Automatic quantification of disgust reactions in mice using machine learning

Shizuki Inaba, Naofumi Uesaka*, Daisuke H. Tanaka*

Department of Cognitive Neurobiology, Graduate School of Medical and Dental Sciences, Institute of Science Tokyo, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8519 Japan.

*Authors for correspondence

Naofumi Uesaka; Department of Cognitive Neurobiology, Graduate School of Medical and Dental Sciences, Institute of Science Tokyo, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8519 Japan; Tel +81-3-5803-5445; e-mail: uesaka.cnb@tmd.ac.jp

Daisuke H. Tanaka; Department of Cognitive Neurobiology, Graduate School of Medical and Dental Sciences, Institute of Science Tokyo, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8519 Japan; Tel +81-3-5803-5445; e-mail: dhtanaka.mphm@tmd.ac.jp

Abstract

Disgust is a primary negative emotion that is crucial for avoiding intoxication and infection. Disgust in rodents has been quantified as the score of disgust reactions triggered by an unpleasant taste. Disgust reactions were video-recorded and manually quantified, requiring significant time and effort for analysis. Here we developed a method to automatically count disgust reactions in mice by using machine learning. The disgust reactions were automatically tracked using DeepLabCut as the coordinates of the nose and both front and rear paws. The automated tracking data were split into test and training data, and the latter were combined with manually labeled data on whether a disgust reaction was present and, if so, which type of disgust reaction was present. Then, a random forest classifier was constructed, and the performance of the classifier was evaluated in the test dataset. The total number of disgust reactions estimated by the classifier highly correlated with those counted manually (Pearson's $r = 0.98$). The present method will decrease the time and effort required to analyze disgust reactions, thus facilitating the implementation of the taste reactivity test in large-scale screening and long-term experiments that necessitate quantifying a substantial number of disgust reactions.

Introduction

Disgust is elicited by distaste for food and contact with feces. Disgust has been recognized as a primary negative emotion since Darwin [1]. Disgust is crucial for avoiding intoxication and infection in animals [2]. Abnormal manifestations of disgust have been observed in patients receiving chemotherapy for cancer [3-4] and have been proposed in the pathogenesis of various intractable psychiatric disorders, including eating disorders and obsessive-compulsive disorder [5-7]. Therefore, elucidating the neural mechanisms underlying disgust is essential for understanding the biological basis of emotion, reducing drug side effects, and effectively managing various psychiatric disorders.

Since disgust is a conscious experience and thus cannot be directly observed [8-12], an observable readout is required to elucidate its neural mechanisms. Taste reactivity, consisting of orofacial and somatic behavioral reactions in response to taste stimuli, expresses the hedonic evaluation of taste, but not taste quality [13-15]. Several specific behaviors, such as gape, in taste reactivity have been established as reliable indicators of disgust [15] and are referred to as disgust reactions [16-19].

The quantification of disgust reactions is so complex that manual analysis has been carried out and automated analysis methods have not been developed. In previous analyses of disgust reactions, orofacial and somatic behavioral reactions were video-recorded, and the relevant behaviors were subjectively distinguished while the video was played frame-by-frame. Thus, the quantified data of a particular test can be somewhat variable between the people analyzed, and a significant amount of time and effort is required for the analysis [20].

To overcome these limitations and reproducibly and quickly assess disgust reactions in mice, we constructed a classifier to automatically count them. To accomplish this, we utilized DeepLabCut as a pose estimation system based on transfer learning with deep neural networks that track defined body parts of animals without physical markers [21] and random forest as a supervised learning model with an ensemble learning method [22].

Results

Taste reactivity test for innate and learned disgust reactions

Disgust reactions can be observed under two different experimental conditions: one is innate, typically elicited by intraoral quinine stimulation in mice, and the other is learned through associative learning between a particular taste, such as sweet taste, and subsequent nausea, typically elicited by intraperitoneal LiCl injection [18]. To assess disgust reactions in the innate and learned experimental conditions, mice were divided into an innate group (n = 13) and a learned group (n = 14). The same surgical procedures to implant the intraoral tube were performed in all mouse groups. The saccharin solution was infused intraorally through the tube to allow the mice in the learned group to learn conditioned disgust. This was followed by intraperitoneal injection of LiCl solution on days one and 4. On the test day, the mice received an intraoral infusion of saccharin solution. Orofacial and somatic behaviors during the infusions were video-recorded in a bottom-up view at 30 frames per second (Fig. 1). Mice in the innate group received an intraoral infusion of quinine solution followed by an intraperitoneal injection of saline on day one and day 4. The mice received an intraoral infusion of quinine solution on the test day, and their behavior was video recorded.

Manual analysis of disgust reactions

First, we manually and quantitatively analyzed the video-recorded disgust reactions. We played the recorded videos frame by frame to identify behaviors that could be distinguished as known disgust reactions, such as chin rubs, face washes, forelimb flails, gapes, and head shakes. A total of 97200 frames from all 27 videos were labeled as five types of disgust reactions and others (Fig. 2A). For the five types of disgust reactions, we analyzed consecutive frames containing the same reaction as one bout (Fig. 2B). Consistent with a previous study [18], disgust reactions were observed in both the innate and learned groups, with similar frequencies of occurrence for each type of disgust reaction (Fig. 2). Regarding each type of disgust reaction, forelimb flails, face washes, and chin rubs were more common than gapes and head shakes (Fig. 2B, Wilcoxon signed-rank sum test; gape vs. chin rub: $p=0.049$, gape vs. face wash: $p<1e-6$, gape vs. forelimb flail: $p<1e-7$, head shake vs. chin rub: $p=0.031$, head shake vs. face wash: $p<1e-7$, head shake vs. forelimb flail: $p<1e-7$).

Automated body part tracking

Next, we automatically estimated the mouse posture to perform automatic quantification of

video-recorded disgust reactions. Because the best number of body parts to be labeled for posture estimation is yet to be well established, we referred to the occurrence rate of each type of disgust reaction (Fig. 2) to determine the number of markers. The movements that can be tracked by focusing on the nose and the front and rear paws (i.e., chin rubs, face washes, forelimb flails, and head shakes) comprised the majority of disgust reactions. In contrast, movements that required attention to mouth movements (i.e., gapes) were rarely observed. Thus, we defined the five body parts of mice, the nose, and the front and rear paws (Fig. 3A) were sufficient for the present study. DeepLabCut2.2rc1 [23] was used to develop a tracking framework with five mouse body parts. A total of 2674 frames from 27 single-trial video recordings were manually annotated. The pose-estimation errors were evaluated after the model training procedure. The results showed that the mean error for each body part was less than 1 mm (Fig. 3B), which was sufficiently small compared to previous studies using DeepLabCut (1-8 mm) [24-25] to confirm that the training was complete.

Outlier frame detection and correction

Next, we corrected for mouse posture estimation because our posture estimation often failed in specific postures (e.g., standing up with its paws against a wall) and rapid movements that could not be captured on a recording. We introduced a rule-based correction procedure that refers to the likelihood of preventing incorrect estimations in such cases. Specifically, the following two corrections were made for all frames in the video posture-estimated by DeepLabCut when the estimated likelihood of a particular body part in a specific frame was less than the predetermined threshold value: (1) If the likelihood of the body part in the next frame exceeded the threshold value, the position of the body part in the frame was corrected to the middle of the frame before and after. (2) If the likelihood of the body part in the next frame was below the threshold, the body part's position in the frame was corrected to the same position as that in the previous frame. In this study, we set the threshold value to 0.6, which is the default value for an outlier correction method inside DeepLabCut.

Random forest classifier

We developed a random forest classifier to automate the image classification task of determining whether each video frame contained the behaviors of interest. Before developing the classifier, the data were divided into training (n = 19; innate = 9, learned = 10) and test (n = 8; innate = 4, learned = 4) datasets and randomized equally between the innate and learned

groups. The classifier was constructed considering the imbalanced presence of frames containing movements of interest and was cross-validated for each mouse in the training dataset to prevent overfitting. After training, we evaluated the performance of the classifier in the test dataset. Predictions of the total number of individual disgust reactions were compared to values manually counted by humans. We found a high correlation between prediction and human count (Pearson's r = 0.98, Fig. 4, see Supplementary Movie).

We evaluated the classifier's frame-by-frame performance in discriminating each type of disgust reaction by calculating its positive predictive value (PPV) and sensitivity. PPV, also called precision, is expressed by the following equation:

$$\frac{true\ positive}{true\ positive\ +\ false\ positive}$$

Sensitivity, also called recall, is expressed by the following equation:

$$\frac{true\ positive}{true\ positive\ +\ false\ negative}$$

In the test set, face washes (sensitivity = 0.69, PPV = 0.85) (Fig. 5A), and forelimb flails (sensitivity = 0.56 and PPV = 0.74) (Fig. 5B) were detected frame by frame with reasonable accuracy. In contrast, chin rubs (sensitivity = 0.16, PPV = 0.58) (Fig. 5C), gapes (sensitivity = 0), and head shakes (sensitivity = 0) were almost undetectable (Table 1). Others (i.e., movements other than the behaviors of interest) were detected with 0.99 sensitivity and 0.97 PPV. Overall, the developed classifier was capable of performing sufficiently accurate video analysis at the frame level in terms of detecting the presence or absence of disgust reactions.

Discussion

We constructed a classifier using DeepLabCut and random forest to automatically count innate and learned disgust reactions in the taste reactivity test. The total predicted values of the classifier showed a strong correlation with human observations. The present study is the first to automatically assess disgust reactions in taste reactivity tests.

Our classifier was constructed according to the published taste reactivity test protocols. Therefore, the scores analyzed by our classifier can be compared with manually counted scores. We found a very high correlation in this comparison ($r = 0.98$, Fig. 4). The accuracy of our classifier is comparable to that of previous automated methods: the pup retrieval behaviors in mice were automatically assessed with 0.51-0.90 correlations using DeepLabCut and machine learning [24], and the number of chemical-induced scratching in mice was predicted with 0.98 correlations using a convolutional recurrent neural network [26].

Our classifier can be built with minimal resources. An inexpensive and commercially available handheld video camera is sufficient for recording mouse behavior. Our classifier required only 30 frames per second for analysis, which is equal to or less than that used in previous studies (30-60 frames/sec) [26-27]. Using videos with lower frame rates saves time for pose estimation and reduces the storage size of the original videos, ultimately improving the model [28]. In addition, the present method required only five tracking points, fewer than those in previous studies using DeepLabCut (8-13 points/mouse) [25, 29]. This reduces the burden of manual annotation and decreases the machine training costs.

The present study used DeepLabCut for markerless pose estimation because a previous study demonstrated its excellent performance [29]. Nonetheless, alternative open-source software programs such as DeepPoseKit [30] and SLEAP [31] allow for comparable estimations. Our present classifier may also assess disgust reactions, even if these programs are used instead of DeepLabCut. The research field of markerless pose estimation has made remarkable progress in recent years, and more accurate programs are likely to become available. Accordingly, automatic behavior evaluation procedures that perform pose estimation in advance, such as our present classifier, are expected to improve in accuracy because at least some of the failures of our classifier were a result of pose estimation failures (Fig. 5).

Regarding each type of disgust reaction, our classifier discriminated the presentation of forelimb flails and face washes with moderate accuracy but could not detect gapes and head shakes (Table 1). Despite this limitation, the low prevalence of gapes and head shakes in disgust reactions (Fig. 2) resulted in an overall accuracy comparable to that of human observations

(Fig. 4). Unfortunately, our present classifier may not be directly applicable to disgust reactions in rats, in which gapes are more frequently observed [32], compared with the present study in mice. To detect specific behaviors more accurately, increasing the number of tracking points, increasing the frame rate of the video, and capturing videos from multiple angles for 3D pose estimation may be effective.

A recent study evaluated facial expressions in response to taste stimuli in mice using a machine learning-based method [33]. Facial expressions in response to bitter stimuli were automatically distinguishable from those in response to the other stimuli. Bitter-induced facial expressions were also induced by the intake of saccharin solution paired with intraperitoneal LiCl injection and by the intake of NaCl solution at high concentrations. However, it remains unclear what specific brain functions are reflected in the observed facial expressions because the taste stimuli used may evoke multiple brain functions, including disgust and motivation to avoid the stimulants. In addition, the facial expressions detected by this method in mice were challenging to interpret as relevant facial expressions in other species, including humans, making it difficult to interpret the mouse's facial expressions as a readout of emotion. In contrast, all disgust reactions expressed by mice are shared by several different species, such as rats, monkeys, and human neonates [15]. This fact supports the idea that a mouse's disgust reactions and human disgust may share a common neural basis. In support of this, taste reactivity may reflect different brain functions and neural activities from facial expressions because the numbers of disgust reactions and facial expressions do not always fit in mice [34]. In addition to facial expression, intake and licking behaviors have been widely examined in taste avoidance and evaluation studies, but the results from these behaviors do not always match those of taste reactivity [20, 35]. Thus, taste reactivity will continue to be a valuable readout for future research to understand the neural basis of disgust.

In the present study, we developed an automated assessment method for mouse disgust reactions. The present method allows for the evaluation of disgust reactions independent of the skill level and bias of the human evaluator. In addition, the present method will facilitate the implementation of taste reactivity tests in large-scale screening and long-term experiments that require counting numerous disgust reactions, which are challenging to perform manually. Thus, the present method may help researchers implement taste reactivity tests in a broader range of studies than previously thought [20]. The present method is expected to accelerate our understanding of the neural basis of disgust, contribute to understanding the pathophysiology of various psychiatric disorders, and advance the development of preventive and therapeutic

interventions.

Materials and methods

Animals

The adult (6-7 weeks-old) male C57BL/6J mice were obtained from Japan SLC Inc. (Shizuoka, Japan). Mice were maintained at 23 ± 1 ˚C under a 12 h light/dark cycles (lights on at 8:00 am) and given *ad libitum* access to food and water. All the animal experiments were approved (No. 0150384A, 0160057C2, 0170163C, A2017-194A, and A2018-138C4) by the Institutional Animal Care and Use Committee of the Institute of Science Tokyo, and performed according to the relevant guidelines and regulations.

Surgery for intraoral tubing

Intraoral tube implantation was performed as previously described [18]. Briefly, a mixture of midazolam (4 mg/kg body weight (BW)), butorphanol (5 mg/kg BW), and medetomidine (0.3 mg/kg BW) was used to anesthetize mice. A curved needle attached to an intraoral polyethylene tube was inserted from the incision site and advanced subcutaneously posterior to the eye to exit lateral to the first maxillary molar on the right side of the mouth.

Taste reactivity tests

The taste reactivity tests were performed as previously described [18]. The experimental procedure is briefly described below. In the present study, videos obtained from five different experiments were used for the analysis (Table 2).

Experiment 1 for the innate group: On an experimental day (Day) 1, mice (n=5) received intraoral infusions of 50 μL of 3 mM quinine solution twice for 1 min each with a 1 min interval in between to adapt to the test procedure. On Day 2, the mice received the same infusions as those administered on day 1, and their behaviors during the infusions were recorded using a digital video camera (HDR-PJ800; Sony, Tokyo, Japan) for subsequent video analysis.

Experiment 2 for the innate group: On Day 1, mice (n=2) continuously received infusions of 80 μL of 3 mM quinine solution for 2 min without an interval to adapt to the test procedure. On Day 2, the mice received the same infusions as those performed on Day 1, and their behaviors during the infusions were video recorded.

Experiment 3 for the innate group: On Day 1, mice (n=6) received intraoral infusions of 50 μL of 3 mM quinine solution twice for 1 min each with a 1 min interval in between. Shortly thereafter, the mice received an intraperitoneal (i.p.) injection of physiological saline (10 mL/kg body weight [BW]) as a control treatment in the learned group. The same procedure

was repeated on Day 4. On Day 6, the mice received the same infusions as those performed on Day 1, and their behaviors during the infusions were video recorded.

Experiment 4 for the learned group: On Day 1, mice(n=11) received infusions of 50 μL of 5.4 mM saccharin solution twice for 1 min each with a 1 min interval in between. Shortly thereafter, the mice received an i.p. injection of LiCl solution (0.3 M, 10 mL/kg BW) (1st conditioning). The same procedure was repeated on Day 4 (2nd conditioning). On Day 6, the mice received the same infusions as those performed on Day 1, and their behaviors during the infusions were video recorded.

Experiment 5 for the Learned group: On Day 1, mice (n=3) continuously received infusions of 80 μL of 5.4 mM saccharin solution for 2 min without an interval. Shortly thereafter, the mice received an i.p. injection of LiCl solution. The same procedure was repeated on Day 4. On Day 6, the mice received the same infusions as those performed on Day 1, and their behaviors during the infusions were video recorded.

Manual count of disgust reactions

Manual video analysis was performed as previously described [18]. To be sure that each component of disgust reactions contributes equally to the final scores, reactions that occurred in continuous bouts were scored in time bins [15, 36]. Components characterized by long-duration bouts, such as chin rubs, were scored in five-second bins (successive repetitions within five seconds scored as one occurrence), and face wash was scored in two-second bins. The other three reactions that could occur as a single behavior were scored as separate occurrences (e.g., one forelimb flail equals one occurrence). The total disgust reactions were quantified as the sum of all five reactions during the two minutes of intraoral infusion of the test solutions.

Automated body part tracking

The videos used as part of the training or test datasets in the present study (n = 13; innate = 6, learned = 7) were obtained from the same videos from an earlier report in which automated tracking was not performed [18]. Initially, 20 frames were selected from each video for training purposes. The provisional coordinates of the mouse body part were estimated according to the temporarily created model, and 40 candidate frames for false predictions were selected for each video and manually annotated. This procedure was repeated until the error was sufficiently small (Fig. 3). Frame choice and training were performed using the default settings of

DeepLabCut and trained for 500000 iterations each, achieving a loss of <0.001.

Random forest classification

We extracted 25 per-frame geometrical features consisting of five distances, ten displacements, and ten displacement changes. In addition, five consecutive frames are considered the "segment," so each frame had 125 features. We then constructed a classifier using the caret [37] and Rborist [38] packages in R.

Computer software and hardware

A desktop computer equipped with an Intel Core i7-10700 CPU, an NVIDIA GeForce GTX 1660 SUPER GPU, and 16 GB of RAM was used for mouse posture estimation in DeepLabCut and for data processing in R.

References

[1] Darwin, C. The Expression of the Emotions in Man and Animals, 3rd edition, Fontana Press, London UK (1872).

[2] Rozin, P. & Fallon, A. E. A perspective on disgust. *Psychol. Rev.* **94**, 23–41 (1987).

[3] Bernstein, I. L. Learned taste aversions in children receiving chemotherapy. *Science* **200**, 1302–1303 (1978).

[4] Jacobsen, P. B. *et al*. Formation of food aversions in cancer patients receiving repeated infusions of chemotherapy. *Behav. Res. Ther.* **31**, 739–748 (1993).

[5] Phillips, M. L., Senior, C., Fahy, T. & David, A. S. Disgust – the forgotten emotion of psychiatry. *Br. J. Psychiatry* **172,** 373–375 (1998).

[6] Knowles, K. A., Jessup, S. C. & Olatunji, B. O. Disgust in anxiety and obsessive-compulsive disorders: Recent findings and Future Directions. *Curr. Psychiatry Rep*. **20,** 68 (2018).

[7] Anderson, L. M., Berg, H., Brown, T. A., Menzel, J. & Reilly, E. E. The role of disgust in eating disorders. *Curr. Psychiatry Rep*. **23,** 4 (2021).

[8] Nagel, T. What is it like to be a bat? *Philos. Rev.*, **4,** 435–450 (1974).

[9] Chalmers, D. J. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press. (1996).

[10] Chalmers, D. J. First-person methods in the science of consciousness. http://consc.net/papers/firstperson.html (1999).

[11] Velmans, M. Heterophenomenology versus critical phenomenology. *Phenomenol. Cogn. Sci.*, **6,** 221–230 (2007).

[12] Tanaka D. H. & Tanabe T. CHANging Consciousness Epistemically (CHANCE): An empirical method to convert the subjective content of consciousness into scientific data, *J. Mind Behav.* **40,** 177-190 (2019).

[13] Steiner J. E. The gustofacial response: observation on normal and anencephalic newborn infants. *Symp. Oral Sens. Percept*. **4,** 254–278 (1973).

[14] Grill, H. J. & Norgren, R. The taste reactivity test. I. Mimetic responses to gustatory stimuli in neurologically normal rats. *Brain Res*. **143,** 263–279 (1978).

[15] Berridge, K. Measuring hedonic impact in animals and infants: Microstructure of affective taste reactivity patterns. *Neurosci. Biobehav. Rev.* **24,** 173–198 (2000).

[16] Parker, L. A. Taste avoidance and taste aversion: Evidence for two different processes. *Learn. Behav.* **31**, 165–172 (2003).

[17] Ho, C. Y. & Berridge, K. C. Excessive disgust caused by brain lesions or temporary inactivations: mapping hotspots of the nucleus accumbens and ventral pallidum. *Eur. J. Neurosci.* **40**, 3556-3572 (2014).

[18] Tanaka, D. H. *et al.* Genetic access to gustatory disgust-associated neurons in the interstitial nucleus of the posterior limb of the anterior commissure in male mice. *Neurosci.* **413,** 45–63 (2019).

[19] Tanaka, D. H. *et al.* Genetic recombination in disgust-associated bitter taste-responsive neurons of the central nucleus of amygdala in male mice. *Neurosci. Lett.* **742**, 135456 (2021).

[20] Dwyer, D. M. Licking and liking: The assessment of hedonic responses in rodents. *Q. J. Exp. Psychol.* **65**, 371–394 (2012).

[21] Mathis, A. *et al.* Deeplabcut: Markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289 (2018).

[22] Breiman, L. Random Forests. *Machine Learning* **45**, 5–32 (2001).

[23] Nath, T. *et al.* Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nat. Protoc.* **14,** 2152–2176 (2019).

[24] Winters, C. *et al.* Automated procedure to assess pup retrieval in laboratory mice. *Sci. Rep.* **12,** 1663 (2022).

[25] Nilsson, S. R. O. *et al.* Simple Behavioral Analysis (SimBA) – an open source toolkit for computer classification of complex social behaviors in experimental animals. Preprint at https://www.biorxiv.org/content/10.1101/2020.04.19.049452v2 (2020).

[26] Kobayashi, K. *et al.* Automated detection of mouse scratching behaviour using convolutional recurrent neural network. *Sci. Rep.* **11,** 658 (2021).

[27] Jhuang, H. *et al.* Automated home-cage behavioural phenotyping of mice. *Nat. Commun.* **1,** 68 (2010).

[28] von Ziegler, L., Sturman, O. & Bohacek, J. Big behavior: challenges and opportunities in a new era of deep behavior profiling. *Neuropsychopharmacol.* **46**, 33–44 (2021).

[29] Sturman, O. *et al.* Deep learning-based behavioral analysis reaches human accuracy and is capable of outperforming commercial solutions. *Neuropsychopharmacol.* **45**, 1942–1952 (2020).

[30] Graving, J. M. *et al.* DeepPoseKit, a software toolkit for fast and robust animal pose estimation using Deep Learning. *eLife* **8,** e47994; 10.7554/eLife.47994 (2019).

[31] Pereira, T. D. *et al.* SLEAP: A deep learning system for multi-animal pose tracking. *Nat. Methods* **19**, 486–495 (2022).

[32] Berridge, K., Grill. H. J. & Norgren. R. Relation of consummatory responses and preabsorptive insulin release to palatability and learned taste aversions. *J. Comp. Physiol. Psychol.* **95**, 363-382 (1981).

[33] Dolensek, N., Gehrlach, D. A., Klein, A. S. & Gogolla, N. Facial expressions of emotion states and their neuronal correlates in mice. *Science* **368,** 89–94 (2020).

[34] Kawai, H *et al*. Median raphe serotonergic neurons projecting to the interpeduncular nucleus control preference and aversion. *Nature Communications*. **13**, 7708 (2022).

[35] Lin, J.-Y., Arthurs, J. & Reilly, S. Conditioned taste aversion, drugs of abuse and palatability. *Neurosci. Biobehav. Rev.* **45,** 28–45 (2014).

[36] Koizumi, M., Cagniard, B. & Murphy, N. P. Endogenous nociception modulates diet preference independent of motivation and reward. *Physiol. Behav.* **97,** 1–13 (2009).

[37] Kuhn, M. caret: Classification and Regression Training. R package version 6.0-93. https://CRAN.R-project.org/package=caret (2022).

[38] Seligman, M. Rborist: Extensible, Parallelizable Implementation of the Random Forest Algorithm. R package version 0.3-2. https://CRAN.R-project.org/package=Rborist (2022).

Acknowledgments

This work was supported by JST SPRING, Grant Number JPMJSP2120 to S.I., JSPS KAKENHI Grant Number JP19K06938, and TMDU Priority Research Areas grant to D.H.T. We thank Dr. Tsutomu Tanabe at TMDU for his helpful comments and discussion.

Author contributions

N.U. supervised this project. S.I. and D.H.T. conceived the project and designed experimental and analytical strategies. D.H.T. conducted experiments. S.I. analyzed the data and wrote the code. S.I. drafted the manuscript, which was discussed and edited by all the co-authors.

Data availability

The videos, code, and generated datasets used in the present study can be provided by the corresponding authors upon reasonable request.

Additional Information

Competing Interests

The authors declare that they have no conflicts of interest.

Figure legends

Figure 1

Experimental setup for taste reactivity test in mice

(A) This scheme shows the Experimental timeline.

Experiment 1,2 (Innate): On the experimental day (Day) 1, mice received intraoral infusions of Quinine solution as habituation. On Day 2, the mice received the same infusions as those performed on Day 1, and the behaviors of mice during the infusions were video recorded. In Experiment 1, infusions were performed twice for a total of 1 min each, with a 1 min interval in between. In Experiment 2, the infusions were continuous for 2 min, without intervals.

Experiment 3 (Innate): On Day 1, mice received intraoral infusions of quinine solution, and shortly thereafter received an i.p. injection of physiological saline as a control treatment for the learned groups described below. The same procedure was repeated on Day 4. On Day 6, the mice received intraoral infusions of quinine solution, and the behaviors of the mice during the infusions were video-recorded. In Experiment 3, infusions were performed twice for a total of 1 min each, with a 1 min interval in between.

Experiment 4,5 (Learned): On Day 1, mice received intraoral infusions of saccharin solution, and shortly after that, they received an i.p. injection of LiCl solution (1st conditioning). The same procedure was repeated on Day 4 (2nd conditioning). On Day 6, the mice received intraoral infusions of saccharin solution, and the behaviors of the mice during the infusions were video-recorded. In Experiment 4, infusions were performed twice for a total of 1 min each, with a 1 min interval in between. In Experiment 5, the infusions were continuous for 2 min without intervals.

(B) The solution was infused intraorally through the implanted tube. Orofacial and somatic behaviors during infusions were video-recorded in a bottom-up view.

Figure 2

The number of each type of disgust reaction

(A) Number of frames that included any disgust reactions confirmed manually.

(B) The number of bouts for each type of disgust reaction was confirmed manually.

Figure 3

Annotation of five body parts for pose estimation

(A) Definition of body parts used to estimate mouse posture using DeepLabCut. Only those

visible in the bottom-up view were annotated. Scale bar, 1 cm.

(B) Root mean square error (RMSE) for each body part during the training.

Figure 4

The classifier's predictions were highly correlated with human observations

Comparison of disgust reaction counts for each video between observations and predictions in the test dataset. Pearson's r = 0.98.

Figure 5

Typical examples of positive and false classifications by the present random forest classifier

The nose, right front paw, and left front paw of the mouse are highlighted with green, yellow, and blue crosses, while the right and left rear paws are marked with yellow and blue dots, respectively. Successive images were recorded at intervals of 33 milliseconds (ms) (1/30th of a second). Scale bar, 1 cm.

(A) Examples of the classifier's positive and false classifications of face wash. (Top) A series of frames in which the classifier correctly classifies the face wash. (Bottom) In this series, the classifier did not detect face wash, and there was a pose estimation error: at t = 0 and 33 ms, the right front paw positions were correctly estimated (red arrow), but at t = 67 ms, they were incorrectly estimated (red arrowhead).

(B) Examples of the classifier's positive and false classifications of forelimb flails. (Top) A series of frames in which the classifier correctly classified the forelimb flails. (Bottom) A series of frames in which the classifier did not detect forelimb flails and there was no pose estimation error.

(C) Examples of the classifier's positive and false classifications of chin rubs. (Top) A series of frames in which the classifier correctly classifies chin rubs. In this series, the left front paw of the mouse was moving so quickly that it could not be accurately recorded on the video; at t = 33 and 67 ms, its position was correctly estimated (red arrow), but at t = 0 ms, its position was incorrectly estimated (red arrowhead). However, pose estimation was successful. (Middle) In this series, the classifier did not detect chin rub, and there was a pose estimation error: at t = 0 and 67 ms, the right front paw positions were correctly estimated (red arrow), but at t = 33 ms, they were incorrectly estimated (red arrowhead). (Bottom) The classifier did not detect chin rubs in this series, and there was no pose estimation error.

Supplementary Movie

Typical example of video recording in a mouse expressing disgust reactions.

The nose, right front paw, and left front paw of the mouse are highlighted with green, yellow, and blue crosses, while the right and left rear paws are marked with yellow and blue dots, respectively. The left area shows the total number of disgust reactions and each type of disgust reaction, with the upper half counted by human observers and the lower half automatically determined by the classifier. A video was recorded at 30 fps.

Figure 1

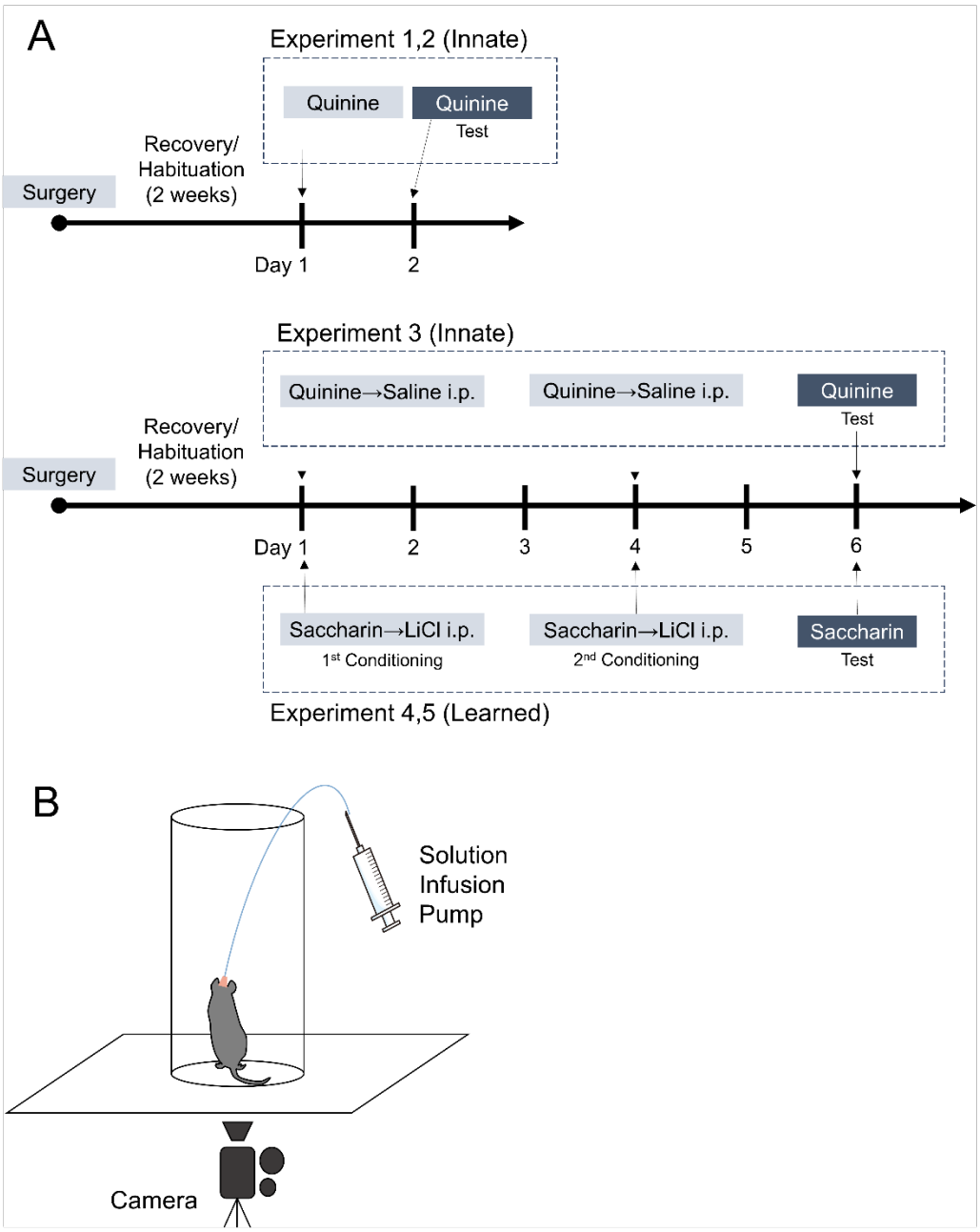
A
Experiment 1,2 (Innate)
Quinine
Quinine
Test
Recovery/
Habituation
(2 weeks)
Surgery
Day 1
2
Experiment 3 (Innate)
Quinine→Saline i.p.
Quinine→Saline i.p.
Quinine
Test
Recovery/
Habituation
(2 weeks)
Surgery
Day 1
2
3
4
5
6
Saccharin→LiCl i.p.
1st Conditioning
Saccharin→LiCl i.p.
2nd Conditioning
Saccharin
Test
Experiment 4,5 (Learned)
B
Solution
Infusion
Pump
Camera

Figure 2

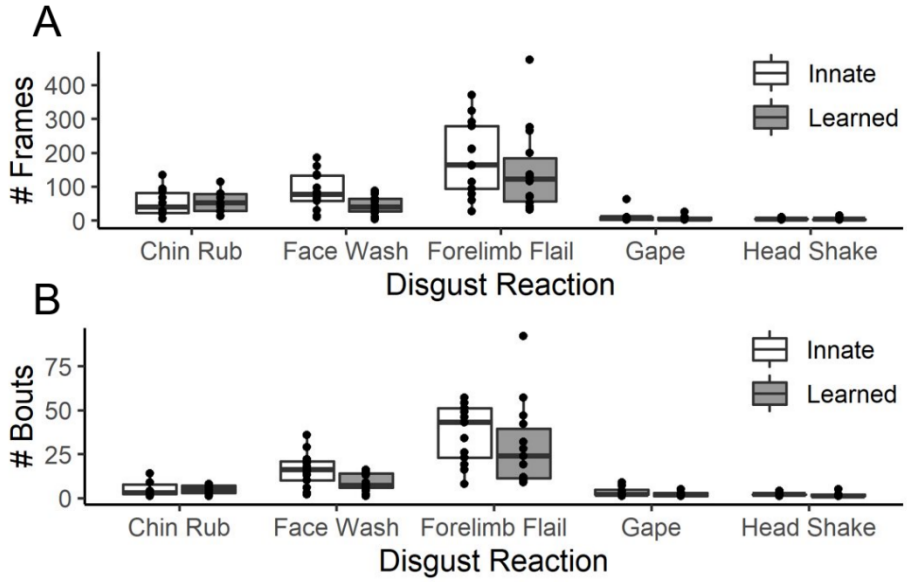
A
Frames
400
300
200
100
0
Innate
Learned
Chin Rub
Face Wash
Forelimb Flail
Gape
Head Shake
Disgust Reaction
B
Bouts
75
50
25
0
Innate
Learned
Chin Rub
Face Wash
Forelimb Flail
Gape
Head Shake
Disgust Reaction

Figure 3

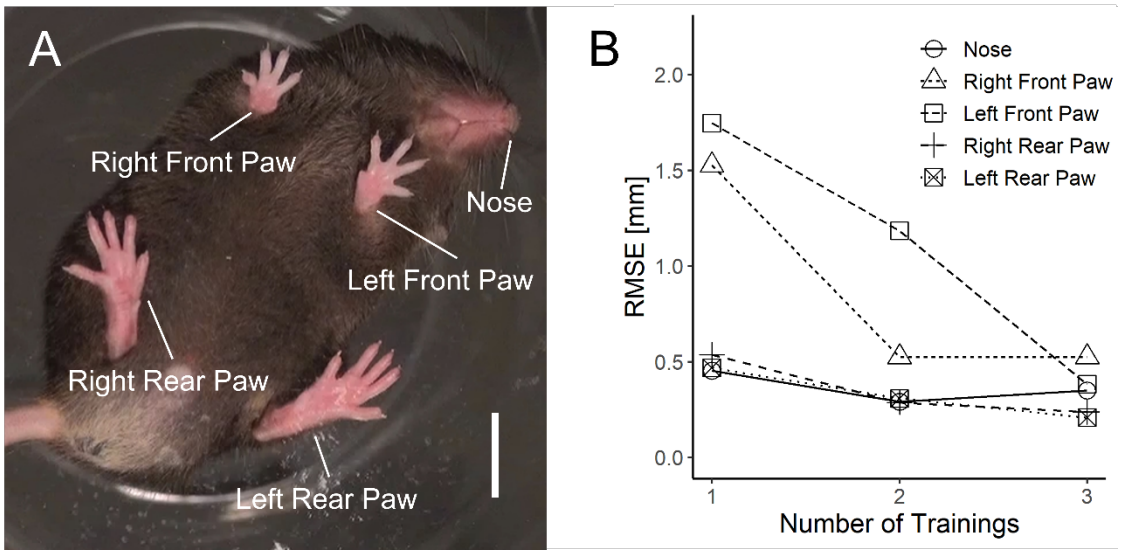

Figure 4

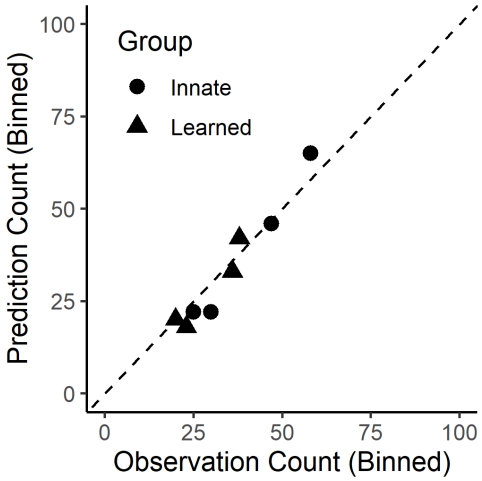

Figure 5

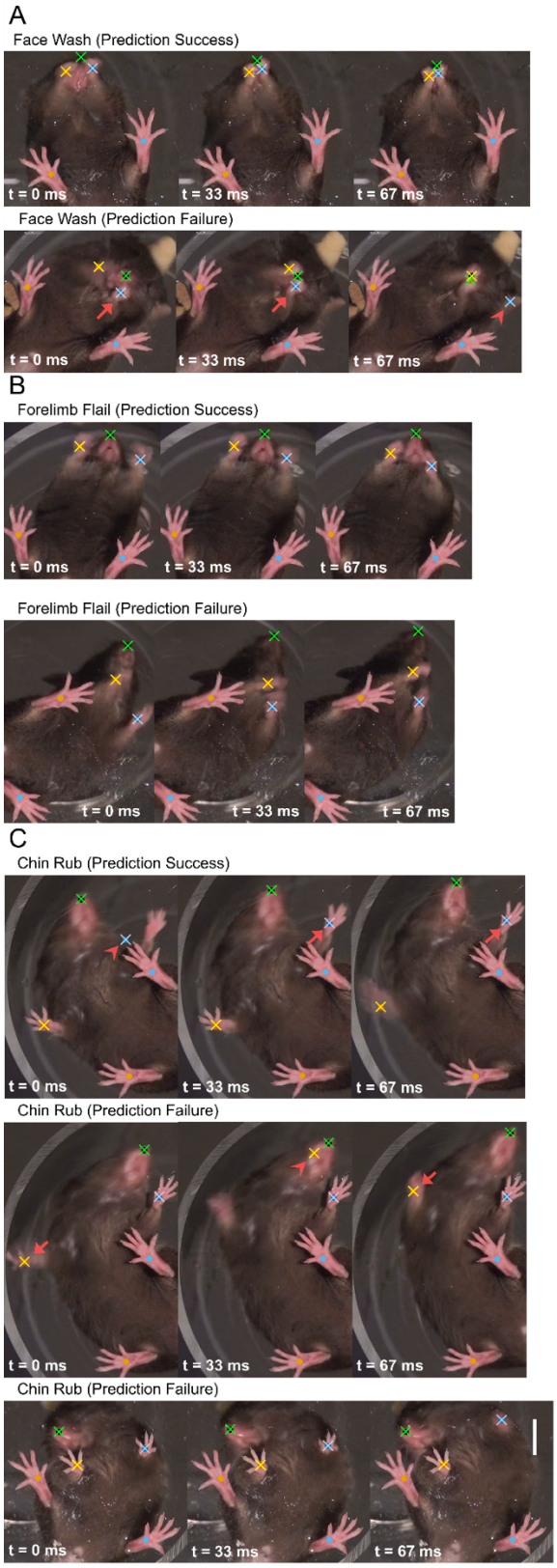
A
Face Wash (Prediction Success)
t = 0 ms
t = 33 ms
t = 67 ms
Face Wash (Prediction Failure)
t = 0 ms
t = 33 ms
t = 67 ms
B
Forelimb Flail (Prediction Success)
t = 0 ms
t = 33 ms
t = 67 ms
Forelimb Flail (Prediction Failure)
t = 0 ms
t = 33 ms
t = 67 ms
C
Chin Rub (Prediction Success)
t = 0 ms
t = 33 ms
t = 67 ms
Chin Rub (Prediction Failure)
t = 0 ms
t = 33 ms
t = 67 ms
Chin Rub (Prediction Failure)
t = 0 ms
t = 33 ms
t = 67 ms

Table 1

Observation and prediction of the number of frames that include each type of disgust reaction and other movement

	Observation						
Prediction		Chin Rub	Face Wash	Forelimb Flail	Gape	Head Shake	Others
	Chin Rub	45	0	0	0	0	33
	Face Wash	0	283	25	0	0	26
	Forelimb Flail	0	49	496	0	1	121
	Gape	0	0	0	0	0	0
	Head Shake	0	0	0	0	0	0
	Others	236	77	362	22	18	27006

Table 2

Taste reactivity test schedule

	Innate		Learned
	Expt. 1,2	Expt. 3	Expt. 4,5
	Surgery		
	Recovery / Habituation		
Day 1	Q	Q +Saline	S +LiCl
Day 2	Q (Test)		
Day 3			
Day 4		Q +Saline	S +LiCl
Day 5			
Day 6		Q (Test)	S (Test)

This table presents an experimental timeline. All mice underwent the same surgeries, recoveries, and habituations. On the final experimental (Test) day, all mouse behaviors were video-recorded for analysis.

Q: Intraoral quinine infusion, S: Intraoral saccharin infusion, +Saline: Saline i.p. immediately after intraoral infusion, +LiCl: LiCl i.p. immediately after intraoral infusion.