1 **Title**

2 **High-throughput automatic training system for odor-based cognitive behaviors**
3 **in head-fixed mice**
4
5 **Authors:**
6 Zhe Han[1,2], Xiaoxing Zhang[1], Jia Zhu[1,2], Yulei Chen[1], and Chengyu T. Li[1*]
7
8 [1] Institute of Neuroscience, State Key Laboratory of Neuroscience, Key Laboratory of
9 Primate Neurobiology, CAS Center for Excellence in Brain Science and Intelligence
10 Technology, Chinese Academy of Sciences, Shanghai 200031, China.
11 [2] University of Chinese Academy of Sciences, Beijing 100049, China.
12
13 [*] Corresponding author: tonylicy@ion.ac.cn.
14

Abstract

Understanding neuronal mechanisms of cognitive behaviors requires efficient behavioral assays. We designed a high-throughput automatic training system (HATS) for olfactory behaviors in head-fixed mice. The hardware and software were constructed to enable automatic training with minimal human intervention. The integrated system was composed of customized 3D-printing supporting components, an odor-delivery unit with fast response, Arduino based hardware-controlling and data-acquisition unit. Furthermore, the customized software was designed to enable automatic training in all training phases, including lick-teaching, shaping, and learning. Using HATS, we trained mice to perform delayed non-match to sample (DNMS), delayed paired association (DPA), Go/No-go (GNG), and GNG reversal tasks. These tasks probed cognitive functions including sensory discrimination, working memory, decision making, and cognitive flexibility. Mice reached stable levels of performance within several days in the tasks. HATS enabled an experimenter to train eight mice simultaneously, therefore greatly enhanced the experimental efficiency. Combined with causal perturbation and activity recording techniques, HATS can greatly facilitate our understanding of the neural-circuitry mechanisms underlying cognitive behaviors.

Introduction

Behavioral design and analysis are critical for understanding neural mechanism of cognition (Gomez-Marin et al., 2014), including working memory (Fuster, 1997; Baddeley, 2012), decision making (Gold and Shadlen, 2007; Lee et al., 2012), and reversal of learnt rules (Bunge and Wallis, 2008). Combined with novel neural-circuitry technologies, such as optogenetics (Fenno et al., 2011), chemogenetics (Armbruster et al., 2007), and imaging methods (Deisseroth and Schnitzer, 2013), well-disigned behavioral paradigms can greatly facilitate the ciruitry level understanding of behaivor. Reliable behavioral paradigms are also useful in pre-clinic studies such as target identification and mechanistic studies for brain diseases (Fernando and Robbins, 2011, Gotz and Ittner, 2008, Nestler and Hyman, 2010).

Optimally, behavioral training systems should be automatic, ready to scale up, blind in design, and flexible in changing paradigms. Automatic training systems (Schaefer and Claridge-Chang, 2012) met well with these criteria. There was a long history of designing automatic behavior-training systems, for example in studies of operant conditioning (*e.g.*, Davidson et al., 1971). Automatic training systems are composed of monitoring and feedback controlling of behavior. In free-moving mice, automatic measurement has been implemented in characterizing visual performance (Benkner et al., 2013, de Visser et al., 2005, Kretschmer et al., 2013), evaluation of pain sensitivity(Kazdoba et al., 2007, Roughan et al., 2009), freezing behavior during fear conditioning(Anagnostaras et al., 2010, Kopec et al., 2007), home-cage

phenotyping(Balci et al., 2013, Hubener et al., 2012, Jhuang et al., 2010), anxiety(Aarts et al., 2015), diurnal rhythms(Adamah-Biassi et al., 2013), and social behavior(Hong et al., 2015, Ohayon et al., 2013, Weissbrod et al., 2013). With feedback controlling components, automatic training systems have been successfully implemented in multiple behavioral domains, including memory assessment (Reiss et al., 2014), operant learning (Remmelink et al., 2015), and training limb function (Becker et al., 2016). Automatic training systems with multiple cognitive behaviors requiring memory, attention, and decision making have been developed previously in free-moving rats (Erlich et al., 2011, Poddar et al., 2013) and mice (Gallistel et al., 2014, Romberg et al., 2013). Moreover, such systems were successful in dissecting neural-circuitry mechanisms underlying cognitive behaviors (*e.g.*, Brunton et al., 2013, Erlich et al., 2011, Hanks et al., 2015). Head-fixed mice (Dombeck et al., 2007; Guo et al., 2014) renders great flexibility in recording (Harvey et al., 2009; Boyd et al., 2012; Fukunaga et al., 2012; Kollo et al., 2014) and imaging (Boyd et al., 2015; Chu et al., 2016, Dombeck et al., 2007, Komiyama et al., 2010; Yamada et al., 2017) technologies. Moreover, free-moving and head-restrianted mice exhibit similar ability of olfactory discrimination (Abraham et al., 2012). However, automatic training systems in head-fixed mice were not developed previously.

Olfaction is an important sensory modality for cognitive behavior (Doty, 1986; Ache and Young, 2005). Previous studies have demonstrated that rodents are very good at olfactory discrimination, memory, and decision (Abraham et al., 2004, Barnes et al., 2008, Cleland et al., 2002, Haddad et al., 2013, Hubener and Laska, 2001, Kepecs et al., 2007, Komiyama et al., 2010; Liu et al., 2014, Lu et al., 1993, Mihalick et al., 2000, Passe and Walker, 1985, Petrulis and Eichenbaum, 2003, Rinberg et al., 2006, Slotnick et al., 1991, Uchida and Mainen, 2003). Automatic behavioral systems have been developed for studying innate olfactory behaviors (Qiu et al., 2014). Olfactory behavioral testing has been developed in head-fixed rodents and greatly facilitates the understanding of neural circuits underlying olfaction (Verhagen et al., 2007; Wesson et al., 2008; Shusterman et al., 2011; Kato et al., 2013; Boyd et al., 2015) and odor-based cognition (Komiyama et al., 2010; Liu et al., 2014; Gadziola et al., 2015). However, fully automatic training systems for odor-based cognitive behaviors were not available for head-fixed mice.

We therefore designed a high-throughput automatic training system (HATS) for olfactory behaviors in head-fixed mice. Using the automatic step-by-step training procedures, we trained mice to perform olfactory delayed non-match to sample (DNMS), delayed paired association (DPA), Go/No-go (GNG), and GNG reversal tasks. Mice reached stable levels of performance within several days in the tasks. HATS can be an important tool in our understanding of the neural-circuitry mechanisms underlying odor-based cognitive behaviors.

Material and Methods

Animals

Male adult C57BL/6 mice (SLAC, as wild-type) were used for the current study (8-40 weeks of age, weighted between 20 to 30 g). Wild-type mice were provided by the Shanghai Laboratory Animal Center (SLAC), CAS, Shanghai, China. Mice were group-housed (4-6/cage) under a 12-h light-dark cycle (light on from 5 a.m. to 5 p.m.). Before behavioral training, mice were housed in stable conditions with food and water ad libitum. After the start of behavioral training, the water supply was restricted. Mice could drink water only during and immediately after training. Care was taken to keep mice body weight above 80% of a normal level. The behavioral results reported here were collected from a total of 25 wild-type mice. All animal studies and experimental procedures were approved by the Animal Care and Use Committee of the Institute of Neuroscience, Chinese Academy of Sciences, Shanghai, China.

Animal surgery

Mice were anesthetized with analgesics (Sodium pentobarbital, 10mg/mL, 80 mg/kg body weight) before surgery. All surgery tools, materials, and experimenter-coats were sterilized by autoclaving. Surgery area and materials that cannot undergo autoclaving were sterilized by ultraviolet radiation for more than 20 minutes. Aseptic procedures were applied during surgery. Anesthetized mice were kept on a heat mat to maintain normal body temperature. Scalp, periosteum, and other associated soft tissue over skull were removed. Skull was cleaned by filtered artificial cerebrospinal fluid (ACSF) with cotton applicators. After skull was dried out, a layer of tissue adhesive was applied on the surface of the skull. A steel plate was placed on the skull and then fixed by dental Cement.

Behavior setups

HATS was composed of a mouse containing, head-fix, odor delivery and reward delivery, Arduino based control, and data acquisition units (diagram in **Figure 1A**, photo in **Figure 1B**). All valves and motors were controlled by Arduino based processors and customized software. The 3d printing files, a step by step instruction for hardware assembling, the source code for behavior training and the data acquisition source code were publically available (https://github.com/wwweagle/serialj ; https://github.com/jerryhanson/frontiers)/

Three-dimensional printing technique was used to generate the small components in the system (**Figure 1C**). The training tube was used to maintain the relative position of mouse body to the water- and odor-delivery ports. The motor slot held a direct-current motor to move the water port forward or backward. The water

tube slot held a metal needle with a blunt tip, from which mice obtained water as a reward. The odor-tube slot connected the odor tube from the odor-delivery unit.

A movable water port was connected to a peristaltic pump, which was controlled by an Arduino board. The volume of water reward was controlled by changing the duration of the output signal to peristaltic pump from the Arduino board. Peristaltic pumps of different setups were calibrated for the stable volume of water delivery in each trial (5 ± 0.5 μL).

Water- and odor-delivery units were both controlled by an Arduino board. During behavior training, detailed timing information of events was sent back to the computer via the USB-simulated serial-port interface and stored by a customized Java program. The stored events included an odorant valve on/off, peristaltic pump on/off, and licking start/end. Licking event was detected by a capacity detector. Infrared LED-based licking detectors were used for electrophysiological recording if required. An infrared camera was placed under the water port to monitor behavioral states of mice.

Olfactometer

The olfactometer was designed to efficiently and reliably mix and deliver odor. Air source was a pump that provided air flow with the flow rate of ~120 L/min. The filter was applied to eliminate moisture and dust. Eight training setups shared one set of pump and filter. For each setup, pure air with the flow rate of 2 L/min is constantly delivered to mice during the entire process. The air input to each air route could be turned on and off by a manual valve (labeled as “M” in **Figure 2A** and **2B**). The flow rate was adjusted by a needle valve (labeled as “V” in **Figure 2A** and **2B**). As shown in the **Figure 2B**, one type of odorant in liquid state was stored in one airtight bottle. The air-in tube was placed right above the surface of the liquid odorant. Two-way solenoid valves were used to switch the odor to either mouse or flow mater. In the standby state (no odor was delivered, **Figure 2A**), the valve to odorant bottle (labeled as “O”) was closed, and that to the flow meter (labeled as “F”) was opened. Therefore, no odor will be mixed with pure air and delivered to the mouse. In the working state that odor was delivered **(Figure 2B)**, “O” was open and “F” was closed. Therefore odor was mixed with constant air and delivered to the mouse. Four kinds of odorants were used in the behavior tasks, 1-Butanol, Methyl butyrate, Hexanoic acid, and Octane. The relative volume ratios of these odorants in the pure air were 10%, 2.5%, 15% and 5%, respectively. The difference was due to the distinct evaporation pressure of different odorant molecules at room temperature (see **Table 1** for detailed rising/decay and residual time of the odorants). The odor tubes after “O” valves and before mixture chamber had an inner diameter of 0.5 mm. The odor tube for constant air before mixture chamber had an inner diameter of 2.5 mm.

Behavior training

Water restriction

Mice were allowed at least seven days for recovery after surgery for head-plate implantation. Before the start of formal training, mice were water restricted for 48~72 hours, in which licking for water was allowed (less than 1.0 ML per day, exact amount was not monitored). Throughout the training, the daily intake of water was at least 0.6 ML per day (as in Guo et al., 2014) and typically 1.0 ML per day. Body weight was closely monitored and a steady increase in body was observed after initial decrease following 24 hour restriction.

Habituation phases

The habituation phase started thirty minutes before the start of the training phase and only occurred once. A training tube was placed into the home cage. Mice could explore the tube freely to be familiar with it. This step was designed to decrease the stress level of mice on the first day.

Automatic licking teaching phase

This phase was designed to teach mice to lick freely from the water tube. A mouse was fixated on the head plate to a holding bar connected to the training tube. The animals were transferred from homecages to the apparatus and headfixed manually by experimenters. The total time spent in transition was less than a minute. Then the training tube was placed into and fixated to sliding sockets in the sound-attenuated box (the typical decrease from background noise was 15dB). Initially, the tip of the water port was placed five millimeters away from the mouse mouth. By using a program-controlled movable water port, the initiation of a teaching bout was associated with the forward movement of the water port. During each day, this phase was divided into three bouts to facilitate the association between movement of the water port and delivery of water. In each bout, water port moved forward firstly to seduce mouse to lick. After two seconds, water port will be reset back to the original place. Once mouse licked, one water drop (volume of ~5 μL) was delivered for every three licks. This bout ended when mice did not lick continuously for two seconds, or rewarded size is larger than 200 μL from this bout. The daily reward size could vary between each mouse (typically 0.6 ML and less than 1.0 ML). This phase lasted for three days. Mice stayed in training apparatus for 1-2 hours per day in all training phases.

Automatic shaping phase

This phase was designed to teach mice to lick for water only in the response window, which was from 0.5 to 1.5 sec after the offset of the second odor delivery. Only rewarded condition was applied, which were non-matched pairs for DNMS task,

paired odors for DPA task, or go cue for GNG task, respectively. Mice could lick in response window to trigger water reward from every trial. During this phase, water port may or may not move while water was delivered. If mice missed several trials, lick-teaching would resume, in which the water port was moved forward and water was delivered during the response window. The reward in lick-teaching was program-controlled and was not triggered by lick. Two types of trials were defined for this phase, the self-learning (**Figure 5, left) and** program-teaching (**Figure 5, right)** trials, which switched automatically under the condition introduced below. The water port was moved forward during the response window in the program-teaching trials, in which the water delivery was automatic without triggered by licks. In the self-learning trials, however, reward delivery was licking-triggered, and water port did not move. The condition for switching from self-learning to program-teaching trials was that mice missed five times within 30 trials or missed during the last program-teaching trial. The condition for switching from program-teaching trial to self-learning trial is that mice licked in response window and obtained a reward from the last teaching trial. Daily shaping phase ended when mice performed 100 hit trials in total. This phase lasted for three days.

Full task training phase

DNMS task training

In the DNMS task, a sample odor was delivered at the start of a trial, followed by a delay period (4-5 seconds) and then a test odor, same to (matched) or different from (non-matched) the sample (**Figure 6**). Two kinds of odorants were used in DNMS task, 1-Butanol, and Methyl butyrate. The relative volume ratios in the pure air were 10% and 2.5%, respectively. Odor-delivery duration was one second. Mice were trained to lick in the response window in non-match trials. The response window was from 0.5 to 1.5 sec after the offset of the second odor delivery. Licking events detected in the response window in the non-match trials were regarded as Hit and will trigger instantaneous water delivery (a water drop around 5 μL). The false choice was defined as detection of licking events in the response window in the match trials. Mice were not punished in the False Choice trials. Mice were neither punished nor rewarded for the Miss (no-lick in a non-match trial) or the Correct rejection (CR, no-lick in a matching trial) trials. Behavioral results were binned in blocks of 24 trials. There was a fixed inter-trial interval of 10 seconds between trials. After training ended each day, mice were supplied with water of at least 300 μL and up to 1 mL daily intake. This phase lasts for four to five days. The well-trained criterion was set to the existence of three continuous correct-rates larger than 80%, calculated using a sliding window of 24 trials. The reason to use 24 trials as a block is to maintain the consistency of different trial types between different tasks, with the need to be commonly divided by four and eight types of odor sequence for different tasks (4 for DNMS, 4 for DPA). It was intended to facilitate the compairsion of the performance in the different tasks in the the current study. It can be easily modified according to the needs of .

DPA task training

For the DPA task, a sample and a test odor were delivered, separated by a delay period (**Figure 7**). Four kinds of odorants were used, 1-Butanol (S1), Methyl butyrate (S2), Hexanoic acid (T1) and Octane (T2). The relative volume ratios in pure air were 10%, 2.5%, 15% and 5%, respectively. Odor delivery duration was one second. Delay period between two odors in a trial was 8-9 seconds. Response window was set to 0.5-1 second after the offset of the test odor in a trial. Mice were trained to lick to obtain water reward only after the paired trials (S1-T1 or S2-T2). Licking events detected in the response window in paired trials were regarded as Hit and will trigger instantaneous water delivery. The false choice was defined as detection of licking events in the response window in non-paired trials (S1-T2 or S2-T1), and mice were not punished in False Choice trials. Mice were neither punished nor rewarded for Miss (no-lick in the paired trial) or Correct rejection (CR, no-lick in a non-paired trial) trials. Behavioral results were binned in blocks of 24 trials. There was a fixed inter-trial interval of 16 seconds between trials. After training ended each day, mice were supplied with water of at least 300 μL and up to 1 mL daily intake. This phase lasts for four to five days. The well-trained criterion was set to the existence of three continuous correct-rates larger than 80%, calculated using a sliding window of 24 trials.

GNG and GNG reversal task training

For the GNG task, mice were trained to lick for water only after the Go cue but not No-go cue. Hexanoic acid and Octane were used as Go and No-go cues, respectively. The relative volume ratios in the pure air were 15% and 5%, respectively. Odor-delivery duration was one second. Response window was 0.5-1.5 second after the offset of a cue. Licking events detected in the response window in Go trials were regarded as Hit and triggered instantaneous water delivery. The false choice was defined as the detection of licking events in the response window in No-go trials. Mice were not punished in the False Choice trials. Mice were neither punished nor rewarded for the Miss (no-lick in a Go trial) or the Correct rejection (CR, no-lick in a No-go trial) trials. Behavioral results were binned in blocks of 24 trials. There was a fixed inter-trial interval of 5 seconds between trials. After training ended each day, mice were supplied with water of at least 300 μL and up to 1 mL daily intake. This phase lasts for three days. The well-trained criterion was set to the existence of three continuous correct-rates larger than 80%, calculated using a sliding window of 24 trials.

In the third day of training, the GNG reversal task began, in which the odor-reward relationship was reversed.

Data analysis

The performance of the correct rate (referred to as "performance" in labels of figures) of each bin was defined by:

Performance correct rate = (num. hit trials + num. correct rejection trials) / total number of trials

Hit, False choice, and Correct rejection rates were defined as follows:

Hit rate = num. hit trials / (num. hit trials + num. miss trials)
False choice rate = num. false choice trials / (num. false choice trials + num. correct rejection trials)
Correct rejection rate = num. correct rejection trials/ (num. false choice trials + num. correct rejection trials)

Mean correct rate (CR rate/ FA rate) was calculated as an averaged correct rate (CR rate / FA rate) between different mice.

Error bars from the mean value of the correct rate (CR rate / FA rate) was calculated by the standard error of the mean. N represents the number of mice.

The licking rate was calculated as lick numbers within each time bin (bin size:100 ms). The curve was smoothed by smooth function from Matlab with a span size of 5 bin.

Discriminability (*d'*) was defined by:

d' = norminv (Hit rate)- norminv (False choice rate). The norminv function was the inverse of the cumulative normal function. Conversion of Hit or False choice rate was applied to avoid plus or minus infinity (Macmillan and Creelman, 2005). In conversion, if Hit or False choice rate was equal to 100%, it was set to [1-1/(2n)]. Here, n equals to a number of all possible Hit or False choice trials. If Hit or False choice rate was zero, it was set to 1/(2n).

Licking efficiency = rewarded licking number / (rewarded licking number + unrewarded licking number).

A number of trials to criterion was calculated as the trial numbers before reaching 80% correct rate for 24 consecutive trials. "NRC" in **Figure 6-8** represented Not Reaching Criterion, which indicated that mice did not reach the above criterion for that day.

Results

Overview of Hardware, Software, and protocol

In our previous study (Liu et al., 2014), mice were manually taught to lick for water and shaped for a DNMS task. The goal of the current study was to allow fully automatic training. The only things human operators need to perform were to fixate mice onto head-fix bars, close doors of training boxes, and run computer software controlling training protocols. The current study fulfilled the goal by designing HATS for olfactory and odor-based cognitive behavior in head-fixed mice. HATS was composed of a mouse containing, head-fix, odor- and water- delivery, Arduino based control, and data acquisition units (diagram in **Figure 1A**, photo in **Figure 1B**, 3D-printed parts in **Figure 1C**). Optogenetic, chemogenetic, recording, and imaging methods can be easily integrated into HATS. All valves and motors were controlled by Arduino based processors and customized software. The daily routine was composed of system adjustment, head-fixation of mice, choosing a protocol, and training mice a given behavior (**Figure 1D**).

Fast odor delivery

In studying olfactory behaviors, it is critical to have fast rise and decay for odor delivery.

Our olfactometer exhibited fast response and stable performance. The reaction time constant for the onset of these odors was between 11-71 ms (**Table 1**), measured with a photoionization detector (PID). Another key parameter was the time constant for decay after the offset of the odor-delivery unit, which was especially important in working memory-related tasks. The current odor-delivery unit exhibited fast decay (time constant: 20-41 ms, **Figure 2C**, **Table 1**). Moreover, odor concentration remained stable following more than 200 trials of odor delivery (**Figure 2F**), which was important for behavioral and recording experiments.

Automatic training protocol

To achieve fully automatic training, we developed a step-by-step training protocol. The protocol was separated into two preparatory steps (water deprivation and habituation) and three training phases (lick-teaching, shaping, and learning, **Figure 3A**).

The first step of training was to automatically teach licking freely from water tube (**Figure 4A-B)**. Moveable water port (**Figure 4C-E**) was located 5 mm away from the mouth of a mouse. The flow chart of the lick-teaching protocol was plotted in **Figure 4A**. At the start of a teaching bout, water port would deliver 10 μL water and then moved forward until contacting the mouth, thus encouraging the licking. If mouse licked, 4 μL water would be rewarded for every three licks. After no licking was detected for consecutive 2 sec or water of 200 μL was delivered, one bout of

teaching was completed, and the water port was moved back to the initial position. The teaching bout was repeated for several times until water of 400 μL was rewarded in total. The volume of water rewarded in each day was plotted in **Figure 4B**.

The second step of training was shaping for a specific task. This phase was designed to allow mice to be familiar with the temporal structure of the tasks and the involved sensory stimuli, without experiencing the full task. In shaping, only the trials with water reward were applied. Specifically, for DNMS task, only non-matched odor pairs were applied to mice (**Figure 5A-B**). For DPA task, only paired trials were applied. For GNG task, only Go cue was applied. Two types of trials were designed, self-learning and teaching trials. In self-learning trials, water delivery was triggered by licking in the response window (**Figure 5C** left box). In teaching trials, water port moved forward and delivered water automatically during response window (**Figure 5C** right box). These two types of trials were designed to switch automatically. The condition for switching from learning to teaching trial was that mice missed five trials in 35 trials. The condition for switching from teaching to learning trial was that mice licked within the response window in the last teaching trial. Daily shaping phase ended when mice performed 100 hit trials in total. This phase lasted for three days.

Training the DNMS task

We trained eight head-fixed mice to perform an olfactory DNMS task(Liu et al., 2014) (**Figure 6A**). In this design mouse needed to temporally maintain information during the delay period before behavioral choices and motor planning. After the shaping protocol, we added the non-rewarded matched trials, which induced false choice and reduced performance to chance level (**Figure 6B**). Gradually the performance, correct rejection, and discriminability (d') progressively increased, whereas the hit rate remained at a ceiling level (**Figure 6B-E**). After the training of five days (600 trials), the performance showed significant increase (ANOVA, $p<0.0001$, $F=775.89$). Mice experienced a certain level of relearning each day, with a decreased number to criteria (defined as a correct rate above 80% in 24 consecutive trials) each day through learning (**Figure 6F**). Most of the licking responses were associated with non-match odor and expectation of water reward (**Figure 6G**). There were licks associated with the first odor delivery in the early phase of learning (**Figure 6G** black curve), which were declined through learning (**Figure 6G** blue curve). Also, the licking efficiency (defined as the ratio of successful licks resulting water reward) was increased progressively through learning (**Figure 6H**).

Training the DPA task

The second set of head-fixed mice was trained to perform an olfactory DPA task (**Figure 7A**). As in the DNMS task, the performance, correct rejection, and discriminability (d') progressively increased, whereas the hit rate remained at ceiling level (**Figure 7B-E**). After the training of five days (600 trials), the performance showed significant increase (ANOVA, $p<0.0001$, $F=1139.03$). Mice also experienced

a certain level of relearning each day (**Figure 7F**). Most of the licking responses were associated with paired odor and expectation of water reward (**Figure 7G**). There were licks associated with the first odor delivery in the early phase of learning (**Figure 7G** black curve), which were declined through learning (**Figure 7G** blue curve). Such early licks associated with the sample odor were lower than that in the DNMS task. The licking efficiency also increased progressively through learning (**Figure 7H**).

Training the GNG and reversal tasks

The third set of head-fixed mice was initially trained to perform an olfactory GNG task (**Figure 8A** above), then subsequently sensory-cue reversal task (**Figure 8A** below). The performance, correct rejection, and discriminability (*d'*) progressively increased, whereas the hit rate remained at ceiling level (**Figure 8B-E**). After the training of two days (200 trials), the performance showed significant increase (ANOVA, $p<0.0001$, $F=3455.17$). Mice also experienced a certain level of relearning each day (**Figure 8F**). Most of the licking responses were associated with paired odor-pair and expectation of water reward (**Figure 8G**). The licking efficiency also increased progressively through learning (**Figure 8H**). After two-days of GNG training, the odor-reward relationship was reversed (**Figure 8A** below). The performance, correct rejection, discriminability (*d'*), and licking efficiency were decreased initially, and then progressively increased (**Figure 8B-E**). The hit rate remained at ceiling level (**Figure 8D**) and relearning was evident from the number of trials to criteria (**Figure 8F**).

Discussion

Automated, quantitative, and accurate assessment of behaviors is critical for understanding mechanisms underlying cognition. Here we presented HATS, a new integrated hardware and software system that combined fast olfactometer, 3D-printed components, step-by-step automatic training, for automatic training of cognitive behaviors in head-fixed mice. The robustness of the system was validated in multiple olfactory and odor-based tasks. The involved tasks require cognitive abilities including working memory (Fuster, 1997; Baddeley, 2012), decision making (Gold and Shadlen, 2007; Lee et al., 2012), and reversal of learnt rules (Bunge and Wallis, 2008), all of which are required in more naturalistic environment and vital for survival.

An obvious limitation is that free-moving mice cannot be trained with HATS. Another limitation is that HATS only monitor the lick as behavioral readouts, therefore is more suited for large-scale screening of optogenetic. Although the head-movement was restrained in the current design, one would like to monitor the muscles controlling head or chewing movement to further eliminate the potential artifacts in electrophysiological recording. To obtain deep understanding of neural circuit underlying these behavior, one would also like to integrate more monitoring systems for behavioral events, such as sniffing (Kepecs et al., 2007; Verhagen et al., 2007; Wesson et al., 2008; Shusterman et al., 2011; Deschenes et al., 2012; McAfee et

al., 2016), pupil size (Reimer et al., 2014; McGinley et al., 2015; Vinck et al., 2015; Bushnell et al., 2016; Reimer et al., 2016), and whisker movement (Orbach et al., 1985; Friedman et al., 2006; Birdwell et al., 2007; O'Connor et al., 2010; Deschenes et al., 2012; Petreanu et al., 2012; Moore et al., 2013).

In designing HATS, we tried to fasten the training history, therefore aiding the dissection of neural circuit. However, this fast training in animals would only sufficiently model fast learning in humans. Indeed, many human behaviors and human learning are slow in learning and require extensive training, such as fine motor skill (*i.g.*, driving, playing piano) and sensory discrimination (*i.g.*, wine tasting). Thus, automations achieved in HATS have limitations to what kinds of behavioral and neural processes are being effectively modeled.

Nevertheless, HATS allowed for rapid, automated training of cognitive behaviors across diverse experimental designs. Our approach can also support high-throughput behavioral screening. In summary, the newly developed HATS is well-suited for circuitry understanding of odor-based cognitive behavior.

Acknowledgments

The work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDB02020006), the Chinese 973 Program (2011CBA00406), the Instrument Developing Project of the Chinese Academy of Sciences (Grant No. YZ201540), the Key Research Project of Frontier Science of the Chinese Academy of Sciences (Grant No. QYZDB-SSW-SMC009), China – Netherlands CAS-NWO Programme: Joint Research Projects, The Future of Brain and Cognition (153D31KYSB20160106), the Key Project of Shanghai Science and Technology Commission (No.15JC1400102, 16JC1400101), the National Science Foundation for Distinguished Young Scholars of China (31525010, to C.T.L.), the General Program of Chinese National Science Foundation (31471049), the State Key Laboratory of Neuroscience, and CAS Hundreds of Talents Program (2010OHTP04, to C.T.L.).

Reference

Aarts, E., Maroteaux, G., Loos, M., Koopmans, B., Kovacevic, J., Smit, A.B., et al. (2015). The light spot test: Measuring anxiety in mice in an automated home-cage environment. *Behav Brain Res* 294**,** 123-130. doi: 10.1016/j.bbr.2015.06.011.

Abraham, N.M., Guerin, D., Bhaukaurally, K., and Carleton, A. (2012). Similar odor discrimination behavior in head-restrained and freely moving mice. *PLoS One* 7(12)**,** e51789. doi: 10.1371/journal.pone.0051789.

Abraham, N.M., Spors, H., Carleton, A., Margrie, T.W., Kuner, T., and Schaefer, A.T. (2004). Maintaining accuracy at the expense of speed: stimulus similarity defines odor discrimination time in mice. *Neuron* 44(5)**,** 865-876. doi: 10.1016/j.neuron.2004.11.017.

Ache, B.W., and Young, J.M. (2005). Olfaction: diverse species, conserved principles. *Neuron* 48(3), 417-430. doi: 10.1016/j.neuron.2005.10.022.

Adamah-Biassi, E.B., Stepien, I., Hudson, R.L., and Dubocovich, M.L. (2013). Automated video analysis system reveals distinct diurnal behaviors in C57BL/6 and C3H/HeN mice. *Behav Brain Res* 243, 306-312. doi: 10.1016/j.bbr.2013.01.003.

Anagnostaras, S.G., Wood, S.C., Shuman, T., Cai, D.J., Leduc, A.D., Zurn, K.R., et al. (2010). Automated assessment of pavlovian conditioned freezing and shock reactivity in mice using the video freeze system. *Front Behav Neurosci* 4. doi: 10.3389/fnbeh.2010.00158.

Armbruster, B.N., Li, X., Pausch, M.H., Herlitze, S., and Roth, B.L. (2007). Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. *Proc Natl Acad Sci U S A* 104(12), 5163-5168. doi: 10.1073/pnas.0700293104.

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annu Rev Psychol* 63, 1-29. doi: 10.1146/annurev-psych-120710-100422.

Balci, F., Oakeshott, S., Shamy, J.L., El-Khodor, B.F., Filippov, I., Mushlin, R., et al. (2013). High-Throughput Automated Phenotyping of Two Genetic Mouse Models of Huntington's Disease. *PLoS Curr* 5. doi: 10.1371/currents.hd.124aa0d16753f88215776fba102ceb29.

Barnes, D.C., Hofacer, R.D., Zaman, A.R., Rennaker, R.L., and Wilson, D.A. (2008). Olfactory perceptual stability and discrimination. *Nat Neurosci* 11(12), 1378-1380. doi: 10.1038/nn.2217.

Becker, A.M., Meyers, E., Sloan, A., Rennaker, R., Kilgard, M., and Goldberg, M.P. (2016). An automated task for the training and assessment of distal forelimb function in a mouse model of ischemic stroke. *J Neurosci Methods* 258, 16-23. doi: 10.1016/j.jneumeth.2015.10.004.

Benkner, B., Mutter, M., Ecke, G., and Munch, T.A. (2013). Characterizing visual performance in mice: an objective and automated system based on the optokinetic reflex. *Behav Neurosci* 127(5), 788-796. doi: 10.1037/a0033944.

Birdwell, J.A., Solomon, J.H., Thajchayapong, M., Taylor, M.A., Cheely, M., Towal, R.B., et al. (2007). Biomechanical models for radial distance determination by the rat vibrissal system. *J Neurophysiol* 98(4), 2439-2455. doi: 10.1152/jn.00707.2006.

Boyd, A.M., Kato, H.K., Komiyama, T., and Isaacson, J.S. (2015). Broadcasting of cortical activity to the olfactory bulb. *Cell Rep* 10(7), 1032-1039. doi: 10.1016/j.celrep.2015.01.047.

Boyd, A.M., Sturgill, J.F., Poo, C., and Isaacson, J.S. (2012). Cortical feedback control of olfactory bulb circuits. *Neuron* 76(6), 1161-1174. doi: 10.1016/j.neuron.2012.10.020.

Brunton, B.W., Botvinick, M.M., and Brody, C.D. (2013). Rats and humans can optimally accumulate evidence for decision-making. *Science* 340(6128), 95-98. doi: 10.1126/science.1233912.

Bunge, S.A., and Wallis, J.D. (2008). *Neuroscience of rule-guided behavior.* Oxford ; New York: Oxford University Press.

Bushnell, M., Umino, Y., and Solessio, E. (2016). A system to measure the pupil response to steady lights in freely behaving mice. *J Neurosci Methods* 273, 74-85. doi: 10.1016/j.jneumeth.2016.08.001.

Chu, M.W., Li, W.L., and Komiyama, T. (2016). Balancing the Robustness and Efficiency of Odor Representations during Learning. *Neuron* 92(1), 174-186. doi: 10.1016/j.neuron.2016.09.004.

Cleland, T.A., Morse, A., Yue, E.L., and Linster, C. (2002). Behavioral models of odor similarity. *Behav Neurosci* 116(2), 222-231.

Davidson, A.B., Davis, D.J., and Cook, L. (1971). A rapid automatic technique for generating operant key-press behavior in rats. *J Exp Anal Behav* 15(1), 123-127.

de Visser, L., van den Bos, R., and Spruijt, B.M. (2005). Automated home cage observations as a tool to measure the effects of wheel running on cage floor locomotion. *Behav Brain Res* 160(2), 382-388. doi: 10.1016/j.bbr.2004.12.004.

Deisseroth, K., and Schnitzer, M.J. (2013). Engineering approaches to illuminating brain structure and dynamics. *Neuron* 80(3), 568-577. doi: 10.1016/j.neuron.2013.10.032.

Deschenes, M., Moore, J., and Kleinfeld, D. (2012). Sniffing and whisking in rodents. *Curr Opin Neurobiol* 22(2), 243-250. doi: 10.1016/j.conb.2011.11.013.

Dombeck, D.A., Khabbaz, A.N., Collman, F., Adelman, T.L., and Tank, D.W. (2007). Imaging large-scale neural activity with cellular resolution in awake, mobile mice. *Neuron* 56(1), 43-57. doi: 10.1016/j.neuron.2007.08.003.

Doty, R.L. (1986). Odor-guided behavior in mammals. *Experientia* 42(3), 257-271.

Erlich, J.C., Bialek, M., and Brody, C.D. (2011). A cortical substrate for memory-guided orienting in the rat. *Neuron* 72(2), 330-343. doi: 10.1016/j.neuron.2011.07.010.

Fenno, L., Yizhar, O., and Deisseroth, K. (2011). The development and application of optogenetics. *Annu Rev Neurosci* 34, 389-412. doi: 10.1146/annurev-neuro-061010-113817.

Fernando, A.B., and Robbins, T.W. (2011). Animal models of neuropsychiatric disorders. *Annu Rev Clin Psychol* 7, 39-61. doi: 10.1146/annurev-clinpsy-032210-104454.

Friedman, W.A., Jones, L.M., Cramer, N.P., Kwegyir-Afful, E.E., Zeigler, H.P., and Keller, A. (2006). Anticipatory activity of motor cortex in relation to rhythmic whisking. *J Neurophysiol* 95(2), 1274-1277. doi: 10.1152/jn.00945.2005.

Fukunaga, I., Berning, M., Kollo, M., Schmaltz, A., and Schaefer, A.T. (2012). Two distinct channels of olfactory bulb output. *Neuron* 75(2), 320-329. doi: 10.1016/j.neuron.2012.05.017.

Fuster, J.M. (1997). *The prefrontal cortex : anatomy, physiology, and neuropsychology of the frontal lobe.* Philadelphia: Lippincott-Raven.

Gadziola, M.A., Tylicki, K.A., Christian, D.L., and Wesson, D.W. (2015). The olfactory tubercle encodes odor valence in behaving mice. *J Neurosci* 35(11), 4515-4527. doi: 10.1523/JNEUROSCI.4750-14.2015.

Gallistel, C.R., Balci, F., Freestone, D., Kheifets, A., and King, A. (2014). Automated, quantitative cognitive/behavioral screening of mice: for genetics, pharmacology, animal cognition and undergraduate instruction. *J Vis Exp* (84), e51047. doi: 10.3791/51047.

Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. *Annu Rev Neurosci* 30, 535-574. doi: 10.1146/annurev.neuro.29.051605.113038.

Gomez-Marin, A., Paton, J.J., Kampff, A.R., Costa, R.M., and Mainen, Z.F. (2014). Big behavioral data: psychology, ethology and the foundations of neuroscience. *Nat Neurosci* 17(11), 1455-1462. doi: 10.1038/nn.3812.

Gotz, J., and Ittner, L.M. (2008). Animal models of Alzheimer's disease and frontotemporal dementia. *Nat Rev Neurosci* 9(7), 532-544. doi: 10.1038/nrn2420.

Guo, Z.V., Hires, S.A., Li, N., O'Connor, D.H., Komiyama, T., Ophir, E., et al. (2014). Procedures for behavioral experiments in head-fixed mice. *PLoS One* 9(2), e88678. doi: 10.1371/journal.pone.0088678.

Haddad, R., Lanjuin, A., Madisen, L., Zeng, H., Murthy, V.N., and Uchida, N. (2013). Olfactory cortical neurons read out a relative time code in the olfactory bulb. *Nat Neurosci* 16(7), 949-957. doi: 10.1038/nn.3407.

Hanks, T.D., Kopec, C.D., Brunton, B.W., Duan, C.A., Erlich, J.C., and Brody, C.D. (2015). Distinct

relationships of parietal and prefrontal cortices to evidence accumulation. *Nature* 520(7546), 220-223. doi: 10.1038/nature14066.

Harvey, C.D., Collman, F., Dombeck, D.A., and Tank, D.W. (2009). Intracellular dynamics of hippocampal place cells during virtual navigation. *Nature* 461(7266), 941-946. doi: 10.1038/nature08499.

Hong, W., Kennedy, A., Burgos-Artizzu, X.P., Zelikowsky, M., Navonne, S.G., Perona, P., et al. (2015). Automated measurement of mouse social behaviors using depth sensing, video tracking, and machine learning. *Proc Natl Acad Sci U S A* 112(38), E5351-5360. doi: 10.1073/pnas.1515982112.

Hubener, F., and Laska, M. (2001). A two-choice discrimination method to assess olfactory performance in pigtailed macaques, Macaca nemestrina. *Physiol Behav* 72(4), 511-519.

Hubener, J., Casadei, N., Teismann, P., Seeliger, M.W., Bjorkqvist, M., von Horsten, S., et al. (2012). Automated behavioral phenotyping reveals presymptomatic alterations in a SCA3 genetrap mouse model. *J Genet Genomics* 39(6), 287-299. doi: 10.1016/j.jgg.2012.04.009.

Jhuang, H., Garrote, E., Mutch, J., Yu, X., Khilnani, V., Poggio, T., et al. (2010). Automated home-cage behavioural phenotyping of mice. *Nat Commun* 1, 68. doi: 10.1038/ncomms1064.

Kato, H.K., Gillet, S.N., Peters, A.J., Isaacson, J.S., and Komiyama, T. (2013). Parvalbumin-expressing interneurons linearly control olfactory bulb output. *Neuron* 80(5), 1218-1231. doi: 10.1016/j.neuron.2013.08.036.

Kazdoba, T.M., Del Vecchio, R.A., and Hyde, L.A. (2007). Automated evaluation of sensitivity to foot shock in mice: inbred strain differences and pharmacological validation. *Behav Pharmacol* 18(2), 89-102. doi: 10.1097/FBP.0b013e3280ae6c7c.

Kepecs, A., Uchida, N., and Mainen, Z.F. (2007). Rapid and precise control of sniffing during olfactory discrimination in rats. *J Neurophysiol* 98(1), 205-213. doi: 10.1152/jn.00071.2007.

Kollo, M., Schmaltz, A., Abdelhamid, M., Fukunaga, I., and Schaefer, A.T. (2014). 'Silent' mitral cells dominate odor responses in the olfactory bulb of awake mice. *Nat Neurosci* 17(10), 1313-1315. doi: 10.1038/nn.3768.

Komiyama, T., Sato, T.R., O'Connor, D.H., Zhang, Y.X., Huber, D., Hooks, B.M., et al. (2010). Learning-related fine-scale specificity imaged in motor cortex circuits of behaving mice. *Nature* 464(7292), 1182-1186. doi: 10.1038/nature08897.

Kopec, C.D., Kessels, H.W., Bush, D.E., Cain, C.K., LeDoux, J.E., and Malinow, R. (2007). A robust automated method to analyze rodent motion during fear conditioning. *Neuropharmacology* 52(1), 228-233. doi: 10.1016/j.neuropharm.2006.07.028.

Kretschmer, F., Kretschmer, V., Kunze, V.P., and Kretzberg, J. (2013). OMR-arena: automated measurement and stimulation system to determine mouse visual thresholds based on optomotor responses. *PLoS One* 8(11), e78058. doi: 10.1371/journal.pone.0078058.

Lee, D., Seo, H., and Jung, M.W. (2012). Neural basis of reinforcement learning and decision making. *Annu Rev Neurosci* 35, 287-308. doi: 10.1146/annurev-neuro-062111-150512.

Liu, D., Gu, X., Zhu, J., Zhang, X., Han, Z., Yan, W., et al. (2014). Medial prefrontal activity during delay period contributes to learning of a working memory task. *Science* 346(6208), 458-463. doi: 10.1126/science.1256573.

Lu, X.C., Slotnick, B.M., and Silberberg, A.M. (1993). Odor matching and odor memory in the rat. *Physiol Behav* 53(4), 795-804.

Macmillan, N.A., and Creelman, C.D. (2005). *Detection theory : a user's guide.* Mahwah, N.J.: Lawrence Erlbaum Associates.

McAfee, S.S., Ogg, M.C., Ross, J.M., Liu, Y., Fletcher, M.L., and Heck, D.H. (2016). Minimally invasive highly precise monitoring of respiratory rhythm in the mouse using an epithelial temperature probe. *J Neurosci Methods* 263**,** 89-94. doi: 10.1016/j.jneumeth.2016.02.007.

McGinley, M.J., Vinck, M., Reimer, J., Batista-Brito, R., Zagha, E., Cadwell, C.R., et al. (2015). Waking State: Rapid Variations Modulate Neural and Behavioral Responses. *Neuron* 87(6)**,** 1143-1161. doi: 10.1016/j.neuron.2015.09.012.

Mihalick, S.M., Langlois, J.C., Krienke, J.D., and Dube, W.V. (2000). An olfactory discrimination procedure for mice. *J Exp Anal Behav* 73(3)**,** 305-318. doi: 10.1901/jeab.2000.73-305.

Moore, J.D., Deschenes, M., Furuta, T., Huber, D., Smear, M.C., Demers, M., et al. (2013). Hierarchy of orofacial rhythms revealed through whisking and breathing. *Nature* 497(7448)**,** 205-210. doi: 10.1038/nature12076.

Nestler, E.J., and Hyman, S.E. (2010). Animal models of neuropsychiatric disorders. *Nat Neurosci* 13(10)**,** 1161-1169. doi: 10.1038/nn.2647.

O'Connor, D.H., Clack, N.G., Huber, D., Komiyama, T., Myers, E.W., and Svoboda, K. (2010). Vibrissa-based object localization in head-fixed mice. *J Neurosci* 30(5)**,** 1947-1967. doi: 10.1523/JNEUROSCI.3762-09.2010.

Ohayon, S., Avni, O., Taylor, A.L., Perona, P., and Roian Egnor, S.E. (2013). Automated multi-day tracking of marked mice for the analysis of social behaviour. *J Neurosci Methods* 219(1)**,** 10-19. doi: 10.1016/j.jneumeth.2013.05.013.

Orbach, H.S., Cohen, L.B., and Grinvald, A. (1985). Optical mapping of electrical activity in rat somatosensory and visual cortex. *J Neurosci* 5(7)**,** 1886-1895.

Passe, D.H., and Walker, J.C. (1985). Odor psychophysics in vertebrates. *Neurosci Biobehav Rev* 9(3)**,** 431-467.

Petreanu, L., Gutnisky, D.A., Huber, D., Xu, N.L., O'Connor, D.H., Tian, L., et al. (2012). Activity in motor-sensory projections reveals distributed coding in somatosensation. *Nature* 489(7415)**,** 299-303. doi: 10.1038/nature11321.

Petrulis, A., and Eichenbaum, H. (2003). "Olfactory memory," in *Handbook of olfaction and gustation,* ed. R.L. Doty. 2nd ed (New York: Marcel Dekker), 409-438.

Poddar, R., Kawai, R., and Olveczky, B.P. (2013). A fully automated high-throughput training system for rodents. *PLoS One* 8(12)**,** e83171. doi: 10.1371/journal.pone.0083171.

Qiu, Q., Scott, A., Scheerer, H., Sapkota, N., Lee, D.K., Ma, L., et al. (2014). Automated analyses of innate olfactory behaviors in rodents. *PLoS One* 9(4)**,** e93468. doi: 10.1371/journal.pone.0093468.

Reimer, J., Froudarakis, E., Cadwell, C.R., Yatsenko, D., Denfield, G.H., and Tolias, A.S. (2014). Pupil fluctuations track fast switching of cortical states during quiet wakefulness. *Neuron* 84(2)**,** 355-362. doi: 10.1016/j.neuron.2014.09.033.

Reimer, J., McGinley, M.J., Liu, Y., Rodenkirch, C., Wang, Q., McCormick, D.A., et al. (2016). Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nat Commun* 7**,** 13289. doi: 10.1038/ncomms13289.

Reiss, D., Walter, O., Bourgoin, L., Kieffer, B.L., and Ouagazzal, A.M. (2014). New automated procedure to assess context recognition memory in mice. *Psychopharmacology (Berl)* 231(22)**,** 4337-4347. doi: 10.1007/s00213-014-3577-3.

Remmelink, E., Loos, M., Koopmans, B., Aarts, E., van der Sluis, S., Smit, A.B., et al. (2015). A 1-night operant learning task without food-restriction differentiates among mouse strains in an

automated home-cage environment. *Behavioural Brain Research* 283**,** 53-60. doi: http://dx.doi.org/10.1016/j.bbr.2015.01.020.

Rinberg, D., Koulakov, A., and Gelperin, A. (2006). Speed-accuracy tradeoff in olfaction. *Neuron* 51(3)**,** 351-358. doi: 10.1016/j.neuron.2006.07.013.

Romberg, C., Horner, A.E., Bussey, T.J., and Saksida, L.M. (2013). A touch screen-automated cognitive test battery reveals impaired attention, memory abnormalities, and increased response inhibition in the TgCRND8 mouse model of Alzheimer's disease. *Neurobiol Aging* 34(3)**,** 731-744. doi: 10.1016/j.neurobiolaging.2012.08.006.

Roughan, J.V., Wright-Williams, S.L., and Flecknell, P.A. (2009). Automated analysis of postoperative behaviour: assessment of HomeCageScan as a novel method to rapidly identify pain and analgesic effects in mice. *Lab Anim* 43(1)**,** 17-26. doi: 10.1258/la.2008.007156.

Schaefer, A.T., and Claridge-Chang, A. (2012). The surveillance state of behavioral automation. *Curr Opin Neurobiol* 22(1)**,** 170-176. doi: 10.1016/j.conb.2011.11.004.

Shusterman, R., Smear, M.C., Koulakov, A.A., and Rinberg, D. (2011). Precise olfactory responses tile the sniff cycle. *Nat Neurosci* 14(8)**,** 1039-1044. doi: 10.1038/nn.2877.

Slotnick, B.M., Kufera, A., and Silberberg, A.M. (1991). Olfactory learning and odor memory in the rat. *Physiol Behav* 50(3)**,** 555-561.

Uchida, N., and Mainen, Z.F. (2003). Speed and accuracy of olfactory discrimination in the rat. *Nat Neurosci* 6(11)**,** 1224-1229. doi: 10.1038/nn1142.

Verhagen, J.V., Wesson, D.W., Netoff, T.I., White, J.A., and Wachowiak, M. (2007). Sniffing controls an adaptive filter of sensory input to the olfactory bulb. *Nat Neurosci* 10(5)**,** 631-639. doi: 10.1038/nn1892.

Vinck, M., Batista-Brito, R., Knoblich, U., and Cardin, J.A. (2015). Arousal and locomotion make distinct contributions to cortical activity patterns and visual encoding. *Neuron* 86(3)**,** 740-754. doi: 10.1016/j.neuron.2015.03.028.

Weissbrod, A., Shapiro, A., Vasserman, G., Edry, L., Dayan, M., Yitzhaky, A., et al. (2013). Automated long-term tracking and social behavioural phenotyping of animal colonies within a semi-natural environment. *Nat Commun* 4**,** 2018. doi: 10.1038/ncomms3018.

Wesson, D.W., Carey, R.M., Verhagen, J.V., and Wachowiak, M. (2008). Rapid encoding and perception of novel odors in the rat. *PLoS Biol* 6(4)**,** e82. doi: 10.1371/journal.pbio.0060082.

Yamada, Y., Bhaukaurally, K., Madarasz, T.J., Pouget, A., Rodriguez, I., and Carleton, A. (2017). Context- and Output Layer-Dependent Long-Term Ensemble Plasticity in a Sensory Circuit. *Neuron* 93(5)**,** 1198-1212 e1195. doi: 10.1016/j.neuron.2017.02.006.

Figure and Table Legends

FIGURE 1 | Components and operational processes for HATS.

(A) Schematics showing the components of HATS. **(B)** Photos of HATS hardware. a, sound-attenuated box; b-c, odor containers; d, flow meter; e, needle valve; f, training tube for restraining mouse body; g, camera; h, holder for the odor- and water-delivery unit and motors; i, capacitance detector for licking; j, 3D-printed odor and water delivery unit; k, ventilator. **(C)** 3D-printed components. 1, mouse-body tube; 2, a

socket for odor-delivery tubes; 3-4, slots for the motor of the moveable water port; 5, a holder for the odor- and water-delivery unit and motors; 6, a socket for the training tubes. **(D)** Schematics of the operational processes for HATS.

FIGURE 2 | Design, implementation, and reaction time of the olfactometer.

(A) Schematic showing the "standby" condition of the olfactometer. Diagram for the two-odor delivery unit was shown. The flow meter was designed for monitoring potential system failure. The flow rate was labeled as the numbers with the unit "L/min." Arrows indicated for the direction of air flow. **(B)** Schematic showing the "working" condition when one odor was delivered (through "r2"). Reduction of the readout from the flow meter indicated for normal operation. **(C)** Photo of the flow-controlling unit for the olfactometer. **(D)** Photo of the tubing unit and mixing chamber. Thin tubes were used for fast reaction for odor delivery. Mixing chamber was designed for a maximal mixture of pure air (from "r1" in **B**) and the delivered odor (from "r2" in **B**). **(E)** Fast response of the olfactometer. Readout from PID was plotted in the log scale for main figure and linear scale for inset (Mean ± SEM, standard error of the 0mean, unless stated otherwise; calculated from odor application of 200 trials). Rising/decay time constant and time with residual-odor were shown in **Table 1**. **(F)** Odor stability across trials.

FIGURE 3 | Step-by-step automatic training procedure.

(A) Step-by-step automatic training procedure. Duration in a given step was labeled to the right.

FIGURE 4 | Automatic lick-teaching protocol.

(A) Flow chart for the automatic lick-teaching protocol. **(B)** Daily consumed water volume in the lick-teaching phase. **(C)** Diagram of the moveable water port. **(D-E)** Diagram showing relative position between water port and mouse mouth in self-learning (**D**) and teaching (**E**) phases.

FIGURE 5 | Automatic shaping protocol.

(A) Design paradigm and time line for the DNMS shaping. Only non-matched trials were applied. **(B)** Licking performance in the shaping phase. **(C)** Flow chart for the DNMS shaping. Left: self-learning trials. Right: teaching trials.

FIGURE 6 | Automatic DNMS training protocol and behavioral results.

(A) Design paradigm and time line for the DNMS training. Both non-matched and matched trials were applied. **(B)** Performance of mice in the DNMS training phase. Bin size: 24 trials. **(C-E)** Correct rejection (CR) rate, hit rate, and *d'* in the DNMS training, respectively. **(F)** Re-learning in each day of the DNMS training, measured by the number of trials to criterion (defined as more than 80% performance in 24

consecutive trials). NRC, not reaching criteria. Mice that were NRC in the 2nd and 3rd days were not included. **(G)** Licking rates for training day 1 and 5. **(H)** Licking efficiency in the DNMS training. Licking efficiency was defined as the ratio of successful licks resulting water reward.

FIGURE 7 | Automatic DPA training protocol and behavioral results.

(A-H) As in Figure 6 A-H.

FIGURE 8 | Automatic GNG and reversal training protocol and behavioral results.

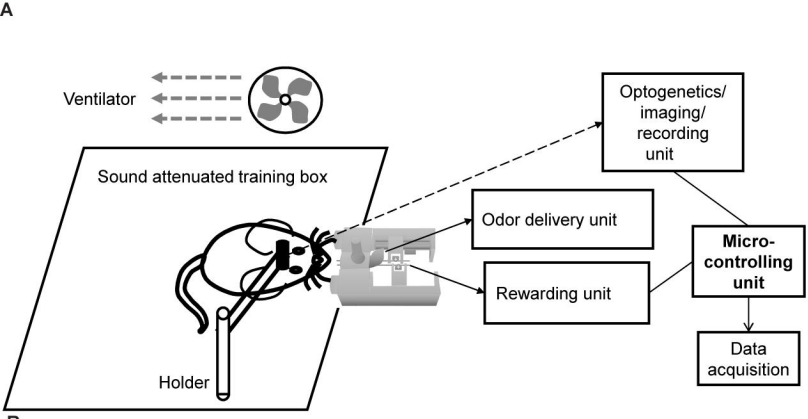
(A-H) As in Figure 6 A-H.

TABLE | 1. Rising and decay properties of odorants. Time was in a millisecond.

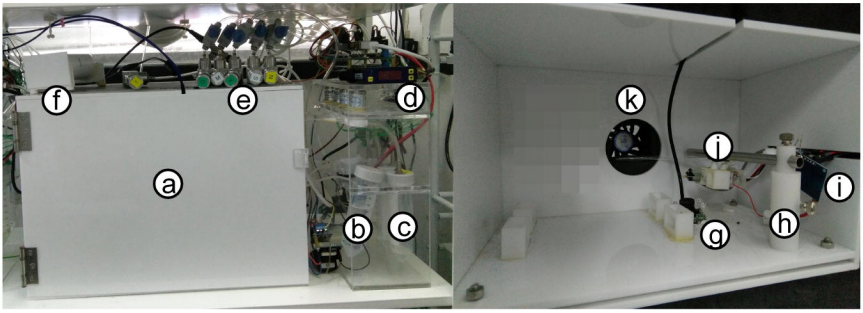
Odorant Name	Relative Volume Ratio in Air (%)	Rising Latency (95% of peak)	Decay Constant Time (1/e of peak)
1-Butanol	10	18 ± 1	20 ± 1
Methyl butyrate	2.5	17 ± 1	22 ± 1
Hexanoic acid	15	31 ± 1	41 ± 1
Octane	5	71 ± 1	31 ± 1

(mean ± standard error of the mean)

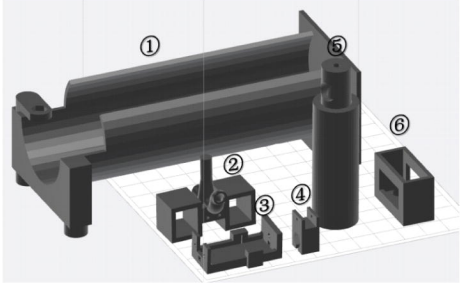
A


Ventilator
Sound attenuated training box
Holder
Optogenetics/ imaging/ recording unit
Odor delivery unit
Rewarding unit
Micro-controlling unit
Data acquisition

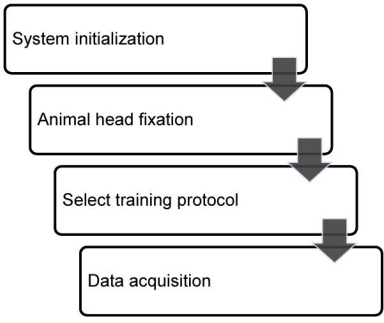
B

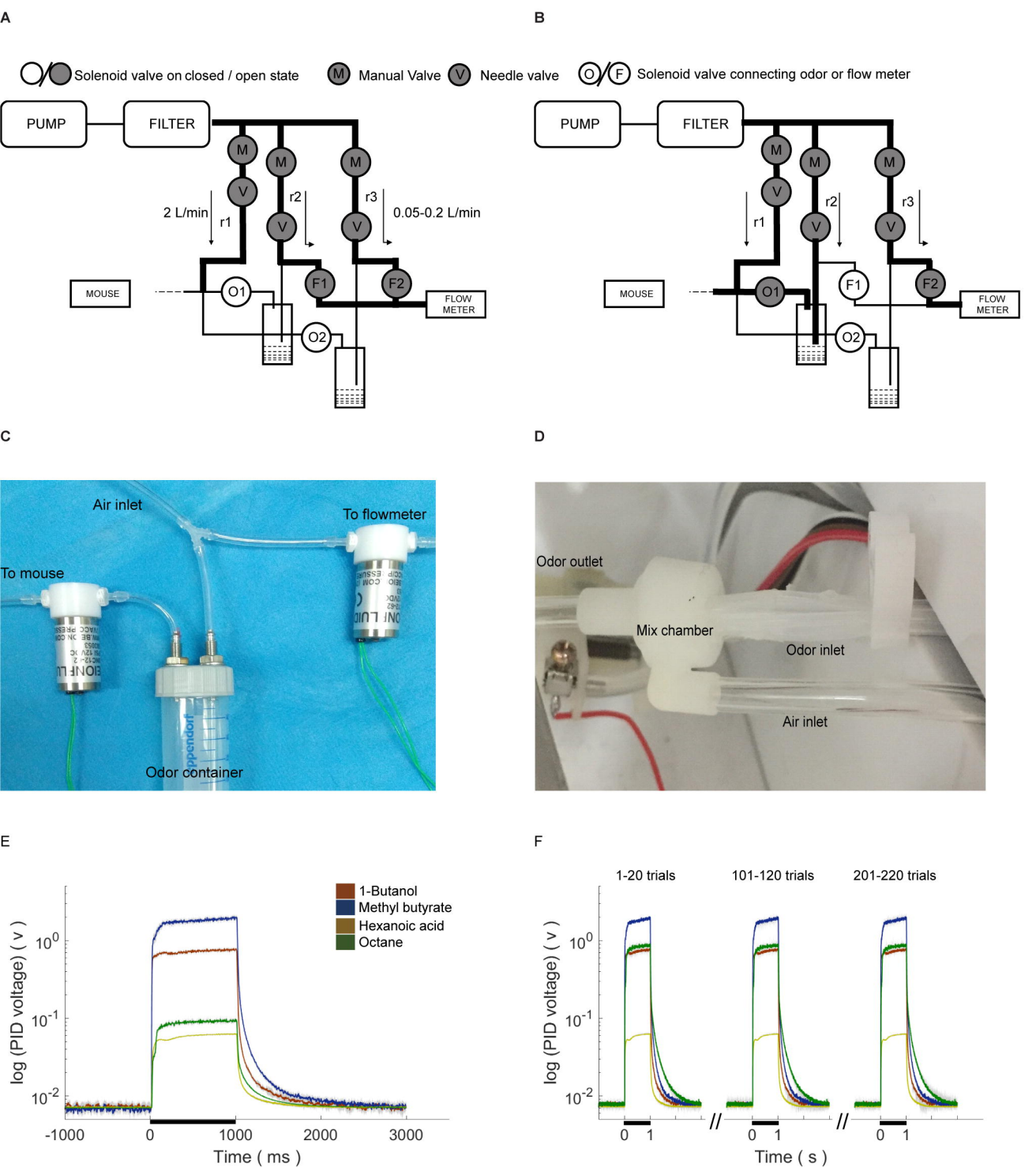

a
b
c
d
e
f
g
h
i
j
k

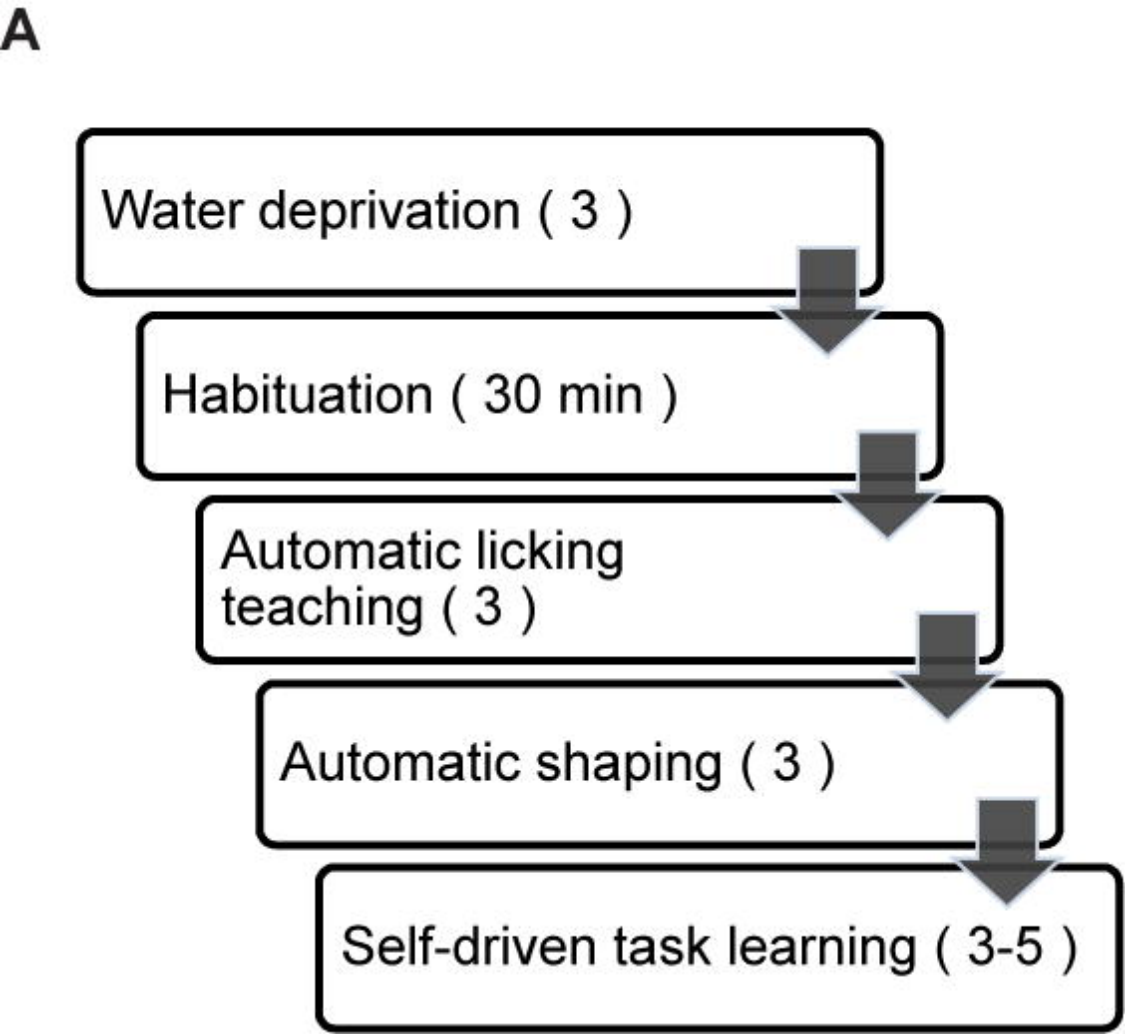
C

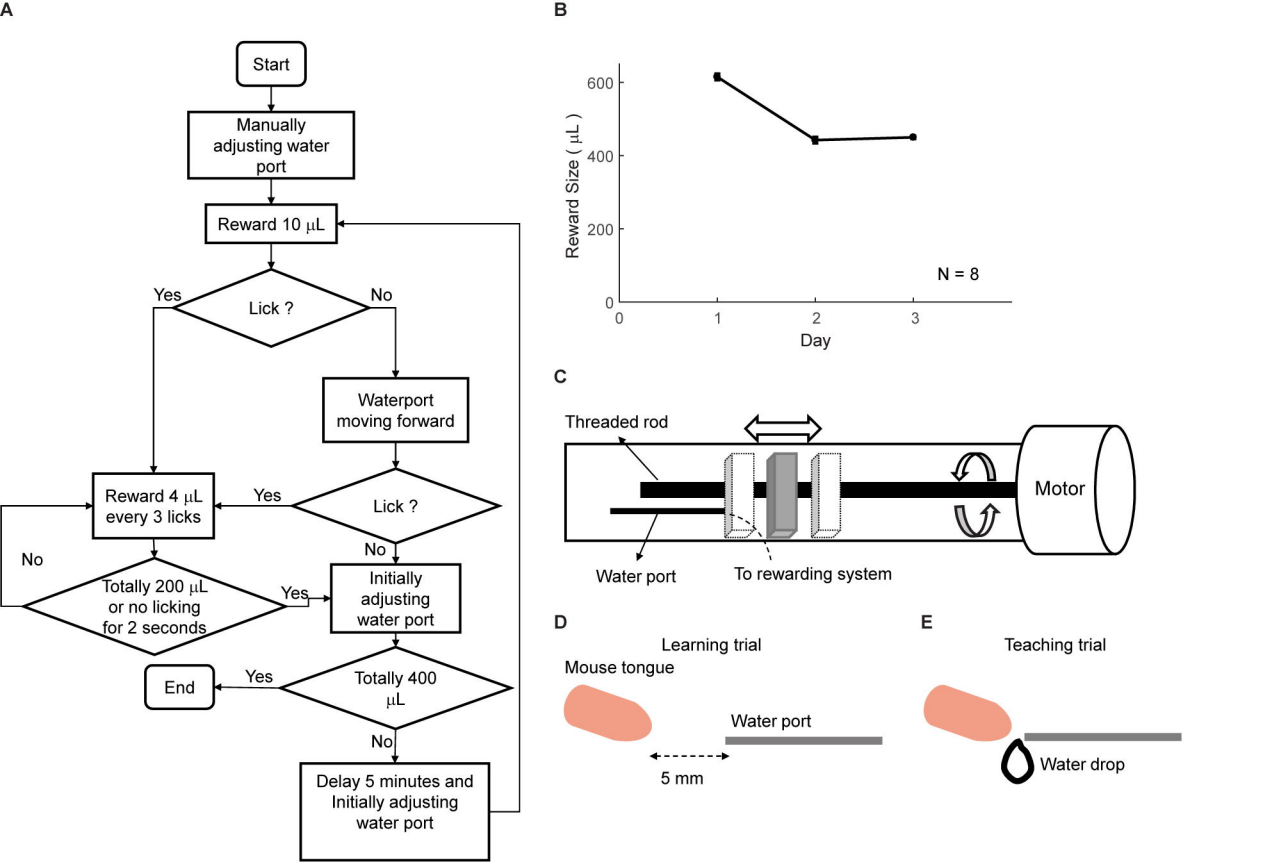

①
②
③
④
⑤
⑥

D

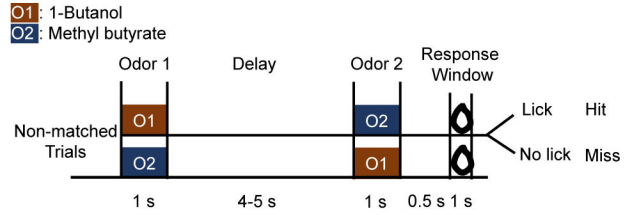

System initialization
Animal head fixation
Select training protocol
Data acquisition


A
Solenoid valve on closed / open state
M
Manual Valve
V
Needle valve
O/F
Solenoid valve connecting odor or flow meter
PUMP
FILTER
2 L/min
r1
r2
r3
0.05-0.2 L/min
MOUSE
O1
O2
F1
F2
FLOW METER
B
PUMP
FILTER
r1
r2
r3
MOUSE
O1
O2
F1
F2
FLOW METER
C
Air inlet
To flowmeter
To mouse
Odor container
D
Odor outlet
Mix chamber
Odor inlet
Air inlet
E
1-Butanol
Methyl butyrate
Hexanoic acid
Octane
log (PID voltage) (v)
10^0
10^{-1}
10^{-2}
-1000
0
1000
2000
3000
Time (ms)
F
1-20 trials
101-120 trials
201-220 trials
log (PID voltage) (v)
10^0
10^{-1}
10^{-2}
0 1
0 1
0 1
Time (s)

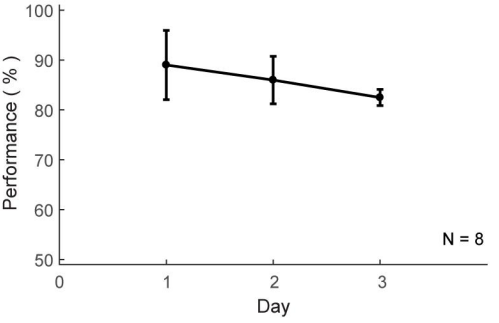

A
Water deprivation (3)
Habituation (30 min)
Automatic licking teaching (3)
Automatic shaping (3)
Self-driven task learning (3-5)


A
Start
Manually adjusting water port
Reward 10 μL
Lick ?
Yes
No
Waterport moving forward
Lick ?
Yes
Reward 4 μL every 3 licks
No
Totally 200 μL or no licking for 2 seconds
Yes
Initially adjusting water port
Totally 400 μL
Yes
End
No
Delay 5 minutes and Initially adjusting water port
B
Reward Size (μL)
600
400
200
0
0
1
2
3
Day
N = 8
C
Threaded rod
Motor
Water port
To rewarding system
D
Learning trial
Mouse tongue
Water port
5 mm
E
Teaching trial
Water drop

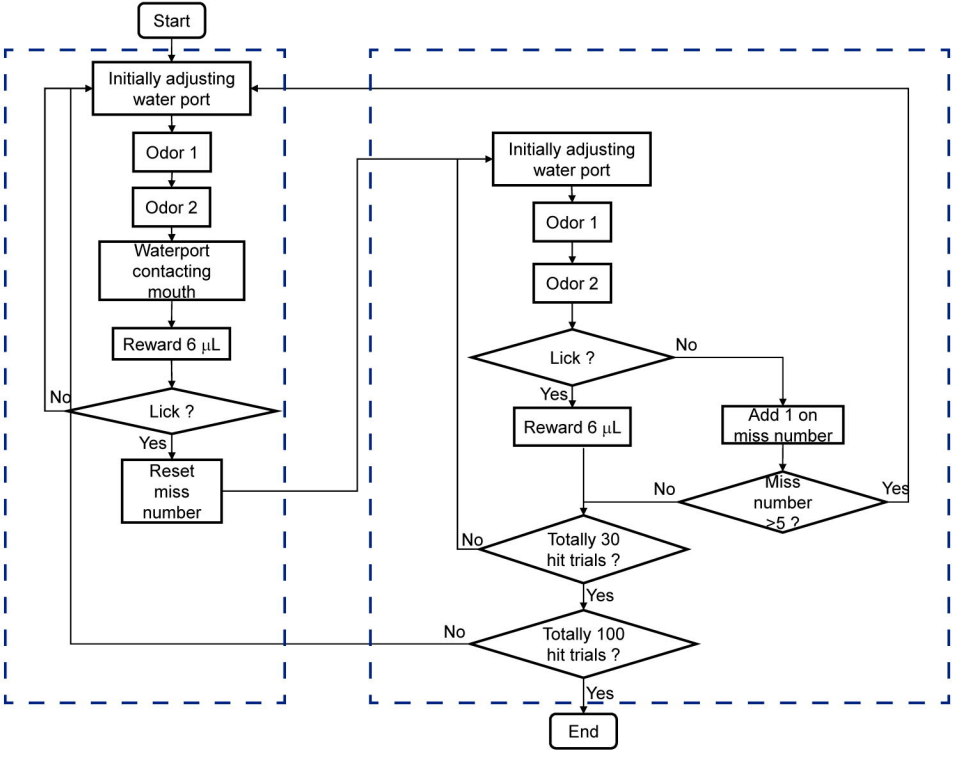
A


O1: 1-Butanol
O2: Methyl butyrate
Odor 1
Delay
Odor 2
Response Window
Non-matched Trials
O1
O2
O2
O1
Lick
Hit
No lick
Miss
1 s
4-5 s
1 s
0.5 s 1 s

B


Performance (%)
100
90
80
70
60
50
0
1
2
3
Day
N = 8

C


Start
Initially adjusting water port
Odor 1
Odor 2
Waterport contacting mouth
Reward 6 μL
Lick ?
No
Yes
Reset miss number
Initially adjusting water port
Odor 1
Odor 2
Lick ?
No
Yes
Reward 6 μL
Add 1 on miss number
Miss number >5 ?
No
Yes
Totally 30 hit trials ?
No
Yes
Totally 100 hit trials ?
No
Yes
End

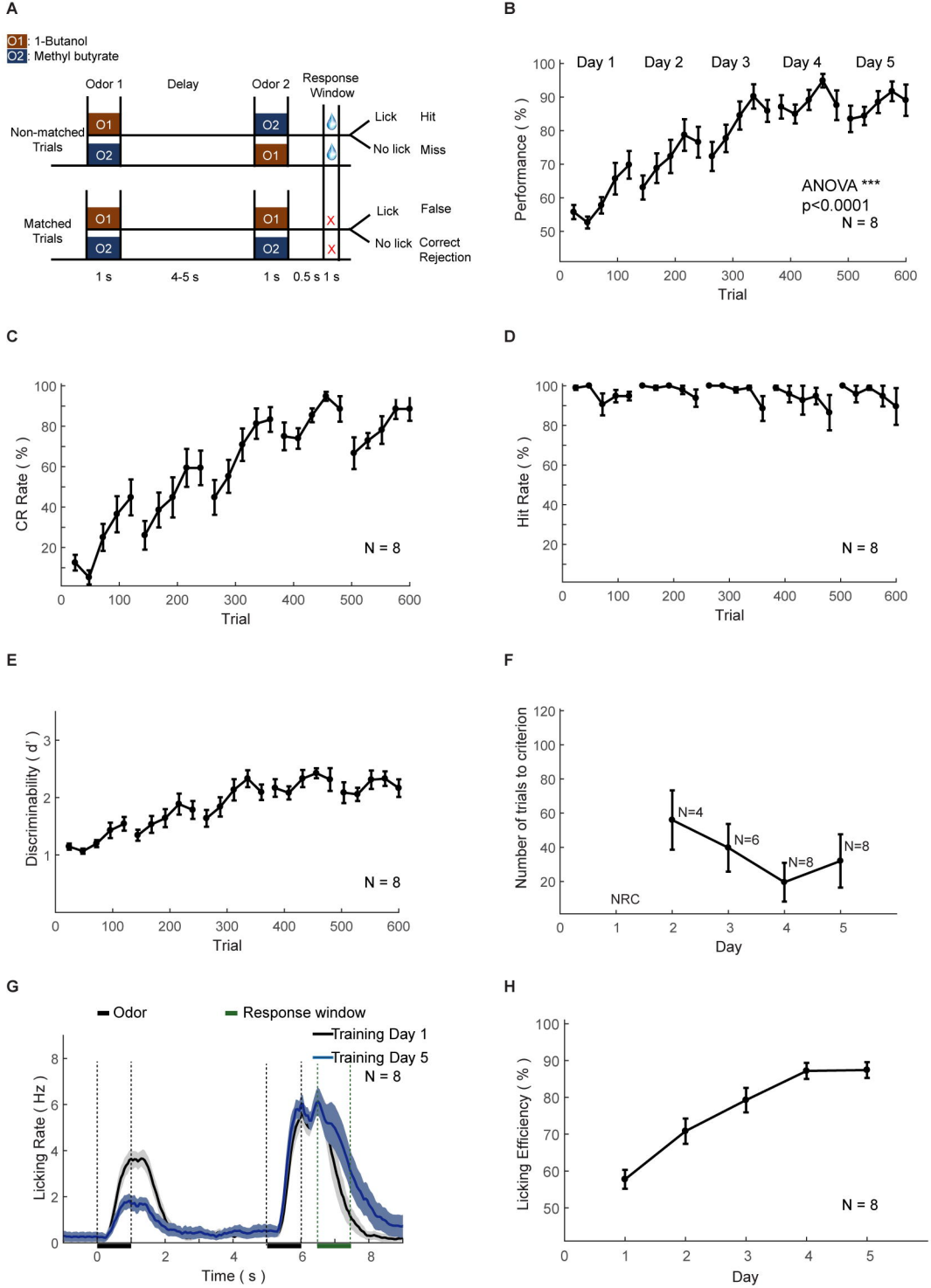


Figure 6 (Han et al., 2017)

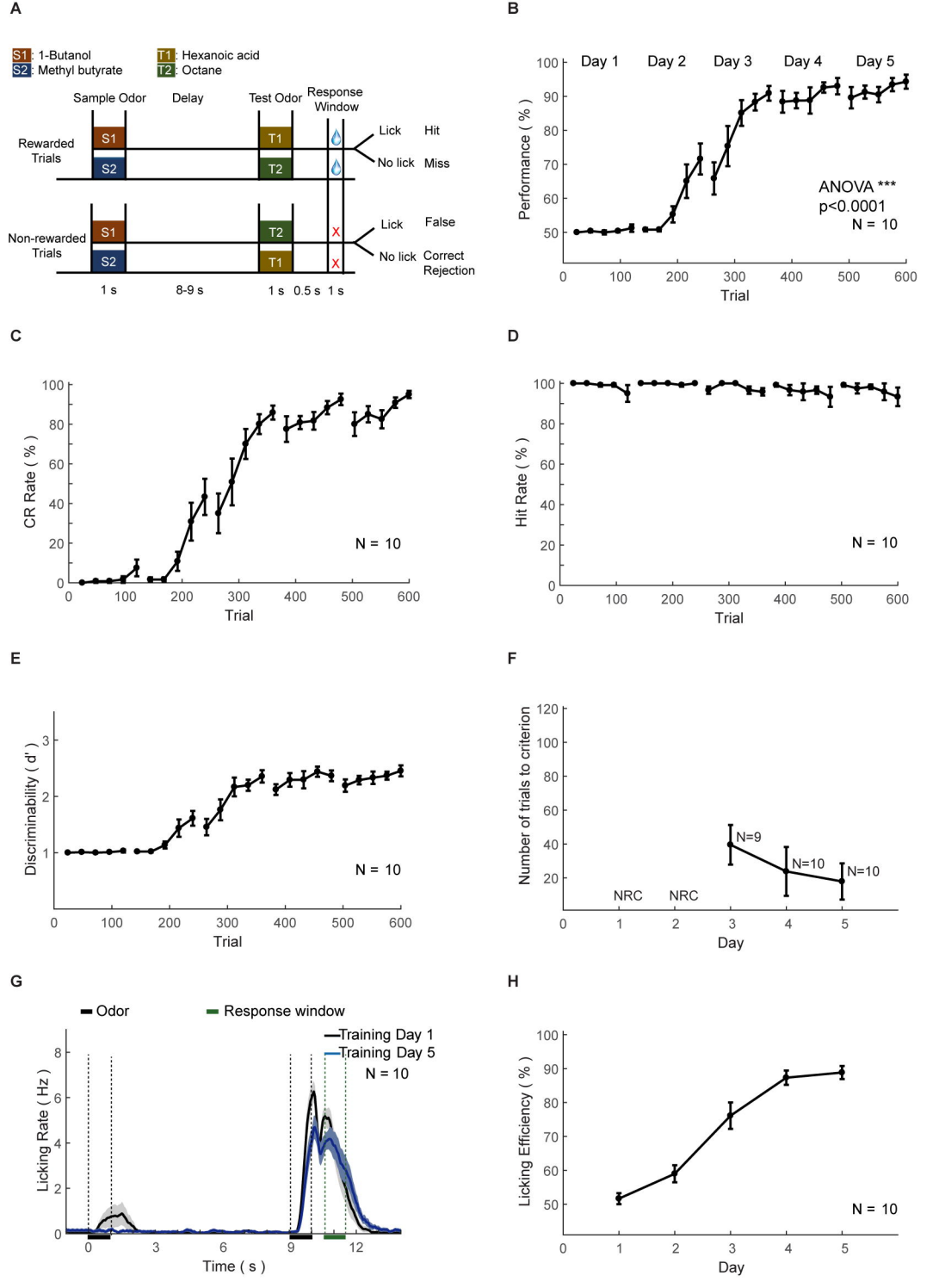


Figure 7 (Han et al., 2017)

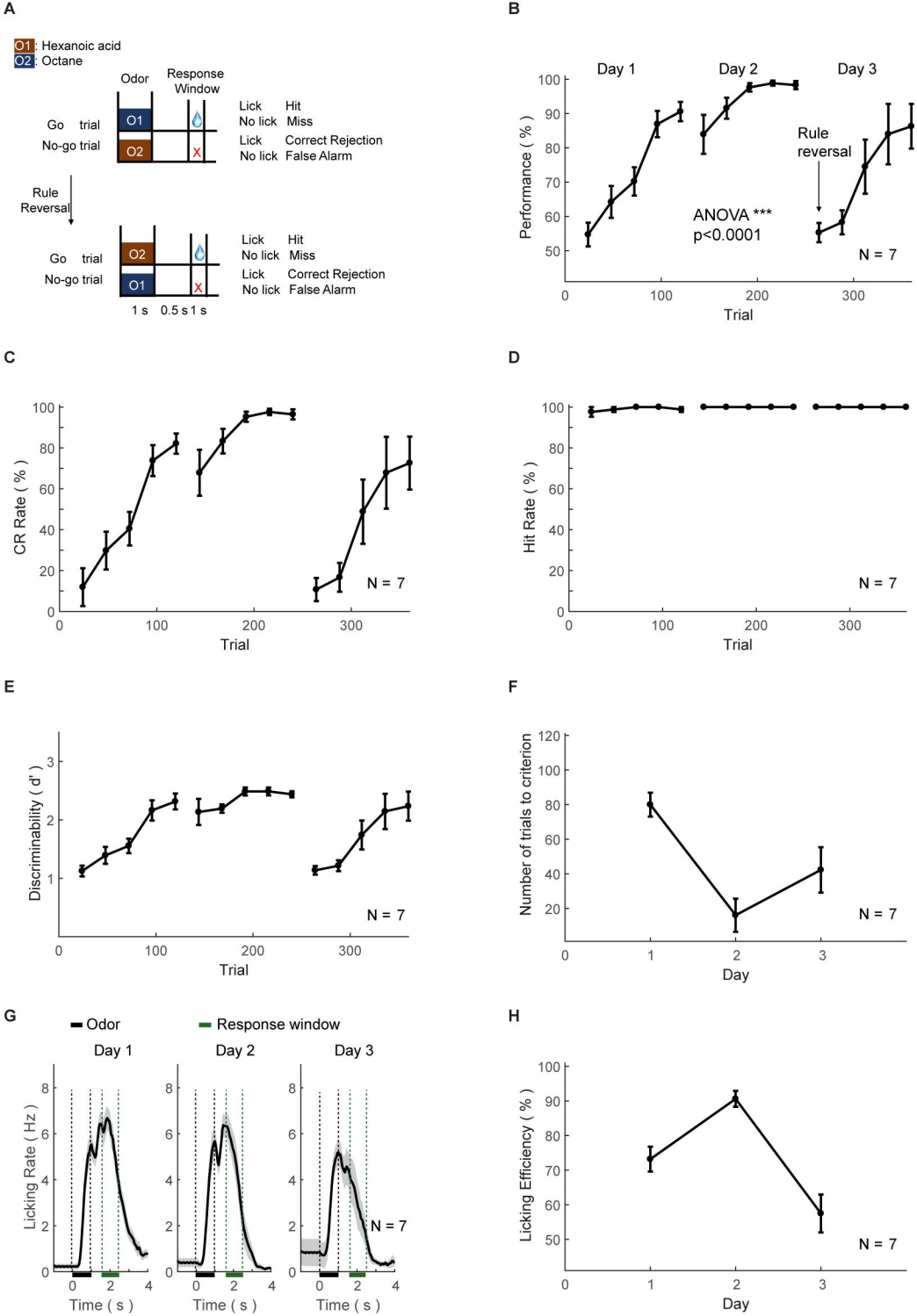


Figure 8 (Han et al., 2017)