable, the shorter CS duration, which are barely capable of eliciting one response per trial, may be preferable for eliciting human-like responses without repetition.

METHODS

Subjects

Twenty-four naïve male C57BL/6 mice, aged 7 weeks old, were obtained from the Experimental Animal Center at Seoul National Dental University, with 8 mice assigned to each group to participate in the experiment. Prior to delay conditioning, each mouse was housed in an individual cage and subjected to a food restriction schedule to maintain their weight at 85% of their free-feeding weight throughout the experiment. Water was provided ad libitum. The mice were kept in cages in a room with a light-dark cycle of 12 hours each (8 am – 8 pm). Training commenced when the mice reached 8 weeks of age and concluded when they were 10 weeks old. The care and use of all animals were approved by the Seoul National University Institutional Animal Care and Use Committee (SNU IACUC).

Apparatus

The experiments were conducted using operant chambers (interior dimensions: 21.6 × 17.8 × 12.7 cm; model EBV-307W-CT, Med Associates), each equipped with a stainless steel grid floor consisting of 24 rods measuring 0.32 cm in diameter (17.8 × 15.2 × 5.7 cm, model ENV-307W-GF), and enclosed within MDF sound-attenuating cubicles (model ENV-022MD). A house light (28 V DC, 1000 mA, model ENV-315W) was illuminated from the beginning to the end of each experiment. A sound cue greater than 75 dB[51] was generated by a white noise amplifier (10 – 25,000 Hz, model ENV-325SW)

and delivered through a cage speaker (model ENV-324W). The noise emitted by fans (model ENV-025F) located in each cubicle was approximately 65 dB. Food reward, consisting of condensed milk, was delivered into a single 0.5 cc stainless steel cup via a liquid pipe (model ENV-303LPHD) using a syringe pump (model PHM-100A-EURO). HEs were counted whenever mice broke the beam in the detector (model ENV-303HDW) inside the liquid pipe. The experiments were controlled using MED-PC V software (model SOF-735).

Stimulus

Two stimuli of CS and US were used in the experiment. The CS consisted of white noise cues of varying durations: 10 s, 5 s, and 2 s (Figure 1). The choice of CS duration of 10 s was informed by literatures[28,44,52,53] related to CS duration. The 5-s and 2-s durations were determined based on the peak latencies of HE and FHE observed in our preliminary tests with a 10-s CS duration. The FHE and HE peaks' patterns were also observed in the present data (Figure S2C). To ensure that the CS durations were within ranges where mice could respond effectively, we analyzed the responses to the 10-s CS. We selected the 2-s duration based on the peak latency of FHE and the 5-s duration based on the peak latency of HE observed during the late training sessions. The US consisted of condensed milk. Following the principles of delay conditioning, the offset timings of the US matched those of the CSs (Figure 1). Therefore, each US was presented for 1 s, occurring from 9 to 10 s after the onset of the 10-s CS, from 4 to 5 s after the onset of the 5-s CS, and from 1 to 2 s after the onset of the 2-s CS.

Procedure

Prior to the main experiment's delay conditioning phase, a pre-conditioning phase spanned eight days, during which mice were subjected to food restriction until their weight reached 85% of their baseline, and were also acclimated to handling and the US. On the final day of the pre-conditioning phase, the mice underwent a session in the chamber without the CS but with the presence of the US. Following pre-conditioning, the delay conditioning experiments were conducted daily for a total of 12 days (Figure 1). Each day, ten trials were administered throughout the experiment. Inter-trial intervals varied and were randomly set to 90, 120, or 250, as shown in Figure S2A. The total experimental time per subject was approximately 1 hour per day, including preparation time. During the experiments, pumps were used to deliver condensed milk from syringes to food cups, with the pumps activated during the US duration. Mice were free to move within the chamber during the experiments. In this study, behavior directed towards the food cup for consuming the reward food was referred to as HE. The timing of HEs into the food cup was recorded with a temporal resolution of 10 ms.

Analysis

The data analysis was conducted on three groups, each with durations of 10 s, 5 s, and 2 s, comprising

a total of 24 mice. For the analysis of HE frequency, HEs were extracted from the time windows of the CS and its pre-CS period for 10 trials, respectively, and the number of HEs was normalized per s. We carefully selected the time window of Pre-CS as a baseline of off-stimulus to compare with the CS of on-stimulus. See Figure S3B for the details of Pre-CS segmentation. In the analysis of HE accuracy, the presence or absence of HE during each CS duration per trial was counted, and then converted into percentages (where 1 correct trial equates to 10 %). The number of HE was computed based on the total number of HEs during each CS duration. The FHEs were extracted from the first HE per trial (Figure S1A). The FHE and HE peaks were determined by extracting from maximum values from a matrix where the numbers of HEs occurring within a 1-s epoch were recorded as frequency levels (Figure S1B and S3). The maximum frequency value for each peak was 10 of the trial number per day. The latency of the peaks was time point of the maximum frequency. Based on this, latencies of the two peaks did not include the incorrect HEs like RT typically based on the correct answers[54,55]. The time windows for estimating FHE and HE peaks were each set from 0 s to 10 s after the onset of CS across all groups (Figure S4). Baseline-corrected HE frequency and peak were calculated by subtracting the mean HE in the pre-CS period from the HE in the CS (Figure S5 and S6).

During the statistical analysis, it was observed that all data did not follow a normal distribution. Therefore, nonparametric analyses were chosen for all statistical procedures. Nonparametric paired t-tests between FHE and HE peaks were performed using the Wilcoxon Signed Ranks test. Nonparametric correlation analyses between each pair of FHE and HE peaks, and between the number of HEs and accuracy, were estimated using the Spearman test. A correlation coefficient value of 1, indicating that the FHE peaks coincide with the HE peaks, was observed across all three groups (Table S3). Notably, the process of converting a decimal to an integer when extracting the two peaks from HEs affected the highest correlation coefficient values significantly (Figure S3). Nonparametric ANOVAs were conducted using the Kruskal-Wallis test. Post-hoc analyses for ANOVAs were performed using the Mann-Whitney test. Type I error due to multiple comparisons between the three groups in the Mann-Whitney test was adjusted using the Bonferroni test. The corrected significant alpha levels were set as follows: *, $P < 0.0167$; **, $P < 0.0033$; and ***, $P < 0.00033$ (Table S5 – S15). All analyses were conducted using MATLAB 9.12.0.2039608 (Math Works Inc., Natick, MA, USA) and SPSS 25.0 software (IBM, Armonk, NY, United States).

Availability of Data and Materials

The datasets used and/or analyzed during the current study available from the corresponding author on reasonable request.

REFERENCES

1 Killeen, P. R. & Fetterman, J. G. A behavioral theory of timing. *Psychol Rev* **95**, 274-295 (1988). https://doi.org:10.1037/0033-295x.95.2.274

2 Schouten, J. F. & Bekker, J. A. Reaction time and accuracy. *Acta Psychol (Amst)* **27**, 143-153 (1967). https://doi.org:10.1016/0001-6918(67)90054-6

3 Luce, R. D. *Response times: Their role in inferring elementary mental organization*. (Oxford University Press, 1986).

4 Baayen, R. H. & Milin, P. Analyzing reaction times. *International journal of psychological research* **3**, 12-28 (2010).

5 Welford, W., Brebner, J. M. & Kirby, N. *Reaction times*. (Stanford University, 1980).

6 Welford, A. T. Reaction time, speed of performance, and age. *Ann N Y Acad Sci* **515**, 1-17 (1988). https://doi.org:10.1111/j.1749-6632.1988.tb32958.x

7 Krajbich, I., Bartling, B., Hare, T. & Fehr, E. Rethinking fast and slow based on a critique of reaction-time reverse inference. *Nat Commun* **6**, 7455 (2015). https://doi.org:10.1038/ncomms8455

8 Ahveninen, J. *et al.* Evidence for distinct human auditory cortex regions for sound location versus identity processing. *Nat Commun* **4**, 2585 (2013). https://doi.org:10.1038/ncomms3585

9 Rao, S. M., Mayer, A. R. & Harrington, D. L. The evolution of brain activation during temporal processing. *Nat Neurosci* **4**, 317-323 (2001). https://doi.org:10.1038/85191

10 Luck, S. J. *et al.* The speed of visual attention in schizophrenia: electrophysiological and behavioral evidence. *Schizophr Res* **85**, 174-195 (2006). https://doi.org:10.1016/j.schres.2006.03.040

11 Kaiser, S. *et al.* Intra-individual reaction time variability in schizophrenia, depression and borderline personality disorder. *Brain Cogn* **66**, 73-82 (2008). https://doi.org:10.1016/j.bandc.2007.05.007

12 Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A. & Ward, B. D. Neural correlates of sensory and decision processes in auditory object identification. *Nat Neurosci* **7**, 295-301 (2004). https://doi.org:10.1038/nn1198

13 Lo, C. C. & Wang, X. J. Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat Neurosci* **9**, 956-963 (2006). https://doi.org:10.1038/nn1722

14 Wig, G. S., Grafton, S. T., Demos, K. E. & Kelley, W. M. Reductions in neural activity underlie behavioral components of repetition priming. *Nat Neurosci* **8**, 1228-1233 (2005). https://doi.org:10.1038/nn1515

15 Stefanics, G. *et al.* Phase Entrainment of Human Delta Oscillations Can Mediate the Effects of Expectation on Reaction Speed. *Journal of Neuroscience* **30**, 13578-13585 (2010). https://doi.org:10.1523/Jneurosci.0703-10.2010

16 Khamechian, M. B., Kozyrev, V., Treue, S., Esghaei, M. & Daliri, M. R. Routing information flow by separate neural synchrony frequencies allows for "functionally labeled lines" in higher primate cortex. *Proc Natl Acad Sci U S A* **116**, 12506-12515 (2019). https://doi.org:10.1073/pnas.1819827116

17 Etzel, B. C. & Wright, E. S. Effects of Delayed Reinforcement on Response Latency and Acquisition Learning under Simultaneous and Successive Discrimination-Learning in Children. *Journal of Experimental Child Psychology* **1**, 281-293 (1964). https://doi.org:Doi 10.1016/0022-0965(64)90043-8

18 Mehraei, G. *et al.* Auditory Brainstem Response Latency in Noise as a Marker of Cochlear Synaptopathy. *Journal of Neuroscience* **36**, 3755-3764 (2016). https://doi.org:10.1523/Jneurosci.4460-15.2016

19 Bale, M. R. *et al.* Learning and recognition of tactile temporal sequences by mice and humans. *Elife* **6** (2017). https://doi.org:10.7554/eLife.27333

20 Esteves, M., Moreira, P. S., Sousa, N. & Leite-Almeida, H. Assessing impulsivity in humans and rodents: taking the translational road. *Frontiers in Behavioral Neuroscience* **15**, 647922 (2021).

21 Johnson, S. A. *et al.* Rodent age-related impairments in discriminating perceptually similar objects parallel those observed in humans. *Hippocampus* **27**, 759-776 (2017).

22 Koelsch, S., Schmidt, B. H. & Kansok, J. Effects of musical expertise on the early right anterior negativity: an event-related brain potential study. *Psychophysiology* **39**, 657-663 (2002). https://doi.org:10.1017.S0048577202010508

23 Poulton, E. C. Perceptual Anticipation and Reaction Time. *Quarterly journal of experimental psychology* **2**, 99-112 (1950). https://doi.org:Doi 10.1080/17470215008416582

24 Butz, M. V. & Hoffmann, J. Anticipations control behavior: Animal behavior in an anticipatory learning classifier system. (2002).

25 Altieri, N. & Townsend, J. T. An assessment of behavioral dynamic information processing measures in audiovisual speech perception. *Front Psychol* **2**, 238 (2011). https://doi.org:10.3389/fpsyg.2011.00238

26 Eshel, N. & Roiser, J. P. Reward and punishment processing in depression. *Biological psychiatry* **68**, 118-124 (2010). https://doi.org:10.1016/j.biopsych.2010.01.027

27 Harris, J. A. & Carpenter, J. S. Response rate and reinforcement rate in Pavlovian conditioning. *J Exp Psychol Anim Behav Process* **37**, 375-384 (2011). https://doi.org:10.1037/a0024554

28 Austen, J. M. & Sanderson, D. J. Cue duration determines response rate but not rate of acquisition of Pavlovian conditioning in mice. *Quarterly journal of experimental psychology* **73**, 2026-2035 (2020). https://doi.org:10.1177/1747021820937696

29 Jennings, D. & Kirkpatrick, K. Interval duration effects on blocking in appetitive conditioning. *Behav Processes* **71**, 318-329 (2006). https://doi.org:10.1016/j.beproc.2005.11.007

30 Groves, P. M. & Thompson, R. F. Habituation: a dual-process theory. *Psychol Rev* **77**, 419-450 (1970). https://doi.org:10.1037/h0029810

31 de Boer, E. *et al.* Habituation and attention in the auditory system. *Auditory system: Clinical and special topics*, 343-389 (1976).

32 Thompson, R. F. Habituation: a history. *Neurobiol Learn Mem* **92**, 127-134 (2009). https://doi.org:10.1016/j.nlm.2008.07.011

33 Cevik, M. O. Habituation, sensitization, and Pavlovian conditioning. *Front Integr Neurosci* **8**, 13 (2014). https://doi.org:10.3389/fnint.2014.00013

34 McSweeney, F. K., Hinson, J. M. & Cannon, C. B. Sensitization-habituation may occur during operant conditioning. *Psychological Bulletin* **120**, 256-271 (1996). https://doi.org:Doi 10.1037/0033-2909.120.2.256

35 Delamater, A. R. & Holland, P. C. The influence of CS-US interval on several different indices of learning in appetitive conditioning. *Journal of Experimental Psychology-Animal Behavioral Processes* **34**, 202-222 (2008). https://doi.org:10.1037/0097-7403.34.2.202

36 Odling-Smee, F. J. Background stimuli and the inter-stimulus interval during Pavlovian conditioning. *Q J Exp Psychol* **27**, 387-392 (1975). https://doi.org:10.1080/14640747508400498

37 Dickinson, A. & Mackintosh, N. J. Classical conditioning in animals. *Annu Rev Psychol* **29**, 587-612 (1978). https://doi.org:10.1146/annurev.ps.29.020178.003103

38 Velazquez-Sanchez, C. *et al.* High trait impulsivity predicts food addiction-like behavior in the rat. *Neuropsychopharmacology* **39**, 2463-2472 (2014). https://doi.org:10.1038/npp.2014.98

39 Celma-Miralles, A. & Toro, J. M. Non-human animals detect the rhythmic structure of a familiar tune. *Psychonomic bulletin & review* **27**, 694-699 (2020). https://doi.org:10.3758/s13423-020-01739-2

40 Johnson, A. W. Examining the influence of CS duration and US density on cue-potentiated feeding through analyses of licking microstructure. *Learning and Motivation* **61**, 85-96 (2018). https://doi.org:10.1016/j.lmot.2017.07.001

41 Rosas, J. M. & Alonso, G. Temporal discrimination and forgetting of CS duration in conditioned suppression. *Learning and Motivation* **27**, 43-57 (1996). https://doi.org:DOI 10.1006/lmot.1996.0003

42 Meltzer, D. Positive conditioned suppression after shifts in CS duration and US probability. *Bulletin of the Psychonomic Society* **26**, 565-568 (1988).

43 Holland, P. C. CS-US interval as a determinant of the form of Pavlovian appetitive conditioned responses. *J Exp Psychol Anim Behav Process* **6**, 155-174 (1980).
44 Thrailkill, E. A., Todd, T. P. & Bouton, M. E. Effects of conditioned stimulus (CS) duration, intertrial interval, and I/T ratio on appetitive Pavlovian conditioning. *J Exp Psychol Anim Learn Cogn* **46**, 243-255 (2020). https://doi.org:10.1037/xan0000241
45 Clark, R. E. & Squire, L. R. Classical conditioning and brain systems: the role of awareness. *Science* **280**, 77-81 (1998). https://doi.org:10.1126/science.280.5360.77
46 Balsam, P. Relative time in trace conditioning. *Ann N Y Acad Sci* **423**, 211-227 (1984). https://doi.org:10.1111/j.1749-6632.1984.tb23432.x
47 Prinzmetal, W., McCool, C. & Park, S. Attention: reaction time and accuracy reveal different mechanisms. *J Exp Psychol Gen* **134**, 73-92 (2005). https://doi.org:10.1037/0096-3445.134.1.73
48 Coddington, L. T. & Dudman, J. T. The timing of action determines reward prediction signals in identified midbrain dopamine neurons. *Nature neuroscience* **21**, 1563-1573 (2018).
49 Eshel, N., Tian, J., Bukwich, M. & Uchida, N. Dopamine neurons share common response function for reward prediction error. *Nature neuroscience* **19**, 479-486 (2016).
50 Anderson, C., Von Keyserlingk, M., Lidfors, L. & Weary, D. Anticipatory behaviour in animals: A critical review. *Animal Welfare* **29**, 231-238 (2020).
51 Tosun, T., Gur, E. & Balci, F. Mice plan decision strategies based on previously learned time intervals, locations, and probabilities. *Proc Natl Acad Sci U S A* **113**, 787-792 (2016). https://doi.org:10.1073/pnas.1518316113
52 Trott, J. M., Hoffman, A. N., Zhuravka, I. & Fanselow, M. S. Conditional and unconditional components of aversively motivated freezing, flight and darting in mice. *Elife* **11**, e75663 (2022). https://doi.org:10.7554/eLife.75663
53 Strickland, J. A., Austen, J. M., Sprengel, R. & Sanderson, D. J. Knockout of NMDARs in CA1 and dentate gyrus fails to impair temporal control of conditioned behavior in mice. *Hippocampus* **34**, 126-140 (2024). https://doi.org:10.1002/hipo.23593
54 Kail, R. Controlled and automatic processing during mental rotation. *J Exp Child Psychol* **51**, 337-347 (1991). https://doi.org:10.1016/0022-0965(91)90081-3
55 Novikov, N. A. *et al.* Slow and Fast Responses: Two Mechanisms of Trial Outcome Processing Revealed by EEG Oscillations. *Front Hum Neurosci* **11**, 218 (2017). https://doi.org:10.3389/fnhum.2017.00218

Funding

This research was supported by Seoul National University Research Grant in 2021 and the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2022R1I1A1A0107380012).

Competing interests

The author declares no competing interests.

FIGURES

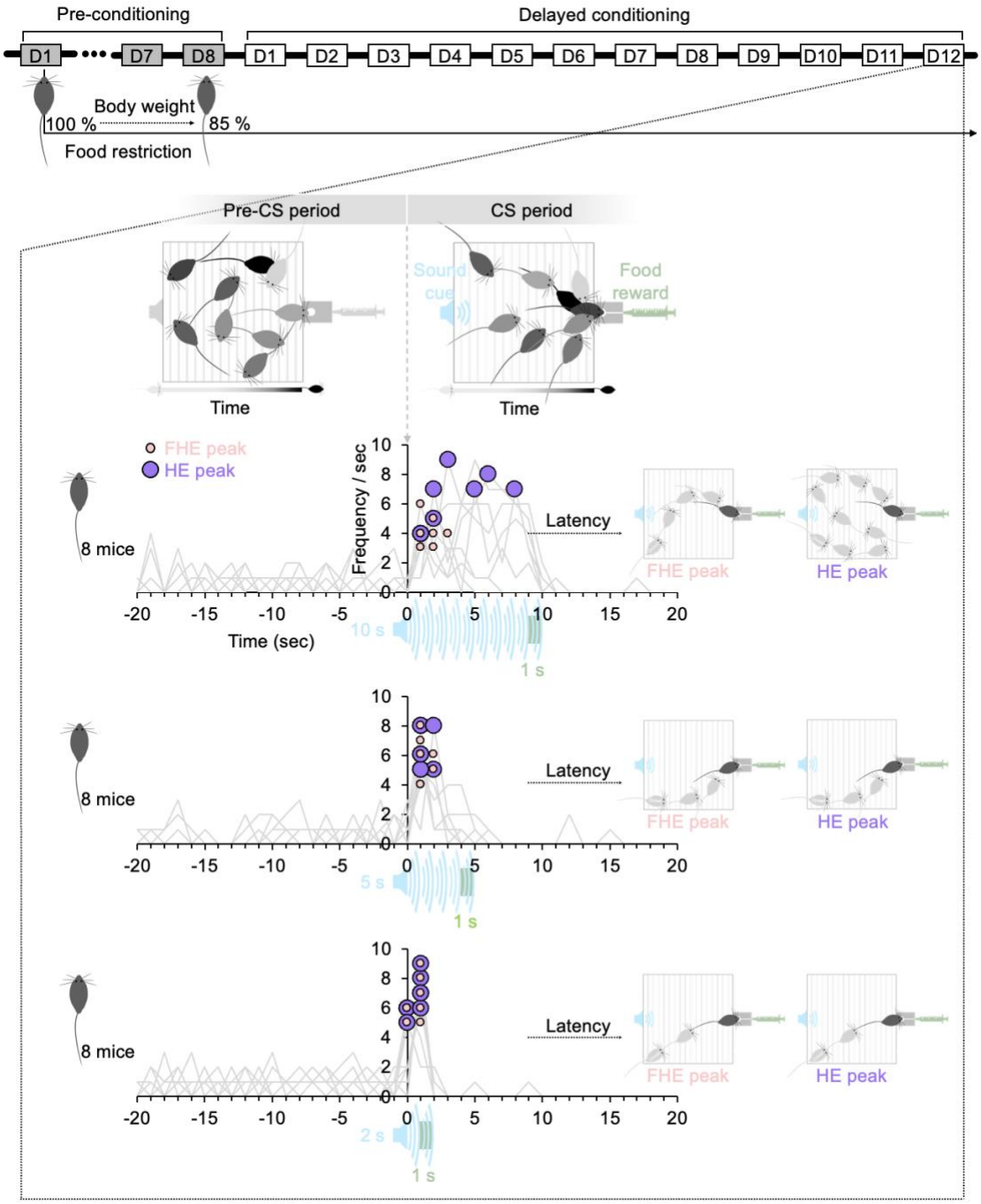

Figure. 1. Model of FHE and HE peaks for human-like RT. Mice in all groups reduced their body weight by up to 85% through food restriction during the pre-conditioning periods before delay conditioning. In CS period, the duration of the sound cue followed by food reward differed between the 10-s, 5-s, and 2-s groups. On the final training day (D12), the HE latency in the CS period after the onset of the sound cue was enhanced compared to that in the pre-CS period before the onset of the sound cue. The FHE peaks preceded the HE peaks in the 10-s group, while in the 5-s and 2-s groups, the two peaks were almost synchronized. The latencies of the two peaks demonstrate how mice in each group react differently to sound cues of varying durations.

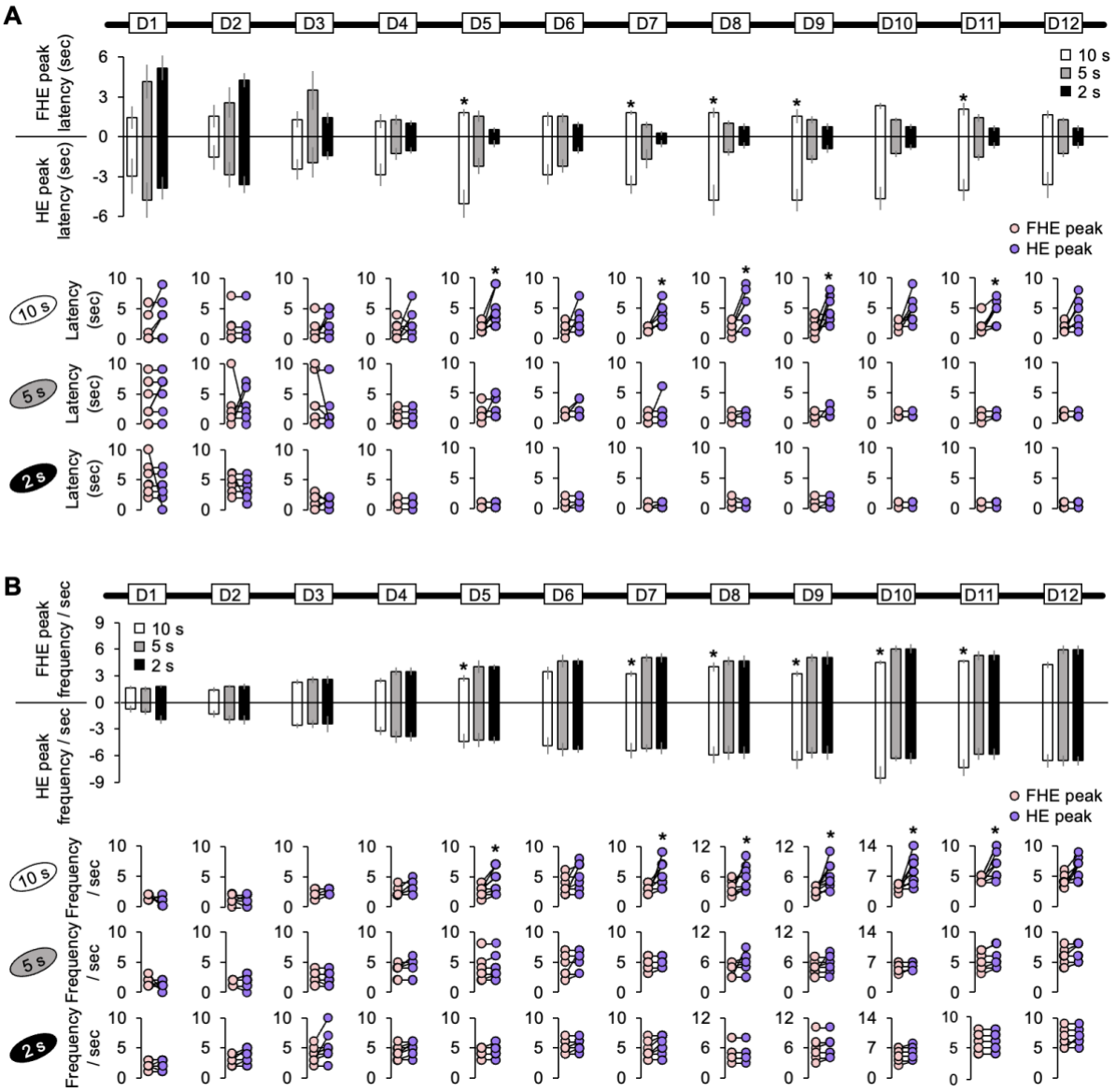

Figure. 2. Difference between FHE and HE peaks in latency and frequency. (A) Significant differences in latencies between FHE and HE peaks were observed only in the 10-s group (Wilcoxon Signed Ranks test, $P < 0.05$). The latency for the HE peak was longer than for the FHE peak. However, significance was not observed in the 5-s and 2-s groups (Wilcoxon Signed Ranks test, $P > 0.05$). (B) Significant differences in frequencies between FHE and HE peaks were also observed only in the 10-s group (Wilcoxon Signed Ranks test, $P < 0.05$). The frequency for the HE peak was higher than for the FHE peak. *, $P < 0.05$, **, $P < 0.01$, and ***, $P < 0.001$.

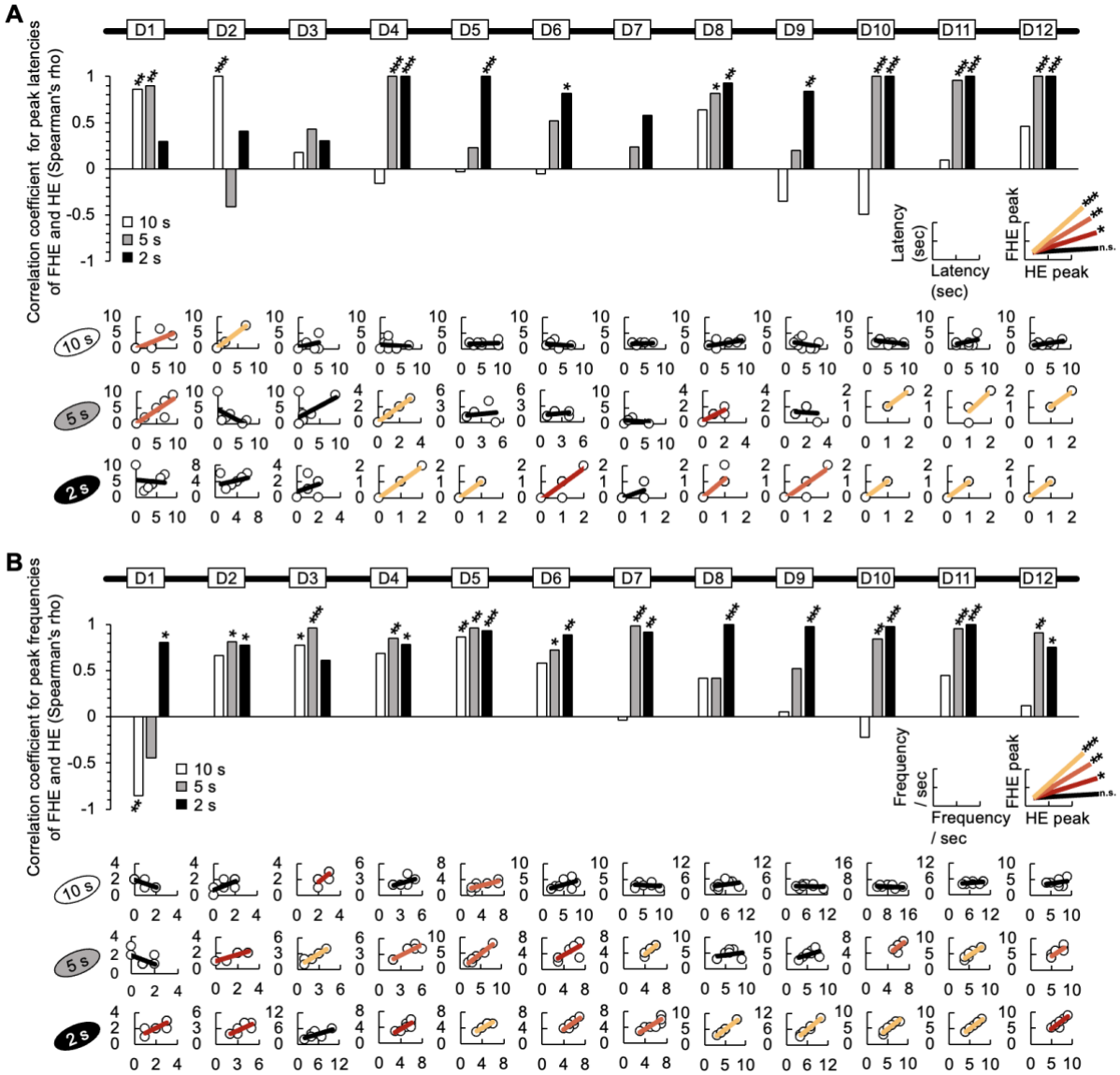

Figure. 3. Correlation between FHE and HE peaks in latency and frequency. (A) The latencies of FHE and HE peaks were strongly and positively correlated in the later training days for the 2-s group (D4 – D6 and D8 – D12). Positive correlations were observed in the later days of the 5-s group (D1, D4, D8, D10, D11, D12), whereas they were only observed on D1 and D2 for the 10-s group. (B) Strong positive correlations in the frequencies of FHE and HE peaks were observed in the 2-s group across all 12 days, except for D3. Significant positive correlations were also observed in the 5-s and 10-s groups. In the 5-s group, positive correlations were observed on all days except for D1, D8, and D9. However, in the 10-s group, positive correlations were observed only in the early period of training (D3 and D5). The non-parametric correlation was tested using Spearman's test. *, $P < 0.05$, **, $P < 0.01$, and ***, $P < 0.001$.

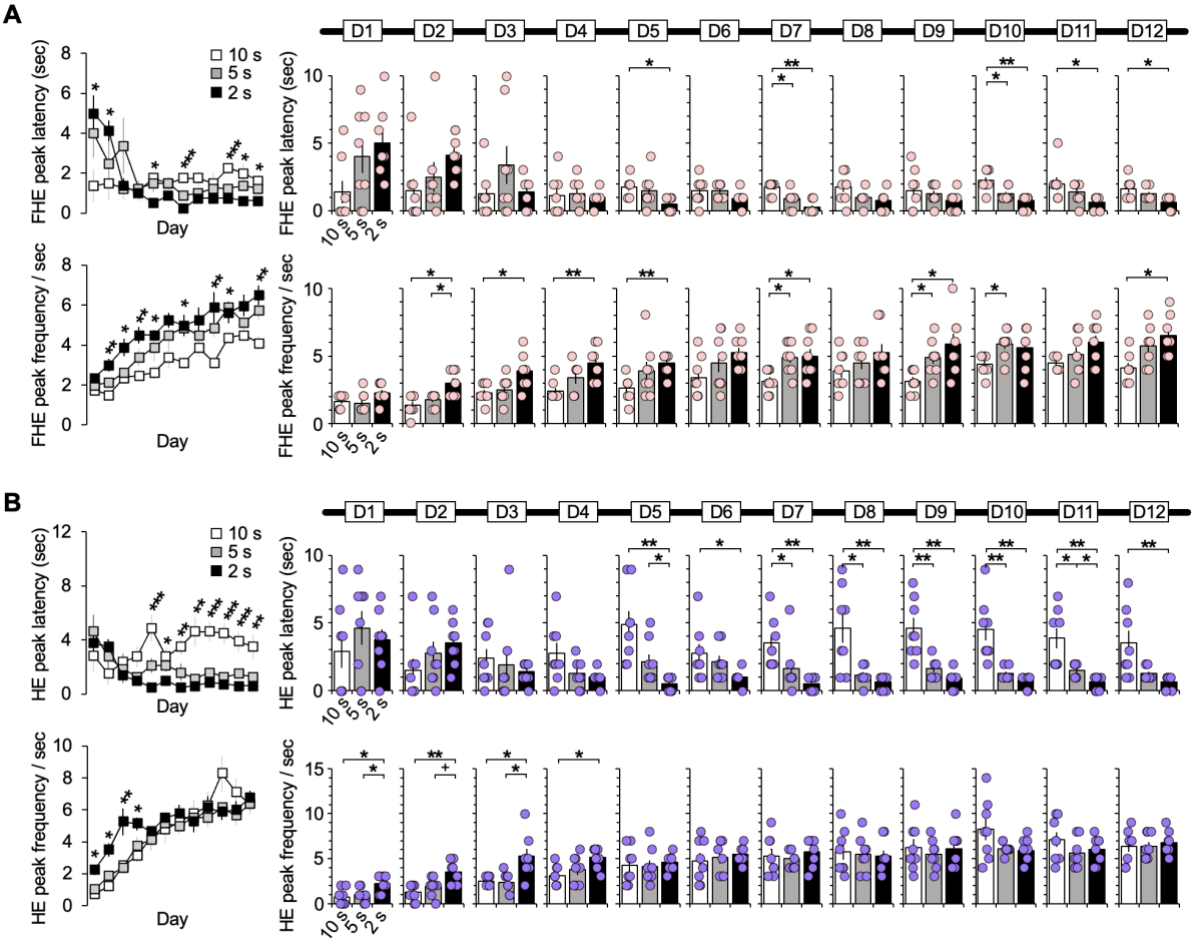

Figure. 4. Difference between groups in latency and frequency of FHE and HE peaks. One-way ANOVA results for FHE and HE peak latency and frequency showed the significant effect of Group factor (10 s, 5 s, and 2 s). (A) The FHE peak showed a significant Group effect across the training days for both latency (D1, D2, D5, D7, D10, D11, and D12) and frequency (D2, D3, D4, D5, D7, D9, D10, and D12). In post-hoc tests, significant differences were mostly observed between the 10-s and 2-s groups or between the 10-s and 5-s groups for both latency and frequency. (B) One-way ANOVA results for the HE peak revealed a significant Group effect for both latency (D5–D12) and frequency (D1–D4). In post-hoc analysis, the latency of the 10-s group differed significantly from other groups from D5 onwards. In post-hoc analysis for frequency, group differences were highest in the 2-s group compared to other groups from D1 to D4. In the Kruskal-Wallis test of one-way ANOVA, *, $P < 0.05$, **, $P < 0.01$, and ***, $P < 0.001$. In the Mann-Whitney test of post-hoc analysis, *, $P < 0.016$, and **, $P < 0.0033$ (Bonferroni corrected).

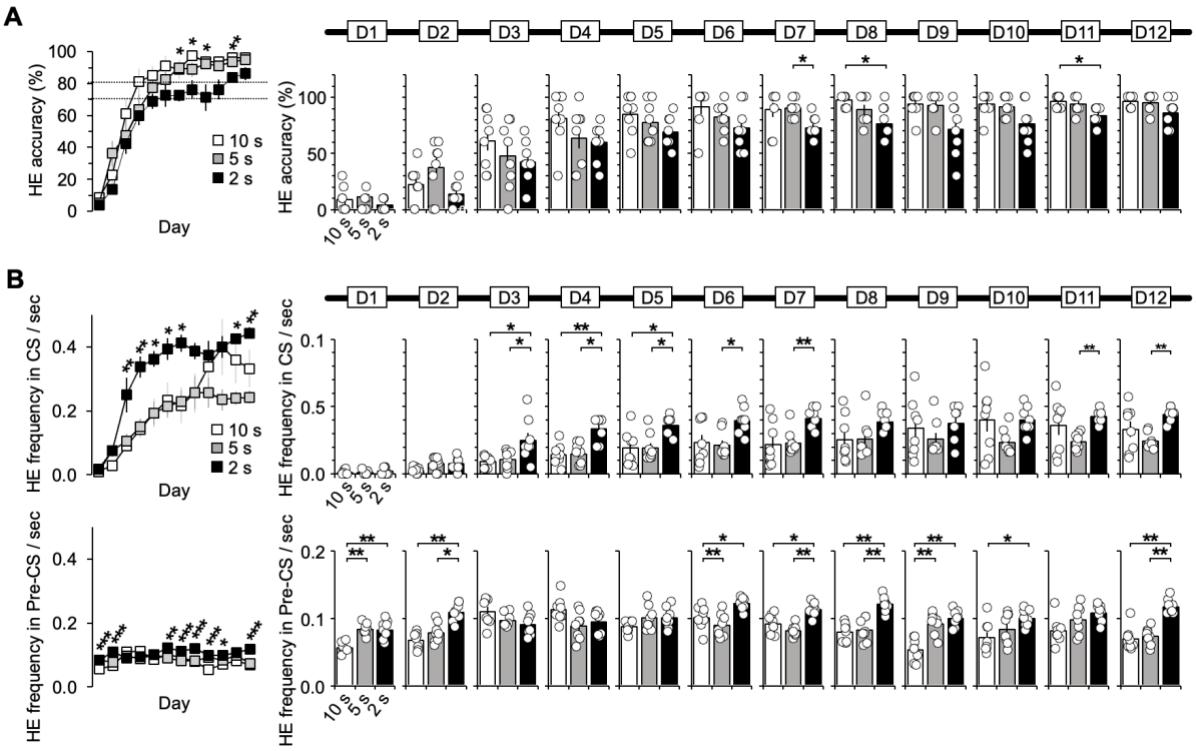

Figure. 5. Basic properties of HE. (A) One-way ANOVA results indicated that the Group effect for HE accuracy was only observed on D7, D8, and D11. However, all subjects eventually performed the tasks well, achieving an accuracy over the mean of 80%, regardless of CS duration. (B) One-way ANOVA revealed a significant Group effect for HE frequency during the CS period (D3–D7, D11, and D12). In post-hoc analysis, the frequency of the 2-s group was significantly higher than that of the 5-s and 10-s groups. For the Pre-CS period, the Group effect was significant across eight days except for D3, D4, D5, and D11. Post-hoc analysis showed higher frequencies in the 2-s group than in the 10-s group. Differences between the 10-s and 5-s groups were observed on D1 and D9, while differences between the 5-s and 2-s groups were observed on D2, D6, D7, D8, and D12.

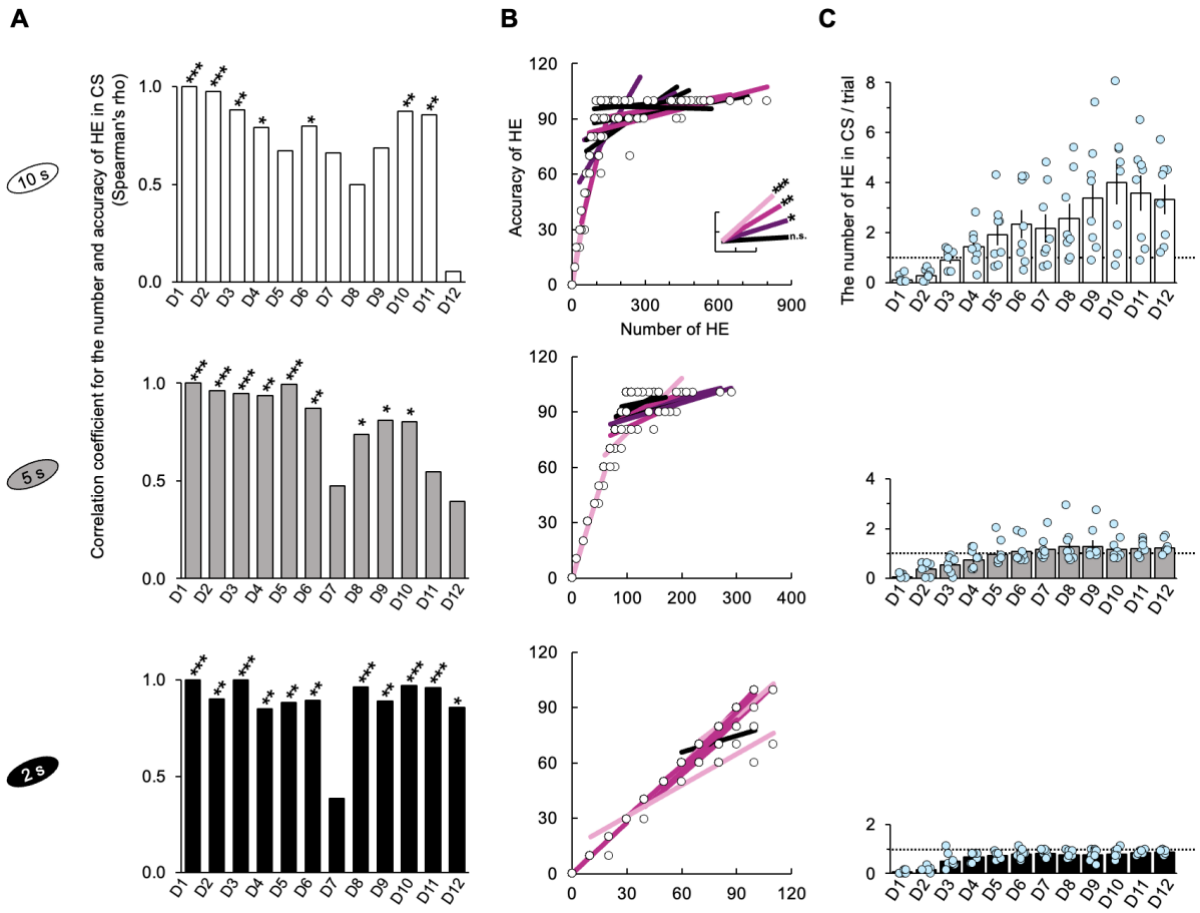

Figure. 6. The number of HE and correlation with accuracy. (A) Significant positive correlations between the number of HEs and accuracy were observed in all groups. For the 2-s group, these significances were consistent on all training days except for D7, but not in the other groups. The non-parametric correlation was tested using Spearman's test. (B) The lines of best fit and data points, which relate to the correlation coefficients in (A) and display the trend in the data for each day, were all included in the graphs (Figure S6). The scale of the x-axis was set based on the maximum value across all training days. In the case of the 2-s group, many data points overlap, resulting in a small representation among the 96 points collected from 8 subjects over 12 days. (C) The number of HEs was highest in the 10-s group. In the 2-s group, HE occurrences were below 1, while in the 5-s group, HE slightly exceeded 1. The dotted line denotes the number of HE corresponding to the ono-to-one matched relation of stimulus and response. *, $P < 0.05$; **, $P < 0.01$; and ***, $P < 0.001$.

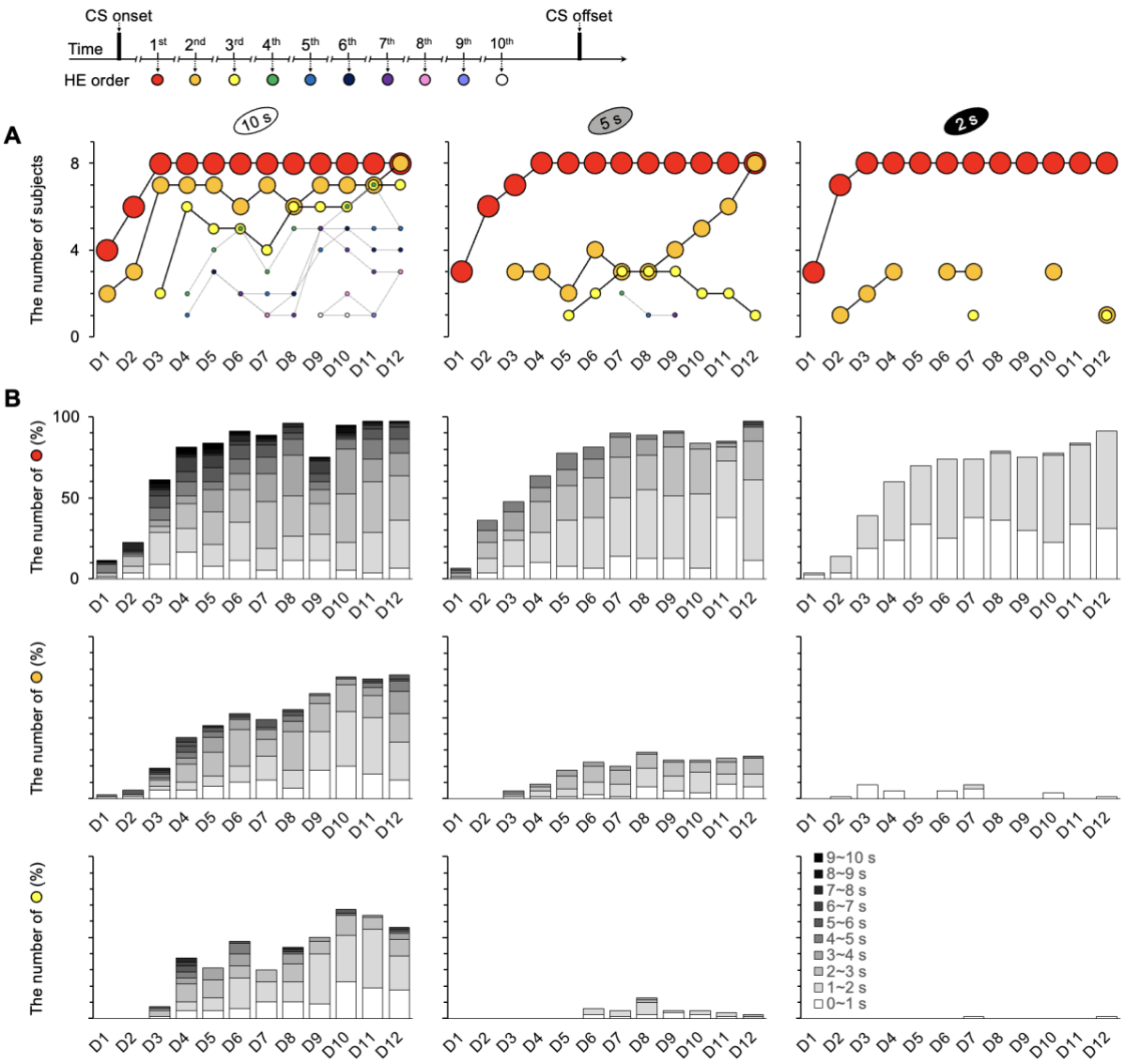

Figure 7. The number of repeating HEs and participating subjects per trial. Repetitions of HEs were observed up to 10 across all trials for all training days. (A) In the 10-s group, responses were observed up to the tenth HE. Meanwhile, the maximum observed in the 5-s group was the seventh HE, and in the 2-s group, it was the third HE. The pattern of the first HE was similar across all groups; however, differences emerged in the second and third HEs. Specifically, the second HE increased toward 8 mice on D12 for both the 10-s and 5-s groups, but not for the 2-s group. The second HE of the 2-s group from D3 to D7, similar with 5-s group in the number of participating subjects, was fast than 5-s group, as intervals mostly between 0 and 1 s in (B); however, it did not increase like the 5-s group toward D12. (B) The latencies between HEs varied according to CS durations. The interval latency in the 10-s group was the longest, while it was the shortest in the 2-s group. For results from the fourth to tenth HEs connected by light gray lines (Figure S8B).