Evaluating Clicker Training as a Novel Refinement Method for Reducing Stress in Laboratory Mice During and After Subcutaneous Injection

Veterinary Medicine Utrecht University
Department Population Health Sciences; Division: Animals in Science and Society (DWM)

Author: S.J. Budel (5854997)
Tutor: Dr. E.M.A. Langen
Deadline: 11-02-2024

Abstract

Because laboratory mice are essential to scientific research and are extensively utilized, refining research protocols and methods is imperative to safeguard the welfare of these animals, maintain ethical standards, and improve the reliability and applicability of research outcomes. Accordingly, the main aim of this study was to investigate if clicker training laboratory mice to stand still during a subcutaneous (SC) injection with minimal restraint could effectively reduce stress both during and after the injection, potentially serving as a novel refinement method. In order to investigate this, a (trained) group (n=8) underwent 19 sessions of three minutes each, during which they were trained to stand still on a training pad while having a skin fold lifted, allowing for a potential SC injection. Additionally, to distinguish between the effects of training versus exposure, as solely exposure would eliminate the challenging aspects associated with clicker training, an exposed control group (n=8) was created. This group was exposed to the training environment, reward and researcher for the same amount of sessions as the trained group, but without training. Finally, an unexposed control group was created (n=8) that was never exposed to the training environment and had only limited exposure to the researcher. Ultimately, all 24 mice were injected with 0,1 ml NaCl 0,9% to assess the effect of training on the behavioral response of the animals during and after the injection, specifically regarding stress-related behaviors. During injection, the mice's ability to stand still was assessed, as a measure of trainability and potentially their stress level. Following the injection, their reward acceptance was tested for one minute to determine if they remained receptive to receiving a reward from the researcher or would refuse the reward possibly attributed to anxiety-induced anorexia. Following the reward acceptance test, a voluntary approach test (VAT) of two minutes was conducted, during which various stress-related behaviors were assessed. The analysis showed a difference between treatment groups regarding standing still during injection, with trained mice standing still longer than unexposed mice, indicating that trained mice potentially experience less stress during injection than unexposed mice. A trend was found between the exposed and unexposed mice, with the exposed mice standing still longer, indicating that exposure alone may affect the stress response of mice during a SC injection. Another significant correlation was found between treatment groups and reward acceptance, with the exposed mice eating longer than unexposed mice, again highlighting the potential impact of exposure on reducing stress during injection. No other significant differences between the treatment groups were found. Several limitations in the study, like small sample size, lack of naivety in the unexposed mice, and their advanced age may have influenced the results of the remaining behaviors. Clicker training appears to have been effective, as evident by trained mice standing still longer during injection. However, no effects of training on any stress-related behaviors were observed yet, suggesting exposure alone might yield stronger effects. Further research on clicker training mice is warranted to understand its comprehensive impact on stress levels and whether exposure is as effective or even more so.

Table of contents

Introduction

Laboratory animals have been indispensable in scientific research, serving as crucial models for investigating innumerable aspects of biology, including physiology and disease (van Zutphen, 2002). Laboratory mice in particular have emerged as a vital asset, playing a pivotal role in advancing our understanding of fundamental biological processes and the development of therapeutic interventions (Bryda, 2013; Phifer-Rixey & Nachman, 2015). Recently our understanding of animal welfare and ethical considerations has evolved (Broom, 2011; Broom, 2017; Parker & McElligott, 2023). Recognizing that mice are sentient beings capable of experiencing pain and discomfort, it is imperative to treat them accordingly. In the Netherlands, as in other EU countries, the law on animal testing emphasizes the importance of researchers avoiding causing unnecessary pain and distress, endorsing the principles of the 3 Rs –replacement, reduction and refinement– introduced in 1959 by William Russel (Russell & Burch, 1992; Wet op de dierproeven, 2023). Replacement advocates for avoiding the use of laboratory animals when not inherently necessary, striving to employ alternative methods whenever possible. Reduction emphasizes minimizing the number of animals used in scientific research while still achieving statistically significant results. Lastly, refinement involves adjusting and fine-tuning experimental methods and housing to reduce pain, suffering and stress by, for example, realizing sufficient cage enrichment or using non-aversive handling methods (e.g. handling with a tunnel) (Gouveia & Hurst, 2013; Russell & Burch, 1992; Rusche, 2003; Davies et al., 2022). This aspect of the 3Rs will be the focus of this study.

The laboratory setting imposes inherent restrictions for the mice, such as possible aversive experiences that arise from the necessity to control experimental variables. Experimental conditions within the laboratory, including exposure to novel and occasionally threatening stimuli like unfamiliar odors, sounds, or handling procedures, can contribute to aversive experiences for the mice, potentially leading to acute and, in some cases, chronic stress (Kramer, et al., 2004; Meijer, et al., 2005; Morgan & Tromborg, 2006). In a physiological context, stress functions as an adaptive response vital for survival. Stress hormones such as glucocorticoids and (nor)adrenaline are released, enhancing alertness and energy mobilization to cope with the situation/stressor and maintain homeostasis (Dhabar, 2018; Haykin & Rolls, 2023). However, prolonged or excessive stress can potentially result in maladaptive responses, adverse health effects and compromised welfare (Ketchesin, et al., 2017). The definition of welfare in the current study aligns with Dutch animal law, requiring animals to be housed appropriately and free from abuse, neglect, and unnecessary pain (Wet op de dierproeven, 2023). Additionally, animals must be able to actively adapt to their living conditions (Wet dieren, 2022; Arndt et al., 2022; Ohl & Hellebrekers, 2009). Stress not only impacts mice's quality of life but may also introduce confounding variables, potentially compromising research validity and reliability (Gouveia & Hurst, 2017; Neely et al., 2018). Therefore, it is crucial for both laboratory animals and future science to explore refinement methods for minimizing unnecessary stress.

In light of these challenges in the laboratory setting and the potential impact on mice welfare, this study seeks to address the need for refinement by investigating a specific method –clicker training– to reduce stress during and after receiving a subcutaneous (SC) injection. As indicated by prior research, clicker training emerges as a promising cognitive enrichment tool and refinement measure for mice, with Leidinger et al.'s (2017) study establishing a strong correlation between clicker training and diminished fear responses in mice during human-mice interactions. BALB/c mice were trained to follow a target stick in their home cage, culminating in the 10th and final session where the mice were required to follow the target stick onto the researcher's hand. Notably, the study observed a reduction in anxiety-related behaviors in the trained mice, such as defecation, vocalization, and urination after grasping the scruff of the neck or base of the tail with one hand. Given the evidence establishing clicker training as a refinement for mice, the current study incorporated this training method instead of conventional training. Furthermore, considering that mouse handling has been shown to induce an elevated heart frequency, indicative of stress (Meijer, et al., 2006; Kramer, et al., 1993; Kramer, et al., 2004), this study explores a training method requiring minimal restraint. It involves only the lifting of a skin fold to administer the SC injection.

More recently, Van Eldik (2021) explored the possibility of clicker training mice to stand still for a SC injection and assessed its potential stress reduction. In their study, three treatment groups were created: a training group, a researcher exposed control group and a researcher unexposed control group. The training group followed a training protocol, while the exposed control mice were exposed to the researcher and received rewards, but were not trained. The unexposed control group remained entirely unexposed to both the researcher and the reward. The formation of these three distinct groups was designed to provide insight into whether habituating mice could be sufficient enough to reduce stress and the degree to which this reduction compared to the trained group. This consideration accounts for the feasibility factor as well, given that exposing mice is more practically attainable than training them. Van Eldik (2021) trained the mice to stand still on a specific side of a small platform. The mice were trained/exposed for 24 sessions of five minutes distributed across several weeks. While not statistically significant, the findings implied that trained mice exhibited a lower frequency of freezing and sought more contact with the researcher post-injection compared to the unexposed control group. This observation suggests that clicker training in mice holds the potential to serve as a refinement measure. Moreover, mice from the exposed control group froze less than the unexposed control group as well, indicating that habituation alone might contribute to the observed outcome.

The study by van Eldik was undoubtedly necessary in the scientific domain as it serves as a valuable pilot study concerning clicker training and stress reduction in laboratory mice. Hence, for the current study, comparable groups and protocols were created and similar behaviors were assessed. However, the experiment did exhibit several limitations. Notably, the practicality of van Eldik's protocol came into question, mainly because of the use of a platform that required mice to stand on a specific side, needlessly complicating the protocol and consequently requiring more sessions. Moreover, certain mistakes were made during training, for example no precautions were taken to ensure that mice not undergoing training were isolated from the audible sound of the clicker within the training room. This circumstance could potentially have hindered the formation of an association between the bridging stimulus and the reward. Hence, the objective of the present study was to develop a simpler and thereby more practical training protocol, with the aim of achieving faster training while concurrently meeting the same refinement goals.

While addressing the limitations identified in van Eldik's study, the current study also delves into evaluating stress levels during and after a SC injection through various behavioral tests. The behaviors that have been assessed mirror those observed in van Eldik (2021). During the injection, the behaviors 'struggling' and 'standing still' were scored. Struggling is defined and scored as any moment the mouse was not standing still during injection. Besides standing still and struggling, the entire duration of the administration of the injection was also measured. The expected result is that trained mice will exhibit a reduced injection duration, will stand still longer and struggle less compared to the two control groups. However, it is possible that mice would freeze out of fear (Eilam, 2005), leading to a longer duration of stillness and shorter duration of struggle, with the unexposed control group expected to exhibit this behavior the most. Given that the injection is perceived as threatening for the mice, the observed struggle during injection can be interpreted as a stress response, triggering the "fight or flight" reaction (Walker *et al*., 2003; Stuart & Robinson, 2015). The discomfort or pain associated with the injection likely contributes to their struggle. The researcher faces increased difficulty in administering the injection without restraint when a mouse struggles, potentially resulting in a longer injection time compared to when the mouse remains still. Consequently, an extended injection time may suggest a higher level of stress for the mice (except in cases where the mice exhibit 'freezing' as previously mentioned).

Following the injection, mice are provided with a reward for one minute, and the duration during which the mice engaged in eating is recorded as accepting the reward. Given that mice exhibit neophobia (Kronenberger & Médoni, 1985) and perceive new food, environments, or individuals as potential threats, declining the reward may be indicative of stress (Misslin & Cigrang, 1986; Qiao, *et al*., 2020; Francois, *et al*., 2022; Lezak, *et al*., 2022). While habituation to the reward may minimize this concern regarding food, unexposed mice will still have to accept the reward in an unfamiliar environment from

an unfamiliar individual. Furthermore, research indicates that female mice, in particular, may exhibit stress-induced anorexia, providing another potential reason for them to decline the reward (Yamada *et al.*, 2020). The trained and exposed control mice, having familiarity with the environment, researcher and reward, are expected to accept the reward for a longer duration compared to unexposed control mice.

Another behavioral test conducted in this study was a voluntary approach test (VAT) after injection. During this VAT, several contact seeking behaviors were scored, such as: 'sniffing the researcher', 'touching the researcher', 'spending time on the researcher', 'nibbling the researcher' and 'sitting next to the researcher'. The reluctance to initiate contact with the researcher may be a sign of wariness, particularly when accompanied by other stress indicators (e.g. the behaviors mentioned before or an increased heart frequency and raised body temperature) (Kramer, *et al.,* 2004). Again, considering the neophobic tendencies of the mice, it is probable that they may be hesitant to initiate contact with the researcher, and a diminished duration of seeking contact could be indicative of stress. Due to the familiarity of the trained group with the researcher, environment, and injection procedure, the expectation is that these mice will exhibit more contact-seeking behaviors in contrast to the control groups. The exposed control group, acquainted with the researcher, reward, and environment but not the injection procedure, is expected to display fewer contact-seeking behaviors than the trained mice. Conversely, the unexposed control group, being novel to all elements, is expected to demonstrate the least amount of contact-seeking behavior. Two behaviors also observed in this study that can be linked to stress and fear directly are flinching and freezing (Lezak, *et al.,* 2022; Walker, *et al.*, 2003). Additionally, the duration of mice engaging in self-grooming was recorded. In rodents, grooming serves purposes beyond hygiene maintenance and thermoregulation; it can also be indicative of an adaptive response to stress (Mu, *et al.*, 2020). However, lower durations of self-grooming may also be observed in stressful situations, such as when mice exhibit freezing or fleeing (Song, *et al.*, 2016). Another study by the same researchers shows that the bout length of the behavior and their grooming pattern may also be correlated with the underlying cause of the manifestation of the behavior. (Kalueff, *et al.*, 2015). These matters contribute to the complexity of drawing definitive conclusions to the exhibition of this behavior. However, the behavior was still observed to potentially identify differences between treatment groups, as it is a behavior that was expected to be exhibited a lot. Lastly, various exploratory behaviors were also scored, including 'rearing', 'stretch attend', 'stretched walk' and 'looking over the edge'. It must be noted that certain contact-seeking behaviors may also be considered exploratory behaviors, e.g. sniffing the researcher. Research suggests that stress (and the triggered anxiety-like behavior) and fear often coincide with a decrease in exploration (Heinz, *et al.*, 2021; Ahumada, *et al.*, 2022). Building on this premise, the expectation is that mice from the trained group will exhibit more explorative behavior compared to the two other groups. Though, it is noteworthy that the trained and exposed mice, having acclimated to the environment, might manifest reduced exploratory behavior compared to their response in a novel environment (Heinz, *et al.*, 2021).

In conclusion, the purpose of this study was to investigate if training mice to stand still during an injection or solely exposure, can effectively reduce stress and/or fear in the context of a SC injection. Another objective was to refine and optimize the protocol as described in van Eldik (2021). Overall it is expected that the unexposed animals will exhibit the highest occurrence of stress-related behaviors (more struggling, freezing, flinching and a longer injection time – less standing still, eating, contact-seeking and explorative behavior), followed by the exposed mice displaying fewer stress-related behaviors than the unexposed mice, and the trained mice exhibiting the least. As previously stated, the employed training method emphasizes minimal restraint, providing an additional facet for stress reduction in these animals. The study will lastly also assess the practicality of the proposed protocol and its potential to be implemented in future research endeavors.

Materials and methods

The Animals

To calculate the amount of animals needed, a Power Analysis based on the mean and standard deviation of the contact seeking parameter in van Eldik (2021) was executed. The analysis was performed with G*Power 3.1.9.7 (Faul, *et al.*, 2007). The calculation showed that to keep a power of 80%, the study group needed to consist of 104 mice. Due to practical reasons this study focuses on a first batch, consisting of 24 mice. The animals in this study were female C3H/HeOuJ (n=5) and Balb/cAnNCrl (n=19) mice, born on 24-08-2022. One of the animals (X28Y5) functioned as a sentinel, a role established prior to the beginning of the study. Blood samples were obtained from this mouse on two occasions throughout the experiment to evaluate the general health status of the aging cohort.

Since before and during the study, the mice were kept in makrolon type IV-S cages with four to five mice per cage in the Central Laboratory Animal Research Facility of the University of Utrecht (Gemeenschappelijk Dierenlaboratorium (GDL)) since they arrived there at 7 weeks of age. The cages have the following dimensions: 59 x 38 x 20 centimeters and contain different types of enrichment: two orange tunnels, a transparent tunnel fastened to the lid of the cage that is used to transport the mice (Gouveia & Hurst, 2013), several wooden blocks as gnawing material and two small cardboard boxes they can use to hide and/or rest in. Figure 1 shows the general interior design of one of the cages, however the mice move (and tear up) the objects sometimes, making the design per individual cage a little bit different. The bedding of the cages consisted of wood chips. The animals received water and food (pellets: Rat/Mouse maintenance, Ssniff Spezialdiäten GmbH, DE-59494 Soest) *ad libitum* and were kept in a 12:12 day:night cycle (lights on at 7 am). The temperature of the room where the mice were housed was roughly 22 degrees Celsius, with a humidity of approximately 65%. The radio was on at a low volume during the 12 day hours of the day:night cycle. The cages were cleaned every first Monday of the month. During the cleaning process, the mice were tunneled to a different cage until cleaning was finished. They received clean wood chips as well as new cardboard boxes and paper tissues to tear up and use as bedding after every cleaning session.

Figure 1. On the left, the general interior design of a home cage is visible. On the right, the tunnel connected to the lid of the cage that is used for tunnel handling.

The mice were kept in the same cages as before the study, no pre-existing groups were mixed. The mice have been used/handled previously for educational purposes by Veterinary Medicine and Laboratory Animal Science students in multiple classes. The students practiced different handling and restraining methods on the mice, as well as (subcutaneous and intraperitoneal) injections with saline solution. The total amount of times they had been handled prior to the start of the study differs per

cage and mouse but was approximately 2-3 times a week. During this period, the mice had received 4-5 injections on average. The use of these animals for teaching, and the training of the animals (partly to prepare them for the teaching) has been approved by the Dutch Central Authority for Scientific Procedures on Animals (CCD) and the Animal Ethics Committee of Utrecht University (license: AVD10800202216046) and conducted in agreement with Dutch laws (Wet op de Dierproeven) and European regulations (EU-Guideline 2010/63/EU).

Study design

Habituation phase before training

Before the training phase of the experiment, all mice have been exposed to the researcher and the reward. These habituation sessions lasted five minutes per cage per day and were conducted to familiarize all mice with the researcher and the reward. During these sessions, the researcher would lay their (gloved) hand in the middle of the cage, holding a non-standardized amount of reward. Yogurt drops (ESVE, Knaagdier Drops Yoghurt & Bosvruchten) were utilized as the reward throughout the entire experiment, given the mice's existing familiarity with them. The drops were ground into a powderlike structure, preventing the mice from carrying the reward away from the researcher. This required them to remain in close proximity to the researcher if they wished to obtain the reward. The amount of reward was not standardized for every cage, because not all mice would eat (the same amount of) the reward. The mice were allowed to eat as much of the reward as they liked. This habituation phase ended when more than 90% of the mice would eat from the researcher's hand, which was after four days. It must be noted that the success of habituation was not measured specifically before proceeding to training/exposure. Figure 2 illustrates a timeline of the study, giving a quick overview.
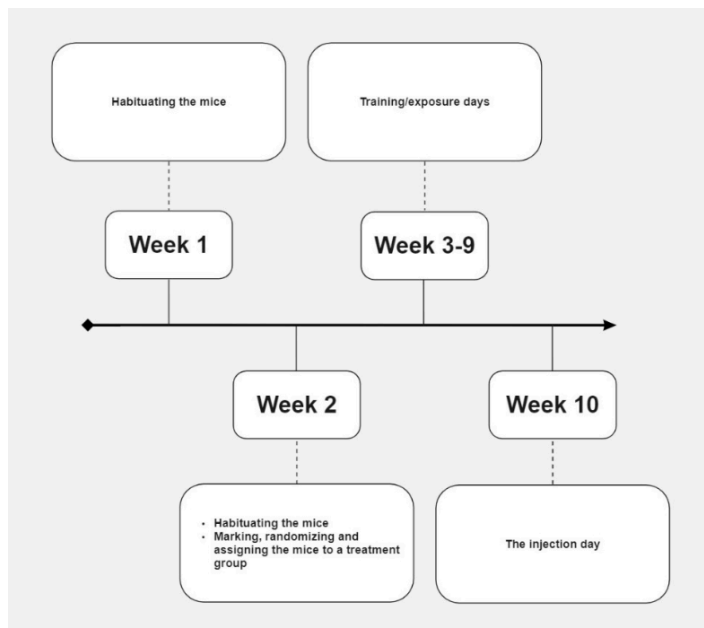

Figure 2. A timeline of the study. During week 1 and 2, four habituation sessions were conducted. During week 3 to 9, multiple training sessions were conducted, typically occurring thrice a week, with the inclusion of one or two "rest" days in between training sessions.

After the habituation phase, each mouse was marked using a permanent marker. They were placed on top of their respective cages via tunnel handling and received a mark (1 to 5 bands) on their tail. After marking the animals they were allocated randomly into three different groups using the RAND function in Excel. The mice were assigned names for verification based on their cage and the amount of bands. For instance, a mouse with three bands on its tail from cage 28 was labeled X28Y3.

The experiment

Three study groups were formed: 1) a training group (mice that were trained on the training pad), 2) an exposed control group (mice exposed to the training pad, researcher and reward, but not trained) and 3) an unexposed control group (mice always remained in their home cage and were never trained or exposed). For more information on the study groups and the different protocols, see "Protocols". The mice were kept in the same cages as before the study to minimize cage effects. After randomization the distribution of the groups per cage was reviewed to ensure every cage included at least one mouse of every group. Furthermore, it was ensured that the training group did not exclusively consist of Y1 mice since they were the first to be marked in their cage, likely indicating they were the easiest to tunnel/handle at this time in the project.

Every training/exposure session lasted three minutes, typically occurring three times per week with one or two 'rest' days between each session. Before each session, all cages were transported from their home room in the mouse housing facility, to the training rooms. This transfer involved placing the cages on a metal cage rack system on wheels, which would always be covered with a cotton sheet during transportation. The distance between the two locations was approximately 20 meters, and the entire process took a few minutes each time. The training and exposure took place in interconnected rooms, comprising a big area designated as the training area and a smaller room referred to as the waiting room. These two spaces were connected by a narrow corridor. The reason this space was selected is because of the pre-existing familiarity of the mice with these environments. The training area was previously employed for student classes (as mentioned before) and was intended for continued use in educational activities post-experimentation.

Before each training day, the sequence in which cages and mice were trained or exposed was determined using the RAND function in Excel. Each cage and mouse had an equal chance to be first and last in the sequence, which minimized the potential of order effects. For practicality, all mice from the same cage would be trained or exposed consecutively. At the initiation of each training day, the cage destined to start was removed from the cage rack system and positioned on a separate table in the waiting room. The researcher removed all enrichment, leaving only the wood chips bedding. The mice did have access to food and water during this time, as the lid would stay on their cage. Given that the sessions per mouse take three minutes, a single cage would remain without enrichment on the table for a duration equal to three minutes multiplied by the number of mice undergoing training or exposure from that cage (the time it takes to walk to the training area from the waiting room and back with the mice not included). The researcher consistently handled the mice with gloved hands, changing gloves only in instances of mice nibbling, causing damage to the gloves.

Upon emptying the cage, the mice were handled and transported to the training area using their tunnel. The general setup that was used during the sessions is shown in Figure 3. As can be seen in this figure, a (training) pad (Absorin Comfort, Medeco, Brandpuntlaan Zuid 14, 2665 NZ Bleiswijk, The Netherlands) of 60x60 centimeters was used, as well as a reward and a clicker (4011905228600, Trixie Amazon). The rewards used during this study were yogurt and forest fruit drops (ESVE, Knaagdier Drops Yoghurt & Bosvruchten), given the mice's existing familiarity with them. To time the sessions, the timer function on the researcher's phone was used. When finally measuring the animals' responses to the injections (see 'Study design' and 'Protocols' for more information), multiple cameras were needed to record the behavior(s) of the animals. Post-training or exposure, the researcher would re-mark the mice on the training pad by holding their tail slightly while drawing one to five bands. When all trained/exposed mice from a single cage had undergone a session, were re-marked and had been brought back to their home cage, the enrichment was transferred back and the cage was returned to its place in the metal cage rack. The rack would always stay in the waiting room, preventing the other mice from hearing the click during the sessions in the training area.
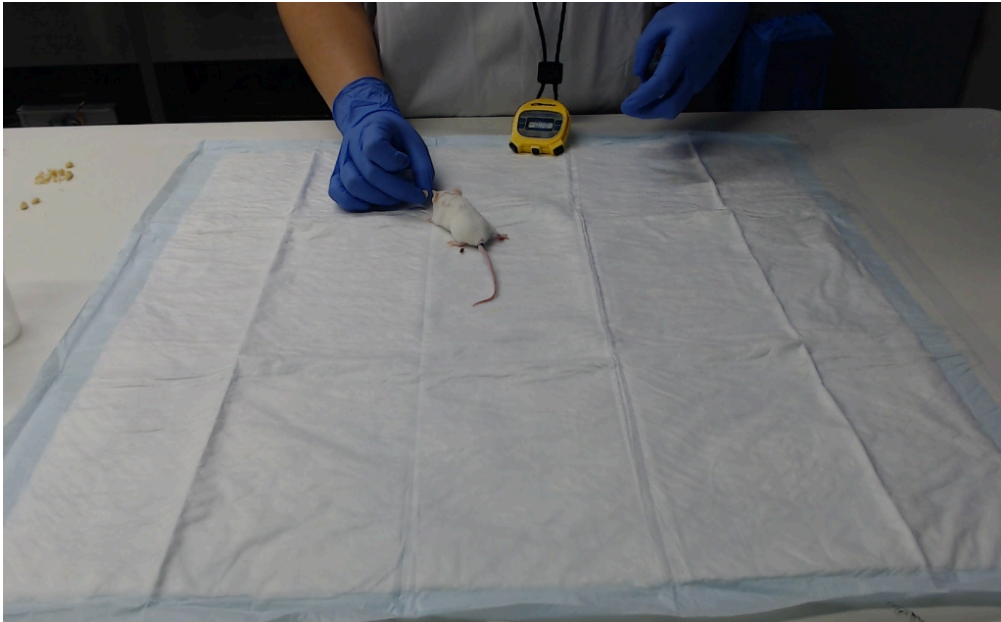
Figure 3: The general setup used during training and exposure. The clicker is in the researcher's left hand (not visible), while the reward is being offered with the researcher's right hand. The partially crushed reward is also visible on the table at the right side (left in the image) of the training pad.

The mice used in this project were housed in the same home room as mice used in another similar project investigating a training protocol for intraperitoneal (IP) injections. All mice would be brought to and from the waiting room as a group, as they share a home room in the housing facility. This means the mice were out of their home room for the duration of the sum of two batches of mice/sessions per training day. Each training day for this project (consisting of: emptying the cages of their enrichment, training all trained mice, exposing all exposed control mice, re-marking every mouse including unexposed control mice and putting back all the enrichment) lasted approximately two hours. A training day for the IP project running parallel to this project lasted two and a half hours on average. This means the cages were out of their home room for circa 4-5 hours on a training day. Training sessions consistently occurred between 10 am and 4 pm in this study, not only for the purpose of standardization but also due to the practical consideration that this time frame aligns with when the mice are commonly used by students for educational purposes. This scheduling decision aimed to ensure consistency and familiarity for the mice, as they are accustomed to being awake and active during this time frame. Additionally, if the protocol leads to refinement, it would be easier to directly implement it in education and teaching.

Protocols
Training protocol
The full training protocol and experimental set up used in the study are visible in Table 1 and Figure 3, respectively. Each session lasted three minutes. The amount of reward offered was not standardized between mice nor groups as not every mouse accepted the (same amount of) reward, especially in the beginning of the study. During the training sessions the mice's progression and reward acceptance were documented using the datasheet visible in Table A1 in the Appendix. The initial level of the mouse at the start of the session was noted in the 'Level' row. Progression of the mice was documented in the 'Level completed?' row. The progression through the levels per mouse during every session is depicted in Table A2 in the Appendix. It is important to note that a mouse could advance through more than just one level within a session (e.g. starting at level 3 but reaching level 6 by the end of the same session).

Level 0 → Linking	Sessions 1-7: Linking (and habituation) phase* → Rewarding after every click on the training pad.
	The first 7 sessions were used as the linking and exposure phase; to establish an association between the click and the reward by presenting the reward after every click. Additionally, if habituation was not entirely achieved during the prior habituation sessions, the possibility remains for it to occur during these (and following) sessions.
Level 1	From session 8 onward: Accepting the reward (at least twice) after a click on the training pad.
	The mouse has to realize that being in the new environment (on the training pad) will result in a click and a reward. The mouse moves on to level 2 when it accepts the reward from the researcher (at least twice).
Level 2	Standing still, while the researcher is holding the tail.
	The mouse has to realize that being on the training pad while standing still will result in a click and a reward. The mouse moves on to level 3 when it does not struggle anymore while her tail is held by the researcher.
Level 3	Standing still, without the researcher holding the tail.
	The mouse has to realize that she has to remain still when her tail is released by the researcher. This will result in a click and a reward. The mouse moves on to level 4 when it stays in the same spot when its tail is released by the researcher.
Level 4	Standing still, while the hand of the researcher is held above its body (within 5 cm), without holding the tail.
	The mouse has to realize that movement of the researcher's hand above its head or body while standing still results in a click and a reward. The mouse moves on to level 5 when it does not attempt to leave the training pad and stays in the same spot when the researcher moves their hand above the mouse.
Level 5	Standing still, while the researcher touches or pets them, without holding the tail.
	The mouse has to realize that when the researcher touches or pets it while it is standing still, results in a click and a reward. The mouse moves on to level 6 when it accepts being touched or petted by the researcher.
Level 6	Standing still, while a skinfold is lifted, without holding the tail.
	The mouse has to realize that lifting a skin fold (anywhere on the body but preferably the flank), while standing still, results in a click and a reward. The mouse moves on to level 7 when it accepts it when the researcher lifts a skin fold.
Level 7	Standing still, while a skinfold is lifted and a capped needle touches the skin, without holding the tail.
	The mouse has to realize that having its skin fold lifted and touched with a needle on a syringe (with the cap on), while standing still, results in a click and a reward. The mouse moves on to level 8 when it remains still when the researcher lifts a skin fold and touches the skin with a capped needle.
Level 8	Standing still, while a skinfold is lifted and an uncapped needle touches the skin, without holding the tail.
	The mouse has to realize that having its skin fold lifted and touched with the needle on a syringe (with the cap off), while remaining still, results in a click and a reward. The mouse moves on to level 9 when it stands still when the researcher lifts a skin fold and touches its skin with a needle.
Level 9	Injection.**
	The mouse has its skin fold lifted and is injected with 0,1 ml NaCl SC. Further details can be found in section "The day of injection".

*Table 1: A summary of the training protocol. *It must be noted that the process of linking may persist throughout each level/session and thus potentially establish at a later point. **This level is executed only once, on the day of the injection.*

The purpose of the linking phase was to make the mice link the sound of the click to the presentation of a reward. It was expected that successful linking would elicit some type of anticipatory behavior, such as sniffing the air, looking around or at the researcher's hand. Nevertheless, at session 7, not all mice consistently exhibited this behavior. Moreover, the anticipation was that all mice would accept the reward after a few sessions, which did not happen either (Table A2 in the Appendix). However, 75% of the mice did accept the reward, prompting the initiation of the training phase of the protocol, as waiting until all mice would eat would take too long. The expectation was that for the mice that were not yet successfully linked or did not yet eat, this would gradually follow in subsequent sessions,

which was the case for two additional animals. Lastly, it is essential to clarify that the mice could not progress through any levels during the linking sessions. These seven sessions exclusively served the purpose of linking and did not involve training yet.

A fixed number of training sessions was not predetermined at the start of the study, given our initial uncertainty regarding the mice's adaptation rate to the training schedule. In total the mice have been trained or exposed for 19 sessions of three minutes. At session 14, four out of the eight training mice had reached level 8, and all but one mouse from the exposure group were accepting the reward rather consistently. At this point, the objective was to have more than 50% of the trained mice complete level 8 before initiating the injection for all mice. This was with the anticipation that one or more trained mice might still progress through a few additional levels. By session 17, one additional mouse reached the final level, resulting in more than 50% of the trained mice achieving and finishing level 8. Moreover, we acknowledged that one mouse (X19Y2) would not progress any further at this rate as she had never accepted the reward. Consequently, this mouse was excluded from the statistical analysis. The injection day was scheduled a week after session 17, allowing for two additional training sessions. In these two sessions, two mice (X19Y5 and X28Y4) advanced several levels, which was beneficial. Mice that completed level 8 earlier than session 19 were still trained and went through levels 6, 7 and 8 again every additional session until the day of injection.

Exposure protocol
Similar to the trained mice, the exposed control mice underwent a total of 19 sessions. They received *ad libitum* rewards from the researcher on the training pad and were allowed to roam free during the entire exposure session. Every exposure session lasted 3 minutes, the same amount of time as the training sessions. The clicker was used during these sessions to expose the mice to the clicks, but it was applied in a randomized manner, deliberately hindering the establishment of an association between the click and the reward. The researcher made efforts to sustain an average click frequency roughly equivalent to that seen in the trained group of mice. As the mice were not restrained, some would leave the training pad to roam on the table. The researcher would then gently scoop them up with their hands or pick them up using the tunnel to transfer them back to the training pad. If a mouse would climb the researcher (via their arm or lab coat) another researcher would safely put them back on the training pad. The reward acceptance of the exposed mice during the exposure sessions is visible in table A3 in the Appendix.

Control protocol
The unexposed control mice never received any rewards during weeks 3 to 9 of the training phase, nor have they spent any time on the training pad. The only instance they were in contact with the researcher was when they had to be re-marked due to them grooming the ink off. The re-marking of the unexposed control mice always occurred during a training day, when the cage was emptied of enrichment. The mice would be re-marked by either the researcher at the end of the sessions of their cage or during the session by the researcher of the parallel running IP project. The re-marking always occurred in the waiting room on the lid of their home cage. They are familiar with the tunnel, as it is used to move them onto the lid.

The day of injection
To assess the efficacy of training as refinement compared to the exposed and unexposed control groups on (stress-related and researcher-directed) behavior(s) in response to a SC injection, all 24 mice received a SC injection of 0,1 ml NaCl 0,9% with 30G needles, in week 10. Firstly, all mice were re-marked a few days prior to the injection by the research supervisor, ensuring that the researcher was blinded to the mice's group allocation while performing the injection. However, it should be noted that some mice (n=5) exhibited distinct physical characteristics, inadvertently revealing their identities to the researcher, such as: coat color, bald spots or an ear cut. The project supervisor randomized the order of the cages and mice within a cage for the injection, while ensuring that the order in which mice

were used was balanced for treatment (for example, not measuring all trained mice as the first mice from their cage, as this could lead to confounding treatment effects with potential order effects). Secondly, one day prior to the injection, the setup in the training area and waiting room were prepared (Figure 4). In both rooms a webcam was present to record the mice's behavior before and after injection in their home cage and after the injection on the training pad.
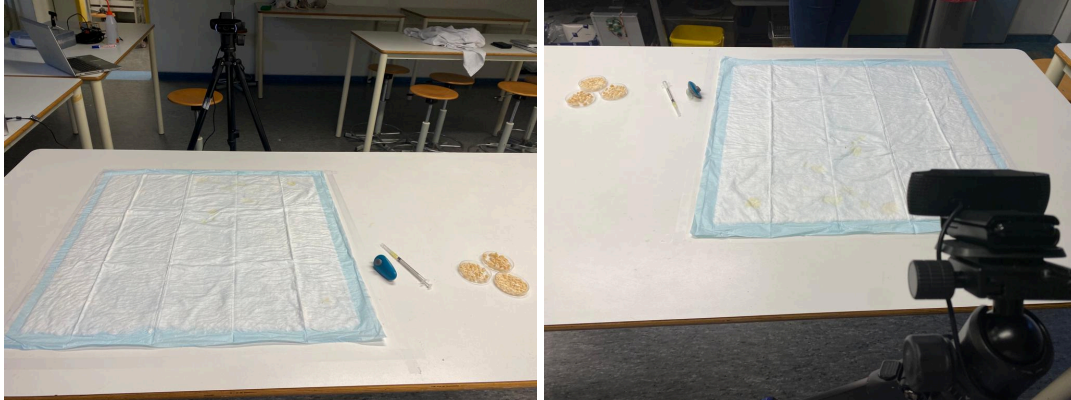

Figure 4: The setup for the injections and recording of the behavioral tests in the training area on the day of the injection.

On the day of the injection, all animals underwent three different tests. First, all animals from one cage would simultaneously undergo a VAT in their home cage in the waiting room. For this initial test, all enrichment was removed and the cage was positioned on a table, as per usual during a training session. The researcher would lay their gloved flat hand flat in the right corner of the cage for five minutes, to be able to observe the behavioral response to the researcher's hands while not yet having been exposed to the injection (a potential stressor). The results from this test will not be discussed in this article. After the initial home cage VAT, the researcher would change their gloves in the corridor with the door closed and get the first mouse that had to be injected according to the sequence. The webcam in the waiting room remained on, continuously recording the mice in their home cages before and after injection.

The mouse was brought to the training pad in the training area with a tunnel as per training protocol. Upon gently placing the mouse on the training pad and picking up the syringe, a one minute timer was started. During this minute the researcher would try to inject the mouse following the same procedure as during training. This involved lifting a skin fold, making contact with the needle on the skin, but this time actually injecting 0,1 ml of NaCl 0,9%, followed by a click and a reward. As expected, not all mice stayed still during the injection attempts, prompting the introduction of the one minute timer. If injecting the mouse was deemed unsafe due to excessive struggling within the one-minute timeframe, the research supervisor intervened to restrain the mouse, enabling the injection to proceed safely. Restraint was necessary for only two mice, namely X19Y1 (unexposed control) and X19Y3 (exposed control). Unfortunately, in the cases of X23Y5 (unexposed), X19Y5 (trained) and X11Y2 (trained), the needle went through the skin, resulting in them not receiving the 0,1 ml of NaCl subcutaneously. However, the mice were still included in the statistical analysis, since the NaCl is expected to have no effect and the needle did pierce their skin, exposing them to the same stressor as the other mice.

After the injection, the reward was offered for one minute during the reward acceptance test. The researcher would follow the mouse with the reward if they decided to walk and explore the training pad. The mice were allowed to have as much of the reward as they desired within this one-minute period.

After the reward acceptance test, the researcher would put away the reward and a second VAT started – this time on the training pad. The researcher laid their hand flat in the corner of the training pad for two minutes, letting the mouse roam around and explore. The behaviors scored during these

two minutes are detailed in Table 3 in the "Ethogram and behavior scoring" section. If the mice ventured off the training pad, exceeding a body length (including the tail) to the left or right, or when they moved out of camera frame in the front, they would be picked up by the researcher and returned to the middle of the training pad. If, however, a mouse returned to the training pad independently before the researcher could scoop them up, they were not handled or put to the center of the pad. Instead, the researcher returned their hand to the corner once the mouse reentered the frame and the VAT continued. Occasionally, a mouse would leave the camera frame while climbing the arm of the researcher, in which case they had to be put back on the pad with the help of a colleague. On average, mice were out of frame 0,7 times (range: 0-3) totaling a mean duration of 3,3 seconds (range: 0-17,744). Notably, a considerable variability among individual mice was observed, prompting the decision to control for the time the mice were out of frame when analyzing the behavioral data (refer to the "Statistics" section for details).

When the two minute VAT ended, the mouse was immediately returned to its home cage. The researcher would change gloves again in the corridor and get the second mouse. This process was repeated as such for all mice. Once the last mouse of a cage was injected, rewarded and observed during the VAT on the training pad, a ten-minute intermission started. This break was implemented to allow the last mouse to have some recovery time following the injection as well, considering that the preceding mice had already had a minimum of five minutes of recovery. After the ten minute intermission, a second home cage VAT was performed to investigate if the mice's reaction to the researcher changed compared to before the injection. Once again, the results of these home cage VATs will be published elsewhere. At the five-minute mark on the stopwatch, the enrichment was returned to the cage. Following this, the cage was put back into the cage rack and the recording in the waiting room was stopped. This sequence of tests (Table 2) was consistently repeated for all cages.

TEST	ANIMALS	ENVIRONMENT	DURATION
Voluntary approach test pre-injection	All mice of one cage	Home cage in waiting room	5 minutes
SC injection	One mouse at a time	Training pad in training area	1 minute*
Reward acceptance test	One mouse at a time		1 minute
Voluntary approach test	One mouse at a time		2 minutes
Voluntary approach test post-injection	All mice of one cage	Home cage in waiting room	5 minutes

*Table 2: The different behavioral tests mice from every cage underwent consecutively, the locations where the tests were performed and their respective durations. *The timer for the injection was set to one minute; however, mice that stood still were injected quicker. Conversely, mice that required restraint after the expiration of the one-minute timer had an injection duration longer than one minute. The VATs in the home cages in the waiting room have been marked gray, as they are not part of the analysis in the current study.*

Ethogram and behavior scoring
As logging software the program BORIS (version 8.20.3 2023-06-04) was used to score the mice's behaviors (Friard & Gamba, 2016). The behaviors scored in BORIS during the injection, the reward acceptance test and the VAT can be found in the ethogram (Table 3). For more details on the start and stop margins of the behaviors, consult Table A4 in the Appendix. In total 24 videos/mice have been scored, during which the allocation remained unknown to the researcher. The intraobserver reliability was 84% on average (0,80-0,87).

DURING THE INJECTION	
Time it takes to inject	The total time it took to safely inject SC from the first attempt at lifting a skin fold to letting the mouse go after injection, even if they required restraint.
Standing still	Scored when the only movement of the mouse was breathing and/or sniffing. Moving was allowed to a certain point, e.g. taking a small side step with one of the paws but making less than a 90 degree turn.
Struggling	Scored whenever 'standing still' was not scored, e.g. when walking or flinching during injection/lifting of the skin fold and moving in a certain way that made injecting unsafe, like turning more than 90 degrees, flinching, or rearing upon lifting the skin fold.
DURING THE REWARD ACCEPTANCE TEST (1 MINUTE)	
Accepting the reward	Scored the entire time the mouse was chewing. *If the mouse continued to chew after the reward acceptance test into the VAT, the scoring continued.
DURING THE VAT ON THE TRAINING PAD (2 MINUTES)	
Touching the researcher	Scored when the mouse would touch the hand/arm of the researcher with any body part (front paws, hind paws, body) excluding the tail. If the mouse made contact with the researcher using all four paws simultaneously, scoring of 'touching the researcher' was stopped, and scoring of 'spending time on the researcher' was started.
Spending time on the researcher	Scored when the mouse would sit on top of the researcher's hand/arm with all four paws.
Sniffing the researcher	Scored when the mouse would sniff the researcher's hand/arm.
Sitting next to the researcher	Scored when the mouse would sit within one half of its body length away from the researcher's hand. Scoring of this behavior stopped if the mouse would exhibit any described behavior other than sitting.
Nibbling the researcher > *Only one mouse (unexposed control) exhibited this behavior (0,6 seconds).*	Scored when the mouse would nibble/bite the researcher's hand.
Stretch attend	Scored when the mouse would move its front paws forward or hind paws backward and leaned in to stretch its body or when the mouse would stretch its body by leaning forward without moving any of its paws.
Stretched walk	Scored when the mouse would walk with their back legs stretched after exhibiting a stretch attend.
Rear	Scored when the mouse would sit on her hinds paws without her front paws touching the training pad. Scoring of this behavior stopped when the mouse would start grooming (see "Grooming").
Flinch > *No mice exhibited this behavior.*	N/A
Freeze > *No mice exhibited this behavior.*	N/A
Sitting elsewhere on the mat	Scored when the mouse would sit on the mat further than half of its body length away from the hand of the researcher. Moving of the head was allowed (e.g. when sniffing or looking around) as well as lifting or replacing solely one of its front paws.
Grooming > *Eight mice exhibited this behavior (two trained, two exposed controls and four unexposed controls), with an average of 2,3 seconds.*	Scored when the mouse would sit back on her hind paws and would clean her snout and/or ears with her front paws. Also scored when the mouse would lick or scratch any other part of its body.
Looking over the edge > *Nine mice exhibited this behavior (three trained and six unexposed controls), with an average of 4,7 seconds.*	Scored when the mouse would look over the edge of the table. This behavior was not scored simultaneously with other behaviors, like stretch attend.

ADDITIONAL MEASURED DURATIONS	
Out of frame	Scored any time the mouse would leave the camera frame and no behavior scoring was possible. This could happen by either leaving the training pad or climbing the researcher's arm. Behavioral scoring continued the first frame the mouse would be back in camera frame while exhibiting any of the described behaviors in the ethogram.

Table 3: Ethogram (shortened) of all behaviors scored during the injection and behavioral tests, including additional measured durations. The cells of the behaviors are marked green if they have been statistically analyzed.

Statistical analyses were only executed on the eight most important and prevalent behaviors scored with BORIS. These behaviors are marked green in Table 3. The decision to exclude certain behaviors from the analysis was influenced by time constraints and the fact that not all of those behaviors were exhibited consistently and/or only briefly by all mice. Analyzing these behaviors would most likely yield unreliable results due to insufficient data. The statistical analysis included both the total injection time and 'standing still' during injection. As detailed in Table 3, struggling is defined as any behavior that was not standing still. Consequently, the focus of the analysis was on standing still, as the mice were specifically trained to exhibit this behavior. Considering the behaviors 'freeze,' 'flinch,' 'grooming,' 'nibbling the researcher,' and 'looking over the edge,' the available data was insufficient; consult Table 3 for additional details. The behavior 'sitting elsewhere on the mat', 'rear' and 'stretched walk' were deemed less important to analyze as they are not specifically stress or anxiety related. However, the raw data including all behaviors of all mice is available, making it possible to investigate these behaviors in a subsequent article at a different time.

The total time it takes to inject all mice was measured in seconds. But, because this duration varied among the animals, 'standing still' was measured as a percentage of the total injection time to allow a fair comparison among the mice. Similarly, because some mice went out of camera frame for a short time during the VAT (see "The day of injection" for details), 'spending time on the researcher', 'touching the researcher', 'sniffing the researcher', 'sitting next to the researcher' and 'stretch attend' were measured as a percentage of the total time the mice were visible. 'Accepting the reward' was measured in seconds, as this behavior was sometimes continued into the VAT.

Statistics
Statistical analyses were executed with R 4.3.2., version Eye Holes (R Development Core Team 2023). Statistical significance was assessed at a significance level (α) of 0,05. P-values less than 0,05 were considered statistically significant, indicating rejection of the null hypothesis.

General linear mixed models (GLMMs) were fitted for the following behaviors: 'time it takes to inject', 'standing still during injection', 'accepting the reward' and 'sitting next to the researcher' (calculated using lmer from the lme4 package (Bates, *et al*., 2015), combined with the lmertest package (Kuznetsova, *et al*., 2017) for extracting F and p values). Initially, the GLMMs incorporated factors such as treatment group, the order of injection, cage and the need for restraint during injection. Through a stepwise process, all non-significant predictors (p>0,05) were systematically eliminated, resulting in the utilization of the most straightforward statistical model. This model included the treatment group as a fixed effect and incorporated the cage as a random effect. To achieve normality, the 'time it takes to inject' behavior was transformed to a natural logarithm. Using a general linear mixed model to analyze the behaviors 'touching the researcher', 'sniffing the researcher' and 'stretch attend' resulted in a 'boundary (singular) fit' error when the random effect 'cage' was included. As the estimation of the random effect 'cage' was not achievable in R, a basis analysis of variance (ANOVA) was used for these behaviors. To assess if the residuals of the behaviors were normally distributed, the Shapiro-Wilk normality test was used.

Given the non-normal distribution of residuals for both 'touching the researcher' and 'time spent on researcher' following natural log and square root transformations, non-parametric statistical analysis, specifically the Kruskal-Wallis test, was used. In case of significant main effects, the function emmeans from the 'pbkrtest' package (Halekoh & Højsgaard, 2014) was used for post-hoc testing.

Mice that never had their skin fold lifted (below level 6) were excluded, as this signifies a substantial lack of training and renders them non-representative for the trained treatment group. This only applied to one mouse (X19Y2), because it consistently refused any reward throughout the entire experiment and therefore never completed level 1 (Appendix Table A2).

Results

The total injection time (in seconds)

Treatment did not affect the time it took to inject the mice ($F_{(2,16.106)}=1,25$; $p=0,12$), meaning that training the mice did not result in a faster procedure (mean=13,0 ; SE=4,16), compared to non-trained mice (either exposed (mean=18,0 ; SE=8,69) or unexposed controls (mean=22,5 ; SE=7,64)).
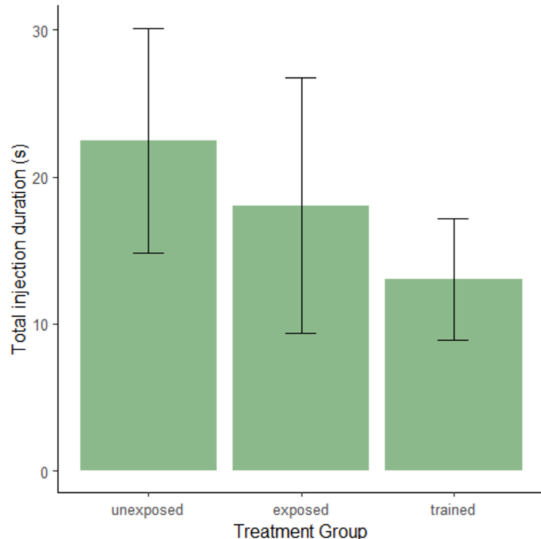

Figure 5: Total injection time (in seconds) for all treatment groups (mean ± 1 SE). (n=7 for trained mice, n=8 for exposed control mice, and n=8 for unexposed control mice).

Standing still during injection (as a percentage of the total injection time)

Treatment did affect the time mice stood still during the injection procedure ($F_{(2,16.169)}=4,21$; $p=0,034$). Post-hoc tests revealed that trained mice (mean=87,0 ; SE=7,13) stood still significantly longer during injection than unexposed control mice (mean=63,3 ; SE=11,3) (p=0,049). Though not significant, exposed control mice (mean=81,3 ; SE=8,13) tended to stand still longer than unexposed control mice (p=0,075), as well. There was no significant difference between the exposed control mice and the trained mice (p=0,95).
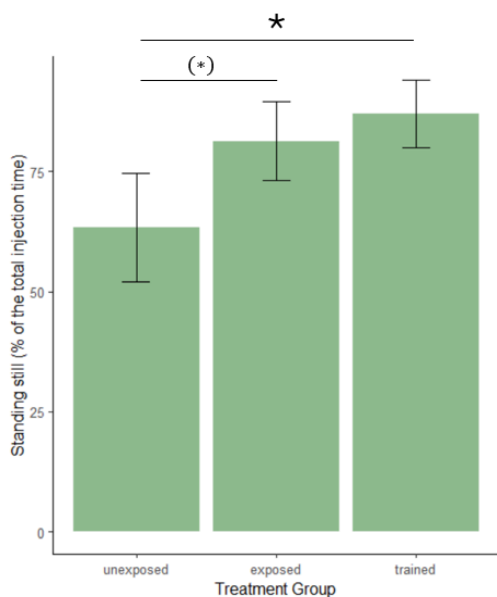

*Figure 6: 'Standing still' duration during the injection (as a percentage of the total injection time) of all treatment groups (mean ± 1 SE). * Indicates a significant difference (p<0,05). (*) Indicates a trend for an effect (0,05<p<0,1). (n=7 for trained mice, n=8 for exposed control mice, and n=8 for unexposed control mice).*

Reward acceptance (in seconds)

The results show that treatment affects the amount of time the mice accepted the reward after the injection ($F_{(2,16.55)}$=4,24 ; p=0,033). Post-hoc testing revealed that the exposed control mice (mean=49,2 ; SE=8,26) ate significantly longer than the unexposed control mice (mean=17,2 ; SE=7,48) (p=0,026). No significant differences were found between the unexposed control and trained group (mean=35,7 ; SE=13,1) (p=0,38) or the exposed control group compared to the trained group (p=0,36).
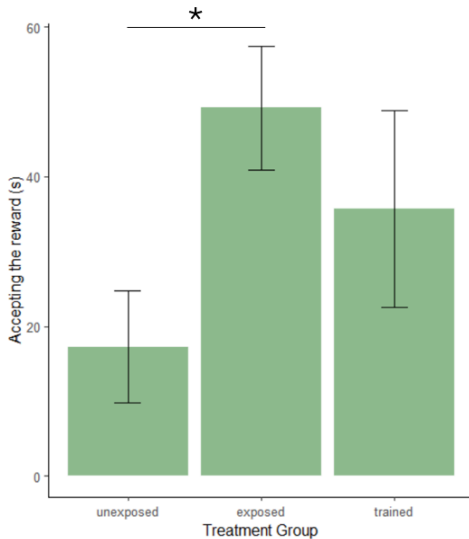

*Figure 7: 'Accepting the reward' duration during the reward acceptance test (in seconds) of all treatment groups. * Indicates a significant difference (p<0,05). (n=7 for trained mice, n=8 for exposed control mice, and n=8 for unexposed control mice).*

Behaviors during the VAT (as a percentage of the total time the mice were in camera frame)

Merely a trend, but no significant difference, was found for 'touching the researcher' ($F_{(2,20)}$=3,38) ; p=0,067) between all treatment groups (trained: median=5,65 ; IQR=11,3), exposed: median=13,4 ; IQR=10,1), unexposed: median=3,2 ; IQR=11,8).
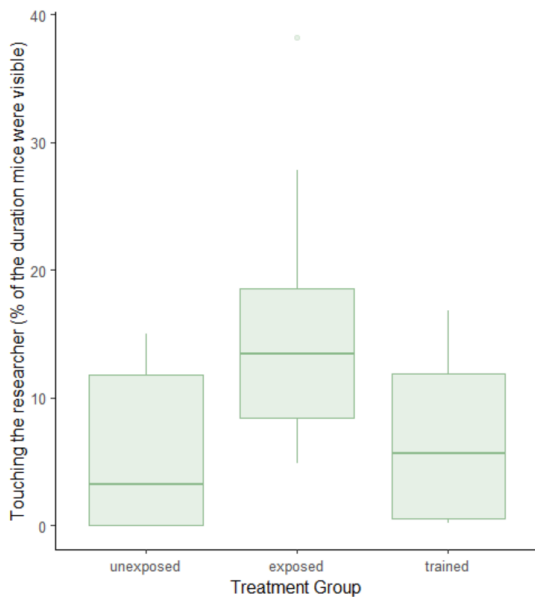

Figure 8: 'Touching the researcher' duration during the VAT (as a percentage of the total time the mice were in camera frame) of all treatment groups. (n=7 for trained mice, n=8 for exposed control mice, and n=8 for unexposed control mice).

Treatment group did not affect the behavior 'spending time on the researcher' ($F_{(2,20)}$=2,3 ; p=0,33), either (trained: median=0 ; IQR=28,6), exposed: mean=10,9 ; IQR=28,9, unexposed: median=0 ; IQR=0,983).
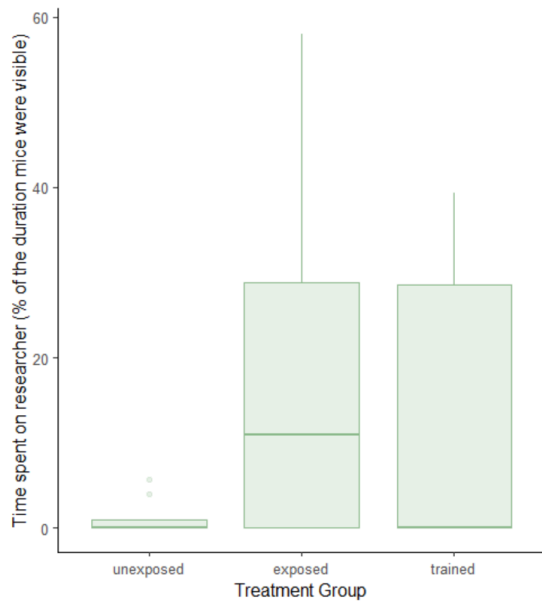

Figure 9: 'Time spent on the researcher' duration during the VAT (as a percentage of the total time the mice were in camera frame) of all treatment groups. (n=7 for trained mice, n=8 for exposed control mice, and n=8 for unexposed control mice).

No significant difference for 'sniffing the researcher' was found among the trained group (mean=8,33 ; SE=1,99), exposed control group (mean=9,75 ; SE=1,74) or unexposed control group (mean=6,97 ; SE=1,79) ($F_{(2,20)}$=0,61 ; p=0,55).
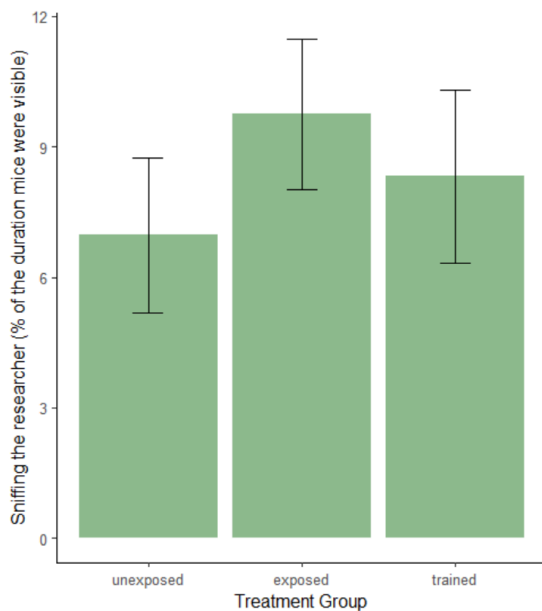

Figure 10: 'Sniffing the researcher' duration during the VAT (as a percentage of the total time the mice were in camera frame) of all treatment groups. (n=7 for trained mice, n=8 for exposed control mice, and n=8 for unexposed control mice).

Considering the 'sitting next to the researcher' behavior also no significant difference was found between the trained mice (mean=16,6 ; SE=4,54), exposed (mean=21,2 ; SE=6,97) and unexposed mice (mean=22,8 ; SE=6,93) ($F_{(2,17.107)}$=0,22 ; p=0,81)
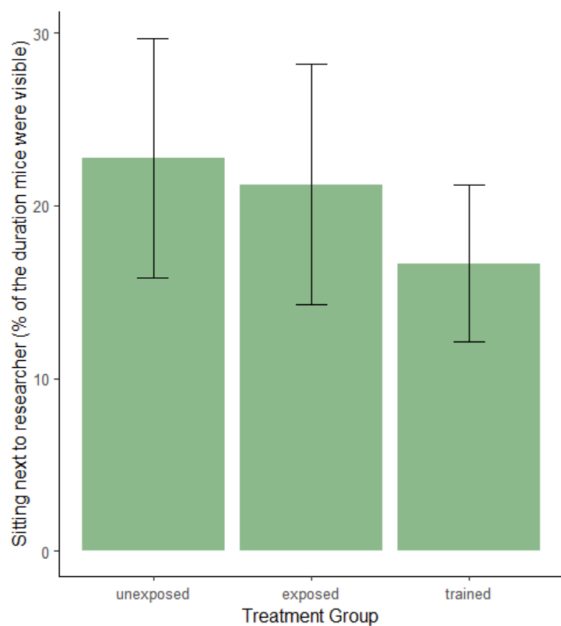
Figure 11: 'Sitting next to the researcher' duration during the VAT (as a percentage of the total time the mice were in camera frame) of all treatment groups. (n=7 for trained mice, n=8 for exposed control mice, and n=8 for unexposed control mice).

Analyzing the behavior 'stretch attend' did not yield any significant results between the trained (mean=19,4 ; SE=3,58), exposed (mean=17,0 ; SE=1,95) and unexposed (mean=24,6 ; SE=2,46) groups either ($F_{(2,20)}$=2,22 ; p=0,13).
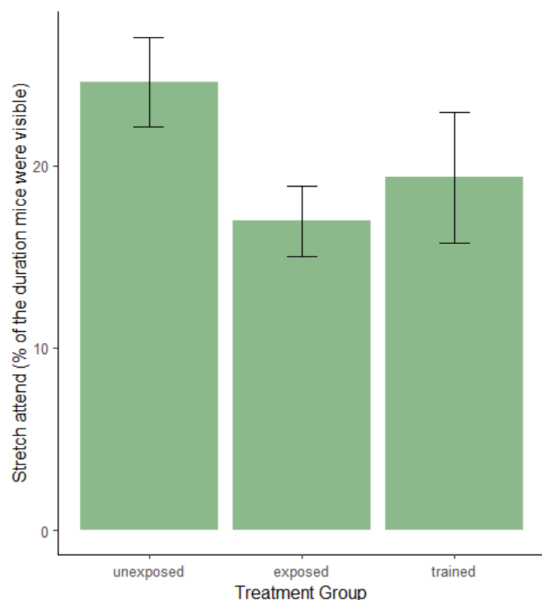
Figure 12: 'Stretch attend' duration during the VAT (as a percentage of the total time the mice were in camera frame) of all treatment groups. (n=7 for trained mice, n=8 for exposed control mice, and n=8 for unexposed control mice).

This means that training the mice did not result in them interacting more with the researcher than exposed or unexposed mice, nor that they exhibited more explorative behavior, like stretch attends, on the training pad compared to the two control groups.

Discussion

Subcutaneous injections are routine procedures in laboratory settings. The injections frequently elicit stress in mice, often associated with the accompanying restraint (Meijer, *et al*., 2006). Recognizing the importance of minimizing stress for both ethical and methodological reasons, this study aimed to investigate the impact of clicker training as a refinement on stress in laboratory mice following a subcutaneous injection with minimal restraint. Of particular interest were the behaviors assessing the mice's inclination to approach the researcher post-injection. To explore the hypotheses – if trained mice would stand still longer during injection, if they would accept the reward longer after injection and if they would exhibit less stress-related behaviors during the VAT – three groups were established: a trained group, an exposed control group, and an unexposed control group. The exposed group was formed primarily to explore whether exposure alone could potentially reduce stress, given its comparatively less time-consuming and less challenging nature compared to training.

Notably, trained mice stood still longer during the injection than unexposed control mice (Figure 6). This result is in line with our expectations, as trained mice were specifically trained to stand still. Although merely a trend, exposed control mice also tended to stand still longer than unexposed control mice. This suggests that exposure alone could be sufficient for a safe and less stressful injection administration. As struggling during injection is often indicative of fear or stress (Walker *et al*., 2003; Stuart & Robinson, 2015), these findings imply that trained mice may experience the least amount of stress during injection compared to the two other groups.

In contrast, it is remarkable that the time it took to complete the injection was not significantly different between treatment groups. This suggests that, despite exposed and unexposed mice struggling more, the researcher was able to administer the injections within a similar timeframe. Nonetheless, as illustrated in Figure 5, from unexposed to exposed to trained mice, there is a decrease in the mean of the total injection duration. Moreover, the trained mice exhibited a greater duration of reward acceptance compared to the unexposed mice. This might indicate that the injection (including the procedure: chasing the mouse with a needle to lift a skinfold due to them not standing still) induced stress, potentially leading to reduced or inhibited eating. Additionally, it is of course safer for both the mice and the researcher if the mice stand still during injection administration.

Curiously, the exposed control group consumed the reward longer than the unexposed control group, with no other significant differences between groups identified (Figure 7). These findings suggest that exposure alone influences the duration mice spend eating or the quantity of reward they are willing to accept. The expectation was that trained mice would accept the reward for a longer duration, as they are supposed to be familiar with all aspects of the injection procedure excluding the actual injection. While not statistically significant, trained mice did exhibit a longer average 'accept the reward' duration than unexposed mice. This discrepancy from the hypothesis may be due to the variations in reward consumption between the two groups during the training/exposure sessions. In contrast to the trained mice, which were required to execute specific behaviors corresponding to their current level before obtaining a reward, the exposed mice were permitted continuous access to the reward throughout a session. Consequently, exposed mice may have developed a greater familiarity with prolonged reward consumption compared to the trained mice. Additionally, during the training sessions, the trained mice received a reward following a click each instance of correctly executed behavior. On the injection day, they heard the click and had to stand still only once, possibly interpreting this as the only time they were allowed to accept the reward. Another reason the results deviated from the hypothesis may be attributed to insufficient analysis of the behavior. Perhaps, the analysis would have been more comprehensive if the latency had been calculated. The latency in this context refers to the time between the researcher presenting the reward and the initial acceptance of the reward by the mouse during the reward acceptance test. There was a variability noticeable in this parameter among different mice during the behavior scoring. Therefore, it is plausible that two mice may have similar 'accepting the reward' durations, yet one mouse might have accepted the reward immediately after

injection, while another mouse may have needed time to recover from the injection before feeling comfortable enough to accept the reward. This suggests the possibility that trained mice may consume less of the reward, but need a shorter recovery period post-injection compared to exposed mice. Unfortunately, the latency could not be directly extracted from BORIS and time constraints precluded manual calculation. However, raw data and footage are available, enabling subsequent studies to utilize the data for latency analysis.

As highlighted in the introduction, research suggests that a mouse's reluctance to eat could be indicative of stress or anxiety (Yamada *et al*., 2020; Qiao, *et al.,* 2020; Misslin & Cigrang, 1986). Based on this insight, these findings imply that exposed mice may experience less stress/anxiety after SC injection than trained or unexposed mice. However, it is important to acknowledge that a mouse's refusal to accept food from a person may not exclusively indicate stress. While stress can indeed impact feeding behavior in mice, several factors influence their willingness to consume food, including environmental conditions, novelty, health status and individual preference (Francois, *et al*. 2020). The first three factors were addressed by standardizing the mice's environment, familiarizing the mice with the reward and researcher, and doing regular health checks on the animals, including the use of sentinels, respectively. Individual preference, while inherently influential, can be partially controlled by the implementation of a food preference test. However, for the sake of maintaining standardized reward acceptance and for practicality, this was not pursued. Moreover, during the habituation phase, over 90% of the mice accepted the reward in their home cage, suggesting a probable preference for the offered reward. Though, a potential avenue for future research could involve testing whether alternative treats of higher value elicit greater acceptance among mice.

Considering the contact-seeking 'touching the researcher' behavior, a trend was observed, with exposure mice touching the researcher for a longer duration than the other two groups (Figure 8). This contradicts the initial hypothesis, suggesting that exposed mice are more inclined to interact with the researcher, potentially indicating lower stress levels. Conversely, it implies that trained mice may be less accustomed to touching the researcher. This could be the case, as they were expected to stand still and perform during the training sessions, unlike the exposed mice. The exposed mice, having more freedom during the sessions, may have had the ability to become more familiar with touching. However, the deviation from the hypothesis could also be attributed to the way the contact-seeking behaviors were scored. As mentioned previously, 'touching the researcher' was not scored concurrently with the time the mice were 'spending time on the researcher', as these were defined as separate behaviors. Combining these behaviors, considering that the mouse is theoretically still touching the researcher even when entirely on them, could yield different results.

The results show that there were no statistical differences between treatment groups considering the remaining behaviors: 'time it takes to inject', 'touching the researcher', 'spending time on the researcher', 'sniffing the researcher', 'sitting next to the researcher' and 'stretch attend'. This indicates that neither training nor exposure had a noticeable impact on the duration of these behaviors. Consequently, based on these behaviors, it remains inconclusive whether mice from any treatment group experienced increased or reduced stress post-injection.

The absence of statistical significance between treatment groups in the remaining behaviors during the VAT prompts an exploration into potential factors contributing to this outcome. It is plausible that several limitations within the study may have influenced the observed results. First of all, the sample size (n=24) was relatively small compared to the size determined by the power analysis (n=112). Combining a second cohort, considering this study focused on the first batch, may enhance the likelihood of validating observed trends and observing more significant results between treatment groups. Secondly, all mice used in the experiment had prior handling and injection experience from students in animal handling classes. As a result, the unexposed control mice were never fully unexposed to handling or injections (with restraint), potentially subjecting them to aversive

experiences. However, each of the 24 mice underwent those classes, ensuring a relatively standardized experience among them. Additionally, the unexposed control mice were exposed to the researcher and reward during the habituation period, and were exposed to the researcher when re-marking their tail was necessary. These factors could potentially have contributed to a less distinct difference between the treatment groups.

Another factor contributing to the training group's lack of significant differences from the other two groups in most behaviors could be attributed to the relatively advanced age of the mice (11 months at the start and 13 months at the end of the experiment). Generally, older animals, as evidenced by studies (Elias & Elias, 1976; Wallis *et al.*, 2016; Head *et al.*, 1995), may experience a decline in learning ability, including cognitive decline affecting (spatial) memory, associative learning, and other cognitive tasks (Matzel *et al.,* 2008), compared to their younger counterparts. However, the extent of this decline can vary among individual animals and may depend on other factors such as genetics, health, and environmental conditions (Nyberg, *et al.*, 2020; Harada, *et al.*, 2014), adding complexity to the relationship between age and learning. For these reasons, it is important to note that individual variations exist, and some older animals may still exhibit strong learning abilities. (Matzel, *et al.* 2008) This was evident in this experiment as well, as trained mice stood still longer compared to the two other groups, and 7 out of 8 trained mice reached at least level 6 by session 19. However, considering the knowledge on cognitive decline in aging animals, employing younger mice might require fewer sessions, enhancing the practical implementability of the protocol.

While the elderly mice in this study demonstrated reasonable trainability, the absence of significant results between treatment groups in the stress-related behaviors assessed during the VAT may still be linked to their age. In a study by Oh, *et al.* (2018) age-related changes in stress responsiveness and coping strategies were explored. The findings indicate an increase in basal serum corticosterone levels, increased sensitivity to stress and a less effective response to stress in older mice (12 and 23 months) compared to younger mice (2 months). Furthermore, research conducted by Lee, *et al.* (2020) delved into stress-induced depression, revealing the manifestation of stress-induced depressive behavior (reduced sociability and locomotion) in aged mice (18 months), absent in young mice (2 months). Considering these findings, it is possible that the age of the mice influenced the observed results of the behaviors scored during the VAT. While the current information is limited and makes predicting the age-related change of the mice's behavior difficult in the context of this experiment, it is essential to acknowledge. More research on clicker training mice, specifically younger mice, is warranted to know if the lack of significance could be partially subjected to the relatively old cohort.

Exclusive use of female mice raises the question of whether employing mixed or solely male mice would have produced different results, specifically regarding trainability and the number of sessions necessary to complete training. Scientific research on sex differences related to memory and learning in rodents suggests that outcomes can vary based on the specific behavioral or cognitive task being assessed (Zorzo *et al.*, 2023; Mifflin, *et al.*, 2021). In some studies, researchers have observed variations in the learning and memory abilities of male and female rodents (Safari, *et al.*, 2021; Duarte-Guterman, *et al.*, 2015). However, the direction and magnitude of these differences varies among studies, and there isn't a universal conclusion that one sex is consistently better trainable than the other across all tasks. Moreover, other studies such as Matzel, *et al.* (2008) and Tsao, *et al.* (2023) used both sexes of mice in their experiments, and no effects of sex were observed in any of the tasks. Considering the complexity of the effect of sex on learning and memory, it is hard to say whether this factor affected the results of the current experiment.

Apart from limitations related to the animals, the study also encompasses constraints concerning the training protocol. As outlined in the "Materials and methods" section, seven sessions were employed as a linking phase, aimed at teaching the mice that a click signifies the forthcoming receipt of a

reward. However, the anticipation was that a clear conclusion of the linking phase would be reached, such as observable indications from the mice, like raising their noses to sniff or looking around apprehensively after a click looking for a reward. This behavior, however, was observed in only a few mice and inconsistently, raising uncertainty about whether all trained mice understood that a click signified the subsequent receipt of a treat. Therefore, the question arises whether they were truly clicker trained or perhaps only conventionally trained.

Addressing this concern, current research in various animal species does not support the notion that the efficacy of clicker training is decisively superior compared to conventional training to learn new behaviors (Chiandetti, *et al*. 2016; Smith & Davis, 2007; Williams *et al*. 2004). Given this, it is not anticipated that potential misuse of the clicker in the first few sessions or a lack of successful linking during the linking phase significantly influenced the study results. Especially since, in a recent study by Swan *et al*. (2023), conventional training has also been shown to reduce distress during injection, by assessing mice's facial expressions. However, there have been animal studies, like Mählis, *et al*. (2023), Verdino (2021) and Jønholt, *et al.* (2021), demonstrating stress-reducing effects of clicker training specifically, as well. Furthermore, studies conducted by Leidinger, *et al.* (2017) and van Eldik, *et al.* (2021) involved clicker training mice, with significant results and trends, respectively, in stress reduction and therefore refinement. This established precedent supports the rationale for the continued use of clicker training in the current study. However, it is important to recognize that clicker training presents several challenges compared to conventional training methods. Clicker training often requires more time, particularly due to the necessity of establishing an association between the click and a reward, with uncertainty surrounding the moment this connection is solidified. Moreover, precise timing of the click is essential for effective conditioning and can be challenging, especially in agile animals like mice. Further investigation is warranted to determine the added value of incorporating a clicker, for both SC injection training specifically and for animal training in general.

Other examples of limitations of the current study are regarding the behavior scoring in BORIS, specifically scoring of the 'struggling' and 'out of frame' behaviors warrant discussion. As described and specified in the ethogram (Table 3), struggling was scored any moment the mouse was not standing still. This includes actions such as walking or running away from the researcher's hand, flinching and/or attempting to escape when trying to lift a skin fold. This definition would closely align with how other research articles typically define struggling. However, behaviors like lifting more than one paw to for example take one small step, or turning more than 90 degrees to for example sniff the researcher's hand, were also categorized as not standing still, so in this case, as 'struggling'. In the case of the last examples, the mice were not 'standing still', but the injection could almost always still be administered safely. The absence of this distinction in scoring was part of the reason for the statistical analysis focusing solely on 'standing still.' For future studies, a more specific analysis of mice exhibiting behaviors other than 'standing still' during injection could be of interest.

Regarding the "out of frame" measurement, it became evident that there was considerable variability in the duration the mice were out of frame and the duration it took for the researcher to return the mouse to the center of the training pad while scoring in BORIS. To correct for this issue, the percentage of behaviors was calculated relative to the time spent on the training pad, as previously mentioned. For future reference, it would be better to demarcate the camera frame boundaries to prevent mice from leaving camera view at all. Additionally, when mice would leave the camera frame by climbing the researcher's arm, this was also scored as "out of frame". Given the fluctuating duration and the invisibility of the exact moment when a colleague would remove the mouse from the researcher, this period was classified as 'out of frame' rather than being considered part of the 'time spent on researcher'. However, when mice were on the researcher, only the behavior 'time spent on researcher' would be scored (see Ethogram). This means it may not have been necessary to remove the mice from the researcher when they left the camera frame while on the researcher, which could be considered for future research. In general minimizing mouse handling during the VAT is

recommended, as some mice appeared a little shocked after being lifted and placed back to the center of the training pad. Some mice would exhibit a variable period of stillness or less explorative behavior. This could have potentially influenced the results of the mice that went out of camera frame by climbing the researcher's arm.

Another aspect, regarding the mice that had to be restrained during injection, must also be addressed. On the injection day, two mice had to be restrained (X19Y1=unexposed and X19Y3=exposure) by scruffing because they struggled too much during the one minute time frame, making injecting unsafe. As handling and restraining can cause stress (Meijer *et al.*, 2006) it is possible the restraining affected the results of the reward acceptance test and the VAT. However, in Gouveia & Hurst (2019) scruffing seemed to have no influence on mice behavior, specifically interaction with the handler, after subcutaneous injection. Moreover, because it was only the case for two mice and they were from different treatment groups it is likely that if it did affect the results, this effect was very minimal. This was confirmed by the analysis, as the need for restraint was included in the initial statistical model as a random effect, but did not change the results.

As evident from the results and the discussion, depending solely on behavioral observation to assess stress and/or anxiety levels in mice is challenging, partly due to a lack of preceding research on the topic. To attain a more thorough assessment of the mice's stress levels during and following a SC injection, future studies could incorporate additional stress assessment methods like measuring physiological parameters to enhance the validity and reliability of research findings. These might include non-invasive approaches like measuring fecal corticosterone metabolites (Touma *et al.*, 2004; Abelson *et al.*, 2016; Rowland & Toth, 2019) or employing infrared thermography (Gjendal, *et al.*, 2018; Blenkuš*, et al.*, 2022).

Lastly, addressing time efficiency is imperative to ensure the practical implementability of the protocol, which was one of the main aims of the study. Although specific expectations regarding the number of required sessions were not established beforehand, the observed count appeared notably high. As addressed previously, using younger mice in a subsequent batch could potentially reduce the required number of sessions to get the trained mice injection-ready. However, this is not the sole factor contributing to the increased amount of sessions required. Firstly, the linking phase may have extended over an excessive number of sessions, as certain anticipative behaviors did not manifest. Secondly, due to the researcher's inexperience in training mice and the lack of scientific precedent for guidance, the researcher may have been too apprehensive or strict in progressing the mice to the next level in the earlier training sessions. As more training sessions went on, the researcher's increased confidence in discerning the (individual) behavior of the mice and their comprehension facilitated a quicker progression through levels. Furthermore, a discernible trend emerged wherein mice progressed rapidly through the training levels once they consistently consumed the reward (see Table A2 in the Appendix for more details), suggesting that the most challenging aspect of training may lay in acclimating to accepting a reward in an unfamiliar environment. This is likely associated with the mice's neophobia, gradually decreasing as they became more acquainted with the reward, environment and training protocol. (Kronenberger & Médoni, 1985; Misslin & Cigrang, 1986). Additionally, the mice achieving level 8 accomplished this prior to session 19, with four mice achieving it in session 14 and one mouse in session 17. This implies that for these mice, the injection day could have taken place after less than 19 sessions. However, due to scheduling constraints, only one injection day was feasible. Nevertheless, this does suggest that for future research, if mice are injected immediately upon reaching level 8, fewer sessions may be sufficient.

In the context of the current study, with the necessity of doing the injections of all mice on the same day, the choice was made to postpone the assessment to observe whether a greater number of mice would achieve level 8, given the sample size. After session 19 the decision was made to stop training (partly due to time constraints) and proceed to the injection day, as 7 out of 8 mice had reached level

6 at least, meaning they were at least familiar with the researcher lifting a skin fold. In the end, the decision was made not to exclude these two mice from analysis, again considering the small sample size. This could have affected the results, although its impact cannot be definitively determined.

Although not analyzed nor further discussed in this study, training and exposing the mice did seem to affect their behaviors during teaching sessions post-injection, as reported by the teachers using the animals (subjectively). In comparison with their behavior before they were trained or exposed, mice seemed to exhibit increased tameness and ease of handling with the tunnel. Additionally, the majority accepted food from unfamiliar students and they were more visible in their cage, standing in the front of it when approached, instead of hiding in their cardboard huts. As no analysis has been done on this observation, no definitive conclusions can be made. Nevertheless, the observation is interesting to mark and might be of further interest to explore in future research.

In conclusion, although the study offers valuable insights into the efficacy of clicker training as a refinement technique for subcutaneous injections with minimal restraint in laboratory mice, further research is essential. In terms of trainability, trained mice stood still longer during injection and a (non-significant) decrease in the mean total injection time from unexposed to trained mice was noticeable. This implies that mice can be effectively clicker trained to stand still, allowing for SC injections with minimal restraint and subsequently reducing stress during the injection. Moreover, as trained mice stood still longer, their reduced struggling suggests lower stress levels during the injection process. In terms of practical applicability, it is expected that, with an experienced trainer, a more optimized protocol and potentially younger mice, fewer than 19 sessions will be required to have trained mice reach the last level of the protocol in the following batch(es). Remarkably, exposed mice exhibited a significantly longer duration of 'accepting the reward' compared to the other two groups, suggesting that exposure alone may play a substantial role in stress reduction. Considering the previously mentioned behaviors, but more importantly regarding the remaining behaviors that lacked statistical significance, addressing identified limitations and incorporating additional observations in future batches and studies is necessary for a more comprehensive understanding of the efficacy of clicker training as a refinement method for SC injections. Nonetheless, this study marks an initial stride in the right direction aimed at reducing stress in laboratory mice and increasing their well-deserved welfare.

Acknowledgements

References

Abelson, K.S.P. Kalliokoski, O. Teilmann, A.C. Hau, J. Applicability of Commercially Available ELISA Kits for the Quantification of Faecal Immunoreactive Corticosterone Metabolites in Mice. In Vivo. 2016;30(6):739-744.

Ahumada, L.H. Morato, S. Lamprea, M.R. Acute stress increases behaviors that optimize safety and decreases the exploration of aversive areas. Learning and Motivation. 2022;80.

Arndt, S.S. Goerlich, V.C. van der Staay, F.J. A dynamic concept of animal welfare: The role of appetitive and adverse internal and external factors and the animal's ability to adapt to them. Frontiers in Animal Science. 2022;3:1-21.

Bates, D. Mächler, M. Bolker, B. Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 2015:67(1);1-48.

Blenkuš, U. Géros, A.F. Carpinteiro, C. Castro Aguiar de, P. Olsson, I.A.S. Franco, N.H. Non-Invasive Assessment of Mild Stress-Induced Hyperthermia by Infrared Thermography in Laboratory Mice. Animals (Basel). 2022;12(2):177.

Broom, D.M. A history of animal welfare science. Acta Biotheoretica. 2011:59;121-137

Broom, D.M. Policy Department C: Citizens' Rights And Constitutional Affairs. Animal Welfare in the European Union. 2017. Retrieved from: https://www.europarl.europa.eu/RegData/etudes/STUD/2017/583114/IPOL_STU(2017)583114_EN.pdf

Browne, C.M. The Effect of Delayed Positive Reinforcement on Learning in Dogs. Thesis, Doctor of Philosophy (PhD). Retrieved from: https://hdl.handle.net/10289/9808

Bryda, E.C. The Mighty Mouse: The Impact of Rodents on Advances in Biomedical Research. Missouri Medicine. 2013:110(3);207-211.

Chiandetti, C. Avella, S. Fongaro, E. Cerri, F. Can clicker training facilitate conditioning in dogs? Applied Animal Behaviour Science. 2016;184:109-116.

Davies, J.R. Purawijaya, D.A. Bartlett, J.M. Robinson, E.S.J. Impact of Refinements to Handling and Restraint Methods in Mice. Animals (Basel). 2022:12(17);2173.

Dhabar, F.S. The short-term stress response – Mother nature's mechanism for enhancing protection and performance under conditions of threat, challenge, and opportunity. Frontiers in Neuroendocrinology. 2018:49;175-192.

Duarte-Guterman, P. Yagi, S. Chow, C. Galea, L.A.M. Hippocampal learning, memory, and neurogenesis: Effects of sex and estrogens across the lifespan in adults. Hormones and Behavior. 2015;74:37-52.

Eilam, D. Die hard: A blend of freezing and fleeing as a dynamic defense—implications for the control of defensive behavior. Neuroscience & Biobehavioral Reviews. 2005;29(1):1181-1191.

Elias, P.K. Elias, M.F. Effects of age on learning ability: Contributions from the animal literature. Experimental Aging Research. 1976;2(2):164-168.

Faul, F. Erdfelder, E. Lang, A.G. Buchner, A-G. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 2007:39(2);175-191.

Francois, M. Delgado, I.C. Shardorodsky, N. Leu, C. Zeltser, L. Assessing the effects of stress on feeding behaviors in laboratory mice. eLife. 2022;11:e70271.

Friard, O. Gamba, M. BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. Methods in Ecology and Evolution. 2016:7(11);1325-1330.

Gjendal, K. Franco, N.H. Ottesen, J.L. Sørensen, D.B. Olsson, I.A.S. Eye, body or tail? Thermography as a measure of stress in mice. Physiology & Behavior. 2018;196:135-143.

Gouveia, K. Hurst, J.L. Reducing Mouse Anxiety during Handling: Effect of Experience with Handling Tunnels. PLOS ONE. 2013;8(6):e66401.

Gouveia, K. Hurst, J.L. Optimising reliability of mouse performance in behavioural testing: The major role of non-aversive handling. Scientific reports. 2017:7(1);44999.

Gouveia, K. Hurst, J.L. Improving the practicality of using non-aversive handling methods to reduce background stress and anxiety in laboratory mice. Scientific Reports. 2019;9:20305.

Halekoh, U. Højsgaard, S. A Kenward-Roger Approximation and Parametric Bootstrap Methods for Tests in Linear Mixed Models – The R Package pbkrtest. Journal of Statistical Software, 2014:59(9);1-30.

Harada, C.N. Natelson Love, M.C. Triebel, K. Normal Cognitive Aging. Clinical Geriatric Medicine. 2013;(29(4);737-752.

Haykin, H. Rolls, A. The neuroimmune response during stress- a physiological perspective. Immunity (manuscript). 2021:54(9);1933-1947.

Head, E. Mehta, R. Hartley, J. Kameka, M. Cummings, B.J. Cotman, C.W. Ruehl, W.W. Milgram, N.W. Spatial learning and memory as a function of age in the dog. Behavioral Neuroscience. 1995;109(5):851-858.

Heinz, D.E. Schöttle, V.A. Nemcova, P. Binder, F.P. Ebert, T. Domschke, K. Wotjak, C.T. Exploratory drive, fear, and anxiety are dissociable and independent components in foraging mice. Translational Psychiatry. 2021;11:318.

Jønholt, L. Bundgaard, C.J. Carlsen, M. Sørensen, D.B. A Case Study on the Behavioural Effect of Positive Reinforcement Training in a Novel Task Participation Test in Göttingen Mini Pigs. Animals Basel). 2021;11(6):1610.

Kalueff, A.V. Stewart, A.M. Song, C. Berridge, K.C. Graybiel, A.M. Fentress, J.C. Neurobiology of rodent self-grooming and its value for translational neuroscience. Nature REviews Neuroscience. 2025;17:45-59.

Ketchesin K.D. Stinnett G.S. Seasholtz A.F. Corticotropin-releasing hormone-binding protein and stress: from invertebrates to humans. Stress. 2017:20(5);449-464.

Kramer, K.van Acker, S.A.B.E. Voss, H.P. Grimbergen, J.A. van der Vijgh, W.J.F. Bast, A. Use of telemetry to record electrocardiogram and heart rate in freely moving mice. Journal of Pharmacological and Toxicological Methods. 1993:30;209–215.

Kramer, K. van de Weerd, H. Mulder, A. van Heijningen, C. Baumans, V. Remie, R. Voss, H.P. van Zutphen, B.F.M. Effect of Conditioning on the Increase of Heart Rate and Body Temperature Provoked by Handling in the Mouse. Alternatives to Laboratory Animals. 2004;32(1A):177-181.

Kronenberger, J.P. Médoni, J. Food neophobia in wild and laboratory mice (Mus musculus domesticus). Behavioural Processes. 1985;11(1):53-59.

Kuznetsova, A. Brockhoff, P.B. Christensen, R.H.B. lmerTest Package: Tests in Linear Mixed Effects Models. Journal of Statistical Software, 2017:82(13);1-26.

Lee, J-E. Kwon, H-J. Choi, J. Seo, J-S. Han, P-L. Aging increases vulnerability to stress-induced depression via upregulation of NADPH oxidase in mice. Communications Biology. 2020;3:292.

Leidinger, C. Herrmann, F. Thöne-Reineke, C. Baumgart, N. Baumgart, J. Introducing clicker training as a cognitive enrichment for laboratory mice. Journal of Visualized Experiments. 2017(121):55415.

Lezak, K.R. Missig, G. Carlezon Jr. W.A. Behavioral methods to study anxiety in rodents. Dialogues in Clinical Neuroscience. 2022:19(2);181-191.

Matzel, L.D. Grossman, H. Light, K. Townsend, D. Kolata, S. Age-related declines in general cognitive abilities of Balb/C mice are associated with disparities in working memory, body weight, and general activity. Learning & Memory. 2008;15(10):733-746.

Meijer, M.K. Spruijt, B.M. Zutphen, L.F.M. Baumans, V. Effects of restraint and injection methods on heart rate and body temperature. Laboratory Animals. 2006:40;382-391

Mifflin, M.A. Winslow, W. Surendra, L. Tallino, S. Vural, A. Velazquez, R. Sex differences in the IntelliCage and the Morris water maze in the APP/PS1 mouse model of amyloidosis. Neurobiology of Aging. 2021;101:130-140.

Misslin, R. Cigrang, M. Does neophobia necessarily imply fear or anxiety? Behavioural Processes. 1986;12(1):45-50.

Morgan, K.N. Tromborg, C.T. Sources of stress in captivity. Applied Animal Behaviour Science. 2006;102(3):262-302

Mu, M. Geng, H. Rong, K. Peng, R. Wang, S. Geng, L. Qian, Z. Yung, W. Ke, Y. A limbic circuitry involved in emotional stress-induced grooming. Nature Communications. 2020;11:2261.

Neely, C. Lane, C. Torres, J. Flinn, J. The effect of gentle handling on depressive-like behavior in adult male mice: Considerations for human and rodent interactions in the laboratory. Behavioural neurology. 2018:2;1-7.

Nyberg, L. Boraxbekk, C.J. Sörman, D.E. Hansson, P. Herlitz, A. Kauppi, K. Ljunberg, J.K. Lövheim, H. Lundquist, A. Adolfsson, A.N. Oudin, A. Pudas, S. Rönnlund, M. Stiernstedt, M. Sundström, A. Adolfsson, R. Biological and environmental predictors of heterogeneity in neurocognitive ageing: Evidence from Betula and other longitudinal studies. Ageing Research Reviews. 2020:64;101184.

Oh, H-J. Song, M. Kim, Y.K. Bae, J.R. Cha, S-Y. Bae, J.Y. Kim, Y. You, M. Lee, Y. Shim, J. Maeng, S. Age-Related Decrease in Stress Responsiveness and Proactive Coping in Male Mice. Frontiers in Aging Neuroscience. 2018;10.

Ohl, F. Hellebrekers, L. "Dierenwelzijn" - de diergeneeskundige positie. Tijdschrift Voor Diergeneeskunde, 2009:134(18);754-755. https://dspace.library.uu.nl/bitstream/1874/202892/1/20090915%20Tijdschrift%20voor%20Diergeneeskunde.pdf

Parker, M.O. McElligott, A.G. Ethical evolutions: navigating the future of animal behaviour and welfare research. Sage Journals. 2023:19(4);369-372

Phifer-Rixey, M. Nachman, M.W. Insights into mammalian biology from the wild house mouse *Mus musculus*. eLife. 2015:4.

Qiao, Y. Zhao, J. Li, C. Zhang, M. Wei, L. Zhang, X. Kurskaya, O. Bi, H. Gao, T. Effect of combined chronic predictable and unpredictable stress on depression-like symptoms in mice. Annals of Translational Medicine. 2020;8(5):942

Rowland, N.E. Toth, L.A. Analytic and Interpretational Pitfalls to Measuring Fecal Corticosterone Metabolites in Laboratory Rats and Mice. Comparative Medicine. 2019;69(5):337-349.

Rusche, R. The 3Rs and animal welfare - conflict or the way forward? ALTEX. 2003:20(1);63-76.

Russel, W.M.S. Burch, R.L. The Principles of Humane Experimental Technique. Universities Federation for Animal Welfare (UFAW). 1992, Special Edition.

Safari, S. Ahmadi, N. Mohammadkhani, R. Ghahremani, R. Khajvand-Abedeni, M. Shahidi, S. Komaki, A. Salehi, I. Karimi, S.A. Sex differences in spatial learning and memory and hippocampal long-term potentiation at perforant pathway-dentate gyrus (PP-DG) synapses in Wistar rats. Behavioral and brain functions. 2021;17(1):9.

Smith, S.M. Davis, E.S. Clicker increases resistance to extinction but does not decrease training time of a simple operant task in domestic dogs (Canis familiaris). Applied Animal Behavior Science. 2007;110(3-4):318-329.

Song, C. Berridge, K.C. Kalueff, A.V. 'Stressing' rodent self-grooming for neuroscience research. Nature Reviews Neuroscience. 2016;17:591.

Stuart, S.A. Robinson, E.S.J. Reducing the stress of drug administration: implications for the 3Rs. Scientific Reports. 2015;5:14288.

Swan, J. Boyer, S. Westlund, K. Bengtsson, C. Nordahl, G. Törnqvist, E. Decreased levels of discomfort in repeatedly handled mice during experimental procedures, assessed by facial expressions. Frontiers in Behavioral Neuroscience. 2023;17:1109886.

Touma, C. Palme, R. Sachser, N. Analyzing corticosterone metabolites in fecal samples of mice: a noninvasive technique to monitor stress hormones. Hormones and Behavior. 2004;45(1):10-22.

Tsao, C. Wu, K. Su, N.C. Edwards, A. Huang, G. The influence of sex difference on behavior and adult hippocampal neurogenesis in C57BL/6 mice. Scientific Reports. 2023;13:17297.

Van Eldik, K.C. Better welfare for mice through handling and training. Studentthesis. 2021.

Van Zutphen, L.F.M.B. Use of animals in research: a science - society controversy? The European perspective. ALTEX. 2002:19(3);140-144.

Verdino, S. The Effects of Clicker Training on Stress Levels in Communally Housed Shelter Cats. Biological Sciences. 2021:3-8.

Walker, D.L. Toufexis, D.J. Davis, M. Role of the bed nucleus of the stria terminalis versus the amygdala in fear, stress, and anxiety. European Journal of Pharmacology. 2003;463(1-3):199-216.

Wallis, L.J. Virányi, Z. Müller, C.A. Serisier, S. Huer, L. Range, Friederike. Aging effects on discrimination learning, logical reasoning and memory in pet dogs. Age. 2016;38(1):6.

Wet dieren (2022). Geraadpleegd van https://wetten.overheid.nl/BWBR0030250/2022-12-22 op 01-01-2024.

Wet op de dierproeven (WOD). Article 1d, paragraph 3. Consulted via https://wetten.overheid.nl/BWBR0003081/2021-07-01 on 06-06-2023.

Williams, J.L. Friend, T.H. Nevill, C.H. Archer, G. The efficacy of a secondary reinforcer (clicker) during acquisition and extinction of an operant task in horses. Applied Animal Behaviour Science. 2004;88(3-4):331-341.

Yamada, C. Iizuka, S. Nahata, M. Hattori, T. Takeda, H. Vulnerability to psychological stress-induced anorexia in female mice depends on blockade of ghrelin signal in nucleus tractus solitarius. British Journal of Pharmacology. 2020;177(20):4666-4682.

Zorzo, C. Arias, J.L. Méndez, M. Are there sex differences in spatial reference memory in the Morris water maze? A large-sample experimental study. Learning & Behavior. 2023.

Appendix

Session	Date	Mouse	Group	Level	Level completed?	Notes
		X11Y2	Treatment			
		X11Y3	Treatment			
		X19Y2	Treatment			
		X19Y5	Treatment			
		X23Y1	Treatment			
		X23Y2	Treatment			
		X25Y2	Treatment			
		X28Y4	Treatment			
		X11Y1	Exposure			
		X19Y3	Exposure			
		X19Y4	Exposure			
		X23Y4	Exposure			
		X25Y1	Exposure			
		X25Y4	Exposure			
		X28Y1	Exposure			

Table A1: Datasheet used during the training/exposure sessions. The 'Level' and 'Level completed?' rows are marked gray for the exposed control mice, as they are not trained.

	X11Y2	X11Y3	X19Y2	X19Y5	X23Y1	X23Y2	X25Y2	X28Y4
Session 1	0	0	0	0	0	0	0	0
Session 2	0	0	0	0	0	0	0	0
Session 3	0	0	0	0	0	0	0	0
Session 4	0	0	0	0	0	0	0	0
Session 5	0	0	0	0	0	0	0	0
Session 6	0	0	0	0	0	0	0	0
Session 7	0	0	0	0	0	0	0	0
Session 8	1	1	0	0	1	1	1	0
Session 9	1	5	0	1	1	2	5	1
Session 10	4	5	0	1	1	2	6	1
Session 11	5	5	0	1	1	6	6	1
Session 12	6	6	0	1	2	6	7	1
Session 13	7	7	0	1	2	7	7	1
Session 14	8	8	0	1	3	8	8	2
Session 15	8	8	0	1	6	8	8	2
Session 16	8	8	0	1	7	8	8	2
Session 17	8	8	0	2	8	8	8	2
Session 18	8	8	0	2	8	8	8	2
Session 19	8	8	0	6	8	8	8	6

Table A2: This table shows the progress (in levels) of each individual mouse in the training group through all 19 sessions. The cell is marked green if that particular mouse accepted the reward from the researcher at least twice during the session (taking/eating the reward from the ground did not count). The level depicted in the cell means this was the mouse's highest level completed during that session, e.g. mouse X11Y2 completed level 1 in session 8 and was trained for level 2 in session 9 but did not complete it (yet). Sessions 1-7 served as linking sessions, designed to associate the sound of the clicker with the reward. During these sessions, the mice could not advance to level 1 yet even if they already accepted the reward, hence the 0.

	X11Y1	X19Y3	X19Y4	X23Y4	X25Y1	X25Y4	X28Y1	X28Y3
Session 1								✓ (32)
Session 2								✓
Session 3	✓						✓	✓
Session 4					✓		✓	
Session 5				✓	✓	✓	✓	✓
Session 6		✓	✓		✓		✓	✓
Session 7	✓	✓	✓		✓	✓	✓	✓
Session 8	✓	✓	✓		✓	✓	✓	✓
Session 9	✓	✓	✓	✓	✓	✓	✓	✓
Session 10	✓	✓	✓		✓	✓	✓	✓
Session 11	✓	✓	✓	✓	✓	✓	✓	✓
Session 12	✓	✓	✓		✓	✓	✓	✓
Session 13	✓	✓	✓	✓	✓	✓	✓	✓
Session 14	✓	✓		✓	✓	✓	✓	✓
Session 15	✓	✓	✓		✓	✓	✓	✓
Session 16	✓	✓	✓		✓	✓	✓	✓
Session 17	✓	✓	✓			✓	✓	✓
Session 18	✓	✓	✓	✓	✓	✓	✓	✓
Session 19	✓	✓	✓	✓	✓	✓	✓	✓

Table A3: This table shows the reward acceptance of all exposed control mice through all 19 sessions. The cell is marked green if that particular mouse accepted the reward from the researcher at least twice during the session.

	START	STOP
DURING THE INJECTION		
Time it takes to inject: The time it took to safely inject SC from the first attempt at lifting a skin fold to letting the mouse go after injection.	The first frame the mouse is visibly touched by the researcher at the first attempt at lifting a skin fold while simultaneously having the syringe in the opposite hand, ready to inject.	The first frame in which the skin fold is visibly let go by the researcher after the mouse received the full injection, even if they had to be restrained by the project supervisor.
Standing still during injection: Behaviors that counted as standing still were when the only movement of the mouse at the time was breathing and sniffing. Moving was allowed to a certain point, like taking a small side step with just one front or rear paw but making less than a 90 degree turn.	The first frame the skin fold was lifted (if the mouse was sitting still) or when a lifted paw would be placed back on the mat (if the mouse was struggling prior). If the mouse turned more than 90 degrees and would sit still in that bent position for more than 5 frames standing still was scored from that frame on.	The first frame a paw would lift off the mat (if the mouse would walk away from the researcher) or when the mouse would start turning (more than) 90 degrees or if the mouse would flinch.
Struggling during injection: Any behavior that was not counted as "standing still". Behaviors that counted as struggling were: walking during injection/lifting of a skin fold, and when the mouse would move in a certain way that made injecting unsafe: e.g. turning more than 90 degrees, rearing and flinching upon lifting the skin fold.	The first frame the mouse would start flinching, rearing or turning (more than 90 degrees). In any of these cases a paw(s) is lifted off the mat, which was when the scoring started.	When the mouse would sit still again, so the first frame all four paws would be on the mat.
DURING THE REWARD ACCEPTANCE TEST (1 MINUTE)		
Accepting the reward: The behavior was scored the entire time the mouse was chewing the reward.* The behavior was also scored when the mouse would be angled with its rear end to the camera (making it very hard to see their jaw move) if the mouse's face was directly in front of the reward while making small movements of their head. *The behavior was also scored during the VAT on the training pad if the mouse kept or was chewing when the 2 minute period of the VAT had begun.	The frame in which the mouse's jaw was open at its widest during the first bite. If the mouse would start chewing on the mat without her taking a bite immediately prior (because e.g. she was sniffing but still had some reward left in her mouth), accepting the reward would start the first frame her jaw would open again to chew.	The first frame the mouse stopped chewing, so when their mouth was entirely closed. If the mouse would turn her head away from the camera, making it unable to observe if she was chewing, there were two options.. 1) The counting of the behavior continued if the mouse was chewing before she turned her head and was still chewing when her head was turned back to the camera. 2) If the mouse would be chewing, then turn her head away from the camera and would have stopped chewing when she turned her head back to the camera, the behavior was stopped at the first frame when her head was fully turned away.
DURING THE VAT ON THE TRAINING PAD (2 MINUTES)		
Touching the researcher: Behaviors that counted as touching the researcher were: touching the researcher with one, two or three paws at a time. Sitting by the researcher with their body against the hand of the researcher was also counted. Behaviors that did not count as touching the researcher were: touching or grazing the researcher with their tail or when it was not completely visible or debatable if the mouse was touching the researcher's hand with their body. When the mouse would be on the researcher entirely (with her whole body and four paws) the behavior would be scored as spending time on the researcher.	The first frame in which one of the mouse's paws would be on the researcher's hand, arm or lab coat. In case of sitting against the researcher with their body, the behavior scoring started when there would be no visible space left between the mouse's body and the researcher's hand or finger. If the mouse was entirely on the researcher and would leap off (meaning, lifting one of their paws to leave the researcher's hand), 'touching the researcher' was started the same frame 'spending time on the researcher' was stopped. If there was any doubt if the	The first frame the last part of the mouse's body (usually a paw) would stop touching the researcher. The scoring of the touching behavior would also be stopped when a mouse would climb the researcher's hand to sit there. For example: The mouse is touching the researcher with one paw, then two, then three and then decides to spend time on the researcher so the mouse puts their last paw on the researcher as well. During the aforementioned time, touching would be scored but stopped when the fourth paw of the mouse would start touching the researcher. This

	mouse was touching the researcher (e.g. bad visibility because of the angle), the behavior was not scored.	means the touching the researcher behavior is stopped the same frame as the spending time on the researcher behavior is started.
Spending time on the researcher: Behaviors that counted as spending time on the researcher were when the mouse would be sitting on the researcher, including all their four paws. When a mouse would spend time on the researcher, no other behaviors would be counted simultaneously. E.g. when a mouse would sniff the researcher or do a stretch attend or reared while on the researcher's hand or arm, only spending time on the researcher would be scored.	The first frame all four paws are on the researcher.	The first frame one of the mouse's paws is lifted from the researcher's hand. Stopping of this behavior is only scored when the intention of the mouse is to leave the hand of the researcher. If the mouse is walking on the hand or arm of the researcher, she is also lifting paws, but not with the intention of leaving the researcher, which means scoring of this behavior is continued.
Sniffing the researcher: When the mouse would sniff any part of the researcher's hand.	When the mouse's head was angled toward the researcher and her nose was in close proximity to / (almost) touching the researcher's hand/finger/arm/lab coat.	When the mouse would start angling her head a different way and her nose would leave the vicinity of the researcher or e.g. when she would climb on the researcher, as that behavior is not scored concurrently with other behaviors.
Sitting next to the researcher: Behaviors that counted were when the mouse would sit within half a mouse's body length of the researcher's hand. Only sitting was allowed, meaning when the mouse would do a stretch attend or would groom within half a body length of the researcher's hand only the aforementioned behaviors would be scored and sitting next to the researcher would be stopped.	The first frame the mouse was sitting next to the researcher within half a body length with all her four paws on the mat.	The first frame a paw would be lifted off the mat to take a step. If the mouse would lift a front paw to place said paw back in the same place without taking a step, the scoring was not stopped. If the mouse would sit next to the researcher and then would exhibit another behavior, like stretch attend, the sitting next to the researcher behavior scoring would stop at the same frame the, in this case stretch attend, would start.
Biting/nibbling the researcher: When the mouse would bite/nibble the hand, finger or arm of the researcher. > Only 1 mouse (unexposed control) displayed this behavior (0,6 seconds).	The frame in which the jaw of the mouse would be at its widest.	The first frame in which the mouse would let go of the researcher.
Stretch attend: Behaviors that counted as stretch attend were when the mouse would move her front paws forward and leaned in, when she would move her hind paws backward stretching her whole body, or when she would not move her front paws but would clearly lean forward, stretching her hind paws. If all paws would stay on the mat and a mouse would sniff the air, stretching only her neck, stretch attend would not be scored.	The frame in which the first paw would leave the mat. If the mouse would only lean in, without lifting her paws, the first frame in which she leaned forward was scored as start.	The frame in which the mouse returned to a normal sitting/standing position (meaning, unstretched). If the mouse would start walking during a stretch attend the scoring would stop when a hind paw would lift off the mat (and at the same time the stretched walk behavior would start).
Stretched walk: The behavior that counted as stretched walk was when a mouse completed a stretch attend and would walk with their back legs stretched.	The frame in which the first paw would lift off the mat after a stretch.	The frame where the mouse would have all four paws on the mat and stopped walking or if the mouse would change to another behavior. In the second case, this frame was the same frame the other behavior started.
Rear: When the mouse would sit on her hind paws without her front paws touching the mat.	The first frame the second front paw would lift off the mat.	The first frame the second front paw would touch the mat again.

Flinch > No mice exhibited this behavior	N/A	N/A
Freeze > No mice exhibited this behavior	N/A	N/A
Sitting elsewhere on the mat: When the mouse would sit/stand still on the mat further than half a body length away from the researcher's hand. Moving of the head (up, left, right, sniffing) was allowed as well as lifting/replacing one front paw, any other movements were scored as different behaviors of this ethogram.	The first frame all four paws are on the mat and the mouse would stay in the same spot for more than 5 frames.	The first frame the mouse would lift a hind paw (which means walking) or when she would change behaviors e.g. start grooming.
Grooming: The mouse sits back on her hind paws and cleans her snout and/or ears with her front paws or licks/scratches any part of her body. > 8 mice exhibited this behavior (2 trained, 2 exposed control, 4 unexposed control), with an average of 2,3 seconds.	The first frame the front paws would lift off the mat or her head would turn to lick/bite her side.	The first frame the mouse stopped cleaning, licking, biting or scratching and her front paws would return back to the mat. Usually both paws would still be in the air for a short moment when the mouse stopped grooming: this was not scored as rear.
Looking over the edge: When the mouse would look over the edge of the table. This behavior was not counted simultaneously with other behaviors like stretch attend. > 9 mice exhibited this behavior (3 trained and 6 unexposed), with an average of 4,7 seconds.	The first frame the mouse's nose would disappear from view from behind the table's edge. If the mouse would look over the edge during a stretch attend, the stretch attend was not scored during the time the looking of the edge behavior was scored.	The first frame the mouse's nose would come back into view from behind the table's edge.
Out of frame: When the mouse would leave the camera frame by either escaping the mat via the table or climbing up the researcher's arm leaving the frame. *This behavior was not displayed by the mice; in contrast, no behaviors could be assessed during this period. The inclusion of "out of frame" in the Ethogram indicates the frames during which the activity was initiated and concluded for scoring, providing clarity for the reader.	The first camera frame the mouse is more than one body length (excl. tail) away from the left or right side of the mat. When the mouse would walk out of frame via the back border of the mat (the border closest to the camera) the out of frame behavior started when the behavior the mouse was exhibiting prior was not visible anymore. E.g. when a mouse would do a stretch attend onto the area that is out of camera frame (so hind paws in frame, front paws out of frame), a stretch attend would still be scored. Though, if the mouse would be sitting still and in frame with only her hind paws, out of frame would be scored as it is not visible if the mouse was exhibiting the sitting elsewhere on the mat or the grooming behavior for example.	The first frame the mouse comes back into the camera frame and another behavior is visible and can be scored or if she walks back into frame. For example: When a mouse would come back into frame with only her head visible for more than one second, out of frame would still be scored as it is not visible if the mouse was exhibiting either the stretch attend or sitting elsewhere on the mat behavior. The scoring of the out of frame behavior was in these cases stopped when the hind paws of the mouse became visible again as well. If she were walking back into frame, the behavior was stopped the first frame the mouse was back into camera frame.

Table A4: Ethogram of all behaviors scored during the injection and behavioral tests, including start and stop margins. Behaviors that have not been scored (because none of the mice exhibited the behavior) are denoted by NA. The cells of the behaviors are marked green if they have been statistically analyzed.