

RESEARCH ARTICLE

Automated facial feature evaluation system to prevent stress of head fixed mice

Anna Nasr^{1,2*}, Gianna Rettinger³, Huibert D. Mansvelder², Robert N.S. Sachdev¹, Matthew E. Larkum¹

1 Institut für Biologie, Neurocore Center for Excellence, Charité Universitätsmedizin Berlin and Humboldt Universität, Berlin, Germany, **2** Vrije Universiteit Amsterdam, Amsterdam, Netherlands, **3** Technische Universität Berlin Institut für Biotechnologie, Berlin, Germany

* anna.nasr87@gmail.com



OPEN ACCESS

Citation: Nasr A, Rettinger G, Mansvelder HD, Sachdev RNS, Larkum ME (2025) Automated facial feature evaluation system to prevent stress of head fixed mice. PLoS One 20(6): e0322530. <https://doi.org/10.1371/journal.pone.0322530>

Editor: Yukiori Goto, Kyoto University Graduate School of Informatics: Kyoto Daigaku Daigakuin Johogaku Kenkyuka, JAPAN

Received: October 22, 2024

Accepted: March 24, 2025

Published: June 23, 2025

Copyright: © 2025 Nasr et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: All data are available at <https://github.com/Nasr-SFB1315/MouseCare>.

Funding: The following funding sources have supported this project: (1) Deutsche Forschungsgemeinschaft (DFG), Grant

Abstract

Head fixation of rodents is a widely utilized and important technique that enables laboratories to measure brain activity during behavior, but head fixation can increase stress which affects both behavior and underlying brain activity, as well as animal welfare. It is therefore critical to keep stress levels low, yet it is particularly challenging to assess stress in immobilized, head-fixed rodents. Conventional stress evaluation methods, such as blood corticosterone analysis, are labor-intensive and conducted post hoc, and In situ approaches require experienced personnel and constant vigilance during experiments. In this study, we present MouseCare, an automated software solution for immediate stress detection by real-time facial feature video analysis in head-fixed mice. MouseCare performs on par with or better than conventional stress measures, and is significantly less labor-intensive. This approach enables objective measures of stress that are needed to determine when an experiment can commence, but also when it should be stopped. We conclude that MouseCare offers a cost-effective and easily implementable strategy to manage stress levels in rodents that can increase data quality and improve animal welfare.

Introduction

Head fixation is a technique widely used in neuroscience laboratories to immobilize the head of an experimental animal [1–8], facilitating various essential experimental approaches [9–16], including 2-photon imaging [17] and patch-clamp recordings in vivo [18]. To immobilize the head, a fixed head-post is secured to the animal's skull to restrict head movement. While head fixation is invaluable for scientific research, prolonged immobilization can be stressful, impacting their welfare and experimental outcomes [19–22]. Additionally, it can be expensive and time-consuming, requiring specialized equipment and skilled personnel. Therefore, it is essential to constantly

Nos. 246731133, 250048060, 267823436, 3442/16-1, and 467613712 – SFB 1315 to ML; (2) Human Brain Project, EU Commission Grant 720270 (SGA1), 785907 (SGA2) and 945539 (SGA3) to ML; (3) Einstein Foundation Berlin EVF-2017-363 to ML; (4) Charité 3R - Universitätsmedizin Berlin, Grant 5000005001 to AN. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

monitor stress levels in mice to ensure their well-being and the reliability of experimental results. Balancing scientific needs and animal welfare is crucial [23].

The first step to mitigating stress is early detection. Habituation and training are necessary for animals to become stress-free under head fixation. Familiarizing the animal with the setup and tasks before recording can help reduce stress [24], but objective assessment of when a mouse is habituated is challenging. A method to systematically and objectively track stress during habituation is important for optimizing the well-being of animals undergoing this procedure.

Detecting stress in head fixed mice is critical for accurate scientific research. Among various methods, measuring stress hormone levels in blood via tail vein is the gold standard due to its precision. However, this method presents a paradox. The procedure of blood sampling induces mild to moderate stress in animals [25–27]. Thus, the most accurate method to detect stress—blood sampling—inevitably is invasive and causes stress. This contradiction poses a significant challenge, as the act of measuring stress can alter the very parameter being measured. Therefore, there is a pressing need to develop an accurate, real-time, non-invasive stress detection method to reduce experimental stress.

One such alternative is the analysis of corticosterone metabolites in feces, which offers a non-invasive approach to stress detection. While this method significantly reduces the stress associated with sample collection, it comes with its own set of challenges, particularly in terms of accuracy. The levels of corticosterone metabolites can be influenced by various factors, including the daily rhythm (circadian cycle), sex, age, and individual metabolic differences. These influences can introduce variability and complicate the interpretation of results. Additionally, corticosterone metabolite analysis requires collecting and processing stool samples, which introduces a significant time lag between the occurrence of stress and the detection of stress [28–31]. This delay is acceptable when assessing overall stress levels in an environment over time, but incompatible with the need for real-time stress monitoring.

Analyzing stress indicators in urine [32], saliva [33], or milk [34], have also been explored. However, these methods often require extra equipment to extract the samples, adding complexity and potential stress to the process. Despite these limitations, corticosterone metabolite analysis represents a promising step towards balancing the need for reliable stress measurement with the imperative to minimize animal distress.

In addition to biochemical methods, observational techniques offer a valuable means of assessing stress in real-time. Researchers can utilize the Mouse Grimace Scale, a validated tool for identifying pain and distress through facial expressions, to evaluate stress on the fly [35–37]. This method allows for immediate assessment and intervention, enhancing the welfare of laboratory animals, without the need for invasive procedures or complex sample analysis. However, the accuracy of this observational approach heavily relies on the skill and experience of the researcher, which can introduce variability and potential bias into the stress assessment [37,38]. Moreover, researchers are typically unable to simultaneously attend to ongoing experimental needs and assess stress.

Recent advances in artificial intelligence, such as deep neural networks, have provided new opportunities for detecting stress in head fixed mice. A deep neural network is a type of machine learning algorithm that can analyze large amounts of data and identify patterns and relationships between variables.

These networks can be trained to track individual pre-defined points in a video, an example of which is DeepLabCut [39]. Using this as a basis, it is possible to track and evaluate changes in different facial features (similar to how the grimace scale works). This information can then be used to update the researcher during the experiment on the well-being of the head fixed mice, independent of the researcher's experience.

Overall, detecting stress in head fixed mice is crucial for ensuring their welfare and the validity of scientific experiments. Therefore, we have developed MouseCare, a cost-effective and easily implementable strategy to evaluate stress levels in head fixed mice.

Materials and methods

All experiments were conducted in accordance with the guidelines of animal welfare of the Charité Universitätsmedizin Berlin and the local authorities (Landesamt für Gesundheit und Soziales, LAGeSo). All mice underwent surgery at the beginning of the experiment.

Surgery

Adult C57BL/6 mice (n = 24) were anesthetized with Ketamine/Xylazine (95 mg/kg and 5 mg/kg), positioned in a stereotaxic apparatus (David Kopf Instruments, California, US) and placed on a heating pad (FHC Inc. Maine US). During surgery, the eyes were covered with ointment (Bepanthen, Bayer, Leverkusen, Germany) to prevent them from drying out. Before making an incision, Lidocaine was injected under the skin and the scalp was disinfected with ethanol. The skull was exposed, the fascia on the bone scraped off with a dental scraper and the skull was air dried. A light-weight aluminum head post was laid on the skull and RelyX (3M, Minnesota, US) cement was used to affix the head post to the skull [40–42]. Black jet acrylic (Ortho Jet, Lang Dental) was used as a second layer to cover the exposed bone and to enhance the cementing of the head post. For analgesic purposes, Buprenorphin (0.1 mg/kg) was injected IP on the day of surgery and Carprofen (5 mg/kg) was injected IP on the day of surgery and for two days post surgery.

In addition, we received video footage of head fixed mice from other projects. This allowed us to evaluate our system with different types of head fixed experiments, without requiring additional mice. The corticosterone metabolite analysis of our stool samples was performed by Prof. Dr. R. Palme from the Vetmeduni Vienna Unit of Physiology, Pathophysiology and Experimental Endocrinology.

Tracking facial features with MouseCare

The automated stress detection, with MouseCare (<https://github.com/Nasr-SFB1315/MouseCare>), uses pre-processing by DeepLabCut (DLC) to track specific facial features of head fixed mice. DLC is an open-source toolbox for detecting features using deep learning that is particularly useful for tracking moving body parts in rodent experiments [39]. In this study, we created a neural network (using the DLC toolbox) and trained it on the facial features of head fixed mice. To do so, we used 24 different 5-minute Videos and extracted 12 frames per video. The training was done after 180000 iterations. Additional setups with different lighting required 2–3 extra videos to further train the network on the new environment. We labeled seven distinct facial features on a set of training images: each ear, each upper and lower eyelid, and the nose (Fig 1A). These features were chosen on the basis of their efficacy for detecting different stress levels and their distinctiveness due to movement, size, and position (versus other features such as cheeks, whiskers, etc. which we found were not as useful for stress detection). The network was trained to accurately identify these features in a variety of video frames, achieving an accuracy rate of 99% (Fig 1B and C). Further examples from different setups are provided in (S1 Fig). This high level of

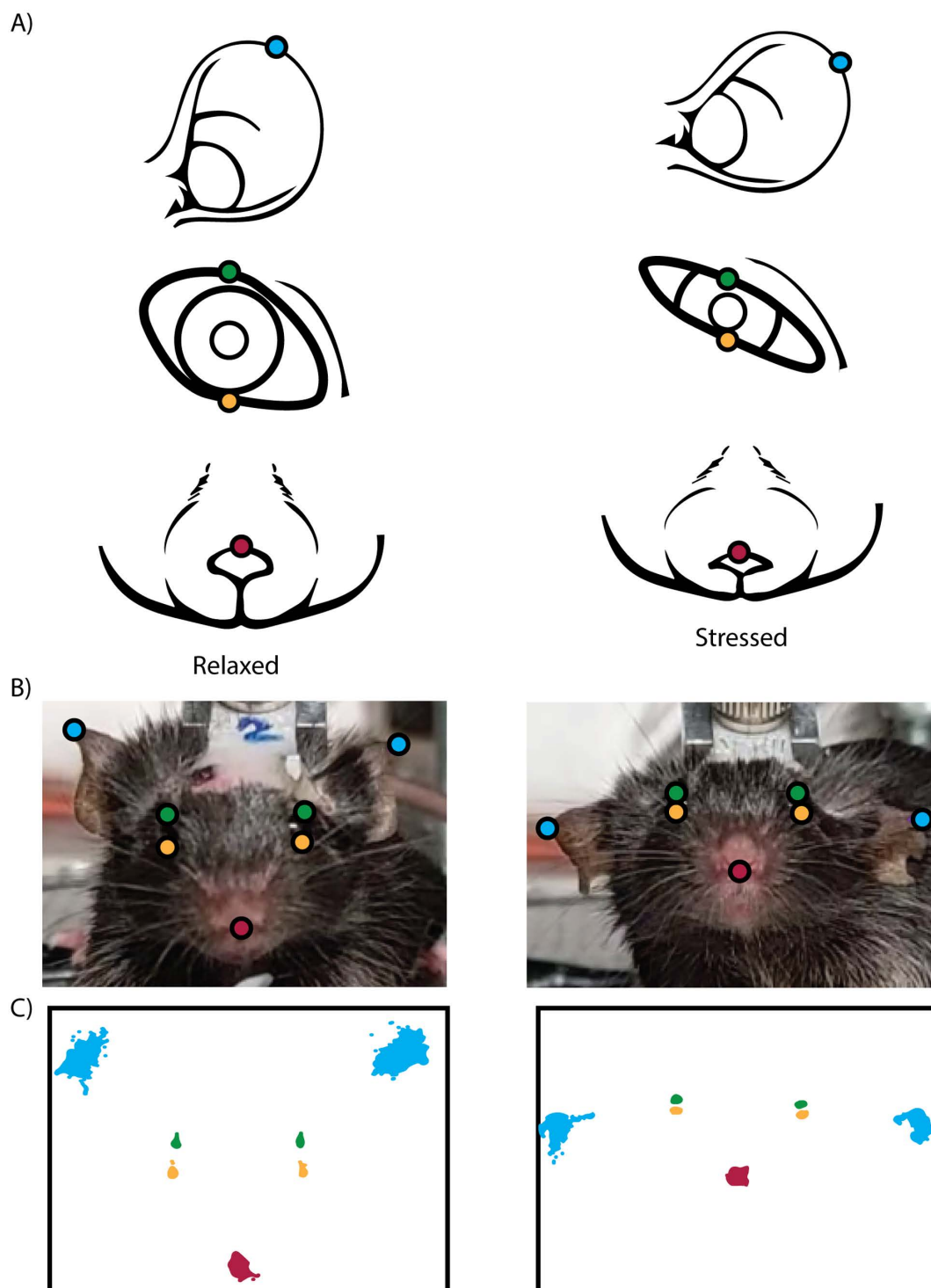


Fig 1. Tracking the facial features of a head fixed mouse. A) Illustration of the different facial features that are tracked and their change and positional change from relaxed to stressed. B) Screen capture of Video footage from a camera facing head fixed mouse with the tracking points overlaid. C) Positional frame summary of the movement range for each facial feature.

<https://doi.org/10.1371/journal.pone.0322530.g001>

accuracy ensures reliable tracking of these features across different video sessions. Once the network training was completed, we applied it to pre-recorded videos to analyze the position of each labeled facial feature over time (Fig 1C). The coordinates of these features were extracted and stored in csv or h5 file formats for further analysis. The stability of MouseCare depends on the balance between camera hardware and computational capacity. While higher frame rates and resolutions can provide more detailed data, they also generate larger data loads that may overwhelm the processing capabilities of the system. We used a frame rate of 30–60 FPS. This was sufficient for tracking facial features accurately in our setups. While higher frame rates (e.g., 200 FPS) are technically feasible, they significantly increase the data load and may cause performance issues depending on the computer's hardware. The resolution of 1080p (1920x1080 pixels) was sufficient for accurate tracking using colored cameras. Similar to the frame rate, higher resolution were possible but not feasible. We tested different cameras and lighting in the setups. Screen captures of those recordings are shown in (S1 Fig).

MouseCare system for stress detection

We developed a Python-based software, to perform the stress detection assessment using our facial feature tracking network (developed with the DLC toolbox) with heuristic based algorithms. This complete system we call MouseCare. The customized software component of MouseCare was designed to analyze the positional data of the tracked facial features and detect stress in the mice. To achieve this, we designed MouseCare to integrate the coordinate data and apply a series of heuristics to assess stress-related behaviors (see below). Here, using the processed DLC output, the rest of the system identifies patterns and indicators of stress, leveraging behavioral analysis techniques, and statistical analyses to provide robust and reliable stress detection. The results could then be compiled and visualized for interpretation.

The individual positions of each facial feature (Fig 1A) as well as the movement of these features were crucial for the analysis and were evaluated over specific periods (epochs) to exclude false positives. Squinting eyes are a known sign of stress and discomfort [43], but can also be triggered by other factors. For instance, in a previous experiment, we observed squinting and rapid blinking when mice received a liquid reward from a licking port [24]. We distinguished these types of squinting by their duration: reward-induced squinting lasted less than one second, while stress-induced squinting persisted for extended periods. This is achieved by calculating the Euclidean distance ($d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$) between the upper and lower eyelid while averaging the x and y over the fps. These values are stored in a dataframe. Subsequently, the dataframe is analyzed to count consecutive entries where the eyelids remain close together, excluding rapid blinking.

Freezing is another stress response. To measure it, we calculated the positional changes of different facial features (ears and nose) over a 2-second time span. Rodents are able to move their nose in 3 dimension while head-fixed [44]. Using a front-facing camera, we recorded nose movements along the left/right (X) and up/down (Y) axes, storing positional data for each frame over 2 seconds. From this data, we calculated the Maximum Range of Motion (The largest displacement of the nose during the 2-second window) and the total Nose Travel Distance (The cumulative Euclidean distance between consecutive nose positions) These metrics were combined to compute the Nose Activity Index, a quantitative measure of nose movement. The same methodology was applied to the ears. In addition to tracking ear movement, we compared positional data between the two ears to detect asymmetry, as unilateral ear movement can indicate discomfort. Finally, we used the maximum range of motion to determine if an ear was folded to one side, another potential indicator of stress.

Live stress detection with MouseCare

Our goal was to generate a user-friendly system for live feedback on the stress state of the mouse, prioritizing the ability of the researcher to react immediately to ongoing changes in stress levels. To achieve this, we used the open source tool-kit DeepLabCut Live (DLC Live) [45], which used our pre-trained DLC deep neural network, to pipeline the detected coordinates to MouseCare. The MouseCare software continuously evaluated the stress of the mouse according to the stress heuristics on a scale of 0–10. To facilitate easy interpretation of stress levels during experiments, MouseCare displays live

stress levels with a colour-coded 3-value system (Fig 2A, okay/green, caution/yellow, stressed/red). Open Broadcaster Software (OBS) was used in conjunction with a local camera to stream video to an external computer (Fig 2B). This has the dual benefit of allowing for centralization of computer resources and also for remote monitoring. By sending only the current status back to the experimental computer, we drastically reduced the burden on local real-time computation.

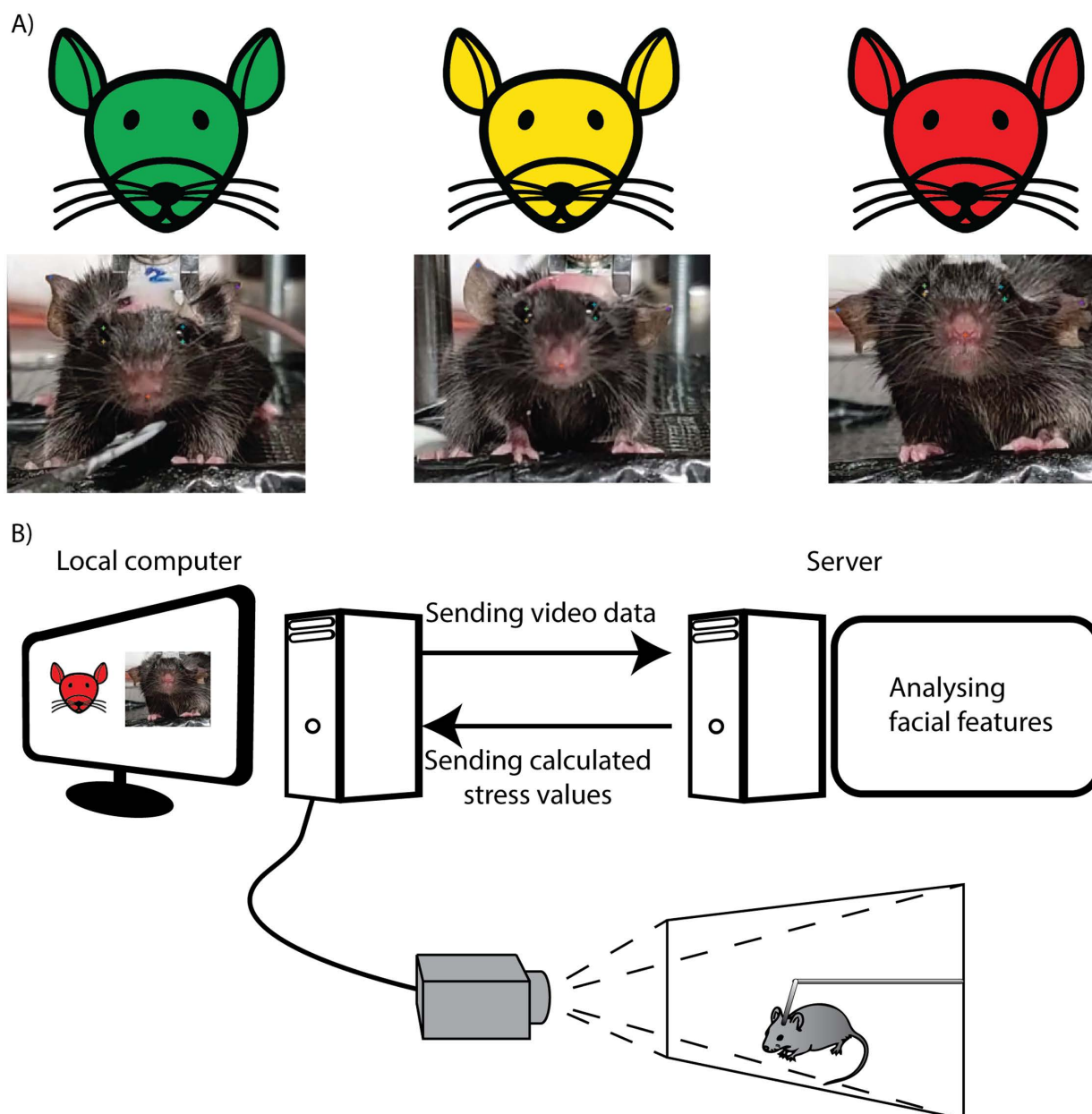


Fig 2. Schematic of the live video transfer from the experimental computer to the analysis computer and the return of the stress evaluation results. A) Depiction of the 3 different stages of stress shown via video capture and the corresponding icon. B) Schematic representation of the camera setup and the data transfer between the experimental computer and analysis computer.

<https://doi.org/10.1371/journal.pone.0322530.g002>

Importance of weight adjustment for facial feature analysis

The variability in head fixation setups, particularly head posts, can significantly impact the movement of facial features in mice, and may also obscure them to a camera. For that, it is important to determine the effect it can have on the stress evaluation. We have separated them into three different categories: All facial features are visible, temporarily obstructed facial feature and permanent obscured a facial feature (Fig 3A).

MouseCare was designed to be flexible across multiple systems, we implemented a dynamic weighting system, which allows the user to adjust the evaluation of facial features at the start of each session. This system ensures that re-training was not required for different head posts, which might obstruct facial features [1–8]. The visibility and tracking accuracy for each of our setup was tested and evaluated. This was tested in six different setups with different cranial attachments, four of which are shown in (S1 Fig). It is important to highlight that while in some setups, all facial features might be visible but

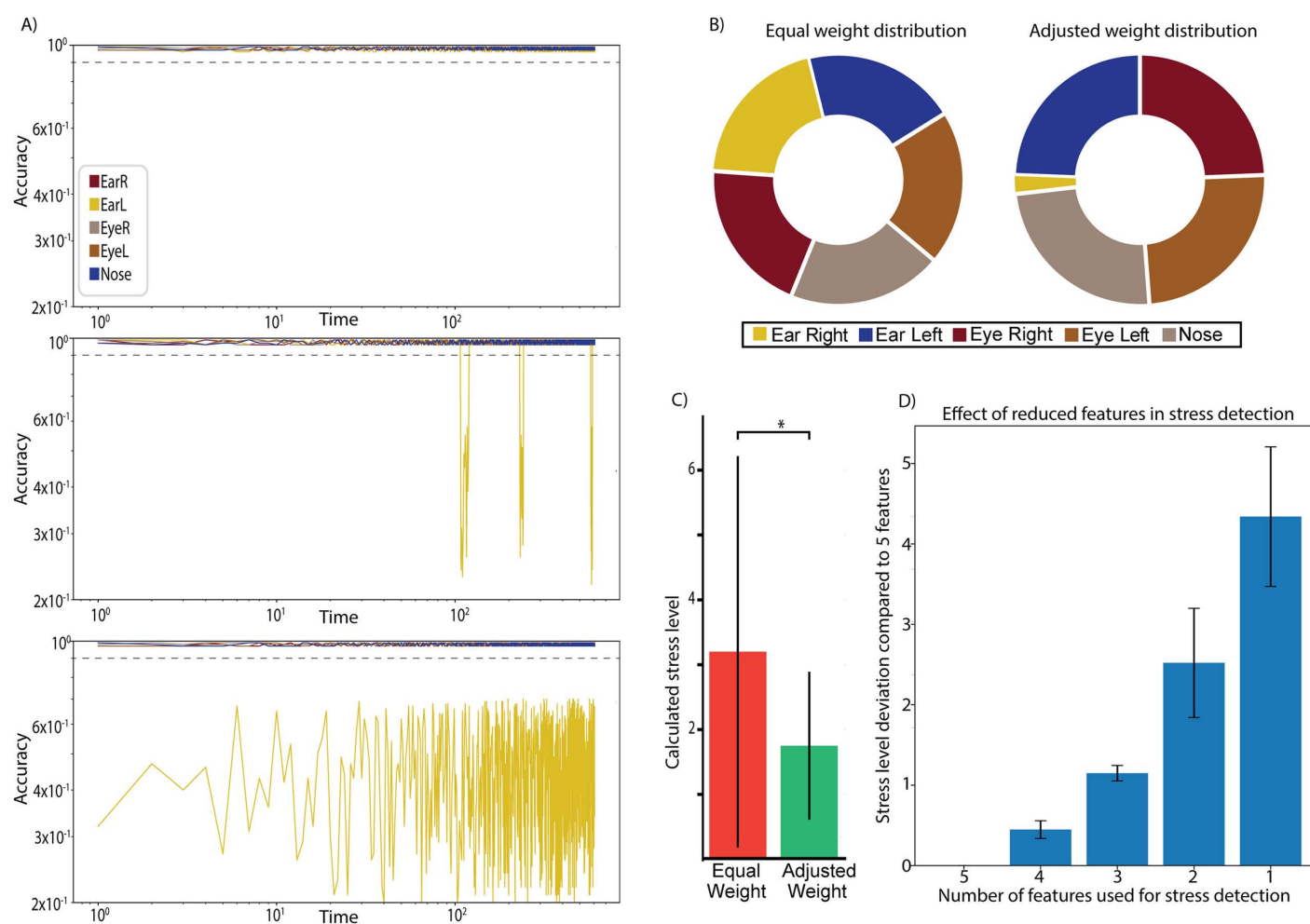


Fig 3. Evaluation of the effect of facial feature weight adjustment. A) Three Logarithmic graphs of the accuracy for each facial feature. Dotted line, accuracy threshold. The accuracy of the Ear Left, Eye Right, Eye Left and Nose are close to each other and overlap in the graph. The top graph represents a mouse with all facial features visible. The middle shows the effect of temporal obstruction, moving objects between the mouse and the camera. The bottom shows the effect of permanent obstruction of a facial feature. B) Comparing the standard weight distribution and the adjusted weight distribution. C) Calculated stress level of the same mouse using equal weight distribution and adjusted weight distribution. D) Effect of removing facial features on stress level evaluation. The Y-axis shows the deviation in stress level compared to the baseline using all five features.

<https://doi.org/10.1371/journal.pone.0322530.g003>

unable to move naturally due to cranial implants. And therefore were treated the same way as obscured facial features. Initially, each facial feature was assigned a weight of 1. As an example, in one set up an ear was pushed down by part of the setup, reducing its weight ([Fig 3B](#)) leading to more stable evaluation with fewer fluctuations in the stress evaluation ([Fig 3C](#)). To further investigate the effect of obfuscation/ removal of facial features, we analyzed recordings from head-fixed mice ($N=3$), setting the weight of individual facial features to 0 to simulate their removal from the evaluation. Using the stress evaluation with all five facial features as baseline, we determined how stress level estimations deviated with a decreasing number of remaining features ([Fig 3D](#)). The Y-axis in ([Fig 3D](#)) represents the absolute deviation from the baseline stress level (same units as in [Fig 3C](#)).

Stool analysis for stress evaluation

To analyze the stool samples of mice during the experiment we kept the mice in single cages and on the day of measurement we replaced the bedding to remove any form of old stool in the cage. After the experiment, the mice were placed back in the cage and 8 hours later, the feces were collected and dried. The dried samples were sent to a laboratory for processing (Palme laboratory at the University of Veterinary Medicine, Vienna). The dried stool samples were diluted, and the corticosterone metabolites were extracted with 1 ml of 80% methanol for 50 mg of feces, or with 0.8 ml for smaller fecal amounts.

Results

Our goal was to develop a software-based system (“MouseCare”) that could quickly, accurately and reliably assess stress in rodents under experimental conditions, providing a superior substitute for current methods. For brevity, we provide the detailed methods of the development and training of the MouseCare algorithm in Methods and [Figs 1–3](#). This involved providing videos of head-fixed mice with a mixture of mild stress-inducing situations (loud noises, vibrations, etc.). Next, we assessed the performance of MouseCare by comparing the stress levels determined by MouseCare with both subjective and objective measures of stress.

Comparison to subjective assessments

The vast majority of animal experiments use observer-assessed stress evaluation, in which facial features are the key indicators [\[43\]](#). We therefore first tuned MouseCare based on stress-scores provided by trained scientists ($n=7$). We first sought to evaluate how closely MouseCare tracked the assessment of typical stress behaviors compared to experimenters. To achieve this, we conducted a survey with two groups of participants: experienced researchers and inexperienced. The primary focus was on the experienced researchers’ assessments, while the results from the inexperienced group were included to provide a broader range of perspectives and are detailed in the supplementary materials ([S1 Fig](#)). Participants were presented with a series of 7-second video clips of head-fixed mice, without further information regarding the stage of habituation to head-fixation. Each participant independently assessed the stress levels of the mice based on observable facial features. Each participant rated the stress on a standardized scale from 1 to 10, with 1 indicating no stress and 10 indicating severe stress, in line with the scoring metric of MouseCare ([Fig 4A](#)).

Comparison to objective assessments

To provide an observer-independent measure of stress, we analyzed corticosterone metabolites in stool samples (analyzed by Palme laboratory at the University of Veterinary Medicine, Vienna). The results were expressed as ng/50 mg of feces. The corticosterone metabolite levels varied, reflecting differences in habituation stages and experimental conditions ([Fig 4B](#)). It is important to mention that there are a lot of influences (daily rhythm, sex, age, diet, etc...) on fecal corticosterone metabolites (FCM) values [\[31\]](#). MouseCare collected the positional data of each facial feature from the same videos as the participants, and then analyzed each facial feature individually and assigned a weighted stress score from 1 to 10

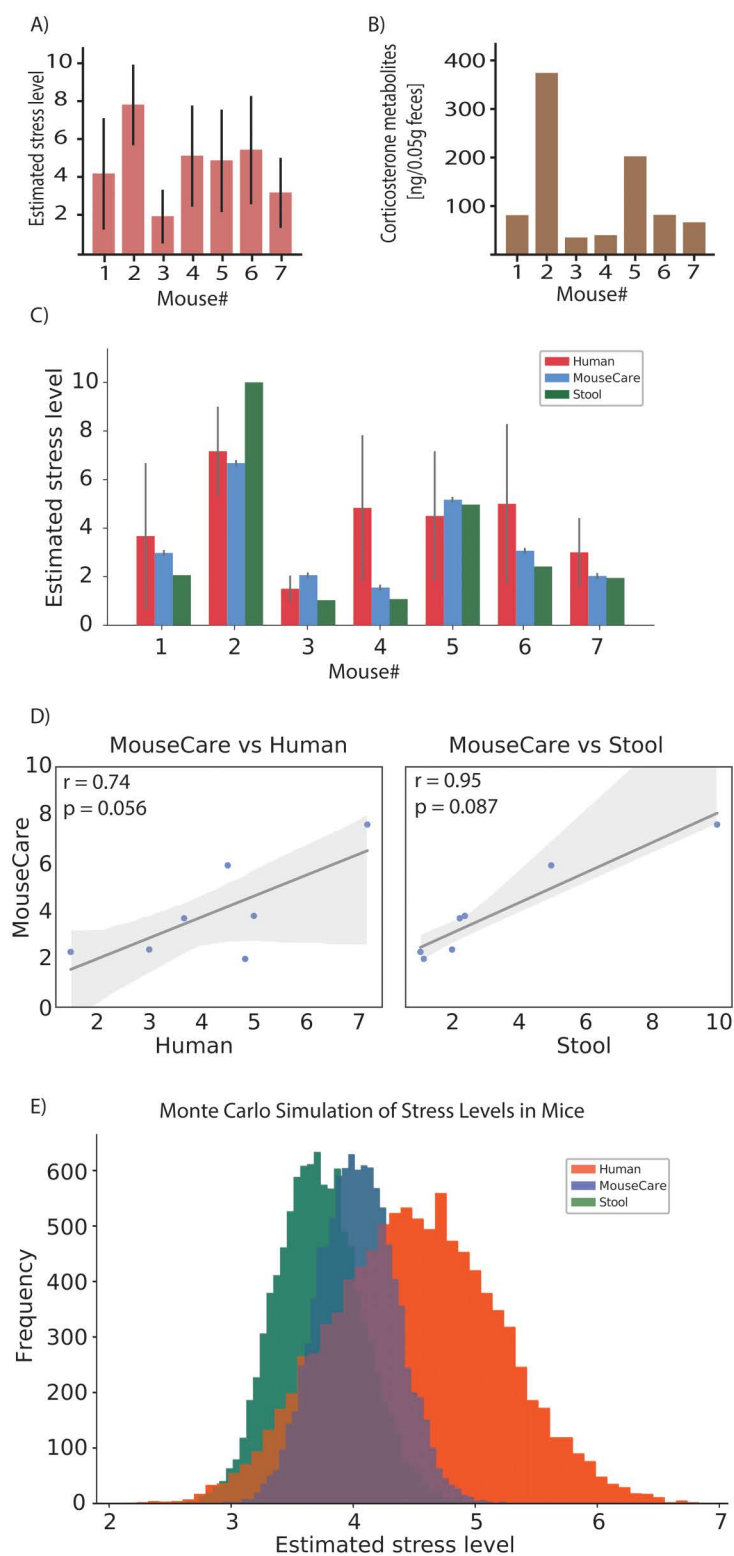


Fig 4. Comparing the stress evaluation results using different methods. A) Estimated stress level based on the facial feature evaluation of the participants. B) Measurement of the corticosterone metabolites for each of the mice shown in the 7 videos. C) Comparing the previous 2 methods with the results from MouseCare. Inter-rater reliability among the senior researchers was assessed using Cohen's kappa coefficient to ensure consistency,

and the Corticosterone metabolites were normalized. D) Calculation of the Pearson correlation coefficients between the three methods for each mouse. The stool samples were normalized. The X and Y axis represent the stress level on a scale from 1–10. In gray is the confidence interval. The Pearson correlation coefficient is labeled “*r*” and the *p* value is labeled as “*p*”. E) Using the Monte Carlo Simulation of the 7 mice and the three different methods of stress evaluation using 10000 simulations.

<https://doi.org/10.1371/journal.pone.0322530.g004>

to each feature. Those individual stress scores were then averaged to generate a total stress level for each video. A final total stress average was calculated for the 7-second duration, this was done for each video individually. By comparing the three methods—manual observation by experienced researchers, corticosterone metabolite levels, and MouseCare—we found that all approaches yielded comparable stress evaluations (Fig 4C). Notably, the experienced researchers’ estimates tended to indicate the highest stress levels, while corticosterone metabolites suggested the lowest, with MouseCare in between. We quantified the consistency between the different approaches by measuring the Pearson coefficient (comparing MouseCare to human reports and MouseCare to FCM) (Fig 4D **additional comparisons are shown in S3 Fig**). Here, we found that MouseCare aligned closer to the stool sample with 0.95 compared to the survey 0.74.

We performed an additional analysis, we performed a Monte Carlo simulation (Fig 4E, **the full list of the parameters used for this simulation is in S4 Fig**) by randomizing the mice IDs for the MouseCare values (see Methods) to estimate the probability of obtaining profiles with this similarity. In general, the MouseCare results more closely followed the FCM estimates, suggesting that MouseCare and FCM are more consistent approaches than subjective human reports. A significant discrepancy was observed only in mouse 4, where the experienced researchers’ stress estimates were substantially higher than those indicated by corticosterone metabolites and MouseCare. The results from the inexperienced group were less self-consistent (S2 Fig), underscoring the need for efficient methods of stress evaluation.

Effect of habituation and training on stress handling

Our final objective was to evaluate the stress levels of head-fixed mice during different stages of their habituation, along with their responses to various external stimuli commonly encountered in laboratory settings, including loud noise, physical contact (e.g., touching the tail of the head fixed animal). By exposing the mice to these three distinct stimuli, we were able to discern their stress effects relative to habituation using MouseCare (Fig 5A).

The reactions of the mice to external stress events were assessed over time (Fig 5B) showing that the effect of the stress stimulus had decreased relatively to the amount of habituation the mice had. This provides valuable insights into the effect of habituation. The visual representation of the network tracking with the original mouse images overlaid displayed the changes in facial expression (Fig 5C) based on the amount of habituation.

Discussion

In this paper, we presented a novel system to efficiently detect stress in head fixed mice. This approach prioritizes speed and accessibility over precision and includes easy-to-use custom-built software (MouseCare). Real-time facial analysis alerts researchers to signs of stress in mice with immediate feedback. Additionally, extraneous sources of stress (e.g., malfunctioning hardware causing irritating noises), can be identified, avoiding negative impact on behavior and training outcomes. We validated the performance of the software by comparing it with the experience of researchers as well as stool samples and showed that it performed equal-to-better compared to the other methods. The MouseCare software provides an overall stress score so that the researcher can easily react to, during the experiment. This approach simplifies stress detection in rodents, enhancing objectivity and reproducibility.

Stress detection in rodents has been relatively understudied. Automated systems using machine learning approaches that exist tend to focus on specific parameters relevant to particular behavioral paradigms [46–48]. MouseCare takes advantage of a preprocessing stage using a network made with DeepLabCut, that allows for body-part tracking for the

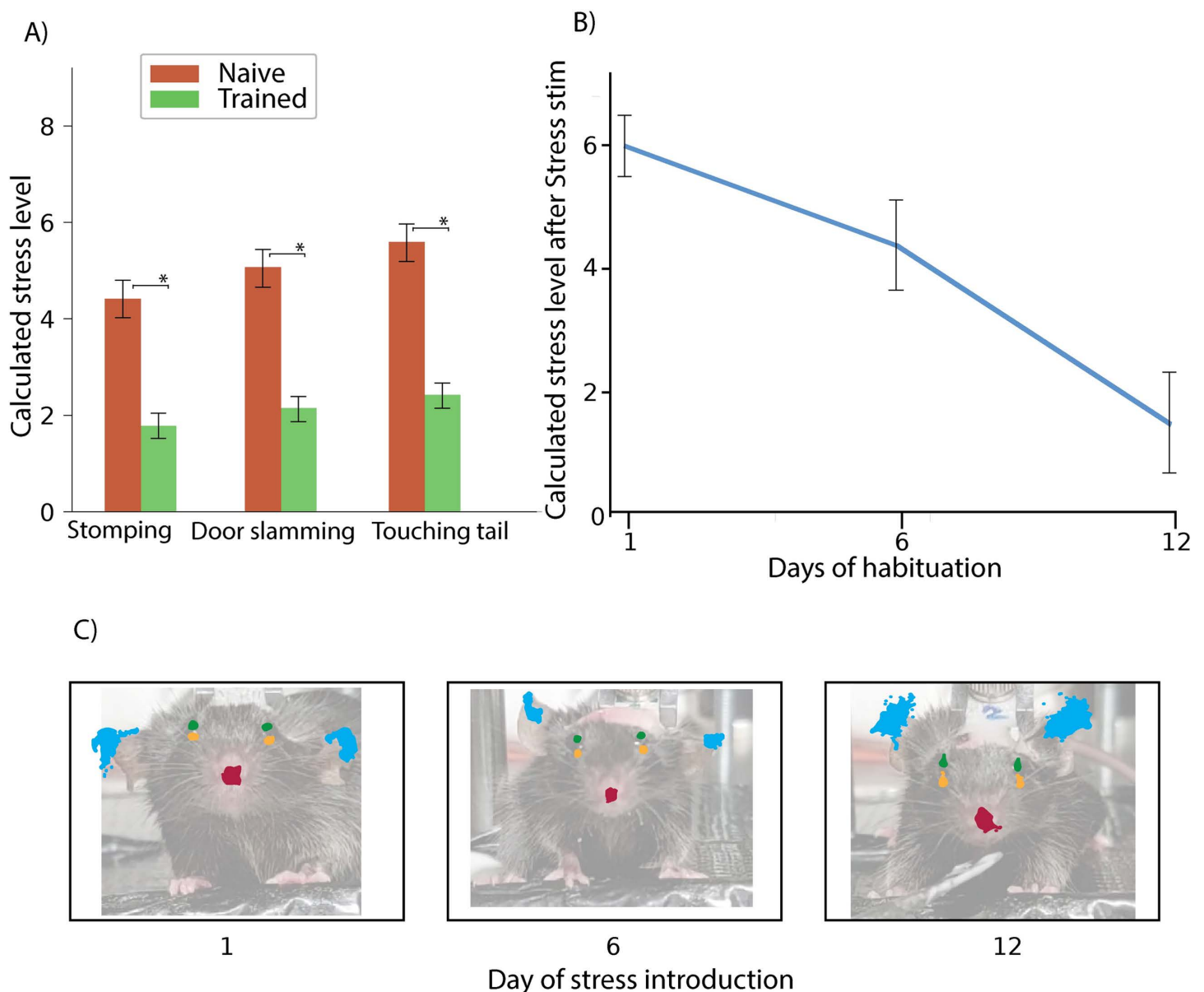


Fig 5. Impact of habituation on the response to external stress stimuli. A) Evaluation of the facial features between naive and trained mice in response to external stress stimuli using MouseCare. B) One time external stress stimulus after different days of habituation. C) Bar chart of the facial area of single mice at the day of stress introduction.

<https://doi.org/10.1371/journal.pone.0322530.g005>

algorithm to assess the stress. This makes the system more flexible and allows for relatively efficient adaptation to a given behavioral paradigm. For instance, this system could be easily adjusted to various setups and head fixation methods, utilizing its open-source nature to facilitate widespread adoption. The weight assigned to each facial feature allows for quick adjustments to accommodate different head posts or other types of obstructions. While this flexibility facilitates rapid adaptation, the removal of facial features impacts the accuracy of stress evaluation. Using five facial features as a baseline shown to provide the best stress evaluation results (Fig 4) we examined the effect of removing individual features to simulate obstructions (Fig 3D). Our findings indicate that the system remains reliable when a single facial feature is

removed but becomes progressively less reliable as additional features are excluded. Therefore, it is recommended to test and optimize the system for each experimental setup to ensure consistent performance.

MouseCare was compared to both objective stress measures, e.g., corticosterone metabolites, and on subjective measures, i.e., human evaluation (Fig 4). To mitigate subjective bias, the human subjects used for the subjective assessment were experienced scientists (i.e., post-docs). As can be observed in Fig 4, the subjective and objective assessments were relatively close, and the training achieved AI-based assessments between these values. Note, subjective assessments could vary considerably based on the experience level of the researcher (S2 Fig), however we found that the experienced group was surprisingly consistent (with one outlying data point in Fig 4). The close correspondence of the assessments made by an objective measure (corticosterone metabolites) to the experienced researchers provides confidence that MouseCare achieves the best possible AI-based score. Nevertheless, it should be noted that the most accurate objective measurement of stress is currently an analysis of levels of cortisol in the blood [25,26]. Corticosterone metabolites are used widely because of their convenience and the fact that they are non-invasive, however they have also been shown to have limitations, particularly regarding gender differences in mice [49]. In general, these objective approaches to stress detection can be difficult and time-consuming, and an automated approach will save researchers a lot of time and/or encourage more laboratories to use continual stress assessment for rodent-based studies. Even more importantly, the live automated approach with MouseCare, is essential for alleviating or mitigating the effects of stress during the experiment itself. An automated system such as MouseCare is also ideal for training new researchers. Here, we suggest that new researchers be trained on pre-evaluated video footage that can be evaluated readily with MouseCare. The use of video footage reduces the number of animals needed for training, which is beneficial from an ethical perspective. In many cases, it will be possible for trainees to use video of previous experiments on the same or similar systems, making the video assessment more relevant. After becoming proficient on videos, the new researcher can transition to live experiments under the continual assistance of MouseCare. Overall, we anticipate that an AI-onboarding process will significantly reduce the total stress for mice.

Our system's ability to continuously monitor mouse facial features over extended periods also offers valuable insights into habituation on an individual basis. Environmental factors, such as laboratory noises, can vary between mice, setups, and experiments. This variability presents a significant hurdle in estimating the ideal duration for habituation in behavioral paradigms involving head fixation. MouseCare can be introduced at different points in the experimental timeline, depending on the researcher's objectives. For individual animals, we recommend starting MouseCare during the first head fixation sessions, as this is when stress levels are typically highest, making early detection crucial for effective habituation. For researchers integrating MouseCare into their experimental protocols, however, we suggest implementing MouseCare after the initial head fixation stages, once mice are acclimated and can tolerate longer periods of fixation. This timeline allows researchers adequate time to calibrate and optimize the system without disrupting studies involving naive mice, which are limited to shorter fixation durations. Introducing MouseCare at this stage ensures the system is fully operational and provides accurate, continuous stress monitoring during critical phases of habituation and training. An AI-driven approach like MouseCare enables a more precise evaluation of habituation, maximizing benefits for the animal (in terms of stress reduction) and for the researcher (by minimizing habituation time). Furthermore, its open-source code allows for modifications and potential expansion to other species, such as rats, or even the detection of emotional states beyond stress, such as joy.

Cost-effectiveness is another crucial consideration. Unlike currently-available systems relying on expensive cameras and computers, MouseCare operates efficiently with more affordable equipment. We recommend using a high-end computer for operating MouseCare, however this can be accessed over a local network minimizing the need for high-end computers in each experimental room, thereby reducing overall costs. In the current design of MouseCare, we used affordable camera technology, however, we would expect potential improvements to pupil tracking and evaluation could be achieved with better cameras which might provide even greater accuracy in stress detection. Machine learning has also opened other avenues for stress detection, since it can also be used to evaluate other bodily functions like breathing [50]. This can be used in experiments where it is not feasible to have a camera recording the face of the mice.

In conclusion, our study demonstrates that MouseCare, an automated system for real-time facial feature analysis, effectively detects stress in head fixed mice. MouseCare performs as well as, or better than, traditional methods such as corticosterone metabolite levels and human observation. MouseCare has the steepest implementation hurdle compared to the other methods, but once it is adjusted to the setup it has the least amount of variability compared to the other two methods, making it more reliable. This system enhances the objectivity, reproducibility, and ethical standards of stress assessment in neuroscience research. By providing immediate feedback, MouseCare enables timely interventions, optimizing both animal welfare and experimental outcomes. Additionally, its adaptability to various experimental setups and cost-effectiveness make it a valuable tool for research laboratories. MouseCare represents a significant advancement in the field, offering a practical and efficient solution for continuous stress monitoring in rodent studies.

Supporting information

S1 Fig. Showing the tracking result of different mice in different setup with different lighting conditions. To test the versatility of MouseCare, we used different mice in different setup to test if the system is capable to track the facial features successfully under different conditions.

(TIF)

S2 Fig. Comparing the results of the stress evaluation between the experienced and inexperienced group. The participants received 7 videos of head fixed mice and were tasked to evaluate the stress of the mice on a scale of 1 (no stress at all) to 10 (very stressed).

(TIF)

S3 Fig. Calculation of the Pearson correlation coefficients between the three methods for each mouse. The stool samples were normalized. The X and Y axis represent the stress level on a scale from 1–10. In gray is the confidence interval. The Pearson correlation coefficient is labeled “r” and the p value is labeled as “p”.

(TIF)

S4 Fig. Performance comparison of Stool samples, humans and MouseCare using a Monte Carlo simulation with additional Numeric values. A) Numeric results of the Monte Carlo simulation shown in [Fig 4D](#). B) Visual representation of the Performance comparison three groups.

(TIF)

Acknowledgments

We thank the Charité Workshop for technical assistance, especially Alexander Schill, Jan-Erik Ode and Daniel Deblitz. We thank Prof. Dr. R. Palme from the Vetmeduni Vienna Unit of Physiology, Pathophysiology and Experimental Endocrinology for the measurement of the corticosterone metabolites from fecal samples. Furthermore, we thank Dr. Tim Zolnik for his useful suggestions and comments on earlier versions of this manuscript.

Author contributions

Conceptualization: Anna Nasr, Robert N. S. Sachdev, Matthew E. Larkum.

Data curation: Anna Nasr.

Funding acquisition: Anna Nasr, Matthew E. Larkum.

Investigation: Anna Nasr, Gianna Rettinger.

Methodology: Anna Nasr.

Project administration: Anna Nasr.

Software: Anna Nasr.

Supervision: Matthew E. Larkum.

Visualization: Robert N. S. Sachdev, Matthew E. Larkum.

Writing – original draft: Anna Nasr, Matthew E. Larkum.

Writing – review & editing: Anna Nasr, Huibert D. Mansvelder, Robert N. S. Sachdev, Matthew E. Larkum.

References

- Guo ZV, Hires SA, Li N, O'Connor DH, Komiyama T, Ophir E, et al. Procedures for behavioral experiments in head-fixed mice. *PLoS One*. 2014;9(2):e88678. <https://doi.org/10.1371/journal.pone.0088678> PMID: 24520413
- Hegedüs P, Velencei A, Belval C-H de, Heckenast J, Hangya B. Training protocol for probabilistic Pavlovian conditioning in mice using an open-source head-fixed setup. *STAR Protoc*. 2021;2(3):100795. <https://doi.org/10.1016/j.xpro.2021.100795> PMID: 34522902
- White JJ, Lin T, Brown AM, Arancillo M, Lackey EP, Stay TL, et al. An optimized surgical approach for obtaining stable extracellular single-unit recordings from the cerebellum of head-fixed behaving mice. *J Neurosci Methods*. 2016;262:21–31. <https://doi.org/10.1016/j.jneumeth.2016.01.010> PMID: 26777474
- Weaver IA, Aryana Yousefzadeh S, Tadross MR. An open-source head-fixation and implant-protection system for mice. *HardwareX*. 2022;13:e00391. <https://doi.org/10.1016/j.hwx.2022.e00391> PMID: 36632413
- Ozgur A, Park SB, Flores AY, Oijala M, Lur G. A Comprehensive, affordable, open-source hardware-software solution for flexible implementation of complex behaviors in head-fixed mice. *eNeuro*. 2023;10(6). <https://doi.org/10.1523/ENEURO.0018-23.2023> PMID: 37286343
- Ortiz-Juza MM, Tormes-Vaquero J, Hegel SM, Curtis VR, Alghorazi RA, Miller NW, et al. Protocol for an open-source system to integrate calcium imaging, pupillometry, and locomotion-estimated tracking in head-fixed mice. *STAR Protoc*. 2024;5(4):103331. <https://doi.org/10.1016/j.xpro.2024.103331> PMID: 39352810
- Furutachi S, Franklin AD, Aldea AM, Mrcic-Flogel TD, Hofer SB. Cooperative thalamocortical circuit mechanism for sensory prediction errors. *Nature*. 2024;633(8029):398–406. <https://doi.org/10.1038/s41586-024-07851-w> PMID: 39198646
- Rupprecht P, Duss SN, Becker D, Lewis CM, Bohacek J, Helmchen F. Centripetal integration of past events in hippocampal astrocytes regulated by locus coeruleus. *Nat Neurosci*. 2024;27(5):927–39. <https://doi.org/10.1038/s41593-024-01612-8> PMID: 38570661
- Jasper HH, Ricci G, Doane B. Microelectrode analysis of cortical cell discharge during avoidance conditioning in the monkey: the Moscow colloquium on electroencephalography of higher nervous activity. *Electroencephalogr Clin Neurophysiol*. 1960;suppl 13:137–55.
- Evarts EV. Relation of pyramidal tract activity to force exerted during voluntary movement. *J Neurophysiol*. 1968;31(1):14–27. <https://doi.org/10.1152/jn.1968.31.1.14> PMID: 4966614
- Lemon RN, Porter R. Afferent input to movement-related precentral neurones in conscious monkeys. *Proc R Soc Lond B Biol Sci*. 1976;194(1116):313–39. <https://doi.org/10.1098/rspb.1976.0082> PMID: 11491
- Svoboda K, Denk W, Kleinfeld D, Tank DW. In vivo dendritic calcium dynamics in neocortical pyramidal neurons. *Nature*. 1997;385(6612):161–5. <https://doi.org/10.1038/385161a0> PMID: 8990119
- Schwarz C, Hentschke H, Butovas S, Haiss F, Stüttgen MC, Gerdjikov TV, et al. The head-fixed behaving rat—Procedures and pitfalls. *Somatosensory Motor Res*. 2010;27(4):131–48. <https://doi.org/10.3109/08990220.2010.513111>
- Jun JJ, Steinmetz NA, Siegle JH, Denman DJ, Bauza M, Barbarits B, et al. Fully integrated silicon probes for high-density recording of neural activity. *Nature*. 2017;551(7679):232–6. <https://doi.org/10.1038/nature24636> PMID: 29120427
- Petersen Carl CH, Hahn Thomas TG, Mehta M, Grinvald A, Sakmann B. Interaction of sensory responses with spontaneous depolarization in layer 2/3 barrel cortex. *Proc Natl Acad Sci*. 2003;100(23):13638–43.
- Brecht M. Barrel cortex and whisker-mediated behaviors. *Curr Opin Neurobiol*. 2007;17(4):408–16. <https://doi.org/10.1016/j.conb.2007.07.008> PMID: 17702566
- Helmchen F, Denk W. Deep tissue two-photon microscopy. *Nat Methods*. 2005;2(12):932–40. <https://doi.org/10.1038/nmeth818> PMID: 16299478
- Margrie TW, Brecht M, Sakmann B. In vivo, low-resistance, whole-cell recordings from neurons in the anaesthetized and awake mammalian brain. *Pflugers Arch*. 2002;444(4):491–8. <https://doi.org/10.1007/s00424-002-0831-z> PMID: 12136268
- Balcombe JP, Barnard ND, Sandusky C. Laboratory routines cause animal stress. *Contemp Top Lab Anim Sci*. 2004;43(6):42–51. PMID: 15669134
- Hurst JL, West RS. Taming anxiety in laboratory mice. *Nat Methods*. 2010;7(10):825–6. <https://doi.org/10.1038/nmeth.1500> PMID: 20835246
- Gouveia K, Hurst JL. Optimising reliability of mouse performance in behavioural testing: the major role of non-aversive handling. *Sci Rep*. 2017;7:44999. <https://doi.org/10.1038/srep44999> PMID: 28322308
- McEwen BS. Mood disorders and allostatic load. *Biol Psychiatry*. 2003;54(3):200–7. [https://doi.org/10.1016/s0006-3223\(03\)00177-x](https://doi.org/10.1016/s0006-3223(03)00177-x) PMID: 12893096
- Barkus C, Bergmann C, Branco T, Carandini M, Chadderton PT, Galiñanes GL, et al. Refinements to rodent head fixation and fluid/food control for neuroscience. *J Neurosci Methods*. 2022;381:109705. <https://doi.org/10.1016/j.jneumeth.2022.109705> PMID: 36096238

24. Nasr A, Dominiak SE, Sehara K, Nashaat MA, Sachdev RNS, Larkum ME. Efficient training approaches for optimizing behavioral performance and reducing head fixation time. *PLoS One*. 2022;17(11):e0276531. <https://doi.org/10.1371/journal.pone.0276531> PMID: 36355714
25. Dölsner A, Hack R, Krüger C, Pils M, Scherer K, Schmelting B. Recommendation for blood sampling in laboratory animals, especially small laboratory animals. July 2017 Available from: https://www.gv-solas.de/wp-content/uploads/2017/03/tie_blutentnahme17_e.pdf
26. Hopster H, van der Werf JT, Erkens JH, Blokhuis HJ. Effects of repeated jugular puncture on plasma cortisol concentrations in loose-housed dairy cows. *J Anim Sci*. 1999;77(3):708–14. <https://doi.org/10.2527/1999.773708x> PMID: 10229368
27. Cook CJ, Mellor DJ, Harris PJ, Ingram JR, Matthews LR. Hands-on and hands-off measurement of stress. In: Moberg GP, Mench JA, editors. *The biology of animal stress*. CABI Publishing; 2000. pp. 123–46. <https://doi.org/10.1079/9780851993591.0123>
28. Schwarzenberger F, Möstl E, Palme R, Bamberg E. Faecal steroid analysis for non-invasive monitoring of reproductive status in farm, wild and zoo animals. *Anim Reprod Sci*. 1996;42(1–4):515–26. [https://doi.org/10.1016/0378-4320\(96\)01561-8](https://doi.org/10.1016/0378-4320(96)01561-8)
29. Kumar A, Mehrotra S, Dangi S, Singh G, chand S, Singh L, et al. Faecal steroid metabolites assay as a non-invasive monitoring of reproductive status in animals. *Vet World*. 2013;6(1):59. <https://doi.org/10.5455/vetworld.2013.59-63>
30. Touma C, Palme R, Sachser N. Analyzing corticosterone metabolites in fecal samples of mice: a noninvasive technique to monitor stress hormones. *Horm Behav*. 2004;45(1):10–22. <https://doi.org/10.1016/j.yhbeh.2003.07.002> PMID: 14733887
31. Möstl E, Palme R. Hormones as indicators of stress. *Domest Anim Endocrinol*. 2002;23(1–2):67–74. [https://doi.org/10.1016/S0739-7240\(02\)00146-7](https://doi.org/10.1016/S0739-7240(02)00146-7) PMID: 12142227
32. Hay M, Mormède P. Urinary excretion of catecholamines, cortisol and their metabolites in Meishan and large white sows: validation as a non-invasive and integrative assessment of adrenocortical and sympathoadrenal axis activity. *Vet Res*. 1998;29(2):119–28. PMID: 9601144
33. Cooper TR, Trunkfield HR, Zanella AJ, Booth WD. An enzyme-linked immunosorbent assay for cortisol in the saliva of man and domestic farm animals. *J Endocrinol*. 1989;123(2):R13–6. <https://doi.org/10.1677/joe.0.123r013> PMID: 2607243
34. Verkerk GA, Phipps AM, Carragher JF, Matthews LR, Stelwagen K. Characterization of milk cortisol concentrations as a measure of short-term stress responses in lactating dairy cows. *Anim welf*. 1998;7(1):77–86. <https://doi.org/10.1017/S0962728600020273>
35. Whittaker AL, Liu Y, Barker TH. Methods used and application of the mouse Grimace Scale in biomedical research 10 years on: a scoping review. *Animals (Basel)*. 2021;11(3):673. <https://doi.org/10.3390/ani11030673> PMID: 35102463
36. Whittaker AL, Howarth GS. Use of spontaneous behaviour measures to assess pain in laboratory rats and mice: How are we progressing? *Appl Anim Behav Sci*. 2014;151:1–12. <https://doi.org/10.1016/j.applanim.2013.11.001>
37. Mogil JS, Pang DSJ, Silva Dutra GG, Chambers CT. The development and use of facial grimace scales for pain measurement in animals. *Neurosci Biobehav Rev*. 2020;116:480–93. <https://doi.org/10.1016/j.neubiorev.2020.07.013> PMID: 32682741
38. Hohlbaum K, Corte GM, Humpenöder M, Merle R, Thöne-Reineke C. Reliability of the Mouse Grimace Scale in C57BL/6JRj Mice. *Animals (Basel)*. 2020;10(9):1648. <https://doi.org/10.3390/ani10091648> PMID: 32937881
39. Nath T, Mathis A, Chen AC, Patel A, Bethge M, Mathis MW. Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nat Protoc*. 2019;14(7):2152–76. <https://doi.org/10.1038/s41596-019-0176-0> PMID: 31227823
40. Dominiak SE, Nashaat MA, Sehara K, Oraby H, Larkum ME, Sachdev RNS. Whisking asymmetry signals motor preparation and the behavioral state of mice. *J Neurosci*. 2019;39(49):9818–30. <https://doi.org/10.1523/JNEUROSCI.1809-19.2019> PMID: 31666357
41. Nashaat MA, Oraby H, Sachdev RNS, Winter Y, Larkum ME. Air-Track: a real-world floating environment for active sensing in head-fixed mice. *J Neurophysiol*. 2016;116(4):1542–53. <https://doi.org/10.1152/jn.00088.2016> PMID: 27486102
42. Ebner C, Ledderose J, Zolnik TA, Dominiak SE, Turko P, Papoutsis A, et al. Optically induced calcium-dependent gene activation and labeling of active neurons using CaMPARI and Cal-Light. *Front Synaptic Neurosci*. 2019;11:16. <https://doi.org/10.3389/fnsyn.2019.00016> PMID: 31178713
43. Langford DJ, Bailey AL, Chanda ML, Clarke SE, Drummond TE, Echols S, et al. Coding of facial expressions of pain in the laboratory mouse. *Nat Methods*. 2010;7(6):447–9. <https://doi.org/10.1038/nmeth.1455> PMID: 20453868
44. Kurnikova A, Moore JD, Liao S-M, Deschênes M, Kleinfeld D. Coordination of orofacial motor actions into exploratory behavior by rat. *Curr Biol*. 2017;27(5):688–96. <https://doi.org/10.1016/j.cub.2017.01.013> PMID: 28216320
45. Kane GA, Lopes G, Saunders JL, Mathis A, Mathis MW. Real-time, low-latency closed-loop feedback using markerless posture tracking. *Elife*. 2020;9:e61909. <https://doi.org/10.7554/eLife.61909> PMID: 33289631
46. Dolensek N, Gehrlach DA, Klein AS, Gogolla N. Facial expressions of emotion states and their neuronal correlates in mice. *Science*. 2020;368(6486):89–94. <https://doi.org/10.1126/science.aaz9468> PMID: 32241948
47. Kuo JY, Denman AJ, Beacher NJ, Glanzberg JT, Zhang Y, Li Y, et al. Using deep learning to study emotional behavior in rodent models. *Front Behav Neurosci*. 2022;16:1044492. <https://doi.org/10.3389/fnbeh.2022.1044492> PMID: 36483523
48. Tanaka Y, Nakata T, Hibino H, Nishiyama M, Ino D. Classification of multiple emotional states from facial expressions in head-fixed mice using a deep learning-based image analysis. 2023. <https://doi.org/10.1501/journal.pone.0288930>
49. Touma C, Palme R, Sachser N. Analyzing corticosterone metabolites in fecal samples of mice: a noninvasive technique to monitor stress hormones. *Horm Behav*. 2004;45(1):10–22. <https://doi.org/10.1016/j.yhbeh.2003.07.002> PMID: 14733887
50. Janke E, Zhang M, Ryu SE, Bhattarai JP, Schreck MR, Moberly AH, et al. Machine learning-based clustering and classification of mouse behaviors via respiratory patterns. *iScience*. 2022;25(12):105625. <https://doi.org/10.1016/j.isci.2022.105625> PMID: 36479148