



OPEN A simple method that doubles learning speed for mice in touchscreen-based visual discrimination

Yuji Kiyama¹, Yusuke Suzuki^{2,5}, Misako Haraguchi¹, Yukimura Oe^{1,3}, Yukihisa Daitoku^{1,3}, Natsumi Higa¹, Yoshihiko Irie^{1,3}, Takeru Suzuki^{1,4}, Itaru Imayoshi^{2,5}, Masaki Kakeyama⁴ & Hiroyuki Okuno¹✉

Touchscreen-based operant learning systems are widely used in behavioral neuroscience to assess cognitive function across various animal models, including mice. However, conventional training protocols for visual discrimination (VD) tasks in mice often require extended training periods. We developed modified touchscreen training protocols that enhance learning efficiency in VD tasks. Our method requires mice to perform two or three consecutive screen touches to confirm their stimulus selection. If inconsistent responses occur during these touches, the trial is reset. This approach significantly improved task acquisition speed. Additionally, we incorporated visual feedback after incorrect choices, which further enhanced learning speed and task accuracy. Using this protocol, mice rapidly acquired both initial VD learning and reversal learning, even with multiple randomly presented stimulus pairs. Furthermore, in serial reversal tasks, mice exhibited improved performance across sessions, indicating rule acquisition and cognitive flexibility. These findings demonstrate that simple protocol adjustments can significantly improve touchscreen-based cognitive assessments in mice, enabling more cognitively demanding behavioral experiments in preclinical research.

Keywords Operant learning, Visual discrimination, Cognitive flexibility

Behavioral analysis using animal models is essential for investigating cognitive functions under both normal and psychiatric or neurological conditions^{1–3}. Among these models, mice are indispensable due to the extensive molecular and anatomical data available and the vast stocks of genetically modified lines. In recent years, touchscreen-based operant task apparatuses have gained popularity for studying mouse behavior⁴. These systems allow for extensive cognitive function tests, including paired-association tasks^{5,6}, position and brightness discrimination⁷, gambling tasks⁸, behavioral sequencing tasks⁹, match-to-position tasks¹⁰, and delayed match-to-position tasks¹¹. They are also commonly used to assess perceptual learning and behavioral flexibility through visual discrimination (VD) learning and subsequent reversal learning^{4,12}.

Several strategies have been explored to improve VD learning efficiency. One approach involves optimizing visual stimulus pairs by increasing shape and contrast differences to enhance discriminability^{13–15}. Additionally, introducing natural scene stimuli has been shown to accelerate learning in VD tasks¹⁶. However, using stimuli with substantial differences in shape, color, or contrast can introduce confounding effects due to the subject's natural preferences, which are difficult for experimenters to control^{11,17}. Another method involves incorporating correction trials following errors, which helps to re-learn the correct choice^{18–20}. However, error correction introduces variability in the number of trials per session, as it depends on the frequency of errors made by the subject, making inter-animal comparison challenging. Therefore, alternative approaches are needed that can enhance VD learning while maintaining controlled experimental conditions.

¹Department of Biochemistry and Molecular Biology, Graduate School of Medical and Dental Sciences, Kagoshima University, 8-35-1 Sakuragaoka, Kagoshima-shi, Kagoshima 890-8544, Japan. ²Laboratory of Brain Development and Regeneration, Graduate School of Biostudies, Kyoto University, Kyoto 606-8507, Japan. ³Department of Anesthesiology and Critical Care Medicine, Graduate School of Medical and Dental Sciences, Kagoshima University, Kagoshima 890-8544, Japan. ⁴Laboratory for Environmental Brain Sciences, Faculty of Human Sciences, Waseda University, Tokorozawa 359-1192, Japan. ⁵Center for Living Systems Information Science, Graduate School of Biostudies, Kyoto University, Kyoto 606-8507, Japan. ✉email: okuno@m.kufm.kagoshima-u.ac.jp

This study presents a simple yet robust method for improving VD task learning efficiency by modifying the task protocol used to assess stimulus choice. We serendipitously discovered that requiring multiple touches, rather than a single touch, for stimulus choice significantly accelerated learning speed and increased the highest performance levels in VD tasks, even without the use of natural scenes or error correction trials. Using this approach, we successfully trained mice to perform multiple reversal VD learning tasks with luminance- and contrast-controlled geometrical stimuli.

Results

Learning of the 1-pair VD task and its reversal in a standard single-touch protocol

To assess the ability of mice to perform VD and reversal learning tasks, they were first trained on a 1-pair VD task using a conventional choice protocol (single-touch choice, 1 T) after pre-training (Fig. S1A). Following the pre-training phase, mice underwent 15 sessions of initial learning training. In each trial, a pair of vertical and horizontal stripe patterns (white and black, repeated ten times) was presented on the screen, with the vertical stripe serving as the rewarded stimulus (S+) and the horizontal stripe as the nonrewarded stimulus (S-) during initial learning (Figs. 1A and S1B). Mice selected one of the two stimuli by touching either stimulus using their nose or forelimbs. Under the 1 T protocol, a single touch to the S+ stimulus resulted in a “Correct” judgment with a reward, while a touch to the S- stimulus led to a “Error” judgment, no reward, and a 10-second intertrial interval (ITI) before the next trial (Fig. 1A). Consistent with previous findings^{4,12}, mice successfully learned the initial VD task, achieving an across-animal average of correct responses exceeding 80% by the 12th session (Fig. S1C). After completing 15 sessions of initial learning, the mice were trained in a reversal learning paradigm, where the S+ and S- stimuli were reversed (Fig. S1B). Mice successfully learned the reversal task, reaching an average of over 80% correct responses by the 17th session after reversal (32nd session, including initial learning sessions) (Fig. S1C). These results indicate that mice are capable of both initial and reversal VD learning using the standard single-touch protocol, but the training duration required to complete the VD tasks was substantial in our training protocol. Previous studies have reported that error correction trials and the use of attractive liquid rewards, both of which were not employed in the current study, can accelerate VD learning^{21,22}.

The multiple-touch requirement improved learning efficiency

During VD training, mice occasionally exhibited impulsive touches or unintended contacts to the screen without carefully observing the stimuli (SI Movie S1). Additionally, mice often nose-poked the reward pod even after making incorrect choices, suggesting poor awareness of failure. These observations suggested that modifying the choice detection algorithm could help reduce these errors and improve the learning efficiency of the VD task. To test this hypothesis, a multiple-touch requirement was introduced in the task algorithm (Fig. 1B and Fig. S1D). In the two-successive-touch protocol (2T), mice were required to touch the same panel twice in succession to confirm their choice (Fig. 1B). Training under this protocol resulted in better learning efficiency compared to the 1 T group (Fig. 1C). To further explore the impact of touch repetition, the three-successive-touch protocol (3T) was introduced, requiring mice to touch the same panel three times in succession before confirming a choice (Fig. S1D). Mice trained with the 3 T protocol exhibited enhanced learning efficiency, similar to those in the 2 T group (Fig. 1C), and reached the performance criterion significantly faster than those trained with the 1 T protocol ($p < 0.05$, Kruskal–Wallis test with Dunn’s multiple comparison) (Fig. 1D). These results indicate that mice trained under the 3 T protocol learned the 1-pair VD task significantly faster than those trained under the 1 T protocol. Additionally, the highest performance levels during the initial training sessions tended to be higher in mice trained under the 3 T and 2 T protocols compared to those under the 1 T protocol, although the difference did not reach statistical significance (Fig. 1E). These findings demonstrate that implementing a multiple-touch requirement for stimulus choice substantially improved the efficiency of initial learning in the VD task.

Introduction of visual feedback further improves learning efficiency

The 3 T protocol was further modified by introducing visual feedback, where white squares appeared on both sides of the windows after a false choice (Fig. 2A). The thickness of the stripes in the vertical and horizontal stimuli was increased to five black-and-white cycles per stimulus. This modified version was termed the modified three-successive-touch (m3T) protocol. Mice were trained in the 1-pair VD task for initial learning, followed by reversal learning, using the m3T protocol (Fig. 2B). During initial learning, the across-animal average of correct responses exceeded 80% by the 6th session (Fig. 2B), showing better performance than the 3 T protocol (Fig. 1C). During reversal learning, the across-animal average of correct responses exceeded 80% by the 7th session, similar to the initial learning phase (Fig. 2B). The number of sessions required to meet the criterion was significantly lower in the m3T group than in the 1 T group for both initial and reversal learning ($p < 0.05$ for initial learning and $p < 0.01$ for reversal learning, Mann–Whitney U test) (Fig. 2C). The highest performance across sessions appeared higher in mice trained with the m3T protocol compared to those trained with the 1 T protocol in the initial learning phase (Fig. 2D left), though the difference did not reach statistical significance. These results indicate that requiring multiple touches for stimulus choice, combined with visual feedback after error choices, improved learning speed in both initial and reversal VD tasks. Both males and females learned the task equally with the m3T protocol (Fig. S2) (main effect of sex, $p = 0.751$; interaction between sex and session, $p = 0.7136$), indicating no sex difference in task performance.

We then analyzed the latency to nose poke following a choice. After correct choices, the average latency was within 3 seconds under both the 3 T and m3T protocols (Fig. S3A). In contrast, following incorrect choices, mice trained with the m3T protocol showed significantly longer latencies compared to those trained with the 3 T protocol (Fig. S3B). Moreover, the proportion of trials in which mice appropriately withheld nose poking during the 10-second ITI after incorrect choices was significantly higher with the m3T protocol than with the

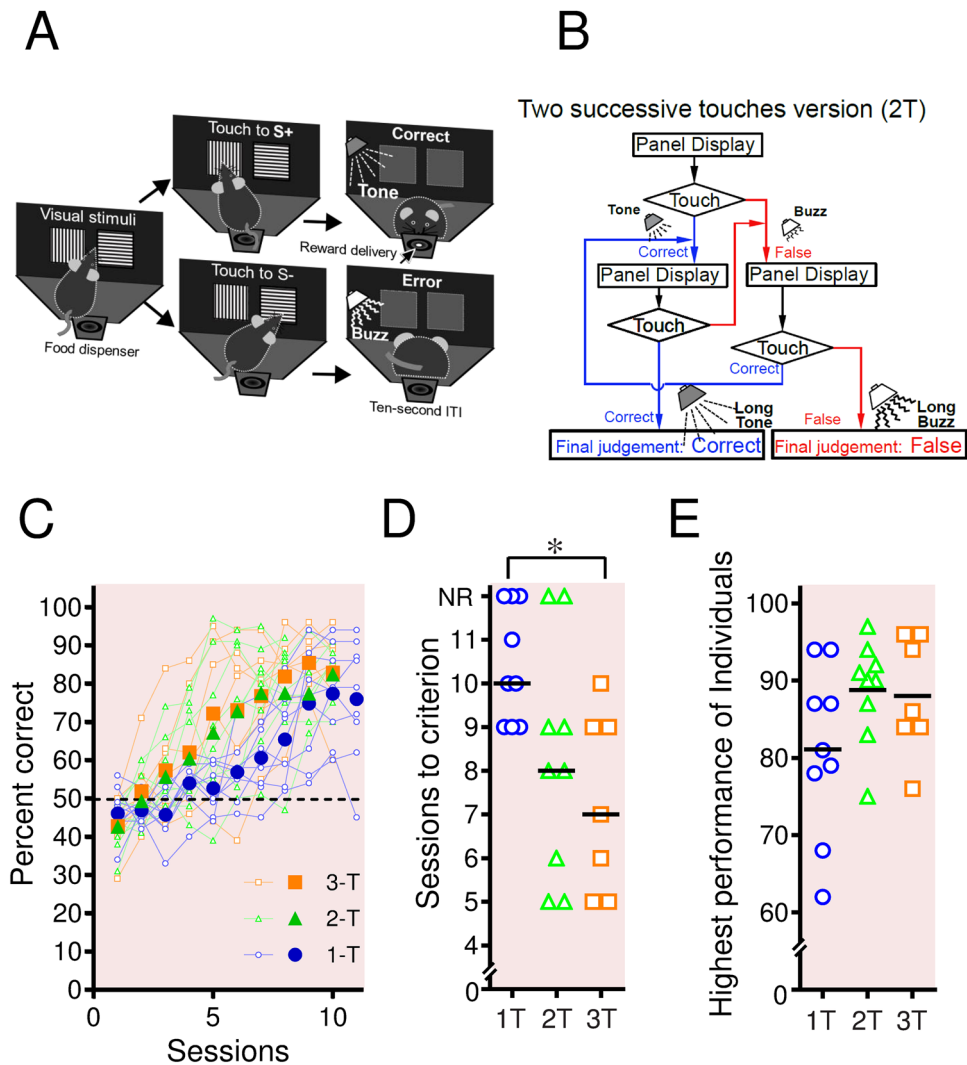


Fig. 1. Improved performance in the 1-pair VD task with multiple-touch protocols. **(A)** Behavioral sequences in a trial of the 1-pair visual discrimination (VD) task. The mouse selects and touches one of two stimulus patterns displayed in separate windows. A correct choice (S+) results in a reward pellet and a tone, while an error choice (S-) triggers a buzzer and a 10-s intertrial interval (ITI) before the next trial. **(B)** The task flowchart for the two-touch (2T) protocol. Mice were required to touch the same stimulus twice in succession to confirm a choice. **(C)** Task performance under 1T, 2T, and 3T protocols, showing the across-animal percentage of correct responses over training sessions. Small pale symbols and lines represent individual animals' performance, while large symbols represent the across-animal average. **(D)** Number of sessions required to reach the criterion (>75% correct for two consecutive sessions) for individual mice. NR denotes "not reaching the criterion" by Session 11. Horizontal bars represent median values. * $p < 0.05$ by Dunn's multiple comparisons test. **(E)** Highest performance achieved by individual mice across training sessions. Horizontal bars indicate mean values. $N = 9$ (1T), 9 (2T), and 7 (3T) mice.

3T protocol (Fig. S3C; see also SI Movie S2). These results suggest that mice trained with the m3T protocol may have been aware of their errors, which likely contributed to their faster learning.

We next evaluated the side bias of stimulus choices in the 1T and m3T protocols (Fig. S4). To quantify individual tendencies toward choosing one side over the other, we calculated the absolute side bias index (ASBI). A session with a completely spatially biased response results in an ASBI of one, while a perfectly spatially balanced session scores zero (see legend of Fig. S4 for details). We observed that ASBI tended to temporarily increase during the early to middle stages of learning, but gradually decreased as task performance improved in the later stage (Fig. S4A). These observations suggest that the mice tended to use a location-biased choice strategy to solve the task in the early stage of learning, but gradually adopted a choice strategy based on visual information as intended. We further assessed whether mice had a general preference for the left or right side in our experimental setup (Fig. S4B). While some individual mice showed a preference for either the left or right side, there was no overall side bias at the group level. Thus, our experimental setup did not cause unintended

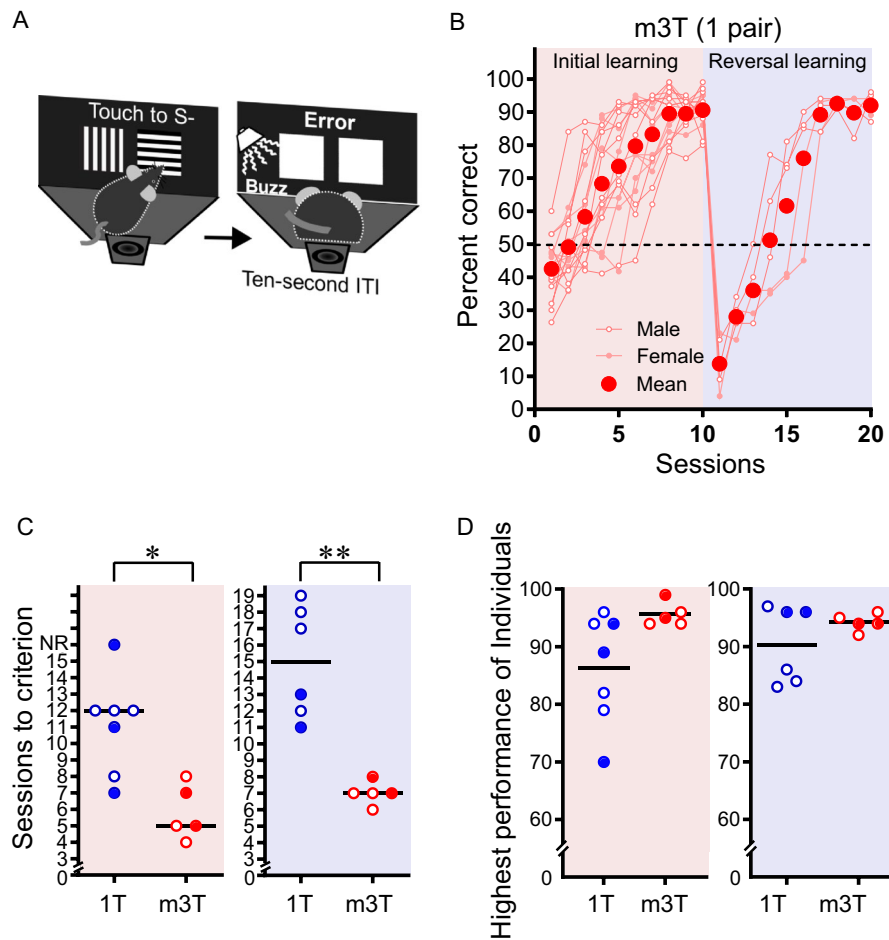


Fig. 2. Enhanced VD learning with the modified three-touch protocol (m3T). **(A)** Visual feedback mechanism introduced in the m3T protocol, displaying white squares on the screen after error trials. A buzzer and ITI were applied as in other protocols. **(B)** Percentage correct in initial and reversal VD tasks with the m3T protocol. Small dots and lines represent individual mice. **(C)** Number of sessions required to reach the criterion for initial (left) and reversal (right) VD learning using 1T and m3T protocols. NR denotes “not reaching the criterion” by Session 15. Horizontal bars indicate median values. ** $p < 0.01$; * $p < 0.05$ (Mann-Whitney U test) **(D)** Highest performance of individual mice during initial (left) and reversal (right) learning. Horizontal bars indicate the mean values. Data for 1T were replotted from Figure S1C. Bordered open small symbols represent individual males, while filled small symbols represent individual females. Large circles indicate the mean across animals.

spatial preferences, suggesting that the animals’ task performance fairly reflected their ability to learn the VD task.

Since the use of the m3T protocol was found to significantly accelerate learning, we attempted to train mice on a transverse patterning task, which requires discrimination of overlapping stimulus pairs based on their relational structure^{23,24} (Fig. S5A–C). In Phase 1, the mice were trained with three geometric pattern stimuli (Fig. S5B). Contrary to our expectation, task performance remained at chance level even after 17 sessions (Phase 1, Fig. S5D). In Phase 2, we replaced the geometric patterns with natural scene images, which have been reported to facilitate VD learning¹⁶ (Phase 2, Fig. S5D) Again, the mean percent correct stayed around 50% over an additional 10 sessions, increasing only after extensive training over 30 sessions. These findings indicate that, even with the m3T protocol, mice were virtually unable to learn the transverse patterning task. Importantly, however, the results of the transverse patterning task served as a control for the auditory touch feedback implemented in our multiple-touch protocols (Figs. 2B, S1D and S5C). In principle, with auditory feedback distinguishing correct from incorrect touches in the m3T protocol, mice could solve the touchscreen VD task without relying on visual information, instead using an auditory win-stay/lose-shift strategy. The above observations, however, indicate that this was not the case.

In parallel, we confirmed that mice did not primarily rely on feedback tones to guide their choices in the m3T protocol. To test this, we modified the m3T protocol so that the same auditory feedback sound was used for both correct and incorrect intermediate touches (Fig. S6A). In this new protocol, termed m3Tn, only visual information was available for solving the VD task. We trained mice on the 1-pair VD task using the m3Tn

protocol (Fig. S6B) and found that the across-animal average of correct responses exceeded 80% by the 5th session (Fig. S6C), showing performance similar to that with the m3T protocol (Fig. 2B). The number of sessions required to reach criterion in the m3Tn protocol (Fig. S6C) was significantly lower than in the 1 T group shown in Fig. 2 ($p = 0.048$, Kruskal–Wallis test with Dunn’s multiple comparisons test) and not significantly different from the m3T group in Fig. 2 ($p > 0.9999$). The highest performance across sessions was also comparable between the m3Tn and m3T protocols (Fig. S6E and Fig. 2D).

Altogether, these results suggest that the impact of auditory feedback for intermediate touches on VD learning acquisition is minimal in our multiple-touch protocols. Nonetheless, because auditory feedback may still offer some beneficial aspects for the VD task, we employed the m3T protocol rather than the m3Tn protocol for further experiments.

Modeling of the modified protocols recapitulates the improvement of VD learning

To determine whether the observed improvements in VD learning could be explained mathematically, a Q-learning model, known to effectively describe animal behavior and dopaminergic neuron activity during reinforcement learning²⁵, was applied to the VD paradigm (see Methods). The simulation results for both initial and reversal learning closely reflected the performance of mice under different training conditions (Fig. 3A).

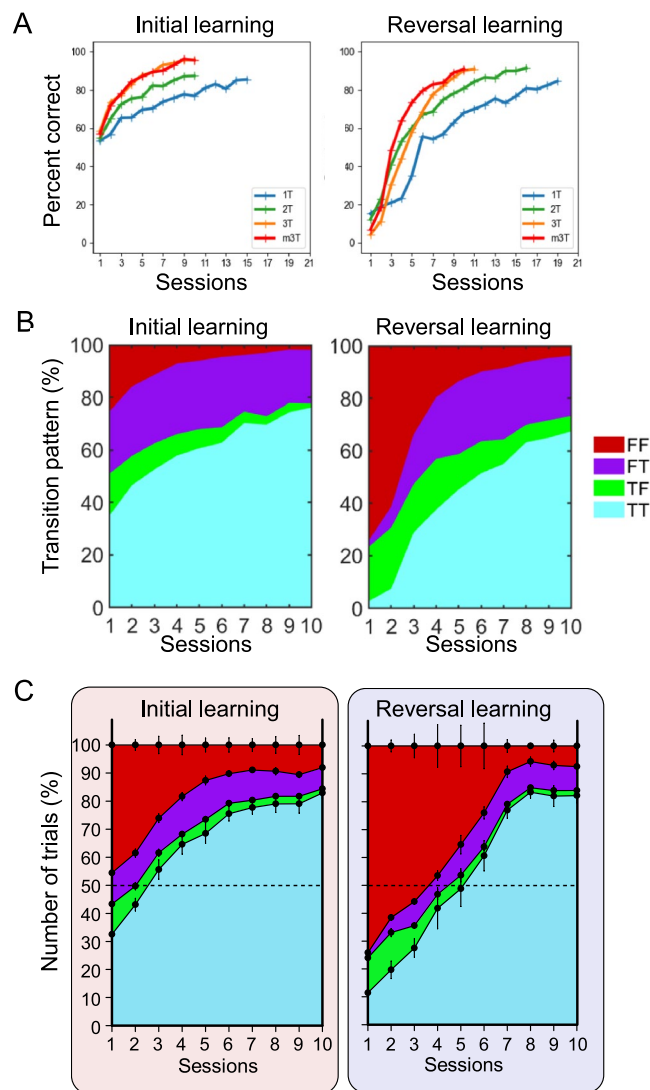


Fig. 3. Q-learning-based modeling and simulation of multiple-touch protocols. **(A)** Simulation of VD learning using the Q-learning model under multiple-touch conditions. **(B)** Changes in stimulus touch patterns during VD learning in the simulation. T-T: Both the first and final stimulus touches were correct. T-F: The first touch was correct but the final touch was incorrect. F-T: The first touch was incorrect but the final touch was correct. F-F: Both the first and final touches were incorrect. **(C)** Observed stimulus choice patterns in the m3T cohort during initial and reversal learning. Data were obtained from the same cohort as in Figure 2B. Color cues are the same as those in **(B)**. Error bars represent SEM.

The simulation also provided insight into the order of stimulus touch patterns in correct and error trials under the multiple-touch protocols. For simplicity, trials were categorized into four patterns: T-T, T-F, F-T, and F-F. The T-T pattern refers to trials where both the first and final touches were correct, while the T-F pattern represents trials where the first touch was correct but the final touch was incorrect. The F-T pattern corresponds to trials where the first touch was incorrect but the final touch was correct, whereas the F-F pattern indicates trials where both the first and final touches were incorrect. Changes in the proportion of these four patterns across sessions were analyzed in the m3T model during initial and reversal learning (Fig. 3B). During initial learning, the T-T pattern gradually increased, while the F-F pattern rapidly decreased. The T-F pattern declined and almost disappeared by the end of the initial learning phase, whereas the F-T proportion remained relatively stable. A similar trend was observed in the reversal learning sessions, where the T-T pattern increased, the F-F pattern decreased, and the T-F and F-T patterns exhibited changes comparable to those seen in initial learning.

A comparable analysis of experimental data from the m3T experiment in Figure 2B revealed that changes in the four touch patterns closely resembled the simulation results (Fig. 3C).

These findings suggest that Q-learning-based neural states may be implemented in the mouse brain to support initial and reversal VD learning under multiple-touch protocols.

Multiple-touch protocols are effective for 3-pair VD learning and its reversal

The improvement of VD learning through multiple-touch protocols in animal experiments and modeling suggests a reduction in the cognitive burdens associated with the task. If these protocols alleviate cognitive demands, mice should be able to perform VD learning efficiently even under more complex conditions. To confirm this, a three-pair VD task (3-pair VD) was introduced, in which three pairs of visual patterns were randomly presented within a session. Mice were first trained in the 1-pair VD task, followed by 3-pair initial VD learning and its reversal (Fig. 4A). Each session in the 3-pair VD task consisted of 150 trials, with 50 trials per stimulus pair. As expected, mice trained with the improved protocols exhibited faster learning during both the initial and reversal phases of the 3-pair VD task (Fig. 4B). The across-animal average of correct responses exceeded 80% by the 10th session for 2T and 3T and by the 6th session for m3T, while mice trained with 1T failed to reach 80% correct responses during the initial learning phase (Fig. 4B, left). A similar trend was observed in reversal learning, where the across-animal average of correct responses exceeded 80% by the 19th session for 1T, the 17th session for 2T, the 12th session for 3T, and the 11th session for m3T (Fig. 4B, right). The number of sessions required for individual mice to meet the performance criterion significantly differed across the four protocols. Mice in the m3T group required significantly fewer sessions than those in the 1T ($p < 0.01$) and 2T ($p < 0.05$) groups during initial learning, as well as fewer sessions than the 2T group ($p < 0.01$) during reversal learning (Fig. S7A, Kruskal–Wallis test followed by Dunn's multiple comparisons test). The highest task performance across sessions in the 3-pair initial learning was significantly higher in the 2T, 3T and m3T groups compared to the 1T protocol. A similar trend was observed in reversal VD learning (Fig. S7B). These findings indicate that incorporating multiple touches and visual feedback into the task protocol enhances learning efficiency and performance plateau levels even under cognitively demanding conditions in the 3-pair VD task.

Serial multiple-reversal paradigms assess different aspects of cognitive flexibility beyond those evaluated in single reversal paradigms. These include the formation of learning sets with original and reversed rules and the ability to switch between learning-rule sets^{26,27}. However, one of the primary challenges in serial reversal-task experiments is the large number of training sessions required when multiple reversals are introduced. To examine whether the efficient m3T protocol could overcome this issue, training was continued with the same mouse cohort that had undergone initial and reversal learning in the 3-pair VD task (Fig. 4C). Following the first reversal (Rev1), the mice proceeded to the second reversal (Rev2), in which the S+ and S- stimuli were reversed back to their original assignments from the initial learning phase. This process was repeated ten times (Rev10), with each reversal occurring once all mice reached the criterion. While Rev1 required 15 sessions, the number of sessions gradually decreased with each subsequent reversal, and by Rev10, the reversal was completed in only eight sessions (Fig. 4C). The total training period, including initial learning and ten reversals, lasted fewer than 120 sessions, whereas in pilot experiments using the conventional one-touch protocol, only four reversals were completed within the same period (data not shown).

The averaged performance in the first session after each reversal improved as the number of reversals increased ($F_{9,63} = 10.09$, $p < 0.0001$; RM 1-way ANOVA) (Fig. S7C). The number of sessions required for individual mice to reach the criterion significantly decreased over successive reversals ($F_{9,63} = 12.77$, $p < 0.0001$; RM 1-way ANOVA) (Fig. S7D). These results indicate that repeated VD reversals using the m3T protocol enabled mice to rapidly switch between task rules, demonstrating improved cognitive flexibility over time.

Divergence of internal states across single- and multiple-touch protocols

The multiple-touch protocols likely altered the “internal states” of mice, leading to the observed improvements. Here, the internal state refers to the model parameters in conjunction with measurable behavioral phenotypes. To examine this, the learning rate (β), decay constant (α), and perseverative traits (τ) were estimated for each protocol using the Q-learning-based mathematical model. The simulated learning curves generated using these parameters closely replicated the observed curves in both initial and reversal learning phases under the 3-pair conditions (Fig. S8A). When comparing parameters across protocols, α remained virtually consistent (Fig. S8B), while β was higher in the 3T and m3T protocols than in the 1T and 2T protocols (Fig. S8C). This suggests that mice trained with 3T and m3T updated state-action values more efficiently, reducing prediction errors. The value of τ followed a decreasing trend from 1T to 2T, 3T, and m3T (Fig. S8D), indicating that feedback in the m3T protocol facilitated faster exploration of new state-action rules after recognizing that previously acquired rules were no longer applicable in the current phase. A lower τ in the m3T protocol suggests increased flexibility in reversal learning, particularly when mice alternate between initial and reversal learning rules multiple times.

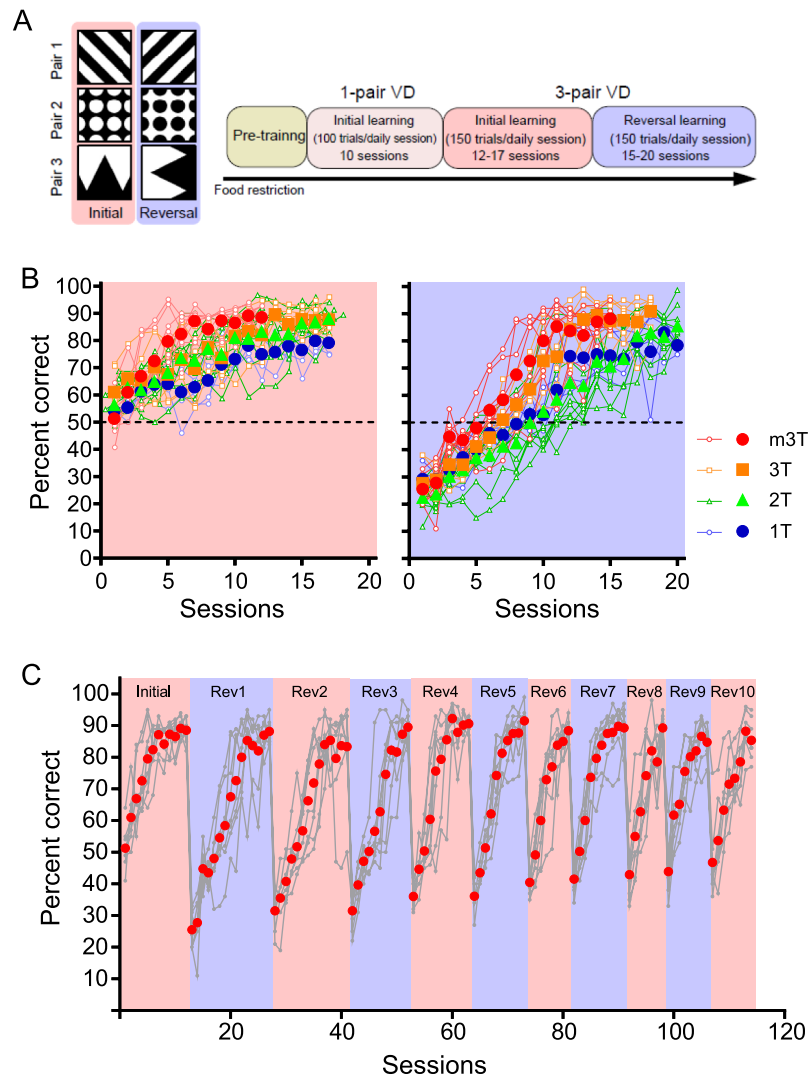


Fig. 4. Performance in the 3-pair VD task under multiple-touch protocols. **(A)** Stimulus pairs used in the 3-pair VD task (left) and training schedule (right), following completion of the 1-pair VD task. **(B)** Task performance under 1 T, 2 T, 3 T, and m3T protocols, showing initial (left) and reversal (right) learning. Small pale symbols and lines represent individual animals' performance, while large symbols represent the across-animal average. $N = 5$ (1T), 9 (2T), 7 (3T), and 8 (m3T) mice. The bordered small dots represent individual male mice. Large symbols indicate the mean across animals. **(C)** Serial reversal learning in the 3-pair VD task with m3T. Percentage correct across ten serial reversals with the m3T protocol. The same cohort in **(B)** was used for serial reversal learning, thus reusing the data for the initial learning and Rev1. Gray lines represent individual mice. $N = 8$ mice.

Discussion

The touchscreen-based operant task system provides flexible and widely used paradigms for evaluating sensory and cognitive functions in both animals and humans^{15,28,29}. Among these paradigms, VD tasks are relatively straightforward but often suffer from slow learning rates with conventional training protocols, limiting their application in mice. This study aimed to accelerate VD learning by developing a robust method to enhance both learning speed (i.e., the number of sessions required to reach the criterion) and performance levels (i.e., highest or plateau performance) through modifications in stimulus choice algorithms and the addition of visual feedback. The multiple-touch protocols, particularly those requiring three successive touches (3T/m3T/m3Tn), significantly improved learning speed compared to the original one-touch (1T) protocol in both 1-pair VD tasks (Fig. 1) and 3-pair VD tasks (Fig. 4). Mathematical modeling and simulation further supported the efficiency of multiple-touch protocols (Figs. 3 and S8). Additionally, serial reversal learning was successfully completed ten times within fewer than 120 training sessions using the m3T protocol (Fig. 4C). These findings demonstrate that the new protocols substantially reduce the time required for behavioral experiments, alleviating both physical and psychological burdens for experimental animals and experimenters.

Notably, no correction trials were applied at any stage of training in the current study. Correction trials are commonly used in touchscreen-based tasks to prevent side bias in touch responses^{21,30–32}, thereby facilitating learning^{10,33}. Although some mice initially exhibited side-biased touch strategies in the multiple-touch protocols, they rapidly adapted to task rules even without correction trials (Fig. S4).

In the current study, we incorporated auditory touch feedback for correct and incorrect intermediate touches in the multiple-touch protocols (Figs. 2B, S1D, and S5C). This auditory feedback could potentially be used by mice as an auditory “win-stay/lose-switch” strategy, independent of the intended task-solving strategy (i.e., visual pattern discrimination). However, our control experiments indicated that it was not the case. Mice were unable to solve the transverse pattern task using auditory cues alone (Figs. S5), but they were still able to solve the 1-pair VD task without correct/incorrect auditory feedback for intermediate touches (Fig. S6). Therefore, in our experiments, mice primarily relied on visual rather than auditory information to solve the touchscreen VD tasks. However, since we conducted these control experiments only during the initial learning phase (Figs. S5 and S6), we cannot rule out the possibility that auditory feedback may have a greater influence during reversal learning. Because eliminating potential auditory-based task-solving strategies is important for evaluating genuine VD learning, the m3Tn protocol, rather than the m3T, should be selected when the study’s objective specifically requires minimizing the impact of auditory cues.

The mechanisms underlying the effectiveness of multiple-touch procedures in VD learning remain unclear, but several possibilities can be considered. One explanation is that requiring multiple touches when making a stimulus choice may reduce error trials caused by unintentional touches, which are common in single-touch protocols. The ability to correct an initial touch error within the same trial allows mice to still obtain a reward, reinforcing task engagement. Another possibility is that making multiple touches extends the stimulus presentation duration, which may enhance the ability to recognize S+ and S- stimuli before they disappear from the display. Additionally, requiring multiple touches may lead to more deliberate and voluntary actions, helping mice acquire task rules more confidently compared to a single-touch protocol. While these factors likely contribute to improved VD learning efficiency, their relative contributions require further investigation.

In addition to multiple touches, the introduction of visual feedback after a false choice appears to have played a significant role in enhancing learning efficiency. The original one-touch protocol used auditory feedback, whereas the m3T protocol incorporated spatially coincident visual feedback. The immediate presentation of visual feedback following an S- choice may have helped mice establish a direct association between their action and the error, unlike tone feedback, which lacks spatial coincidence. Supporting this, mice receiving visual feedback in the m3T group showed a significant decrease in the nose-poking rate after making an error compared to the 3T group (Fig. S3), suggesting increased error awareness. The integration of visual feedback into the trial structure thus likely had a substantial impact on VD learning efficiency.

The selection of visual stimuli, which determines the discriminability of the two patterns in a stimulus pair, may influence VD learning efficiency. Mice exhibit better VD performance when the overall differences in shape and pattern between stimuli are significant³⁴, and natural scenes as visual stimuli can further facilitate VD learning¹⁶. In contrast, the present study used standardized stimuli, in which all stimuli were presented in black and white with a 1:1 proportion, ensuring equal brightness and contrast²¹. Using stimulus pairs with greater differences in saliency, such as natural scenes or colored patterns, may further accelerate VD learning speed.

Cognitive flexibility impairments are reported in various mental disorders, including autism^{35,36} and schizophrenia²⁹. Serial reversal learning paradigms assess multiple aspects of behavioral flexibility, including the formation of original and reversal rule sets and the ability to switch between rule sets through repeated reversals^{37,38}. Touchscreen-based serial VD learning tasks have been conducted in mice²¹, rats³⁹, and marmosets³⁸. Despite their usefulness, extensive serial reversal VD paradigms have been challenging for mice due to the long training periods required²⁶. Using the m3T protocol, the present study demonstrated that ten serial reversals in the 3-pair VD task could be completed in fewer than 120 sessions (Fig. 4). Given that many studies have applied only a few reversals in serial reversal tasks for mice, this protocol enables the completion of multiple serial reversals within a couple of months, greatly expanding the feasibility of cognitive flexibility testing in mice.

While many studies have aimed to identify brain regions responsible for VD learning, reversal, and switching^{40–43}, the critical areas remain ambiguous. This may be due to limitations such as the small number of trials per session, especially in manually operated paradigms, the restricted number of stimulus sets, often composed of only one or two pairs, and the slow learning speed, all of which hinder the efficient recruitment of activated neural ensembles associated with VD learning. Because the improved VD protocols developed in this study address these issues, they may facilitate the identification of brain regions critical for both initial learning and serial reversals of VD tasks. Recent advances in three-dimensional whole-brain analysis and neuronal activity-reporter technologies^{44–47} provide further opportunities to explore neural correlates of decision making. Once key regions are identified, neuronal ensembles can be monitored in free-moving animals using in vivo neuronal recording and imaging⁴⁸, with opto- or chemogenetic techniques used to manipulate their activity^{49–51}. Thus, this work paves the way for investigating VD learning at the cellular and circuit levels in mice.

The improvements developed in this study may also be effective for other types of operant tasks and for species other than mice. Furthermore, the multiple-touch procedures may prove useful in behavioral paradigms employing non-touchscreen devices. While direct evidence remains to be established in future studies, these findings open opportunities for conducting more cognitively demanding tasks, including those currently considered unfeasible.

Materials and methods

Animals

C57BL/6N mice were purchased from Japan SLC, Inc. (Hamamatsu, Japan) and maintained at room temperature (22°C–24°C) under a light-dark cycle, with the light period from 8:00 AM to 8:00 PM. Both male and female

mice were used in this study. The number and sex of the animals used for individual experiments are shown in Supplementary Table 1. All animal care and experimental procedures were approved by the Animal Experiment Committee of Kagoshima University and complied with the U.S. National Institutes of Health guidelines for the care and use of animals. All aspects of this study including the experimental design, methods, analyses, statistical procedures, and interpretations fully comply with the ARRIVE guidelines.

Apparatus for the touchscreen operant task

The touchscreen operant task system for free-moving mice was purchased from O'HARA & Co., Ltd. (TOP-M1, Tokyo, Japan). This system consists of a trapezoidal chamber equipped with a display monitor, touch panel, pellet dispenser, speaker, and water bottles. The touch panel is affixed to the display monitor and covered by a black board with two separate windows (6 cm × 6 cm each), which define the touch-sensitive areas. A reward port, where 10-mg food pellets (Rodent Tablet AIN-76A, TestDiet, St. Louis, MO, USA) are dispensed, is located on the opposite side of the touch panel. The entire system is housed in a sound-isolated box and controlled using Operant Task Studio V2 software (O'HARA & Co., Ltd.).

Behavioral procedure

Pre-training

Mice were at least 10 weeks old at the onset of dietary restriction, which was maintained throughout the study to keep body weight at approximately 85%–90% of that of age- and sex-matched ad libitum-fed control mice. Once the mice reached the target weight, they proceeded to the pre-training phase, which comprised two stages: habituation and shaping. During the habituation stage, mice acclimated to the apparatus and received food pellet rewards. Two habituation conditions were conducted sequentially in a 30-minute session: Habituation 1: The mouse explored the chamber freely and received ten food pellets over a 15-minute period; Habituation 2: The apparatus dispensed one reward pellet every 30 seconds for 15 minutes, accompanied by a tone. The habituation stage was completed once all mice in a training cohort ($n = 5-9$ mice) successfully consumed most of pellets and left only five or fewer (typically within 2–4 days) in a session, and the pre-training phase advanced to the shaping stage. In the shaping stage, mice learned to obtain a reward pellet by touching the screen. The same visual stimuli (black and white vertical stripes) were presented in both stimulus windows. A reward pellet was dispensed when the mouse touched either window, and the next trial began when the mouse nose-poked the reward port. A session ended after either 100 trials or 30 minutes, whichever occurred first. The shaping stage was considered complete when all mice in the cohort performed 90 or more trials in a session.

VD task

In the VD task, a pair of visual patterns was displayed in the stimulus windows on each trial, with their left–right positioning spatially pseudorandomized and evenly distributed throughout the session. The subject received a single food pellet reward with an accompanying one-second tone upon selecting the correct stimulus (S+), and the next trial began when the mouse nose-poked the reward port. Conversely, selecting the error stimulus (S-) resulted in no reward, a one-second buzzer sound, and a 10-second intertrial interval (ITI) before the next trial (Fig. 1A). Each VD session ended when the subject reached the fixed trial limit or time limit, whichever occurred first. The limits were determined by the number of visual stimulus pairs: 100 trials per 30 minutes for one-pair conditions and 150 trials per 45 minutes for three-pair conditions. The performance criterion was defined as achieving $\geq 75\%$ correct responses in two consecutive daily sessions for each individual animal. Once all mice in a cohort met the criterion, the experiment proceeded to the next stages. Cases where mice failed to meet the criterion were recorded individually. No correction trials were applied at any stage of VD training in this study.

Multiple-touch requirements in the VD task

In each trial of the VD task, the task algorithm was modified to require multiple successive touches. In the 2T protocol, mice were required to touch the same panel twice in succession (Fig. 1B). Similarly, in the 3T protocol, mice were required to touch the same panel three times in succession (Fig. S1C). After the first touch in the 2T protocol, and after the first two touches in the 3T protocol, a 0.3-second tone or buzzer was delivered as feedback, indicating that a touch had been registered.

In the modified version of the 3T protocol (m3T), two white squares were presented as visual feedback after the final touch that determined an incorrect choice in an error trial. In the m3T protocol with neutral auditory feedback (m3Tn), the 0.3-second tone and buzzer were replaced with a beep sound distinct from either the tone or the buzz.

Before initiating the VD task with the multiple-touch protocols, an additional shaping stage was conducted, during which reward pellets were dispensed only when the mouse touched either window two or three times in succession.

Single reversal and serial reversal VD tasks

Some of the mice that completed the initial VD task learning were subsequently trained in a reversal VD paradigm. In the reversal phase, the correct (S+) and error (S-) stimuli were reversed within each stimulus pair, relative to the initial learning phase. As in the initial learning, mice were considered to have met the performance criterion when they achieved $\geq 75\%$ correct responses in two consecutive daily sessions. Some mouse cohorts were further subjected to a series of serial reversals. Once all mice in a cohort met the criterion for the first reversal, they proceeded to the second reversal, where the S+ and S- stimuli were reversed back to their original assignment from the initial learning phase. Similarly, upon meeting the criterion in the second reversal, mice underwent a third reversal, in which the stimuli were reversed again. This alternation process was repeated ten times during the reversal phase of the serial reversal VD task.

Modeling and simulation of the VD tasks

The learning process is closely linked to an individual's internal states. To investigate whether the multiple-touch protocols influenced these internal states, we modeled mouse behavior using a Q-learning framework⁵². In brief, a Q-learning agent maintains a table of state-action values (Q) for action a at state s . Q is updated at each time t following,

$$Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \beta \delta_t,$$

$$\delta_t = r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t),$$

where δ is reward prediction error, a' represents a possible action, β and γ denote the learning rate and discount rate, respectively. Action a is either left or right touch in a response, common to all protocols, while state s consists of action states and outcome states. The outcome states are determined by the true or false responses; hence, the former were rewarded while the latter increased γ expressing the prolonged ITI as the punishment. Since rewards were provided immediately following responses, γ was assumed to have a negligible effect and was treated as a constant across protocols; $\gamma = 0.9$ solely in the outcome state for false responses; otherwise, it was set at 0.1. The number of action states depended on the number of required touches in each protocol. For example, in the two-touch protocol, three action states were defined: first touch, second-left touch, and second-right touch. If the left touch was the true (S+) response but the agent initially chose right (S-), the state transitioned to Second-right touch. If the agent touched right again, the final decision was confirmed as false, transitioning to the outcome state with no reward. However, if the agent switched to left, the state reverted to First-touch, allowing another decision attempt. A correct response (two consecutive left touches) transitioned the state to the outcome state with a reward (= 1).

To simulate learning progression, the agent followed an annealing ϵ -greedy policy⁵³, mimicking the balance between exploration and exploitation. The agent selected an action randomly with probability ϵ ; otherwise, it chose the action that maximized Q at each state. As learning progressed, ϵ gradually decreased, reflecting a shift from explorative behavior (early learning) to conservative decision-making (later stages). This decrease followed an exponential decay function:

$$\epsilon = (\epsilon_{\max} - \epsilon_{\min}) \exp(-\alpha i) + \epsilon_{\min},$$

where α and i are the decay constant and trial, respectively. ϵ_{\min} and ϵ_{\max} are the lower and upper boundaries of ϵ , and set at 0.1 and 0.9, respectively. In the reversal phase, we assumed that ϵ increased exponentially to returns ϵ_{\max} over τ trials as a retrograde order of the initial learning, therefore τ can be considered a perseverative tendency for Q acquired at the initial learning, as the agent with larger τ would require several trials to restart action exploration after the reversal phase where no reward is expected under the previous Q . Once ϵ reached ϵ_{\max} at trial τ in the reversal learning phase, then ϵ decreased across the following trials in the same manner as in the initial learning.

The internal state and performance of the Q-learning agent in each trial were determined by the free parameters α (decay constant), β (learning rate), and τ (perseverative tendency). Since internal states were expected to diverge between single- and multiple-touch protocols, these parameters were estimated by minimizing prediction errors between the observed learning curve and the simulated learning curve in the Q-learning model. Estimation was performed using the Markov Chain Monte Carlo method with Metropolis algorithm. Convergence was assessed using the Gelman-Rubin statistic, with a threshold below 1.1. The total sampling number was 5000, with a burn-in period of 2500 and a sampling interval of five, yielding at least 500 final samples for each parameter. The expected value for each parameter was calculated as the median of the sample chain. Simulations were conducted using Python 3.11.11 on Google Colaboratory.

Use of large language models

During the preparation of the manuscript, the authors used ChatGPT solely for language editing and grammatical correction, without altering the meaning of the content. After using the generative artificial intelligence tool, the authors carefully reviewed and, if necessary, further edited the content. The authors take full responsibility for the content of the work.

Statistical analysis

Values are expressed as across-animal averages with individual animal data or standard errors of the mean unless otherwise stated. Comparisons between two groups were analyzed using Student's t-test, while comparisons among more than two groups were evaluated using one-way or two-way ANOVA, followed by Tukey's multiple comparisons test for parametric analysis. For nonparametric data, the Kruskal-Wallis test was applied, followed by Dunn's multiple comparisons test. Repeated measures (RM) ANOVA was used where appropriate. No specific analysis was conducted to determine the minimal population size of animals before starting each experiment, but the number of mice used was consistent with previous studies^{4,12}. A post hoc power analysis confirmed that the sample sizes actually employed for comparisons showing statistically significant differences provided power values above 0.85 except one case, which exhibited a value of 0.72. All statistical evaluations were performed using GraphPad Prism 10 software (Dotmatics, Boston), except for the power analysis, which was performed using G*Power software⁵⁴.

Data availability

All data presented in this paper are available from Hiroyuki Okuno (okuno@m.kufm.kagoshima-u.ac.jp) upon reasonable request. The original code for the mathematical models and simulations is available from Yusuke Suzuki (suzuki.yusuke.7n@kyoto-u.ac.jp) upon request.

Received: 28 April 2025; Accepted: 31 October 2025

Published online: 03 December 2025

References

- Lipp, H. P. et al. IntelliCage: The development and perspectives of a mouse- and user-friendly automated behavioral test system. *Front. Behav. Neurosci.* **17**, 1270538. <https://doi.org/10.3389/fnbeh.2023.1270538> (2023).
- Crawley, J. N. Twenty years of discoveries emerging from mouse models of autism. *Neurosci. Biobehav. Rev.* **146**, 105053. <https://doi.org/10.1016/j.neubiorev.2023.105053> (2023).
- Squire, L. R., Genzel, L., Wixted, J. T. & Morris, R. G. Memory consolidation. *Cold Spring Harb. Perspect. Biol.* **7**, a021766. <https://doi.org/10.1101/cshperspect.a021766> (2015).
- Bussey, T. J., Saksida, L. M. & Rothblat, L. A. Discrimination of computer-graphic stimuli by mice: A method for the behavioral characterization of transgenic and gene-knockout models. *Behav. Neurosci.* **115**, 957–960 (2001).
- Kim, C. H., Heath, C. J., Kent, B. A., Bussey, T. J. & Saksida, L. M. The role of the dorsal hippocampus in two versions of the touchscreen automated paired associates learning (PAL) task for mice. *Psychopharmacology (Berl)* **232**, 3899–3910. <https://doi.org/10.1007/s00213-015-3949-3> (2015).
- Piipponniemi, T. O. et al. Impaired performance of the Q175 mouse model of huntington's disease in the touch screen paired associates learning task. *Front. Behav. Neurosci.* **12**, 226. <https://doi.org/10.3389/fnbeh.2018.00226> (2018).
- Swan, A. A. et al. Characterization of the role of adult neurogenesis in touch-screen discrimination learning. *Hippocampus* **24**, 1581–1591. <https://doi.org/10.1002/hipo.22337> (2014).
- Elsila, L. V., Korhonen, N., Hyytia, P. & Korpi, E. R. Acute lysergic acid diethylamide does not influence reward-driven decision making of C57BL/6 mice in the iowa gambling task. *Front. Pharmacol.* **11**, 602770. <https://doi.org/10.3389/fphar.2020.602770> (2020).
- Tamada, H. et al. Impact of intestinal microbiota on cognitive flexibility by a novel touch screen operant system task in mice. *Front. Neurosci.* **16**, 882339. <https://doi.org/10.3389/fnins.2022.882339> (2022).
- Suzuki, T., Joho, D. & Kakeyama, M. Purposive decision-making task in mice using touchscreen operant apparatus. *Neurosci. Res.* **200**, 34–40. <https://doi.org/10.1016/j.neures.2023.09.007> (2024).
- Kwak, C., Lim, C. S. & Kaang, B. K. Development of a touch-screen-based paradigm for assessing working memory in the mouse. *Exp. Neurobiol.* **24**, 84–89. <https://doi.org/10.5607/en.2015.24.1.84> (2015).
- Ayabe, T., Ohya, R. & Ano, Y. Hop-derived Iso-alpha-Acids in beer improve visual discrimination and reversal learning in mice as assessed by a touch panel operant system. *Front. Behav. Neurosci.* **13**, 67. <https://doi.org/10.3389/fnbeh.2019.00067> (2019).
- Marquardt, K., Cavanagh, J. F. & Brigman, J. L. Alcohol exposure in utero disrupts cortico-striatal coordination required for behavioral flexibility. *Neuropharmacology* **162**, 107832. <https://doi.org/10.1016/j.neuropharm.2019.107832> (2020).
- Clelland, C. D. et al. A functional role for adult hippocampal neurogenesis in spatial pattern separation. *Science* **325**, 210–213 (2009).
- Hataji, Y. & Goto, K. Information-seeking in mice (*Mus musculus*) during visual discrimination: Study using a distractor elimination paradigm. *Anim. Cogn.* **27**, 81. <https://doi.org/10.1007/s10071-024-01920-3> (2024).
- Yu, Y. et al. Mice use robust and common strategies to discriminate natural scenes. *Sci. Rep.* **8**, 1379. <https://doi.org/10.1038/s41598-017-19108-w> (2018).
- Watanabe, S. Preference for and discrimination of paintings by mice. *PLoS One* **8**, e65335. <https://doi.org/10.1371/journal.pone.0065335> (2013).
- Turner, K. M., Simpson, C. G. & Burne, T. H. BALB/c mice can learn touchscreen visual discrimination and reversal tasks faster than C57BL/6 mice. *Front. Behav. Neurosci.* **11**, 16. <https://doi.org/10.3389/fnbeh.2017.00016> (2017).
- Yang, M., Lewis, F. C., Sarvi, M. S., Foley, G. M. & Crawley, J. N. 16p11.2 Deletion mice display cognitive deficits in touchscreen learning and novelty recognition tasks. *Learn. Mem.* <https://doi.org/10.1101/lm.039602.115> (2015).
- Zelesnikow-Johnston, A. M. et al. Touchscreen testing reveals clinically relevant cognitive abnormalities in a mouse model of schizophrenia lacking metabotropic glutamate receptor 5. *Sci. Rep.* **8**, 16412. <https://doi.org/10.1038/s41598-018-33929-3> (2018).
- Dickson, P. E. & Mittleman, G. Visual discrimination, serial reversal, and extinction learning in the mdx mouse. *Front. Behav. Neurosci.* **13**, 200. <https://doi.org/10.3389/fnbeh.2019.00200> (2019).
- Van den Broeck, L., Hansquine, P., Callaerts-Vegh, Z. & D'Hooge, R. Impaired reversal learning in APPPS1-21 mice in the touchscreen visual discrimination task. *Front. Behav. Neurosci.* **13**, 92. <https://doi.org/10.3389/fnbeh.2019.00092> (2019).
- Gracian, E. I., Osmon, D. C. & Mosack, K. E. Transverse patterning, aging, and neuropsychological correlates in humans. *Hippocampus* **26**, 1633–1640. <https://doi.org/10.1002/hipo.22662> (2016).
- Rondi-Reig, L., Libbey, M., Eichenbaum, H. & Tonegawa, S. CA1-specific N-methyl-D-aspartate receptor knockout mice are deficient in solving a nonspatial transverse patterning task. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 3543–3548. <https://doi.org/10.1073/pnas.041620798> (2001).
- Roesch, M. R., Calu, D. J. & Schoenbaum, G. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat. Neurosci.* **10**, 1615–1624. <https://doi.org/10.1038/nn2013> (2007).
- Pais, R. C. et al. Assessing cognitive flexibility in mice using a custom-built touchscreen chamber. *Front. Behav. Neurosci.* **19**, 1536458. <https://doi.org/10.3389/fnbeh.2025.1536458> (2025).
- Roddick, K. M., Schellinck, H. M. & Brown, R. E. Serial reversal learning in an olfactory discrimination task in 3xTg-AD mice. *Learn. Mem.* **30**, 310–319. <https://doi.org/10.1101/lm.053840.123> (2023).
- Brown, T. M. et al. Melanopsin-based brightness discrimination in mice and humans. *Curr. Biol.* **22**, 1134–1141. <https://doi.org/10.1016/j.cub.2012.04.039> (2012).
- Afzal, S. et al. Probing cognitive flexibility in Shank2-deficient mice: Effects of D-cycloserine and NMDAR signaling hub dynamics. *Prog. Neuropsychopharmacol. Biol. Psychiatr.* **134**, 111051. <https://doi.org/10.1016/j.pnpbp.2024.111051> (2024).
- Brigman, J. L. & Rothblat, L. A. Stimulus specific deficit on visual reversal learning after lesions of medial prefrontal cortex in the mouse. *Behav. Brain Res.* **187**, 405–410. <https://doi.org/10.1016/j.bbr.2007.10.004> (2008).
- Brigman, J. L., Ihne, J., Saksida, L. M., Bussey, T. J. & Holmes, A. Effects of subchronic phencyclidine (PCP) treatment on social behaviors, and operant discrimination and reversal learning in C57BL/6J mice. *Front. Behav. Neurosci.* **3**, 2. <https://doi.org/10.3389/fnbeh.2009.08.002.2009> (2009).
- Dickson, P. E. et al. Behavioral flexibility in a mouse model of developmental cerebellar Purkinje cell loss. *Neurobiol. Learn. Mem.* **94**, 220–228. <https://doi.org/10.1016/j.nlm.2010.05.010> (2010).
- Jager, A. et al. Modulation of cognitive flexibility by reward and punishment in BALB/cJ and BALB/cByJ mice. *Behav. Brain Res.* **378**, 112294. <https://doi.org/10.1016/j.bbr.2019.112294> (2020).

34. Bussey, T. J. et al. The touchscreen cognitive testing method for rodents: How to get the best out of your rat. *Learn. Mem.* **15**, 516–523. <https://doi.org/10.1101/lm.987808> (2008).
35. Lin, S. et al. Frontostriatal circuit dysfunction leads to cognitive inflexibility in neuroligin-3 R451C knockin mice. *Mol. Psychiatry* **29**, 2308–2320. <https://doi.org/10.1038/s41380-024-02505-9> (2024).
36. Nakatani, J. et al. Abnormal behavior in a chromosome-engineered mouse model for human 15q11-13 duplication seen in autism. *Cell* **137**, 1235–1246. <https://doi.org/10.1016/j.cell.2009.04.024> (2009).
37. Ritchey, C. M., Gilroy, S. P., Kuroda, T. & Podlesnik, C. A. Assessing human performance during contingency changes and extinction tests in reversal-learning tasks. *Learn. Behav.* **50**, 494–508. <https://doi.org/10.3758/s13420-022-00513-9> (2022).
38. Jackson, S. A. W. et al. Selective role of the putamen in serial reversal learning in the marmoset. *Cereb. Cortex* **29**, 447–460. <https://doi.org/10.1093/cercor/bhy276> (2019).
39. Kumar, G., Talpos, J. & Steckler, T. Strain-dependent effects on acquisition and reversal of visual and spatial tasks in a rat touchscreen battery of cognition. *Physiol. Behav.* **144**, 26–36. <https://doi.org/10.1016/j.physbeh.2015.03.001> (2015).
40. Ragozzino, M. E., Kim, J., Hassert, D., Minniti, N. & Kiang, C. The contribution of the rat prelimbic-infralimbic areas to different forms of task switching. *Behav. Neurosci.* **117**, 1054–1065. <https://doi.org/10.1037/0735-7044.117.5.1054> (2003).
41. Yehene, E., Meiran, N. & Soroker, N. Basal ganglia play a unique role in task switching within the frontal-subcortical circuits: Evidence from patients with focal lesions. *J. Cogn. Neurosci.* **20**, 1079–1093. <https://doi.org/10.1162/jocn.2008.20077> (2008).
42. Cirillo, R. A., Horel, J. A. & George, P. J. Lesions of the anterior temporal stem and the performance of delayed match-to-sample and visual discriminations in monkeys. *Behav. Brain Res.* **34**, 55–69. [https://doi.org/10.1016/s0166-4328\(89\)80090-7](https://doi.org/10.1016/s0166-4328(89)80090-7) (1989).
43. Evenden, J. L. et al. Effects of excitotoxic lesions of the substantia innominata, ventral and dorsal globus pallidus on visual discrimination acquisition, performance and reversal in the rat. *Behav. Brain Res.* **32**, 129–149. [https://doi.org/10.1016/s0166-4328\(89\)80080-4](https://doi.org/10.1016/s0166-4328(89)80080-4) (1989).
44. Niu, M. et al. Claustrum mediates bidirectional and reversible control of stress-induced anxiety responses. *Sci. Adv.* <https://doi.org/10.1126/sciadv.abi6375> (2022).
45. Susaki, E. A. et al. Whole-brain imaging with single-cell resolution using chemical cocktails and computational analysis. *Cell* **157**, 726–739. <https://doi.org/10.1016/j.cell.2014.03.042> (2014).
46. Xin, Q. et al. Deconstructing the neural circuit underlying social hierarchy in mice. *Neuron* **113**, 444–459.e447. <https://doi.org/10.1016/j.neuron.2024.11.007> (2025).
47. Vousden, D. A. et al. Whole-brain mapping of behaviourally induced neural activation in mice. *Brain Struct. Funct.* **220**, 2043–2057. <https://doi.org/10.1007/s00429-014-0774-0> (2015).
48. Diehl, G. W. & Redish, A. D. Differential processing of decision information in subregions of rodent medial prefrontal cortex. *Elife* <https://doi.org/10.7554/eLife.82833> (2023).
49. Rajasethupathy, P., Ferenczi, E. & Deisseroth, K. Targeting neural circuits. *Cell* **165**, 524–534. <https://doi.org/10.1016/j.cell.2016.03.047> (2016).
50. Roth, B. L. DREADDs for neuroscientists. *Neuron* **89**, 683–694. <https://doi.org/10.1016/j.neuron.2016.01.040> (2016).
51. Basu, R. et al. The orbitofrontal cortex maps future navigational goals. *Nature* **599**, 449–452. <https://doi.org/10.1038/s41586-021-04042-9> (2021).
52. Watkins, C. J. & Dayan, P. Q-learning. *Mach. Learn.* **8**, 279–292 (1992).
53. Zhang, S., Peng, H., Nagesh Rao, S. & Tseng, E. Discretionary lane change decision making using reinforcement learning with model-based exploration. In *18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, 844–850 (2019).
54. Faul, F., Erdfelder, E., Lang, A. G. & Buchner, A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Method.* **39**, 175–191. <https://doi.org/10.3758/bf03193146> (2007).

Acknowledgments

We thank Sakurako Inaba, Yukiko Oshiro, and Yoshino Katayama for mouse husbandry, and Maki Kiyama, Yayoi Mori, Yoshiho Sakamoto, Yuya Yamaguchi, Yoshikazu Osako, and Koshiro Otake for technical assistance. We also thank Dr. Shigeru Shinomoto for his thorough reading of the manuscript and constructive suggestions. This work was supported by JSPS KAKENHI (Grant Number 19K07799 to Y.K.; 23H00522 to M.K.; JP25H00964 to I.I.; 23K27281 to H.O.), Transformative Research Areas (A) (23H04680 and 25H02506 to H.O.), JST-CREST (JPMJCR25B1 to I.I.), and the Japan Agency for Medical Research and Development (AMED) Brain/MINDS program (JP19dm0207080 to H.O.; JP19dm0207090 to I.I.), Moonshot Research & Development program (JP-22zf0127007 to I.I.), and Technological Innovation of Regenerative Medicine program (JP21bm0704060, JP-24bm1123049 to I.I.). Additional support was provided by the Takeda Science Foundation (to H.O.), the Co-operative Research Program (Joint Usage/Research Center program) of Institute for Life and Medical Sciences, Kyoto University (to H.O. and I.I.), the Kagoshima University Megumikai Medical Research Promotion Fund (to H.O.), the Core Research Program for the “Neuroscience Core Unit” of Kagoshima University, and the Kodama Memorial Fund for Medical Research (to H.O. and Y.K.).

Author contributions

YK, MK, and HO conceived the project, designed the experiments, and coordinated the study. YK, MH, YO, YD, NH, YI, and TS conducted the animal experiments. YK and MH analyzed the data. YS and II performed the modeling and simulation. YK and HO drafted the manuscript, while YK, YS, MK, and HO finalized it.

Funding

JSPS KAKENHI, 19K07799, 23H00522, 23K27281, JSPS KAKENHI Transformative Research Areas (A), 23H04680, the Japan Agency for Medical Research and Development (AMED) Brain/MINDS program, JP19dm0207080.

Declarations

competing of interest

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-27003-y>.

Correspondence and requests for materials should be addressed to H.O.

Reprints and permissions information is available at www.nature.com/reprints.

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025