TITLE

Harmonization of the fastest and densest responses embodies the 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 an important factor in evaluating delayed latency 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 rather 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 controlling stimulus duration. In delay conditioning with C57BL/6 mice using sound cues of 10, 5, and 2 s, the 2-s group only demonstrated the strongest positive correlations between the two peaks, as well as HE's number and accuracy rate, suggesting coupling of the fastest and densest responses, and a one-to-one relationship between stimulus and response, respectively. We propose that the harmonization of the two peaks, elicited by stimuli that elicit minimal responses, be used as a criterion for designing animal experiments to mimic human-like RT.

KEYWORDS

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

INTRODUCTION

Timing is critical for investigating behavioral response[1]. Reaction time (RT), also known as response time[2-5], is an important 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 is also linked to neural correlation across multiple brain regions[12-14]. Terms such as 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 similarities between human and animal behaviors under different conditions, it is possible that the results cannot be interpreted similarly. The current study proposes theoretical criteria for determining human-like RT from animal data.

RT is typically assessed within a specific time frame for each trial[22], along with an estimate of its time point[23,24]. Researchers must provide clear instructions to human subjects before experiments to ensure RT sensitivity. Human subjects are occasionally asked to respond as-soon-as-possible[25,26]. However, instructing animals to respond quickly may be impractical given communication limitations. Even with clear instructions for animal subjects in experimental paradigm with a long training period, animals may delay their responses intentionally. Furthermore, their behavior during a trial, which includes repeated actions like licking, nose poking, and lever pressing, complicates the interpretation of an RT-like measure. In animal studies, the relationship between stimulus and response may not be one-to-one, as in human studies. It is unclear whether the delayed latencies of repeated responses accurately reflect the animals' decision-making points in the cognitive process. The averaged animals' responses may contain errors in which the intention cannot be accurately determined. Comparing an averaged latency in animals with RT in human studies directly would be difficult. Therefore, the frequency of repeated responses, such as head entry (HE) into the food cup, is critical for assessing the behavioral performance of animals[27-29]. During the training period, repeated HEs for repeated stimulations gradually become the densest point, forming a peak in the HE distribution that reflects subjects' behavioral habituation and sensitization[30-34]. However, in addition to the HE peak latency, the fastest HE (FHE) and its peak may exist within the distribution of HEs (Figure 1 and S1). Time points for RTs in animal data would fundamentally not be a single latency, but rather a subset of both components. The coexistence of distinct two components may raise an issue that requires identifying a true RT in the data. Given one-to-one relationship like that in humans, the FHE peak of the first HE after stimulus onset is sufficient. However, the HE peak is more reliable due to its higher frequency. If the FHE and HE peaks coincide, it indicates a coupling of the fastest and densest responses. This is an ideal option that is closer to humans' RT.

Temporal aspects[35,36] of the experimental paradigm, such as the inter-stimulus interval, intertrial interval, interval between conditioned stimulus (CS) and unconditioned stimulus (US), and CS duration, have an individual and collective influence on the response. This study focuses on investigating the effect of CS duration on determining the CS–US interval in delay conditioning. In classical conditioning[37], the time allowed for animals to respond is determined by the interval between the stimulus and the reward. The US of food following the CS onset of stimulus cue guides the HEs. Some animals can postpone their actions after deciding on long CS–US intervals. A short CS–US interval would be preferable for eliciting more prompt responses from animals, whereas a long CS–US interval would results in acquiring more responses safely and without missing any. A higher level of frequency on delayed latencies may indicate animals' behavior similar to food addiction[38] as well as their deep awareness of task[39]. Latency and frequency are likely to have greater distribution variance in longer CS duration than in short CS duration, reflecting individual differences such as animals' preferences, hesitation, and laziness of subjects, which deviate from human-like RT. Indeed, numerous studies have investigated the effect and significance of the CS–US interval (or CS duration) based on tests with various CS durations[27-29,35,40-44]. However, the specific durations of the stimulus vary depending on the apparatus and study paradigm, so an absolute value cannot be used to determine the appropriate duration for designing an experiment.

RT can be influenced by various experimental conditions. Our objective was to find a temporal factor in experimental design that could combine the two different latencies of FHE and HE peaks into a single cohesive unit of RT, rather than treating them as separate entities. We thought the factor was CS duration. To investigate this, we tested 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 of FHE and HE peaks would vary depending on the CS durations. If FHE and HE peak latencies are synchronized, their frequencies would also align into a single peak, as the relationship between latency and frequency can be interdependent. Furthermore, we hypothesized that the optimal CS duration could be predicted based on the relationship between these two peaks. For this, we recorded the HEs of freely moving mice in an operant chamber based on delay conditioning[45], using white noise and condensed milk. To eliminate additional factors other than training food-reward associations and to observe mice's natural behaviors throughout a sound cue, we did not include no reward or punishment for missed responses in the experimental protocol. Delay conditioning, rather than trace conditioning related to memory-based process[46], was used to elicit a response based on stimulus duration without an interval of silence. In data analysis, to fully evaluate the impact of CS duration, we conducted a detailed analysis of behavioral data at both the temporal and frequency levels. Our analysis took several steps: First, we assessed the difference and correlation between FHE and HE peaks to better understand their relationship. Next, we investigated group differences in FHE and HE peaks separately to better understand their characteristics and implications. Third, we

examined 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 subject performance. We also considered how shorter CS durations affected mice's ability to reach the food cup versus longer CS durations, as well as how CS duration affected HE frequency. Based on these findings, we investigated how subjects' performance was affected by the relationship between FHE and HE peaks. Finally, we examined 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. Finally, we proposed a criterion for determining an appropriate CS duration to elicit 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 that only the 10-s group had differences in latency and frequency between FHE and HE peaks. The significance of these findings varied by training days (D). Specifically, for latency, HE peaks outperformed 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). Similarly, 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, there were no differences in latency or frequency between FHE and HE peaks in the 5-s and 2-s groups ($P > 0.05$ in all cases).

Correlation between FHE and HE peaks for each group

We used the Spearman test to determine the correlation between FHE and HE peaks, based on the nonparametric paired t-test results. As expected, the 2-s group's FHE and HE peak latencies were strongly and positively correlated during the later training days (D4–D6 and D8–D12). Similarly, in the 5-s group, positive correlations were found on D1, D4, D8, D10, D11, and D12. For the 10-s group, D1 and D2 were the only significant variables (Figure 3A and Table S3. Regarding the frequencies of FHE and HE peaks, the 2-s group showed positive correlations across all 12 days, except for D3 (Figure 3B and Table S4). Significant positive correlations were also observed in the 5-s on all days except for D1, D8, and D9. In the 10-s group, those were observed only in the early training period of D3 and D5, with a notable negative correlation on D1. The correlation results show that FHE and HE peaks were similar in groups with short and shorter CS durations, but not in group with long CS duration. Particularly, the two peaks nearly overlapped 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 found 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 in which the significant Group effect for latency on D5–D12 appeared to mirror the significance observed in frequency on D1–D4 (Figure 4B). In post-hoc analysis for frequency, group differences were significantly higher in the 2-s group than in other groups during the early training periods of D1 to D4 (Mann–Whitney test, Bonferroni corrected $P < 0.0167$; Table S7). Post-hoc analysis for latency revealed significant differences 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 had significantly longer latency than the other groups from D5 to D12, and differences between the 5-s and 2-s groups were also found in D5 and D11.

Based on the results of FHE and HE peaks, group differences in FHE peaks appeared sporadically and irregularly across training days, whereas group differences in HE peaks appeared to follow a more consistent pattern. We concluded that the results in HE peak reflected stimulus adaptation after the early training days of D1 to D4, as differences in HE peak latency between the longest and shortest CS groups persisted from D5 to D12. Our results show that HE peaks explain the effect of CS duration better than FHE peaks.

HE accuracy

The HE accuracy was considered correct if the mice's heads touched 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 found on D7, D8, and D11 (Figure 5A). Post-hoc analysis found differences 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 S9). However, no significant difference was discovered on the last day of D12. Overall, subjects achieved high accuracy rates of more than 80%, regardless of CS duration.

HE frequency

HE frequency was calculated for both CS and Pre-CS periods, with HEs normalized per 1 s for the three groups. During the CS period, significant group differences were observed on seven days (D3–D7, D11, and D12) (Figure 5B and Table S10). A post-hoc analysis revealed that the 2-s group had significantly higher frequencies than the 5-s group. Furthermore, the 10-s group had higher

frequencies than the 2-s group only on D3–D7. During the late training, the 5-s and 2-s groups showed a significant difference on D11 and D12 (Mann–Whitney test, Bonferroni corrected $P < 0.0166$). HE frequency decreased in the 5-s group.

For the Pre-CS period, one-way ANOVAs revealed significant group effects on eight days except for D3, D4, D5, and D11 (Figure 5B and Table S11). The post-hoc analysis revealed that the 2-s group had higher frequencies than the 10-s group. The 10-s and 5-s groups showed differences on D1 and D9, while the 5-s and 2-s groups differed on D2, D6, D7, D8, and D12 (Mann–Whitney test, Bonferroni corrected $P < 0.0166$). These findings show how CS duration affects the Pre-CS resting state. Different CS durations most likely resulted in frequent or infrequent HEs from the subjects.

Additional analysis was performed on baseline-corrected HE frequency in CS and HE peak. The mean values of HE frequency and HE peak in Pre-CS were used as the baselines, and each baseline effect was subtracted from HE frequency in CS and HE peak. However, there was no prominent difference from uncorrected results (Figure S2 and Tables S12, S13, and S14). We concluded that responses in the off-stimulus of Pre-CS did not directly impact HE peak and frequency in the on-stimulus of CS.

The number of HE and correlation with accuracy

For each group, we counted the number of HEs in the CS period without normalizing per 1 s, as we did with HE frequency. As expected, longer CS durations elicited a higher response rate from subjects (Figure S3A). In the 10-s group, HE occurrences increased to approximately 3 after D8. In contrast, in the 2-s group, HE occurrences remained below 1, whereas in the 5-s group, HE slightly surpassed 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. There were significant group differences between 10 s and 2 s from D8 to D12, 10 s and 5 s on D4, and 5 s and 2 s on D11 and D12 (Mann–Whitney test, Bonferroni corrected $P < 0.0166$; Table S15).

Next, we investigated the relationship between the accuracy rate and the number of HEs per trial, which was normalized to a percentage value for easier comparison with the 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 S3B). The 2-s group demonstrated a linear trend in which the accuracy rate improved with an increase in the number of HEs across all training days (Figure 6B and Table S16), with the fewest HEs occurring less than once per trial (Figure 6C). However, the correlation results for the 5-s and 10-s groups varied across training days and became less significant near the final day. The lack of significance in the correlation results for the 10-s and 5-s groups could be attributed to the accuracy rate ceiling effect, or to 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 may have emerged due to mice's restrained impulsivity for anticipated food reward, or it may simply reflect their natural tendency to momentarily pause after reaching the food cup. To investigate these possibilities, we organized the HEs for each trial and calculated the time 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 involved in each HE order, as well as the frequency with which responses occurred. If a mouse responded at least once in 10 trials, it was counted as 1 mouse in the order of that HE. The number of HEs for each order was normalized to the percentile, with 80 responses to 80 trials equaling 100%. Due to incomplete successive observations of responses beyond the second HE for all subjects, statistical analysis could not be performed in this study.

HEs were repeated up to 10 times across all trials and training days (Figure 7B and Figure S8). In the 10-s group, one mouse responded up to 10 times in D9 and D10. The seventh HE in the 5-s group and the third HE in the 2-s group were the highest recorded. The number of subjects who produced their first HE increased by D12 across all groups. Furthermore, up to 8 mice in both the 10-s and 5-s groups produced the second HE by D12, indicating a link between the first and second HEs. However, this pattern did not appear in the 2-s group (Figure 7). The second HE of the 2-s group from D3 to D7, which had a similar number of participating subjects to the 5-s group, was faster than the 5-s group, with intervals mostly ranging from 0 and 1 s. However, the number of subjects did not increase as they approached D12, unlike the 5-s group. Despite responding similarly during the 2-s CS period, mice in this group demonstrated fewer repeated HEs than the other groups. These findings indicate that mice in the 2-s group learned the stimulus-response relation more effectively. In contrast, mice in the 10-s and 5-s groups demonstrated more variable responses across training days, as evidenced by fluctuations in the correlation coefficients between the number of HEs and accuracy rate.

DISCUSSION

In the current data, between the two peaks of FHE and HE, which could be considered as RT, it was difficult to determine which one is closer to RT. Positive correlation results between FHE and HE peaks provided a solution to the dilemma of deciding between the two peaks. As CS duration decreased, FHE and HE peak latencies became more aligned, resulting in fewer HE occurrences and increased monotony, similar to human responses. Basic properties of HE accuracy and frequency revealed that mice could reach and respond to the food cup despite varying CS durations. HE accuracy scores exceeded 80 % for all groups of 10 s, 5 s, and 2 s. HE frequency in short and shorter

CS durations was the same as in long CS. Group differences in pre-CS didn't affect the frequency of CS or HE peaks.

In addition to accuracy, the latency of RT accesses subjects' task perception[2]. In human studies[10,12], accuracy and RT are extracted and compared using the same data point. However, in animal studies, subjects may respond repeatedly, resulting in multiple RTs across multiple data points, forming an accuracy. RT may be a more sensitive indicator than accuracy because of their different mechanisms[47], which can fluctuate greatly depending on the researcher's instructions. Subjects can be asked to respond slowly or quickly with little effect on accuracy scores. In our study, all three groups responded well to each stimulus (Figure 5A), but FHE and HE peak latencies differed (Figure 4), implying that different CS durations reflected researchers' instructions on response timing. Longer CS durations gave subjects more freedom in response timing, whereas shorter durations may have caused anxiety about responding within a limited time frame. Thus, the shorter CS duration in our study represented the instruction for a human-like RT, in which FHE and HE peaks were nearly coincident in terms of latency and frequency, and accuracy rate and HE number were positively correlated over the training periods.

FHE and HE peaks demonstrated distinct 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 relationship with animals' habituated behaviors for the stimulus[30-34]. Conversely, the FHE peak showed rarely significant group differences over several days. The frequencies serve as reliability indicators, reflecting the overlapping responses of individual subjects at specific time points[27-29]. The non-significant group difference in HE peak frequency after D5 (Figure 4B) corresponded to saturation of HE frequency (Figure 5B). The results of the HE peak show that all subjects in the three groups adapted to the CS duration and performed well within each CS period. However, the FHE peak results show that the response of the longer CS group was dispersed even within the subject's data due to the low density of responses during the extended CS period when compared to the shorter CS groups. Group differences in FHE and HE revealed dissimilarity in latency and frequency. Based on our findings, we proposed that the 2-s CS duration, which displays synchronized FHE and HE peaks in the current experimental paradigm, is the best CS duration for inducing unbiased results in mice behavior.

The 2-s group had the highest frequency during the Pre-CS period, indicating that they checked the food cup more frequently during shorter CS periods. Furthermore, a linear relationship was found between HE frequency and CS duration during the Pre-CS period. However, it is worth nothing that behavior during the Pre-CS period did not influence behavior during the on-stimulus period, as evidenced by baseline-corrected results for HE peak frequency and latency during CS. Despite having

the highest number of HE occurrences during the CS period, the 10-s group had no significant differences in HE frequency compared to the 5-s or 2-s groups. A linear relationship between CS duration and response[43,44] was only found between the 5-s and the 2-s groups having significantly higher HE frequency than the 5-s group. In contrast to the linear trend in the Pre-CS period, the non-linear trend in the CS period suggests that CS duration plays a different role during on-stimulus and off-stimulus conditions.

HE frequency during the CS period in the 5-s group was significantly lower compared to the 2-s group. This implies that the 5-s and 2-s group's response patterns after the onset of the sound cue would most likely be similar. Interestingly, the latency and frequency of FHE and HE did not differ significantly between 5-s and 2-s groups, except for HE peak latency in D10 (Figure 4). The 5-s and 2-s groups showed a positive correlation between FHE and HE peaks. However, the 2-s group, which had a a stronger correlation between FHE and HE peaks, also showed a correlation between the number of HE and accuracy over training days. Meanwhile, the 5-s group's correlation results show that mice's late-day behaviors were subtle and arbitrary, similar to those of 10-s group. Furthermore, the 2-s group responded less than 1 per trial, whereas the 10-s and 5-s groups responded with more than 1 HE. These findings indicate that the 5-s and 2-s groups performed similarly; however, the 2-s group's mice learned the one-to-one matched relationship between stimulus and response more effectively. Furthermore, in the results of the second HE, mice that produced the first HE repeated it at least once for any trial, but this was seen only in the 2-s group out of all groups (Figure 7). Our findings show that mice in the 2-s group chose to stop responding during the time remaining until the US was presented following the first HE in the CS period. In contrast, the 5-s group demonstrated a steep linear increase in the number of subjects involved in the second HE beginning with D8 and eventually reaching all subjects producing second HEs by D12, as did the 10-s group. The synchrony between FHE and HE peaks, as well as between the number of HEs and accuracy in the 2-s group, was caused by the absence of the second HE. In our data, the prediction for the US included not only enhancement of response[48-50] but also both inhibition and promotion of response based on CS duration. We propose that the second HE, which is influenced by CS duration, can be interpreted as an indicator of impulsivity toward the US in our data.

We wanted to suggest that mice's RT was comparable to humans based on the relationship between FHE and HE peaks. However, it's important to note that results obtained with freely moving C57BL/6 mice may differ depending on the experimental paradigm, apparatus, and specific characteristics of the animals. In our paradigm, the shorter 2-s CS yielded a one-to-one matched relationship with response frequencies less than 0.5 Hz. However, if it involves licking behavior with higher-frequency responses in a head-fixed model, the relationship may change, despite the presence of synchronization between FHE and HE peaks. The one-to-one matched response would be tailored to our paradigm.

Nonetheless, our findings highlight the variety of personalities among animals and how this can be managed in terms of CS duration. Setting a time limit for responses can encourage genuine participation without hesitation, thereby improving the sincerity and reliability of behavioral data. Finally, our findings propose quantifiable criteria for evaluating the duration of stimulus in RT-related experimental paradigms, based on the harmony between FHE and HE peaks, as well as between the number of HEs and accuracy. If animals are capable, the shorter CS durations, which can barely elicit one response per trial, may be preferable for eliciting human-like responses with no repetition.

METHODS

Subjects

Twenty-four naïve male C57BL/6 mice, aged 7 weeks old, were provided by the Experimental Animal Center at Seoul National Dental University. Eight mice assigned to each group.
Prior to delay conditioning, each mouse was housed in its own cage and subjected to a food restriction schedule that kept their weight at 85% of their free-feeding weight throughout the study. There was unlimited access to water. The mice were kept in cages in a room with a light-dark cycle of 12 hours each (8 am–8 pm). Training began when the mice reached 8 weeks old and ended when they were 10 weeks old. The care and use of every animal 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 turned on from the start to the end of each experiment. A sound cue greater than 75 dB[51] was produced by a white noise amplifier (10–25,000 Hz, model ENV-325SW) and delivered via a cage speaker (model ENV-324W). The noise produced by fans (model ENV-025F) 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) connected to a syringe pump (model PHM-100A-EURO). HEs were counted whenever mice broke the beam of the detector (model ENV-303HDW) inside the liquid pipe. The experiments were run using the MED-PC V software (model SOF-735).

Stimulus

The experiment included two stimuli: CS and US. The CS consisted of white noise cues with varying durations: 10 s, 5 s, and 2 s (Figure 1). The selection of a CS duration of 10 s was influenced by

literatures[28,44,52,53] on CS duration. The 5-s and 2-s durations were determined using 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 current data (Figure S5C). To ensure that the CS durations were within ranges where mice could respond effectively, we examined the responses to the 10-s CS. We chose 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 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 lasted 1 s and occurred between 9 and 10 s after the onset of the 10-s CS, 4–5 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 they were also acclimated to handling and the US. On the last day of the pre-conditioning phase, the mice had a session in the chamber without the CS but with the US. Following pre-conditioning, the delay conditioning experiments were repeated daily for a total of 12 days (Figure 1). Each day, ten trials were conducted during the experiment. Inter-trial intervals varied and were randomly set to 90, 120, or 250, as illustrated in Figure S5A. The total experimental time per subject was about an hour per day, including preparation time. During the experiments, pumps were used to deliver condensed milk from syringes to food cups, with the pumps turned on during the US period. The mice were free to move around the chamber during the experiments. In this study, behavior directed toward the food cup to consume the reward food was referred to as "HE." The timing of HEs entering the food cup was recorded with a temporal resolution of 10 ms.

Analysis

The data was analyzed in three groups of 24 mice, each with a duration of 10, 5, and 2 s. For the analysis of HE frequency, HEs were extracted from the CS and pre-CS time windows for 10 trials, respectively, and the number of HEs was normalized per s. We carefully selected the time window of Pre-CS as an off-stimulus baseline to compare to the CS of the on-stimulus. Figure S6B shows 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 converted to percentages (where 1 correct trial equates to 10 %). The number of HE was calculated using the total number of HEs during each CS duration. The FHEs were extracted from the first HE in each trial (Figure S1A). The FHE and HE peaks were calculated by extracting maximum values from a matrix in which the numbers of HEs occurring within a 1-s epoch were recorded as frequency levels (Figure S1B and S6). The maximum frequency value for each peak was 10 times, the trial number per day. The latency of the peaks corresponded to the time of maximum frequency. Based on this, the latencies of the two peaks did not include the

incorrect HEs, as RT is typically based on correct answers[54,55]. The time windows for estimating FHE and HE peaks were each set from 0 to 10 s after the onset of CS in all groups (Figure S7). The baseline-corrected HE frequency and peak were calculated by subtracting the mean HE in the pre-CS period from the HE in the CS (Figures S2 and S8).

The statistical analysis revealed that all data did not follow a normal distribution. Therefore, nonparametric analyses were used for all statistical procedures. Nonparametric paired t-tests between FHE and HE peaks were performed using the Wilcoxon Signed Ranks test. Nonparametric correlations between each pair of FHE and HE peaks, as well as the number of HEs and accuracy, were determined using the Spearman test. All three groups showed a correlation coefficient value of 1, indicating that the FHE and HE peaks coincided (Table S3). Notably, the conversion of a decimal to an integer during extraction of the two peaks from HEs had a significant effect on the highest correlation coefficient values (Figure S6). Nonparametric ANOVAs were conducted using the Kruskal–Wallis test. Post-hoc analyses for ANOVAs were conducted using the Mann–Whitney test. Type I error caused by 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).

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

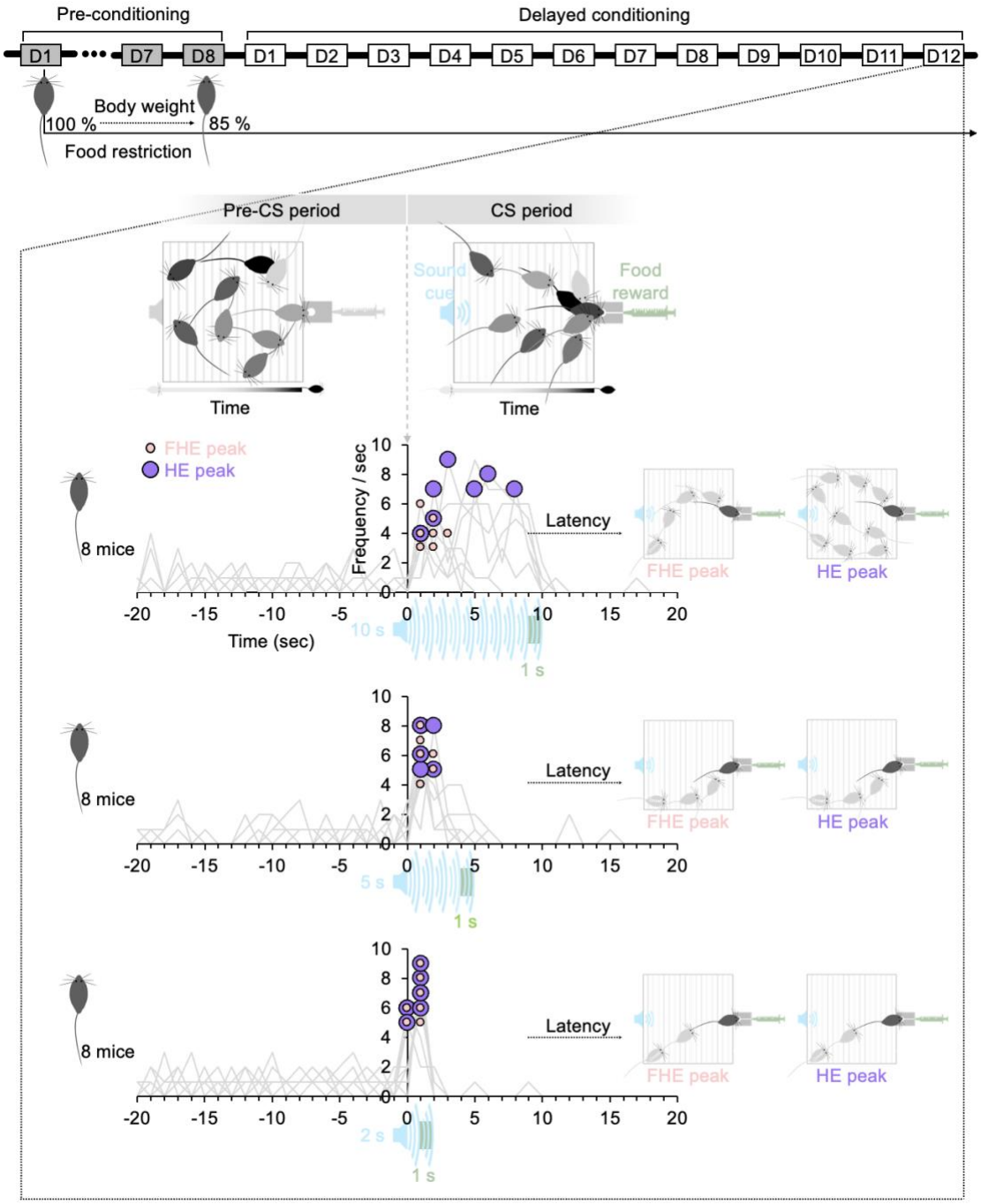

Figure. 1. Model of FHE and HE peaks for human-like RT. Mice in all groups lost up to 85% of their body weight by restricting their food intake before delay conditioning. During the CS period, the duration of the sound cue followed by food reward varied among the 10, 5, and 2-s groups. On the final training day (D12), the HE latency in the CS period following the onset of the sound cue was higher than in the pre-CS period preceding the onset of the sound cue. The FHE peaks occurred before the HE peaks in the 10-s group, whereas in the 5-s and 2-s groups, the two peaks were nearly synchronized. The latencies of the two peaks show how mice in each group respond differently to sound cues with varying durations.

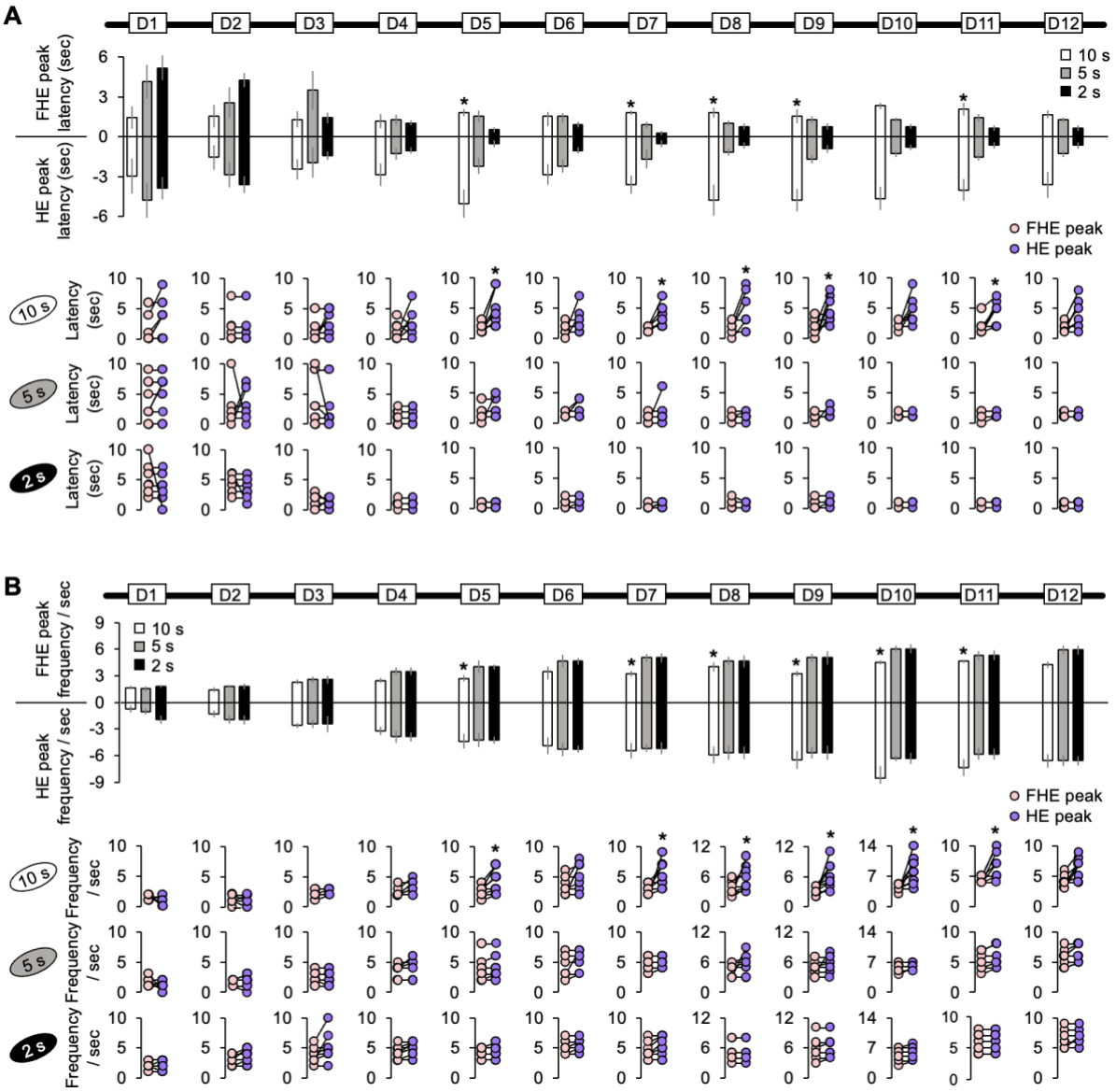

Figure. 2. Difference between FHE and HE peaks in latency and frequency. (A) Only the 10-s group showed significant differences in latencies between FHE and HE peaks (Wilcoxon Signed Ranks test, $P < 0.05$). The latency for the HE peak was greater than for the FHE peak. However, no significance was found in the 5-s and 2-s groups (Wilcoxon Signed Ranks test, $P > 0.05$). (B) Only the 10-s group showed significant differences in the frequencies of FHE and HE (Wilcoxon Signed Ranks test, $P < 0.05$). The frequency of the HE peak was higher than for the FHE peak. *, $P < 0.05$, **, $P < 0.01$, and ***, $P < 0.001$.

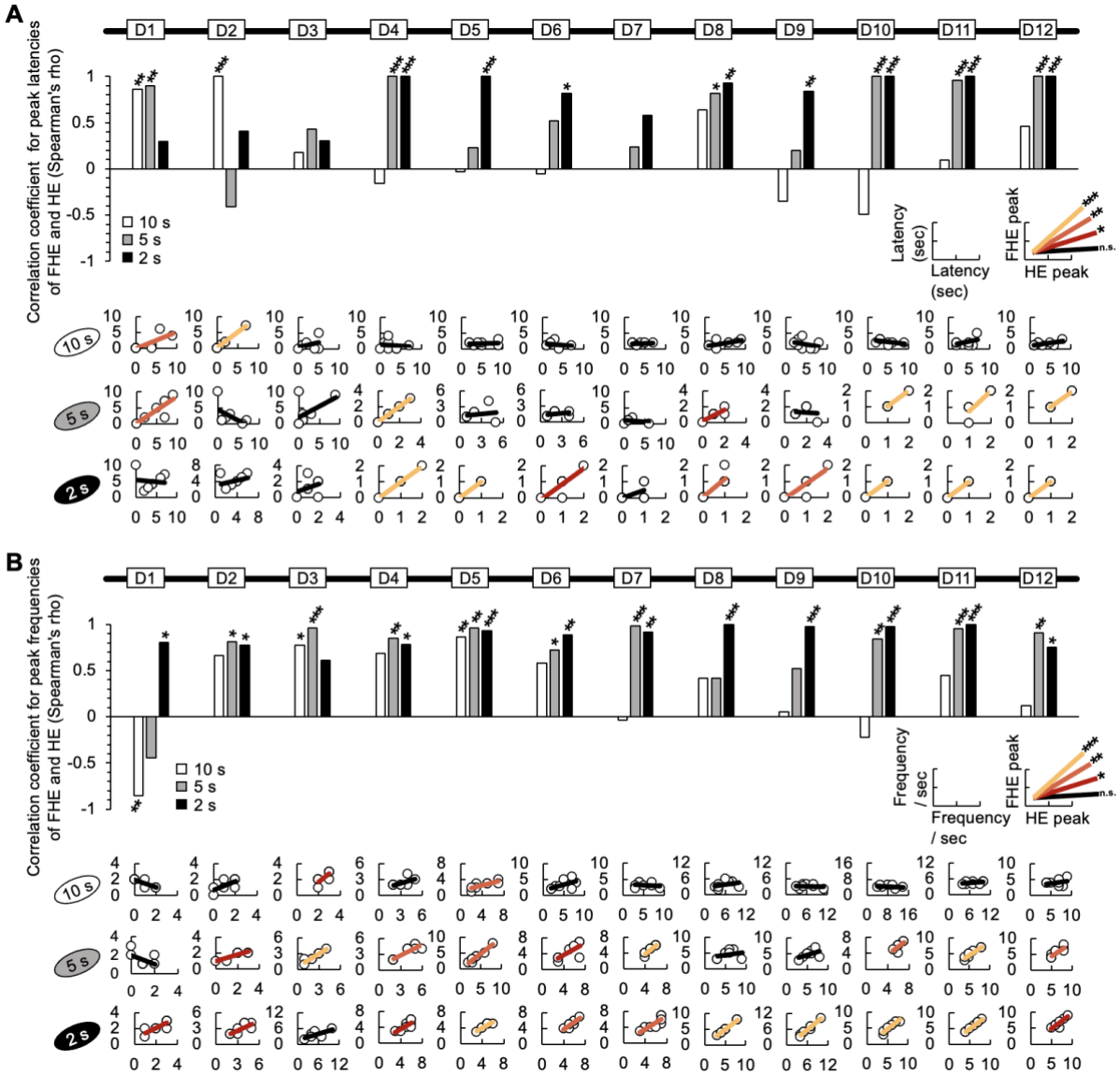

Figure. 3. Correlation between FHE and HE peaks in latency and frequency. (A) The latencies of FHE and HE peaks were strongly and positively correlated in later training days for the 2-s group (D4–D6 and D8–D12). Positive correlations were observed in the later days of the 5-s group (D1, D4, D8, D10,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

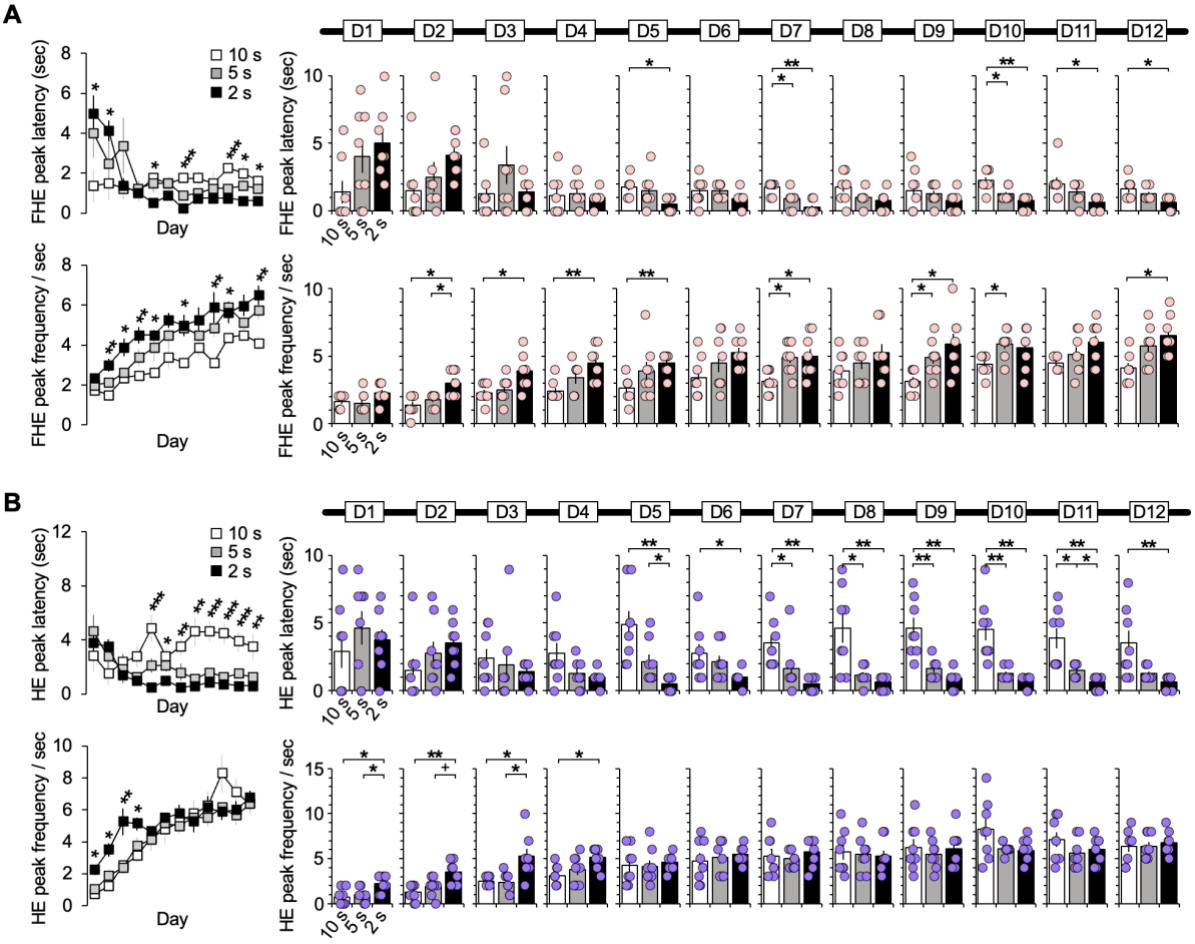

Figure. 4. Difference between groups in latency and frequency of FHE and HE peaks. One-way ANOVA results for FHE and HE peak latency and frequency revealed a significant effect of Group factor (10 s, 5 s, and 2 s). (A) The FHE peak demonstrated a significant Group effect across training days for both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

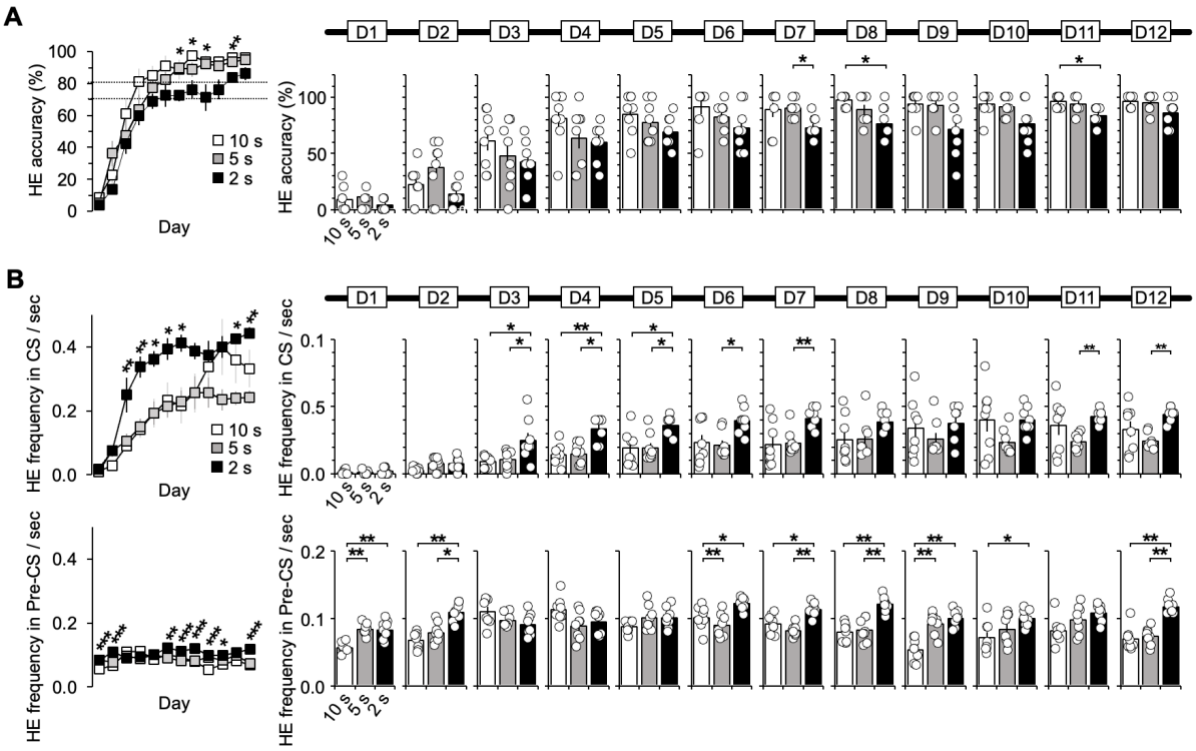

Figure. 5. Basic properties of HE. (A) One-way ANOVA results revealed that the Group effect for HE accuracy was only present on D7, D8, and D11. However, all subjects eventually performed well on the tasks, with accuracy above the mean of 80%, regardless of CS duration. (B) One-way ANOVA found a significant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

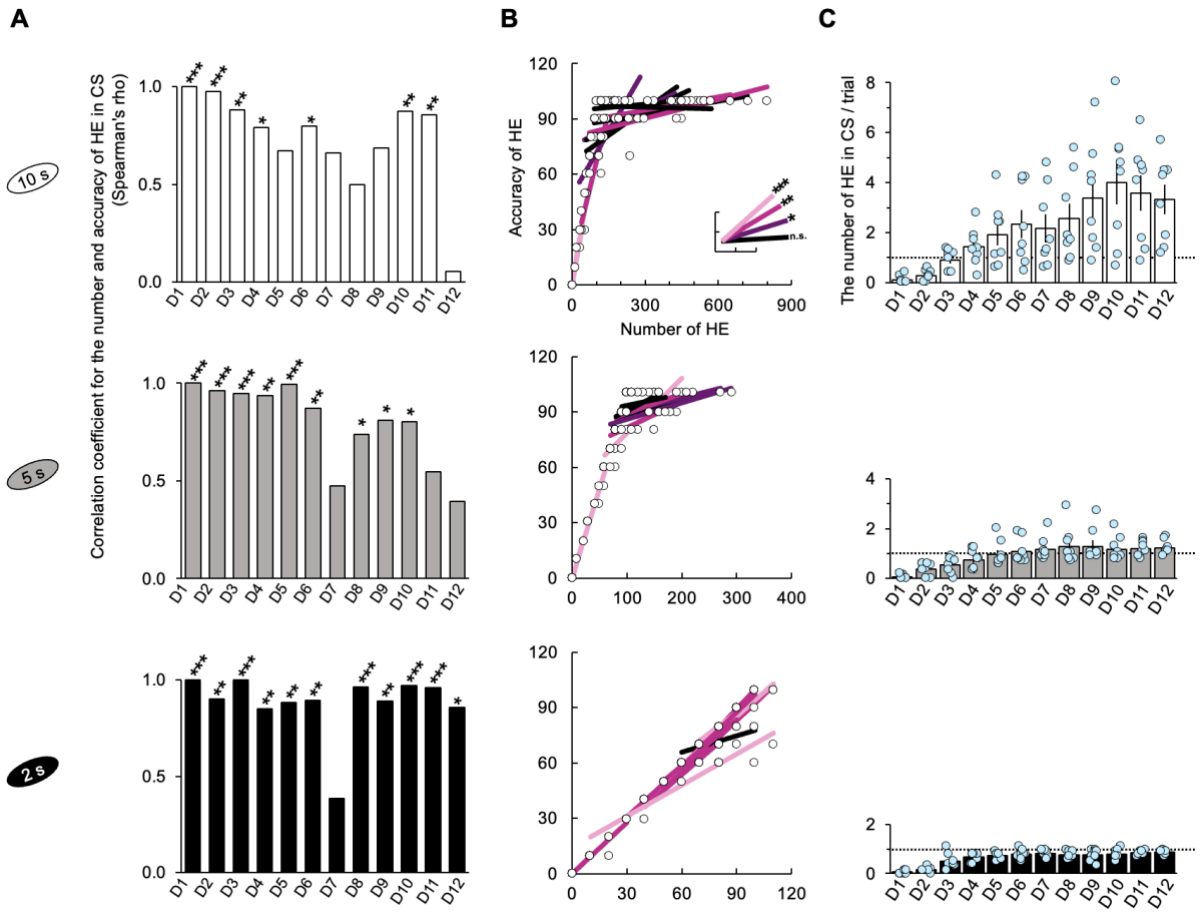

Figure. 6. The number of HE and correlation with accuracy. (A) There were significant positive correlations between the number of HEs and accuracy across all groups. For the 2-s group, these significances were consistent across all training days except D7, but not in the other groups. The nonparametric correlation was assessed using Spearman's test. (B) The graphs included the lines of best fit and data points that correspond to the correlation coefficients in (A) and show the trend in the data for each day (Figure S3). The scale of the x-axis was determined by the maximum value observed across all training days. In the case of the 2-s group, many data points overlap, resulting in a small representation of the 96 points collected from 8 subjects over 12 days. (C) The 10-s group had the highest number of HEs. In the 2-s group HE occurrences were below 1, while in the 5-s group, HE slightly exceeded 1. The dotted line represents the number of HE that corresponds to the ono-to-one matched relationship between stimulus and response. *, $P < 0.05$; **, $P < 0.01$; and ***, $P < 0.001$.

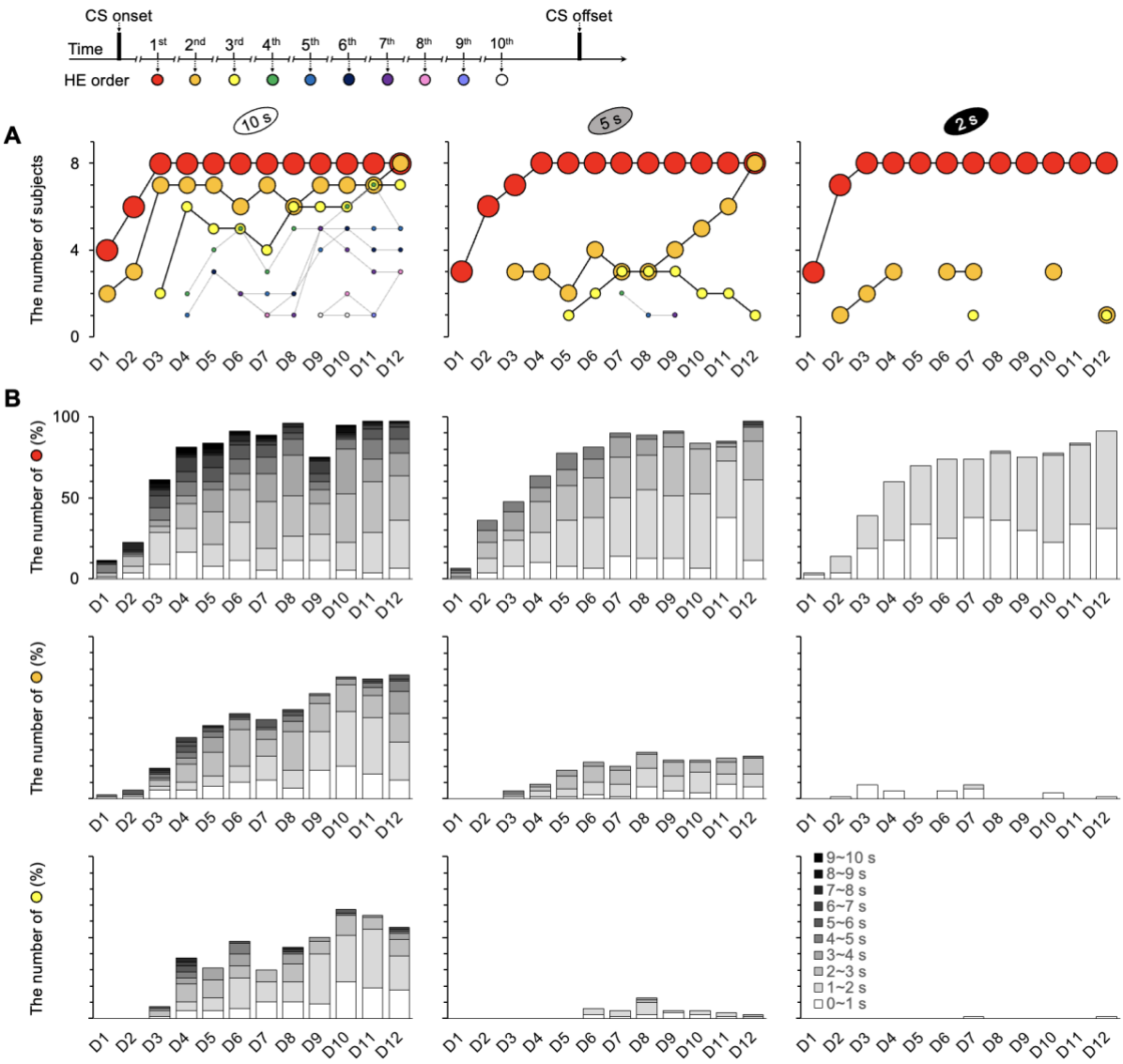

Figure 7. The number of repeating HEs and participating subjects per trial. Repetitions of HEs were observed up to 10 times across all trials and all training days. (A) In the 10-s group, responses were recorded until the tenth HE. Meanwhile, the maximum observed in the 5-s group was the seventh HE, while in the 2-s group, it was the third HE. The pattern of the first HE was consistent across all groups; however, differences emerged in the second and third HEs. Specifically, the second HE increased toward 8 mice on D12 in 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, which had the same number of participating subjects as the 5-s group, was faster than the 5-s group, with intervals ranging from 0 and 1 s in (B); however, it did not increase as the 5-s group toward D12. (B) The latencies between HEs varied with CS durations. The interval latency was longest in the 10-s group and shortest in the 2-s group. For results from the fourth through tenth HEs which are connected by light gray lines (Figure S4B).