Aus dem Institut für Tierschutz, Tierverhalten und Versuchstierkunde
des Fachbereichs Veterinärmedizin
der Freien Universität Berlin

und

aus der Fachgruppe Versuchstierkunde
der Abteilung Experimentelle Toxikologie und ZEBET,
dem deutschen Zentrum zum Schutz von Versuchstieren
des Bundesinstitut für Risikobewertung

The use of an automated and home-cage based test system to improve behavioral experiments for group housed mice

Inaugural-Dissertation
zur Erlangung des Grades eines Doctor of Philosophy (PhD)
in Biomedical Sciences an der Freien Universität Berlin

vorgelegt von
Pia Kahnau
aus Münster

Berlin 2023
Journal-Nr.: 4396

mbvberlin | mensch und buch verlag

Aus dem Institut für Tierschutz, Tierverhalten und Versuchstierkunde
des Fachbereichs Veterinärmedizin
der Freien Universität Berlin

und

aus der Fachgruppe Versuchstierkunde
der Abteilung Experimentelle Toxikologie und ZEBET,
dem deutschen Zentrum zum Schutz von Versuchstieren
des Bundesinstitut für Risikobewertung

The use of an automated and home-cage based test system to improve behavioral experiments for group housed mice

Inaugural-Dissertation zur Erlangung des Grades eines Doctor of Philosophy (PhD)
in Biomedical Sciences an der Freien Universität Berlin

vorgelegt von
Pia Kahnau
aus Münster

Berlin 2023

Journal-Nr.: 4396

Gedruckt mit Genehmigung des Fachbereichs Veterinärmedizin

der Freien Universität Berlin

Dekan:	Univ.-Prof. Dr. Uwe Rösler
Erster Gutachter:	Univ.-Prof. Dr. Lars Lewejohann
Zweiter Gutachter:	Uni.-Prov. Dr. Christa Thöne-Reineke
Dritter Gutachter:	PD Dr. Carola Fischer-Tenhagen

Deskriptoren (nach CAB-Thesaurus):
mice, laboratory animals, animal welfare, animal behaviour, cognition, radio frequency identification

Tag der Promotion: 03.05.2023

Bibliografische Information der *Deutschen Nationalbibliothek*
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

ISBN: 978-3-96729-224-4
Zugl.: Berlin, Freie Univ., Diss., 2023
Dissertation, Freie Universität Berlin
D188

 Choriner Str. 85 - 10119 Berlin
verlag@menschundbuch.de – www.menschundbuch.de

Contents

Contents

Abbreviations

RFID	Radio Frequency Identification
IC	IntelliCage
AG	AnimalGate
CD	Consumer Demand
CB	Cognitive Bias

Chapter 1

Introduction

In science, the principle of the 3 R's (replace, reduce, refine) by Russel and Burch (Russel & Burch 1992 (special edition) is established and required to be applied according to the EU Directive (2010/63/EU). The principle of the 3 R's describes that research should be conducted to replace animal experiments with alternative methods or to optimize existing methods to reduce the number of animals used. In 2020 in Germany, approximately 1.9 million animals were used for scientific research. Thereof, about 71% of the animals were mice (Bundesinstitut für Risikobewertung 2021). Given that it is reasonable to assume that existing animal experiments cannot yet be fully replaced in the near future, animals will continue to be used in scientific experiments. Therefore, it is imperative to refine the conditions of husbandry as well as the experimental conditions to reduce pain, suffering, and harm to the minimum. An important element of such improvements is to include the perspective of the animal. Therefore, methods have to be developed to minimize external influences, e.g. by the experimenter, as much as possible. One possibility to refine experimental conditions is to conduct animal experiments in an automated and home-cage based manner.

1.1 Advantages of home-cage based systems

The use of automated and home-cage based test systems offers many advantages. One obvious advantage is that the animals can be tested in their familiar environment. They do not have to be removed from their home-cage for the experiment and placed in a separate test device. This, in turn, leads to a decreased influence of the experimenter on the animals, since handling of the animals is reduced. Furthermore, the animals do not have to be actively separated from their social group. Both, handling and separation, have been shown to have a negative influence on animal welfare and thus a negative impact on scientific data (Manouze et al. 2019; Gouveia & Hurst 2013; Hurst & West 2010; Krohn et al. 2006). Reducing the influence of the experimenter and performing animal experiments in automated and home-cage based set-ups can also lead to the production of repeatable data since the observer bias is omitted (Voikar & Gaburro 2020; Krackow et al. 2010).

If the experiments are automated and home-cage based, the time required for the experimenter (to carry out the experiments him- or herself) is significantly decreased, since the permanent presence of an experimenter is not necessary. Furthermore, it is not necessary to

adjust the day/night rhythm of the animals to that of humans. Therefore, the animals can be tested in their active phase, which is the dark phase in laboratory mice.

Another advantage is, that the animals enter the test-cage, which is connected to the home-cage through a gate, in a self-determined manner. Therefore, a high motivation of the animals can be assumed, which in turn might have a positive effect on the data itself.

1.2 Potential applications for automated and home-cage based systems

So far, there are some possibilities to record experimental data automatically within the home-cage. Infrared systems, for example, were used to measure activity (e.g., Park et al. 2021; Ticher & Ashkenazi 1995) or video based systems were used to measure home-cage behavior (e.g., Jirkof et al. 2013; Miller et al. 2011; Steele et al. 2007). Transmitters were used to monitor heart rate or temperature over a longer period of time (e.g., Kramer et al. 2004; Späni et al. 2003). Radiofrequency identification (RFID) based systems were used to measure activity (Kahnau et al. 2021; Weegh et al. 2020) or to determine the position of individual animals within their home-cages (Mieske et al. 2021; Freund et al. 2013; Lewejohann et al. 2009) or for preference testing (Habedank et al. 2022; Hobbiesiefken et al. 2021). But also for severity assessment, home-cage based experiments can be carried out. Weegh and colleagues, for example, used an RFID based system to be able to conclude from voluntary wheel running to well-being. They observed group housed female mice of a colitis model and showed a relationship between reduced activity and reduced well-being (Weegh et al. 2020).

Another RFID-based test system is the IntelliCage (IC, TSE-Systems, Germany). The IC is an automated as well as home-cage based test system, in which learning behavior of mice can be investigated (e.g., Kahnau et al. 2021; Voikar et al. 2018; Endo et al. 2011; Krackow et al. 2010). The IC consists of four conditioning corners in which water can be granted or denied, allowing for variety of cognitively challenging tasks. The IC served as a home-cage based test-cage in the experiments presented here.

1.3 Aim of the dissertation

The initial aim of this dissertation was to develop and conduct automated and home-cage based experiments to evaluate the burden of commonly used behavioral tests from the animals' point of view. These findings should serve the researchers but also the authorities to better assess the burden of animal experiments and to keep the burden on the animals as low as possible. In order to include the animal's perspective, it is necessary to investigate not only physiological parameters such as heart rate or stress hormones, but especially the behavior

of the animals. Chapter 2 describes different behavioral methods that can be used for severity assessment.

Chapter 3 presents the use of the IC in a long-term study to investigate the influence of repeated home-cage based cognitive stimulation on physiological parameters and social structure of male mice. The study shows the feasibility of keeping male mice in groups for a long period of time while obtaining individual data.

However, individual mice were observed to push and pull each other out of the IC conditioning corners. Since this may influence the learning performance of the animals, additional compartments were added to the IC.

The next step - the extension of the IC for further home-cage based experiments - required a considerable amount of time. The technical requirements as well as experimental methods had to be developed and validated first, since the experiments presented here have not been performed automated and home-cage based before.

For these experiments, the IC was extended by a gate (AnimalGate, AG, TSE-Systems, Germany) and one more cage. The aim was to develop an automated and home-cage based Consumer Demand and a Cognitive Bias test for mice. The principle of the Consumer Demand test allows to investigate the strength of preferences as well as aversions to let the mice work for access to selected goods (chapter 4). With the Cognitive Bias test, we were able to investigate the emotional state of animals as this test investigates how decisions depend on expectations of future events (chapter 5). Both tests have the potential to assess the wants and needs as well as burden of housing conditions and/or behavioral tests.

Chapter 2

Literature Review

2.1 Publication 1

Behavioral methods for severity assessment

Pia Kahnau, Anne Habedank, Kai Diederich, Lars Lewejohann

Published in journal “Animals” 2020, 10(7), 1136
DOI: https://doi.org/10.3390/ani10071136

animals

MDPI

Review

Behavioral Methods for Severity Assessment

Pia Kahnau [1,*], Anne Habedank [1], Kai Diederich [1] and Lars Lewejohann [1,2]

1 German Federal Institute for Risk Assessment (BfR), German Center for the Protection of Laboratory Animals (Bf3R), 12277 Berlin, Germany; Anne.Habedank@bfr.bund.de (A.H.); Kai.Diederich@bfr.bund.de (K.D.); Lars.Lewejohann@bfr.bund.de (L.L.)

2 Institute of Animal Welfare, Animal Behavior and Laboratory Animal Science, Freie Universität Berlin, 14163 Berlin, Germany

* Correspondence: Pia.Kahnau@bfr.bund.de; Tel.: +49-30-18412-29202

Received: 29 May 2020; Accepted: 30 June 2020; Published: 3 July 2020

Simple Summary: In 2017, 9.4 million animals were used for research and testing in the European Union. Animal testing always entails the potential for harm caused to the animals. In order to minimize animal suffering, it is of ethical and scientific interest to have a research-based severity assessment of animal experiments. In the past, many methods have been developed to investigate animal suffering. Initially, the focus was on physiological parameters, such as body weight or glucocorticoids as an indicator of stress. In addition, the animals' behavior has come more into focus and has been included as an indicator of severity. However, in order to obtain a comprehensive understanding of animal suffering, an animal's individual perspective should also be taken into account. Preference tests might be used, for example, to "ask" animals what they prefer, and providing such goods in turn allows, among other things, to improve housing conditions. In this review, different methods are introduced, which can be used to investigate and evaluate animal suffering and well-being with a special focus on animal-centric strategies.

Abstract: It has become mandatory for the application for allowance of animal experimentation to rate the severity of the experimental procedures. In order to minimize suffering related to animal experimentation it is therefore crucial to develop appropriate methods for the assessment of animal suffering. Physiological parameters such as hormones or body weight are used to assess stress in laboratory animals. However, such physiological parameters alone are often difficult to interpret and leave a wide scope for interpretation. More recently, behavior, feelings and emotions have come increasingly into the focus of welfare research. Tests like preference tests or cognitive bias tests give insight on how animals evaluate certain situations or objects, how they feel and what their emotional state is. These methods should be combined in order to obtain a comprehensive understanding of the well-being of laboratory animals.

Keywords: severity assessment; animal welfare; refinement; preference test; cognitive bias

1. Introduction

In 2017, 9.4 million animals were used for research and testing purposes in the European Union. Mice were the most commonly used experimental animal species (61%), followed by fish (13%) and rats (12%) [1]. These animals were used either in basic research or translational and applied research but also for regulatory use and routine manufacture of medical products [1]. It is acknowledged that all animal research shall be conducted under the premise of the 3Rs (Reduce, Replace, Refine) according to Russell and Burch [2]. In light of the longstanding debate on the ethical acceptability of animal experiments, it is a moral imperative that all experiments, regardless of the species used, be double-checked for opportunities to use alternative methods. In addition, only as few animals as

absolutely necessary shall be used. Finally, all animal research that cannot be reduced or replaced must seek the best possible refinement to be ethically acceptable. In order to minimize the burden laid on animals, it is of ethical and scientific interest to have valid methods for determining animal suffering in animal experiments. Furthermore, and this is important to note, the suffering of the animals can have a profound negative impact on the experimental data. Only if the extent of the suffering is known it is possible to use this information to both strengthen animal welfare and improve the results and validity of future experiments. Animal welfare measures the status of a subjectively perceived quality of life of an individual and is notably hard to access and disentangle [3]. Thereby, animal welfare comprises various aspects, such as animal life quality, health status, biological function, and subjective feelings [4–7]. Apart from objectively measurable deterioration, animal welfare is also affected by the capacity of animals to cope with environmental challenges [8]. Overall, various factors such as social interaction, housing conditions, human handling or laboratory procedures affect animal welfare [9]. It is noteworthy that these different factors can simultaneously influence animal welfare in a non-linear way: Although positive social interaction does not directly influence the perception of pain, it can improve the overall welfare of, e.g., injured animals [3]. All this has to be taken into account for assessing the severity of procedures as well as the potential refinement measures for eliciting positive affective states [10].

In the European Union Directive 2010/63/EU, Article 38, 39, 54 and Annex VIII it is specified that all procedures involving laboratory animals have to be classified into one of four categories describing the severity of the procedure. These categories are "mild", "moderate", "severe" and "non-recovery" [11]. In the European Union in 2017, 51% of all procedures using animals in research and testing were classified as "mild", 32% were classified as "moderate", 11% as "severe" and 6% as "non-recovery" [1]. While "non-recovery" naturally means damage to the animal, paradoxically there is little concern here for the welfare of the animals, since with the death of the animal the capacity for suffering itself is also ended. However, experiments classified in any of the other three categories are under scrutiny regarding the severity of the conditions imposed on the animals so that the defined limits are not exceeded. For the classification of animal suffering, score sheets are used to assess pain, suffering or harm during animal experiments. In planning an animal experiment, all expected burdens have to be defined within these score sheets along with all measures which will be taken to reduce animal suffering. Score sheets should be efficient, easy to follow and adapted to the specific experiment. In addition, researchers and caretakers using score sheets should be well trained to unequivocally recognize and score any changes in animal welfare [12]. Ullmann and colleagues outlined recommendations for the preparation and usage of such score sheets [13]. The score sheets shall include all experiment-specific considerations, for example, van de Meer and colleagues created a score sheet for severity assessment of transgenic mice [14], and Lang and colleagues for osteotomy models in rats and mice [15]. Rix and colleagues used a score sheet for mice, which were given various chemotherapeutic agents, to study the applicability of this score sheet [16]. Only changes in body weight indicated a change in well-being of mice. Since body weight reduction could also be a side effect of chemotherapy the authors suggested to improve score sheets for experiments with chemotherapy trials by including behaviors such as nausea and fatigue into the scoring [16].

Indications of animal suffering can be derived from physiological parameters. Some studies showed that the body weight decreased during distress [17–19]. Rats which were restrained on three consecutive days showed a decreased food intake leading to a decreased body weight compared to non-restrained rats. This reduction was eminent for over 40 days after restraining [18]. However, it should be noted that body weight can be influenced, for example, by tumor growth or fluid accumulation, thus possibly masking any stress-related body weight reduction [20].

Other physiological stress parameters are glucocorticoid stress hormones, which increase in the body as a result of suffering or stress [21]. Glucocorticoids or their metabolites are commonly measured in blood [22], feces [23] or in hair samples [24]. Leenaars and colleagues performed a mapping

review to analyze the frequencies of corticosterone sample types in mice and the different analysis techniques [25].

Such physiological parameters could provide indications of changes in well-being. However, the interpretation does not always seem to be easy. For example, factors such as duration and intensity of changes in physiological parameters must be taken into account [3]. It is also important to record the nature of the situations in which these changes occur. Glucocorticoids, for example, also increase in situations that are not considered to be related to suffering such as mating [26]. Nevertheless, a lack of changes does not necessarily mean that the animal has an unchanged well-being [10]. Therefore, it is deemed useful to extend severity assessment to other parameters like behavior, preferences, or the emotional state.

2. Including the Animal's Behavior

Some experiments, for example, those involving surgical procedures or the application of pharmaceuticals, potentially inflict pain and suffering [27–31]. Treatment-induced suffering can be assessed through a comprehensive behavioral observation. Especially comfort-related behaviors such as nesting and burrowing are used to assess the animal's burden as it is assumed that comfort behaviors decrease in the presence of pain, suffering, or harm [31,32]. Jirkof and colleagues pointed out that nest-building is part of thermoregulation in small rodents, therefore, complex nest-building behavior could be an indication of unfavorable temperature conditions [10]. Häger and colleagues developed a model, in which wheel running was used to assess the severity level for mice in a colitis model. It was shown that the activity in the running wheel is indeed a useful indicator of compromised welfare in mice with a decrease in wheel running associated with increasing severity [19]. Other behaviors like twitching and writhing directly indicate pain [30,33,34]. For example, Roughan and colleagues showed that after surgery, pain behavior was significantly less expressed in rats which were given analgesia compared to rats without analgesia treatment. Based on this knowledge, they developed a pain scoring method for abdominal surgeries [33]. As direct observations are very time consuming and involve the risk of an observer bias, Roughan and colleagues used commercially available software-supported video observations to analyze activity behavior. The software identifies various behaviors such as walking, digging or stretching [35].

Another method for pain assessment is the Grimace Scale, developed first in mice by Langford and colleagues [36]. In this method the facial field of an animal is photographed and evaluated according to certain parameters (e.g., ear position, whiskers, etc.). Overall, this results in a score indicating the level of pain. In other studies, the Grimace Scale was used to assess the effectiveness of analgesics and the influence of repeated anesthesia [24,30,34,37]. The Grimace Scale is a useful method to measure suffering in laboratory animals, although this method is time consuming and there is also the possibility of an observer bias. Therefore, methods are being developed that perform images and video analysis automatically [38–40].

Abnormal behaviors such as stereotypies can also be an indication of animal suffering. Stereotypies are constant and repeated sequences of movements that do not seem to have any obvious utility [41], and can be developed under impoverished environmental conditions, but also as a result of fear or frustration [42]. Powell and colleagues showed that deer mice housed under standard conditions developed stereotyped behaviors earlier and in a higher rate compared to deer mice housed under enriched conditions [43]. Stereotypies may indicate poor well-being but for a profound assessment it is important to consider the frequency of stereotypic behavior, the situations when they occur and the individual characteristics of each animal [42,44].

3. Preference Tests

The physiological and behavioral parameters outlined above are important indicators of animal suffering. However, there is still a large scope for interpretation from the human perspective. It is

therefore necessary to develop methods that include the animal's perspective in order to gain a more comprehensive understanding of severity assessment and animal welfare.

One such animal-centered method is preference testing. Preference tests allow the animal to choose between different goods for a defined period of time. The good that is selected more frequently or for a longer period of time is considered the preferred one. Preference tests have been used frequently and in different ways [45]. However, it should also be noted that choices can be influenced by previous experiences or the current motivational state of the animal [46,47]. For example, Dawkins showed that hens normally preferred litter-floored cages without food rather than wire-floored cages with food. However, if hens previously had no access to food, the hens preferred the wire-floored cage with food [48].

In order to optimize animal husbandry, preference tests can be used to determine which type of cage design or arrangement animals prefer. Among other things, the amount of bedding provided in the home cage was examined: Freymann and colleagues showed by means of preference tests that a larger amount of bedding is preferred by mice over home cages with less bedding. The authors also showed that mice with a large amount of bedding had lower corticosterone titers than mice with less bedding. However, the behavior (e.g., agonistic behavior, locomotion, nest-building, grooming) did not seem to be influenced by the amount of bedding [49]. The preference test was also utilized to determine preference for enrichment items. Lewejohann and Sachser showed that an enriched cage with hiding and climbing possibilities is preferred by male mice over a standard cage without enrichment items [50]. Banjanin and Mrosovsky examined running wheels made of different materials for rodents and showed that mice had a high preference for plastic mesh flooring over metal rods [51].

The Conditioned Place Preference Test (CPP) is mostly used to investigate the effects of drugs [52,53]. The CPP is based on classical (Pavlovian) conditioning, in which a previously neutral stimulus (conditioned stimulus, e.g., floor pattern or odor) is associated with an event eliciting a motivational response (unconditioned stimulus, e.g., drug vs. vehicle). Conditioning itself takes place by confining the animals alternately to two distinct compartments, of which each contains a different neutral stimulus of the same modality (e.g., a pattern of dots vs. a pattern of stripes). In one compartment, the animal is then also exposed to the unconditioned stimulus. In this manner the neutral stimulus is associated with the unconditioned response, and thus becomes a conditioned stimulus. After conditioning, the previously neutral condition should induce the same response as the unconditioned stimulus [54]. Thereby, preference or avoidance can be assessed without using the unconditioned stimuli themselves in order to avoid direct negative effects or habituation to the stimuli. These findings can also be helpful in evaluating animal experiments associated with pain in relation to animal suffering. For example, the CPP has already been used to examine the effect of analgesic drugs [55,56] and has also been used to show, for example, that fish prefer an appetitive stimulus over being chased with a net [57]. In young mice it has been shown that social proximity is rewarding [58]. However, expanding the CPP to a general animal welfare assessment tool has proven to be difficult because results are easily influenced by additional motivations, e.g., spending time in a more familiar environment, or foraging instead of paying attention to the presented stimuli [59].

In addition, if an animal has made a choice and a preferred good has been determined, this does not necessarily mean that this choice is objectively the best choice for the animal. For example, many animals tend to show a strong preference for saccharin despite the lack of caloric gain, or a preference for alcohol regardless of the negative health consequences. A preference for a certain good also does not necessarily imply that if the animal does not have access to this preferred good that the animal will suffer [60]. This is especially true for luxury items or goods that can be easily surrogated by alternative goods. Therefore, it is reasonable to examine the quality of the tested goods more closely, for example, by using the consumer demand test.

Animals **2020**, *10*, 1136

4. Consumer Demand

Consumer demand tests can be used to determine the strength of preference for a preferred good. Vice versa, this test may also be useful to determine the strength of an aversion. The consumer demand test is based on the concept to "work" for access to a preferred good or for avoiding an aversive stimulus. In experimental consumer demand tests animals have to pay a certain price to obtain a good. This can be realized by introducing a workload or obstacles that has to be overcome. Work can be implemented, for example, by pressing a lever or a switch [50,61], or by an obstacle like water or an adjustable weight barrier [62,63].

The derived data can be illustrated as a consumer demand curve with the specified price on the x-axis and the amount consumed on the y-axis (Figure 1). Consumer demand theory predicts that the amount consumed is negatively affected by the price. However, the range of change is influenced by the value of the respective good. For necessary goods, price increases have only a minor effect on the quantity of goods consumed, while for luxury goods, price increases affect largely the consumed quantity. With regard to animal welfare, particular emphasis is placed on the ultimate needs necessary for survival and reproduction. In the language of consumer demand, the ultimate needs would be similar to necessities, with the animal willing to pay almost any price to get this good. On the other hand, lower consumption of a good when the price is raised indicates that such a good is less valued and reflects a luxury, which is less important with regard to animal welfare [26,64]. Importantly, Dawkins pointed out that needs without an obvious influence on survival could still be of significant value to the individual animal [48,64]. As an example, she mentioned a caged bird, whose free-living conspecifics migrate in autumn. In free-living birds, migration increases survival, whereas a caged bird does not need to migrate to survive because it is sufficiently supplied. Nevertheless, the evolutionary developed urge to migrate may be that strong as to cause suffering if the behavior cannot be performed.

Figure 1. Consumer demand curves. The consumption is based on the actual demand and price. While necessities are consumed to a considerable extent regardless of price, luxury goods can easily be dispensed with, if the price becomes too high.

By training the animals to work for the access to certain goods, the preferential strength and the grade of necessity of this good can be determined by increasing the price. Therefore, consumer demand testing is a useful method for animal welfare research and severity assessment [45,48,64].

Sherwin used the consumer demand test to demonstrate the strength of preference for a running wheel or additional space in mice. The mice had to learn to press a switch several times to gain access to a running wheel, an extended tunnel or a complex tunnel system [65]. With increasing

costs, the number of visits to the two tunnel systems decreased. However, the number of visits for the running wheel was unchanged. In the study by Lewejohann and Sachser, mice had to learn to press a lever to access an enriched cage. The mice pressed the lever up to 16 times, showing a high willingness to work for an enriched cage [50].

In order to use the consumer demand test, it has to be noted that the willingness of the animals to work depends on whether or not an adequate alternative is available. In addition, it has to be ensured that the animals have indeed learned how to get the goods, and a decrease in consumption is not due to deficits in associative learning. Notably, the animals have to be trained sufficiently to press a lever or a switch, and this training itself can be very time consuming. The animals are often placed in a separate cage so that they are trained and tested outside their familiar environment. It is also important to consider whether or not the animals are trained and tested during their active phase as this has a profound influence on the motivation for training and testing. Overall, a home cage-based test environment would be preferable as the animals could perform the training phase and consumer demand test during their active phase and in their familiar environment. This in turn could reduce many factors that might negatively affect the data.

5. Cognitive Bias Test

Recently, the cognitive bias test has been developed that promises to be a suitable method for animal welfare research and severity assessment. This test is also a test that allows examining the animal's perspective. In brief, the test investigates the influence of previous experiences on the expectation of future events. Humans and also animals which have experienced negative events tend to have a "pessimistic" expectation regarding future events, meaning they expect additional negative events and react more hesitantly towards new situations. On the other hand, humans and animals are "optimistic" towards future events, if they had more positive experiences or are less worried [66]. The cognitive bias test thus reflects the current emotional state of an individual. Determining the emotional state of laboratory animals can contribute to the improvement of housing and testing conditions. This can lead to more valid data, which also might lead to better transferability of the results.

The emotional state is influenced by cognitive processes and, conversely, the emotional state influences cognitive processes [66,67]. Cognitive abilities enable humans and animals to orient themselves and adapt to their environment. Via a combination of cognition and the emotional components, information is collected and memorized with regard to its valence. This relationship is taken advantage of using the cognitive bias test.

So far, a number of different cognitive bias tests have been presented for different species such as rats, mice, horses, sheep, or honey bees [68–73]. The tests follow the principle of conditioning animals for scalable stimuli like tones or colors (Figure 2). Animals must learn that they receive a reward (e.g., tasty food) for the stimulus at one end of the scale and that they receive a punishment (e.g., air puff) for the other stimulus on the other end of the scale. After the conditioning phase the animals are exposed to experiences potentially influencing their emotional state. Such conditions may be changes in their home cage environment or experiences due to animal experimentation. Thereafter the actual cognitive bias test follows. For this test an ambiguous stimulus, which is calibrated in the middle of the scale between the positive and negative stimuli, is presented and the reaction toward this ambiguous stimulus is measured. If the response to the ambiguous stimulus is fast, the animal seems to anticipate a reward. This behavior is interpreted as an "optimistic" emotional state. If the animal does not response or the response to the ambiguous stimulus is rather reserved, the animal's behavior is interpreted as "pessimistic". This in turn indicates that the animal expects a punishment and the recent experiences seem to have had a negative influence on the emotional state.

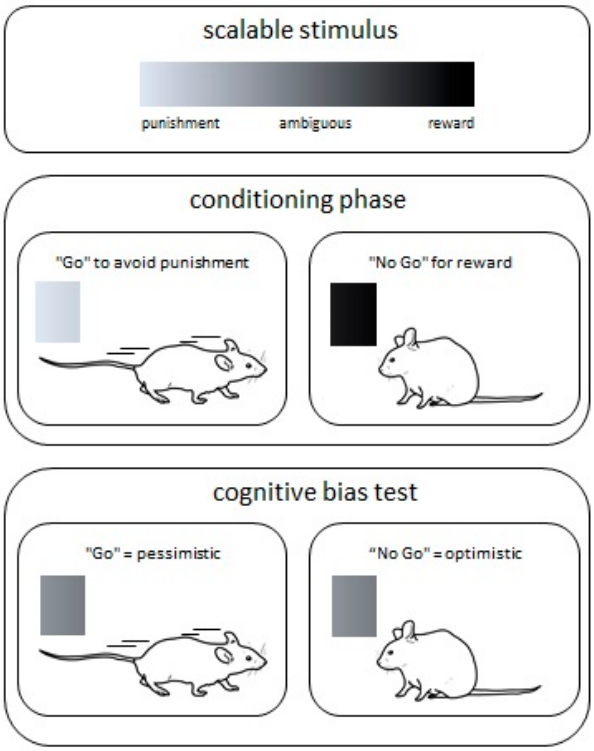

Figure 2. Cognitive bias test. During the conditioning phase the animals learn that one stimulus is associated with a punishment while the other stimulus is associated with a reward. In this Go/No Go example, the animals have to actively avoid a punishment ("Go") or stay ("No Go") to receive a reward. After successful conditioning, an ambiguous stimulus, which is calibrated in the middle of the scale between the positive and negative stimuli, is presented to test the cognitive bias. A "Go" behavior is interpreted as "pessimistic" and a "No Go" behavior as an "optimistic" emotional state.

One often-discussed aspect of the cognitive bias test is whether the test should be carried out according to a Go/No Go or a Go/Go principle, whereby "Go" would require an animal to actively reaching out (e.g., moving towards/away a reward/punishment) and "No Go" would require the animal to passively wait to receive a reward or avoid any action in order to be spared from punishment. Jones and colleagues found that rats reached the learning criterion when they had to actively approach a reward and passively avoid a punishment (i.e., "Go/No Go"). Vice versa, the learning criterion was not reached when applying a paradigm with "Go" to avoid punishment and "No Go" to receive a reward. Interestingly, mice behaved differently. Mice reached the learning criterion for the "Go" to avoid a punishment and "No Go" to receive a reward principle [74]. However, it is discussed whether the "No Go" behavior is less influenced by a negative emotional state, but rather by a lower motivation in general [75,76].

In addition, conclusions about the emotional state or "optimistic"/"pessimistic" behavior have to be made carefully as other factors might influence the animal's behavior. For example, animals which receive an air puff as a punishment for the negative stimulus could become accustomed to it and, as a result, might show less avoidance behavior and react as they would for the rewarding stimulus. This would lead to results indicating a more optimistic behavior, although the cause would not be a positive experience, which such an experiment was meant to evaluate. Moreover, it is also possible that animals which experience a negative situation could perceive the ending of this situation as positive. In the final test (after the negative situation) the results then would also indicate optimistic behavior although the situation itself was negative. Thus, cognitive bias tests always have to be interpreted cautiously and in relation to the context. More information about critical methodological aspects of the cognitive bias test can be found in the reviews of Bethell, Gygax, or Roelofs [77–79].

The cognitive bias test has already been used to examine the influence of housing conditions on the emotional state. Harding and colleagues developed the first cognitive bias test, and conditioned rats to either press or not press a lever when hearing various tones. Rats which were housed under aversive unpredictable housing conditions (e.g., reversing dark/light cycle, damped bedding) pressed the lever less often than rats which were housed under normal conditions. This response to the ambiguous stimulus was interpreted as "pessimistic" and showed that unpredictable housing conditions had a negative influence on the emotional state of rats [68].

In another test, rats had to associate grades of sandpaper (fine or rough) with reward or punishment. The data indicated that rats which were housed first in standard cages without enrichment and then transferred to enriched cages showed an "optimistic" bias compared to rats which were housed permanently in non-enriched cages [80]. Similar results were given in a study with a depression-like phenotype in rats. Rats which were housed unenriched and then transferred to enriched cages showed a shift to an "optimistic" bias [69].

The first cognitive bias test for mice was developed by Boleij and colleagues in 2012. The mice were conditioned to various odor cues [70]. A spatial cognitive bias test for mice was developed by Kloke and colleagues in 2014 [81], showing that mice lacking a functional serotonin transporter tended to be more pessimistic compared to wild type mice.

Past studies have shown that the cognitive bias test is a useful method to examine the emotional state and the expectation regarding future events in animals. Therefore, this test also seems to be a suitable method for animal welfare research and severity assessment. The cognitive bias test can also be used to evaluate housing and experimental conditions of laboratory animals, and allows the animal's point of view to be taken into account for adaptation, refinement and improvement. However, there are also disadvantages within the previous approaches. For example, in the above mentioned test designs, the actual test run could only be carried out once while training proved to be very time-consuming. Therefore, an automated touchscreen-based test design was developed [82]. As more trials per session can be performed in an automated test, the number of ambiguous trials per session can be better balanced. This is important to prevent the animals from learning that there is no reward or punishment for ambiguous stimuli [79,83], and it is possible to repeat the cognitive bias test. In addition, automated data collection avoids an observer bias, allowing neutral data evaluation [84]. However, in this touchscreen-based approach it is still necessary to place the animals in a separate test apparatus for training and testing. Therefore, an automated and home cage-based cognitive bias test would be of great advantage. Both the test itself and the conditioning could be carried out without the influence of handling, during the active phase of the animals and in their familiar environment. This would reduce external influences which could affect the cognitive bias of the animals.

6. Conclusions

The number of animals used for experimental purposes is still alarmingly high. This becomes particularly clear when surplus animals are counted in addition to the pure laboratory animal numbers [85]. It is therefore imperative for all researchers that 3R measures must continue to be used to further reduce these figures. With regard to animal welfare, all animals under human supervision must be taken into account. Especially for all surplus animals, animal welfare can sometimes be improved more easily [3].

Just as researchers can ask themselves what it takes to live a good life, subjective feelings are also of great importance for animals. Subjective experiences are linked to the behavior and physiology of the animal and should not be considered separately [86]. Sandøe states that well-being cannot be assessed by scientific methods alone. He therefore suggests that animal researchers and philosophers should work together to define and evaluate well-being [86].

In the case of animal experiments, animal welfare must be a top priority in addition to the scientific objective. It is therefore essential to assess the severity of procedures involving laboratory animals

as objectively and accurately as possible. However, depending on the nature of the experimental and husbandry conditions, pain, suffering, or harm might be subtle and thus not easy to quantify.

Under laboratory and experimental conditions, animals are restricted in the development of their natural behavioral repertoire, and a wide range of husbandry conditions can be improved to refine the welfare of laboratory animals [3]. Animal experimental research basically involves procedures that, depending on the experiment, are associated with more or less pain, suffering or damage. Therefore, all procedures should be continuously examined for refinement possibilities to minimize suffering. Indeed, continuous monitoring of the health status of laboratory animals can make a huge contribution to reducing animal suffering [19,30,33–36]. All in all, there is still much room for improvement in the welfare of laboratory animals for those animal experiments that cannot be replaced in the foreseeable future. This certainly also includes promoting positive animal welfare in laboratory animals [10] rather than merely avoiding negative impacts.

Improving animal welfare requires methods to assess severity, of which several major approaches are discussed in this article. In addition to objectively measurable parameters, the animals' perspective must be taken into account. Science can only indirectly ask the animals what they want or do not want (preference tests) or how much they want or do not want (consumer demand tests) certain goods. Science can also ask the animals only indirectly how their emotional status is within or after a specific situation (cognitive bias test). However, these approaches offer the possibility to better understand laboratory animals in their entirety, which can also lead to better animal research and results as there is growing evidence that impaired well-being affects the quality of data collected in animal studies [87].

Author Contributions: P.K. conceptualized this review. P.K. and A.H. prepared the figures. P.K., A.H., K.D. and L.L. wrote the manuscript and revised it. L.L. has raised funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by Deutsche Forschungsgemeinschaft (DFG) as part of the research group "Severity Assessment for Animal-Based Research" (FOR 2591, LE2356/5-1).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. European Commission. Report on the Statistics on the Use of Animals for Scientific Purposes in the Member States of the European Union in 2015–2017. Available online: https://ec.europa.eu/transparency/regdoc/rep/1/2020/EN/COM-2020-16-F1-EN-MAIN-PART-1.PDF (accessed on 27 March 2020).
2. Russell, W.M.S.; Burch, R.L. *The Principles of Humane Experimental Technique*; Methuen: London, UK, 1959.
3. Lewejohann, L.; Schwabe, K.; Häger, C.; Jirkof, P. Impulse for animal welfare outside the experiment. *Lab. Anim.* **2020**, *54*, 150–158. [CrossRef] [PubMed]
4. Fraser, D.; Weary, D.M.; Pajor, E.A.; Milligan, B.N. A scientific conception of animal welfare that reflects ethical concerns. *Anim. Welf.* **1997**, *6*, 187–205.
5. Broom, D.M. The use of the concept animal welfare in European conventions, regulations and directives. *Food Chain* **2001**, *2001*, 148–151.
6. Appleby, M.C.; Sandøe, P. Philosophical debate on the nature of well-being: Implications for animal welfare. *Anim. Welf.* **2002**, *11*, 283–294.
7. Broom, D.M. Behaviour and welfare in relation to pathology. *Appl. Anim. Behav. Sci.* **2006**, *97*, 71–83. [CrossRef]
8. Broom, D.M. Animal welfare: Concepts and measurement. *J. Anim Sci.* **1991**, *69*, 4167–4175. [CrossRef]
9. Broom, D.M. Animal welfare defined in terms of attempts to cope with the environment. *Acta Agric. Scand. Sec. A Anim. Sci. Suppl.* **1996**, *27*, 22–28.
10. Jirkof, P.; Rudeck, J.; Lewejohann, L. Assessing affective state in laboratory rodents to promote animal welfare—What is the progress in applied refinement research? *Animals* **2019**, *9*, 1026. [CrossRef]

11. European Parliament. EU Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the Protection of Animals Used for Scientific Purposes. Available online: https://www.bfr.bund.de/cm/343/5_Beratung_Anlage%203_2010-63-EU.pdf (accessed on 27 March 2020).
12. Bugnon, P.; Heimann, M.; Thallmair, M. What the literature tells us about score sheet design. *Lab. Anim.* **2016**, *50*, 414–417. [CrossRef]
13. Ullmann, K.; Jourdan, T.; Kock, M.; Unger, J.; Schulz, A.; Thöne-Reineke, C.; Abramjuk, C. Recommendations for the development and use of Score Sheets as a tool for applied refinement. *Berl. Münch. Tierärztl. Wochenschr.* **2018**. [CrossRef]
14. Van de Meer, M.; Rolls, A.; Baumans, V.; Olivier, B.; van Zutphen, L.F.M. Use of score sheets for welfare assessment of transgenic mice. *Lab. Anim.* **2001**, *35*, 379–389. [CrossRef]
15. Lang, A.; Schulz, A.; Ellinghaus, A.; Schmidt-Bleek, K. Osteotomy models—The current status on pain scoring and management in small rodents. *Lab. Anim.* **2016**, *50*, 433–441. [CrossRef]
16. Rix, A.; Drude, N.; Mrugalla, A.; Mottaghy, F.M.; Tolba, R.H.; Kiessling, F. Performance of severity parameters to detect chemotherapy-induced pain and distress in mice. *Lab. Anim.* **2019**, *0*, 1–9. [CrossRef] [PubMed]
17. Rybkin, I.I.; Zhou, Y.; Volaufova, J.; Smagin, G.N.; Ryan, D.H.; Harris, R.B.S. Effect of restraint stress on food intake and body weight is determined by time of day. *Am. J. Physiol.* **1997**, *273*, R1612–R1622. [CrossRef] [PubMed]
18. Harris, R.B.S.; Zhou, J.; Youngblood, B.D.; Rybkin, I.I.; Smagin, G.N.; Ryan, D.H. Effect of repeated stress on body weight and body composition of rats fed low- and high-fat diets. *Am. J. Physiol.* **1998**, *275*, R1928–R1938. [CrossRef] [PubMed]
19. Häger, C.; Keuler, L.M.; Talbot, S.R.; Biernot, S.; Weegh, N.; Buchheister, S.; Buettner, M.; Glage, S.; Bleich, A. Running in the wheel: Defining individual severity levels in mice. *PLoS Biol.* **2018**, *16*, e2006159. [CrossRef] [PubMed]
20. Ullman-Cullere, M.H.; Foltz, C.J. Body Condition Scoring: A rapid and accurate method for assessing health status in mice. *Lab. Anim. Sci.* **1999**, *49*, 319–323. [PubMed]
21. Dallman, M.F.; Akana, S.F.; Scribner, K.A.; Bradbury, M.J.; Walker, C.D.; Strack, A.M.; Cascio, C.S. Stress, feedback and facilitation in the hypothalamo-pituitary-adrenal axis. *J. Neuroendocrinol.* **1992**, *4*, 517–526. [CrossRef] [PubMed]
22. Tsai, P.; Schlichtig, A.; Ziegler, E.; Ernst, H.; Haberstroh, J.; Stelzer, H.D.; Hackbarth, H. Effects of different blood collection methods on indicators of welfare in mice. *Lab. Anim.* **2015**, *44*, 301–310. [CrossRef]
23. Touma, C.; Palme, R.; Sachser, N. Analysing corticosterone metabolites in fecal samples of mice: A noninvasive technique to monitor stress hormones. *Horm. Behav.* **2004**, *45*, 10–22. [CrossRef]
24. Hohlbaum, K.; Bert, B.; Dietze, S.; Palme, R.; Fink, H.; Thöne-Reineke, C. Severity classification of repeated isoflurane anesthesia in C57BL/6JRj mice—Assessing the degree of distress. *PLoS ONE* **2017**, *12*, e0179588. [CrossRef] [PubMed]
25. Leenaars, C.H.C.; van der Mierden, S.; Durst, M.; Goerlich-Jansson, V.C.; Ripoli, F.L.; Keubler, L.M.; Talbot, S.R.; Boyle, E.; Habedank, A.; Jirkof, P.; et al. Measurement of corticosterone in mice: A protocol for a mapping review. *Lab. Anim.* **2020**, *54*, 26–32. [CrossRef] [PubMed]
26. Dawkins, M.S. From an animal's point of view: Motivation, fitness, and animal welfare. *Behav. Brain Sci.* **2009**, *13*, 1–9. [CrossRef]
27. Lister, R.G. The use of a plus-maze to measure anxiety in the mouse. *Psychopharmacology* **1987**, *92*, 180–185. [CrossRef] [PubMed]
28. Weigle, D.S.; Bukowski, T.R.; Foster, D.C.; Holderman, S.; Kramer, J.M.; Lasser, G.; Lofton-Day, C.E.; Prunkard, D.E.; Raymond, C.; Kuijper, J.L. Recombination of protein reduces feeding and body weight in the ob/ob mouse. *J. Clin. Investig.* **1995**, *96*, 2065–2070. [CrossRef] [PubMed]
29. Baumans, V.; Bouwknecht, J.A.; Boere, H.; Kramer, K.; van Lith, H.A.; van de Weerd, H.A.; van Herck, H. Intra-abdominal transmitter implantation in mice: Effects on behaviour and body weight. *Anim. Welf.* **2001**, *10*, 291–302.
30. Miller, A.L.; Miller, A.L.; Wright-Williams, S.L.; Flecknell, P.A.; Roughan, J.V. A comparison of abdominal and scrotal approach methods of vasectomy and the influence of analgesic treatments in laboratory mice. *Lab. Anim.* **2012**, *46*, 304–310. [CrossRef]

31. Sliepen, S.H.J.; Diaz-Delcastillo, M.; Korioth, J.; Olsen, R.B.; Appel, A.K.; Christoph, T.; Heegaard, A.; Rutten, K. Cancer-induced bone pain impairs burrowing behaviour in mice. *In Vivo* **2019**, *33*, 1125–1132. [CrossRef]
32. Gjendal, K.; Ottesen, J.L.; Olsson, A.S.; Sørensen, D.B. Burrowing and nest building activity in mice after exposure to grid floor, isoflurane or ip injetions. *Physiol. Behav.* **2019**, *206*, 59–66. [CrossRef]
33. Roughan, J.V.; Flecknell, P.A. Evaluation of short duration behaviour-based post-operative pain scoring system in rats. *Eur. J. Pain* **2012**, *7*, 397–406. [CrossRef]
34. Leach, M.; Klaus, K.; Miller, A.; Scotto di Perrotolo, M. The assessment of post-vasectomy pain in mice using behaviour and the mouse grimace scale. *PLoS ONE* **2012**, *7*, e35656. [CrossRef]
35. Roughan, J.V.; Wright-Williams, S.L.; Flecknell, P.A. Automated analysis of postoperative behaviour: Assessment of HomeCageScan as a novel method to rapidly identify pain and analgesic effects in mice. *Lab. Anim.* **2009**, *43*, 17–26. [CrossRef] [PubMed]
36. Langford, D.J.; Bailey, A.L.; Chanda, M.L.; Clarke, S.E.; Drummond, T.E.; Echols, S.; Glick, S.; Ingrao, J.; Klassen-Ross, T.; La Croix-Fralish, M.; et al. Coding of facial expressions of pain in the laboratory mouse. *Nat. Methods* **2010**, *7*, 447–449. [CrossRef]
37. Hohlbaum, K.; Bert, B.; Dietze, S.; Palme, R.; Fink, H.; Thöne-Reineke, C. Impact of repeated anesthesia with ketamine and xylazine on well-being of C57BL/6JRj mice. *PLoS ONE* **2018**, *13*, e0203559. [CrossRef] [PubMed]
38. Andresen, N.; Wöllhaf, M.; Hohlbaum, K.; Lewejohann, L.; Hellwich, O.; Thöne-Reineke, C.; Belik, V. Toward a fully automated surveillance of well-being status in laboratory mice using deep learning. *PLoS ONE* **2020**, *15*, e0228059. [CrossRef] [PubMed]
39. Ernst, L.; Kopaczka, M.; Schulz, M.; Talbot, S.R.; Zieglowski, L.; Meyer, M.; Bruch, S.; Merhof, D.; Tolba, R.H. Improvement of the mouse grimace scale set-up for implementing a semi-automated mouse grimace scale scoring (Part 1). *Lab. Anim.* **2020**, *54*, 83–91. [CrossRef]
40. Ernst, L.; Kopaczka, M.; Schulz, M.; Talbot, S.R.; Struve, B.; Häger, C.; Bleich, A.; Durst, M.; Jirkof, P.; Arras, M.; et al. Semi-automated generation of pictures for the mouse grimace scale: A multi-laboratory analysis (Part 2). *Lab. Anim.* **2020**, *54*, 92–98. [CrossRef]
41. Garner, J.P. Stereotypies and other abnormal repetitive behaviors: Potential impact on validity, reliability, and replicability of scientific outcomes. *ILAR J.* **2005**, *46*, 106–117. [CrossRef]
42. Mason, G.J. Stereotypies and suffering. *Behav. Process.* **1991**, *25*, 103–115. [CrossRef]
43. Powell, S.B.; Newman, H.A.; Pendergast, J.F.; Lewis, M.H. A rodent model of spontaneous stereotypy: Initial characterization of developmental, environmental, and neurobiological factors. *Physiol. Behav.* **1998**, *66*, 355–363. [CrossRef]
44. Mason, G.J. Stereotypies: A critical review. *Anim. Behav.* **1991**, *41*, 1015–1037. [CrossRef]
45. Habedank, A.; Kahnau, P.; Diederich, K.; Lewejohann, L. Severity assessment from an animal's point of view. *Berl. Münch. Tierärztl. Wochenschr.* **2018**. [CrossRef]
46. Dawkins, M.S. Do hens suffer in battery cages? Environmental preferences and welfare. *Anim. Behav.* **1977**, *25*, 1034–1046. [CrossRef]
47. Bloom, H.J.M.; Vanvorstenbosch, C.J.A.H.V.; Baumans, V.; Hoogervorst, M.J.C.; Beynen, A.C.; Vanzutphen, L.F.M. Description and validation of a preference test system to evaluate housing conditions for laboratory mice. *Appl. Anim. Behav. Sci.* **1992**, *35*, 67–82. [CrossRef]
48. Dawkins, M.S. Battery hens name their price: Consumer demand theory and the measurement of ethological 'needs'. *Anim. Behav.* **1983**, *312*, 1195–1205. [CrossRef]
49. Freymann, J.; Tsai, P.; Stelzer, H.; Hackbarth, H. The impact of bedding volumes on laboratory mice. *Appl. Anim. Behav. Sci.* **2016**, *186*, 72–79. [CrossRef]
50. Lewejohann, L.; Sachser, N. Evaluation of different housing conditions for male laboratory mice by means of preference tests. *KTBL Schrift.* **2000**, *391*, 170–177.
51. Banjanin, S.; Mrosovsky, N. Preferences of mice, *Mus musculus*, for different types of running wheel. *Lab. Anim. UK* **2000**, *34*, 313–318. [CrossRef]
52. Tzschentke, T.M. Measuring reward with the conditioned place preference paradigm: A comprehensive review of drug effects, recent progress and new issues. *Prog. Neurobiol.* **1998**, *56*, 613–672. [CrossRef]
53. Schechter, M.D.; Calcagnetti, D.J. Trends in place preference conditioning with a cross-indexed bibliography; 1957–1991. *Neurosci. Biobehav. Rev.* **1993**, *17*, 21–41. [CrossRef]

54. Cunningham, C.L.; Gremel, C.M.; Groblewski, P.A. Drug-induced conditioned place preference and aversion in mice. *Nat. Protoc.* **2006**, *1*, 1662–1670. [CrossRef] [PubMed]
55. King, T.; Vera-Portocarrero, L.; Gutierrez, T.; Vanderah, T.W.; Dussor, G.; Lai, J.; Fields, H.L.; Porreca, F. Unmasking the tonic-aversive state in neuropathic pain. *Nat. Neurosci.* **2009**, *12*, 1364–1366. [CrossRef] [PubMed]
56. Roughan, J.V.; Coulter, C.A.; Flecknell, P.A.; Thomas, H.D.; Sufka, K.J. The conditioned place preference test for assessing welfare consequences and potential refinements in a mouse bladder cancer model. *PLoS ONE* **2014**, *9*, e103362. [CrossRef] [PubMed]
57. Millot, S.; Cerqueira, M.; Casatnheira, M.F.; Øverli, Ø.; Martins, C.I.M.; Oliveira, R.F. Use of conditioned place preference/avoidance tests to assess affective states in fish. *Appl. Anim. Behav. Sci.* **2014**, *154*, 104–111. [CrossRef]
58. Panksepp, J.B.; Lahvis, G.P. Social reward among juvenile mice. *Genes Bain Behav.* **2007**, *6*, 661–671. [CrossRef]
59. Dixon, L.M.; Sandilands, V.; Bateson, M.; Brocklehurst, S.; Tolkamp, B.J.; D'Eath, R.B. Conditioned place preference or aversion as animal welfare assessment tools: Limitations in their application. *Appl. Anim. Behav. Sci.* **2013**, *148*, 164–176. [CrossRef]
60. Dawkins, M.S.; Beardsley, T. Reinforcing properties of access to litter in hens. *Appl. Anim. Behav. Sci.* **1986**, *15*, 351–365. [CrossRef]
61. Sherwin, C.M. The motivation of group-housed laboratory mice, *Mus musculus*, for additional space. *Anim. Behav.* **2004**, *67*, 711–717. [CrossRef]
62. Sherwin, C.M.; Nicol, C.J. Changes in meal patterning by mice measure the cost imposed by natural obstacles. *Appl. Anim. Behav. Sci.* **1995**, *43*, 291–300. [CrossRef]
63. Manser, C.E.; Elliot, H.; Morris, T.H.; Broom, D.M. The use of a novel operant test to determine the strength of preferences for flooring in laboratory rats. *Lab. Anim.* **1996**, *30*, 1–6. [CrossRef]
64. Dawkins, M.S. Behavioural Deprivation: A central problem in animal welfare. *Appl. Anim. Behav. Sci.* **1988**, *20*, 209–225. [CrossRef]
65. Sherwin, C.M. The use and perceived importance of three resources which provide cage laboratory mice the opportunity for extended locomotion. *Appl. Anim. Behav. Sci.* **1998**, *55*, 353–367. [CrossRef]
66. Mendl, M.; Burman, O.H.P.; Parker, R.M.A.; Paul, E.S. Cognitive bias as an indicator of animal emotion and welfare: Emerging evidence and underlying mechanisms. *Appl. Anim. Behav. Sci.* **2009**, *118*, 161–181. [CrossRef]
67. Herrmann, E.; Call, J. Are there geniuses among the apes? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2012**, *367*, 2753–2761. [CrossRef] [PubMed]
68. Harding, E.J.; Paul, E.S.; Mendl, M. Cognitive bias and affective state. *Nature* **2004**, *427*, 2006159. [CrossRef] [PubMed]
69. Richter, S.H.; Schick, A.; Hoyer, C.; Lankisch, K.; Gass, P.; Vollmayr, B. A glass full of optimism: Enrichment effects on cognitive bias in a rat model of depression. *Cogn. Affect. Behav. Neurosci.* **2012**, *12*, 527–542. [CrossRef]
70. Boleij, H.; Klooster, J.; Lavrijsen, M.; Kirchhoff, S.; Arndt, S.S.; Ohl, F. A test to identify judgement bias in mice. *Behav. Brain Res.* **2012**, *233*, 45–54. [CrossRef]
71. Hintze, S.; Melotti, L.; Colosio, S.; Bailoo, J.D.; Boada-Sana, M.; Würbel, H.; Murphy, E. A cross-species judgement bias task: Integrating active trial initiation into a spatial go/no-go task. *Sci. Rep.* **2018**, *8*, 5104. [CrossRef]
72. Verbeek, E.; Ferguson, D.; Lee, C. Are hungry sheep more pessimistic? The effects of food restriction on cognitive bias and the involvement of ghrelin in its regulation. *Physiol. Behav.* **2014**, *123*, 67–75. [CrossRef]
73. Schlüns, H.; Welling, H.; Federici, J.R.; Lewejohann, L. The glass is not yet half empty: Agitation but not verroa treatment causes cognitive bias in honey bees. *Anim. Cogn.* **2017**, *20*, 233–241. [CrossRef]
74. Jones, S.; Paul, E.S.; Dayan, P.; Robinson, E.S.J.; Mendl, M. Pavlovian influences on learning differ between rats and mice in a counter-balanced Go/NoGo judgement bias task. *Behav. Brain Res.* **2017**, *331*, 214–224. [CrossRef]
75. Matheson, S.M.; Asher, L.; Bateson, M. Larger, enriched cages are associated with 'optimistic' response biases in captive European starlings (*Sturnus vulgaris*). *Appl. Anim. Behav. Sci.* **2008**, *109*, 374–383. [CrossRef]

76. Enkel, T.; Gholizadeh, D.; von Bohlen und Halbach, O.; Sanchis-Segura, C.; Hurlemann, R.; Spanagel, R.; Gass, P.; Vollmayr, B. Ambiguous-Cue interpretation is biased under stress and depression like states in rats. *Neuropsychopharmacology* **2010**, *35*, 1008–1015. [CrossRef]
77. Gygax, L. The A to Z of statistics for testing cognitive judgement bias. *Anim. Behav.* **2014**, *95*, 59–69. [CrossRef]
78. Bethell, E.J. A "How-To" guide for designing judgment bias studies to assess captive animal welfare. *Appl. Anim. Welf. Sci.* **2015**, *1*, 18–42. [CrossRef] [PubMed]
79. Roelofs, S.; Boleij, H.; Nordquist, R.E.; van der Staay, F.J. Making decisions under ambiguity: Judgment bias tasks for assessing emotional state in animals. *Front. Behav. Neurosci.* **2016**, *10*, 119. [CrossRef]
80. Brydges, N.M.; Leach, M.; Nicol, K.; Wright, R.; Bateson, M. Environmental enrichment induces optimistic cognitive bias in rats. *Anim. Behav.* **2011**, *81*, 169–175. [CrossRef]
81. Kloke, V.; Schreiber, R.S.R.S.; Bodden, C.; Möllers, J.; Ruhmann, H.; Kaiser, S.; Lesch, K.P.K.; Sachser, N.; Lewejohann, L. Hope for the best or prepare for the worst? Towards a spatial cognitive bias test for mice. *PLoS ONE* **2014**, *9*, e105431. [CrossRef]
82. Krakenberg, V.; Woigk, I.; Garcia Rodriguez, L.; Kästner, N.; Kaiser, S.; Sachser, N.; Richter, S.H. Technology or ecology? New tools to assess cognitive judgement bias in mice. *Behav. Brain Res.* **2019**, *362*, 279–287. [CrossRef]
83. Doyle, R.E.; Vidal, S.; Hinch, G.N.; Fischer, A.D.; Boissy, A.; Lee, C. The effect of repeated testing on judgement biases in sheep. *Behav. Process.* **2010**, *83*, 349–352. [CrossRef]
84. Jones, S.; Neville, V.; Higgs, L.; Paul, E.S.; Dayan, P.; Robinson, E.S.J.; Mendl, M. Assessing animal affect: An automated and self-initiated judgement bias task based on natural investigative behaviour. *Sci. Rep.* **2018**, 812400. [CrossRef]
85. Lewejohann, L.; Grune, B.; Schönfelder, G.; Bert, B. Cut back on surplus laboratory animals [All labortatory animals count]. *Nature* **2020**, *578*, 515. [CrossRef] [PubMed]
86. Sandøe, P.; Simonsen, H.B. Assessing animal welfare: Where does science end and philosophy begin? *Anim. Welf.* **1992**, *1*, 257–267.
87. Pool, T. Happy animals make good science. *Lab. Anim.* **1997**, *31*, 116–124. [CrossRef]

2.2 Publication 2

Lifetime observation of cognition and physiological parameters in male mice

Pia Kahnau, Anja Guenther, Marcus Nicolaas Boon, Janine Denise Terzenbach, Eric Hanitzsch, Lars Lewejohann, Vera Brust

Published in journal “Frontiers in Behavioral Neuroscience” 2021, 15, 709775
DOI: https://doi.org/10.3389/fnbeh.2021.709775

The Supplementary Material for this article can be found online at:
https://www.frontiersin.org/articles/10.3389/fnbeh.2021.709775/full

frontiers in Behavioral Neuroscience

ORIGINAL RESEARCH
published: 01 September 2021
doi: 10.3389/fnbeh.2021.709775

Lifetime Observation of Cognition and Physiological Parameters in Male Mice

Pia Kahnau[1], Anja Guenther[2], Marcus Nicolaas Boon[3,4], Janine Denise Terzenbach[5], Eric Hanitzsch[5], Lars Lewejohann[1,6] and Vera Brust[5]*

[1] Laboratory Animal Science, German Centre for the Protection of Laboratory Animals (Bf3R), German Federal Institute for Risk Assessment (BfR), Berlin, Germany, [2] Department of Animal Behaviour, Bielefeld University, Bielefeld, Germany, [3] Department for Electrical Engineering and Computer Science, Modeling of Cognitive Processes, Technische Universität Berlin, Berlin, Germany, [4] Exzellenzcluster Science of Intelligence, Technische Universität Berlin, Berlin, Germany, [5] Behavioral Phenotyping Unit, University of Osnabrück, Osnabrück, Germany, [6] Animal Behavior and Laboratory Animal Science, Institute of Animal Welfare, Freie Universität Berlin, Berlin, Germany

OPEN ACCESS

Edited by:
Jeansok J. Kim, University of Washington, United States

Reviewed by:
Therese A. Kosten, University of Houston, United States
Tomáš Petrásek, National Institute of Mental Health, Czechia

***Correspondence:**
Pia Kahnau
pia.kahnau@bfr.bund.de

Specialty section:
This article was submitted to Learning and Memory, a section of the journal Frontiers in Behavioral Neuroscience

Received: *14 May 2021*
Accepted: *13 August 2021*
Published: *01 September 2021*

Citation:
Kahnau P, Guenther A, Boon MN, Terzenbach JD, Hanitzsch E, Lewejohann L and Brust V (2021) Lifetime Observation of Cognition and Physiological Parameters in Male Mice. Front. Behav. Neurosci. 15:709775. doi: 10.3389/fnbeh.2021.709775

Laboratory mice are predominantly used for one experiment only, i.e., new mice are ordered or bred for every new experiment. Moreover, most experiments use relatively young mice in the range of late adolescence to early adulthood. As a consequence, little is known about the day-to-day life of adult and aged laboratory mice. Here we present a long-term data set with three consecutive phases conducted with the same male mice over their lifetime in order to shed light on possible long-term effects of repeated cognitive stimulation. One third of the animals was trained by a variety of learning tasks conducted up to an age of 606 days. The mice were housed in four cages with 12 animals per cage; only four mice per cage had to repeatedly solve cognitive tasks for getting access to water using the IntelliCage system. In addition, these learner mice were tested in standard cognitive tests outside their home-cage. The other eight mice served as two control groups living in the same environment but without having to solve tasks for getting access to water. One control group was additionally placed on the test set-ups without having to learn the tasks. Next to the cognitive tasks, we took physiological measures (body mass, resting metabolic rate) and tested for dominance behavior, and attractivity in a female choice experiment. Overall, the mice were under surveillance until they died a natural death, providing a unique data set over the course of virtually their entire lives. Our data showed treatment differences during the first phase of our lifetime data set. Young learner mice showed a higher activity, less growth and resting metabolic rate, and were less attractive for female mice. These effects, however, were not preserved over the long-term. We also did not find differences in dominance or effects on longevity. However, we generated a unique and valuable set of long-term behavioral and physiological data from a single group of male mice and note that our long-term data contribute to a better understanding of the behavioral and physiological processes in male C57Bl/6J mice.

Keywords: laboratory mice, cognition, IntelliCage, lifetime observation, resting metabolic rate

INTRODUCTION

Cognition comprises information processing mechanisms which enable decision making, i.e., perception, memory, and learning (McEwen, 2007; Gruszka et al., 2010; Roelofs et al., 2016). Cognitive abilities thereby enable adaptation to the social and physical environment of an individual. Especially complex or constantly changing environments are better coped with and exploited more successfully by species and individuals with an increased cognitive performance (e.g., Sol and Lefebvre, 2000; Lee, 2003; Dunbar and Shultz, 2007; Boucherie et al., 2019). Accordingly, a higher cognitive ability holds the potential to favor reproductive success and survival (Papini, 2002; Dukas, 2004). However, cognition is also associated with costs as the processing of information requires nervous tissue, which is energetically expensive to develop and maintain (Laughlin et al., 1998; Niven and Laughlin, 2008; Hollis and Kawecki, 2014). Adding up on this constitutive investment are the induced costs of building and maintaining particular memories (Snell-Rood, 2013). The trade-off between costs and benefits of cognition can elegantly be demonstrated by the fact that cognitive abilities are usually not maxed out under natural selection. Indeed, evidence from many different species shows that cognitive abilities can be substantially improved by artificial selection (e.g., Tryon, 1940; Brandes, 1988; Mery and Kawecki, 2002). In most habitats, however, it is generally assumed that natural selection prevents a permanent improvement of cognitive abilities due to trade-offs with other fitness related traits (Buchanan et al., 2013). Thus, an evolved species has a cognitive range within which an individual must incur the corresponding physiological costs depending on the energetic expenditure of brain activity. Studies linking cognition directly to physiological, reproductive, or survival traits are limited (e.g., Cole et al., 2012; Huebner et al., 2018). Therefore, we still know little about how these trade-offs actually shape cognitive abilities or, on an individual level, affect how an individual uses its cognitive abilities. Here, we follow a cohort of male mice throughout life with the aim to study the effects of a cognitively demanding life on different aspects of physiology, reproduction, and survival.

By provisioning an environment which constantly held new cognitive challenges to some but not all of the mice, we wanted to test for the consequences of such different lifestyles. Based on the above mentioned trade-offs, cognitively stimulated mice might have a reduced or slower growth and in addition or alternatively a higher energy turn-over as compared to non-stimulated mice. The energy consumption of an individual can be determined by measuring its metabolic rate, i.e., oxygen uptake and carbon dioxide release, indicating how much energy is produced by aerobic respiration (Brown, 2004). While a direct influence of learning performance or use of cognitive abilities on metabolic rate to our knowledge has not yet been investigated, the metabolic rate in this study was measured four times throughout the mice's life. We wanted to investigate a possible relation between different learning environments suggesting a higher metabolic rate in males which were cognitively stimulated repeatedly throughout their lives.

Besides these possible physiological contrasts between the differently stimulated mice, we assessed an aspect of reproductive success. Only male mice were included in the study and we tested them twice in a female choice task where potential mates were allowed to freely choose to spend time in close proximity with individuals of the different testing groups. Some studies in insects and vertebrates have shown that females choose mates with better cognitive skills reflected in males' courtship behavior, performance in foraging or in diet-dependent morphological traits [reviewed in Boogert et al. (2011)]. Additionally, males of species with a complex and competitive sexual environment may have to process complex sensory information and display learned abilities in courting females (Byrne and Rice, 2006; Dukas, 2006; Griffith and Ejima, 2009). In an experimental evolution study in fruit flies cognitive performance of males declined under the absence of sexual selection (Hollis and Kawecki, 2014). Consequently, we might suggest an impact of the different learning environments applied to our tested mice in relation to their attractiveness toward females. Still, as their potential differences in cognitive abilities are not directly on display in our testing context, it is difficult to predict whether the females are able to include these traits in their decisions. In addition to female choice, male-male competition is a core principle of sexual selection. In order to measure whether or not the cognitive stimulation affected intrasexual competition, we performed direct observations of aggressive behavior within the social groups of male mice.

Measuring longevity is a straightforward way to test if the learning environment has fitness consequences. Under natural conditions, longevity may be directly influenced by differences in foraging success or predator avoidance (e.g., Madden et al., 2018). Accordingly, higher cognitive abilities have repeatedly been found to be positively linked to survival in the wild (see Morand-Ferron, 2017 for an overview). However, under captive conditions, differences in predator avoidance or foraging success are unlikely to occur because animals are usually protected from predation and provided with food *ad libitum*. Nevertheless, the link between physiological condition and cognitive function could affect longevity even under captive conditions, especially if differences in early environments influence physiological and cognitive development through phenotypic plasticity (Pravosudov et al., 2005; Loi et al., 2017). Accordingly, replicate populations of fruit flies selected for an improved learning ability have shown a pronounced reduction in longevity and conversely, lines selected for extended longevity showed a reduction in learning ability (Burger et al., 2008). Our tested mice were kept over their whole lifespan enabling us to directly test for an effect of the applied learning treatment on longevity. In accordance with former laboratory studies, we expect a reduced lifespan in mice kept under cognitively stimulating conditions as compared to the non-stimulated mice. Besides this direct measurement of longevity, we investigated the telomere length of the mice at a later stage in life. Telomeres comprise of repeated and non-coding DNA strands forming the ends of each chromosome and have been linked to rates of aging and age-related diseases in aging human and non-human individuals (Von Zglinicki, 2002; Brouilette et al., 2003;

Benetos et al., 2004; Martin-Ruiz et al., 2006). In addition, several studies such as for example by Yaffe et al. (2011) showed that telomere length can also serve as a marker for cognitive aging in humans with shorter telomeres going along with reduced cognitive abilities. While the effects of stressful environments on telomere length and cognitive decline are well described, it is not yet known if an environment that imposes elevated cognitive processes like applied in our study can affect telomere lengths.

Considering all the above, we hypothesize that cognition in male mice throughout life affects physiology (body mass development, metabolic rate), sexually selected traits (competitive ability, male attractiveness), and longevity (measured by actual survival and telomere length) across life. A unique feature of our study is that mice were kept in groups in a home-cage based test apparatus, the IntelliCage (IC, New Behavior) system. This way, individual testing paradigms could be applied to each animal enabling us to form social groups that contained both, learner and non-learner mice. Previous studies show that mice are able to solve learning tasks within the IC and also other parameters such as activity patterns can be measured within the system (Galsworthy et al., 2005; Mechan et al., 2009; Krackow et al., 2010; Endo et al., 2011; Voikar et al., 2018). Home-cage based test systems offer the advantage of testing animals without daily interference of researchers and in accordance with the animals natural activity phases over a long period of time [reviewed in Voikar and Gaburro (2020)].

Before going into further detail, we wish to emphasize that this study comprises data from three initially independent Master theses which all focus on the costs of cognition. We reverted to the same group of mice in all theses which allowed us to compile a unique lifetime observation data set presented in the current study. In this set-up, we consequently had to face slight variations in experimental procedures due to the changing experimenters whereas the same overarching research question allows the data to be presented as a single long-term study with three consecutive phases. In summary, we present a unique and comprehensive lifetime observation of 48 male mice, which to our knowledge has no comparison in the present literature record.

MATERIALS AND METHODS

Animals and Housing Conditions

48 male C57BL/6J mice (Charles River, Sulzfeld, Germany) arrived at Osnabrück University at the age of 21 days. The mice were randomly separated in four housing groups, 12 mice per cage, which were kept stable over the lifetime of the mice. One day after arrival, the mice received an RFID transponder (radio frequency identification ISO FDX-B 2.12 × 12 mm, Planet ID GmbH, Essen, Germany), which was implanted subcutaneously into the neck area under isoflurane anesthesia. Each social group of 12 mice was further randomly assigned to one of three treatment groups: The learner mice (L), which had to solve various cognitive tasks, the non-learner (NL), which had no tasks to solve, and the equipment control group mice (EC), which also had no tasks to solve, but were exposed to handling similar to that of the L mice during testing. For visual identification, each mouse was assigned a unique two-color code which was applied to the tail skin with lacquer painting pens (edding 750).

During their growth phase up to an age of 73 days, body mass was measured every 3 days. Adult mice were weighted weekly during the process of cage cleaning up to an age of 757 days, except for a non-experimental phase between 192 and 302 days of age. During handling (tail-handling), the tail color-codes were additionally renewed if necessary. The mice were kept at 22 ± 2°C at 56 ± 15% humidity. The dark/light cycle was 12 h each, with light hours between 8:00 am and 8:00 pm. Pellet food (Altromin International, 1314) was available *ad libitum* at all times. At the age of 325 days, the mice moved to another animal facility within the university. The housing conditions were similar but with an additional half hour of sunrise simulation and the last half hour of the light cycle simulating sunset. Within these facilities mice were kept in different cage types throughout their life as described in detail below. Mice were kept in standard home-cages after the last experiment until they died a natural death. However, to keep animal welfare standards mice were euthanized as soon as signs of pain or suffering were detected.

IntelliCage

The IntelliCage (IC, NewBehavior) is a home-cage in which various cognition tasks can be applied to individual animals with minimal invasion by experimenters. Each IC contained bedding (Allspan, Olympia, 2 cm high), paper as nesting material and four red mouse houses ("TheMouseHouse;" Tecniplast) for shelter. The houses were placed directly under a central feeding rack. Water was available in eight dispensers arranged in the four cage corners (**Figure 1**). Each corner is only accessible to one mouse at a time. Each corner contains a presence sensor, one RFID antenna and an airpuff valve for mild punishment (0.5 bar). Access or denial to water can be granted to individual animals identified via an implemented RFID sensor, enabling individual operant conditioning in group housed animals. Each corner contains a nosepoke-sensor and a door per water dispenser. The doors can be opened by a nosepoke detected by the nosepoke-sensor. By assigning mice access to individual corners or water dispensers at certain times or in certain orders, tasks of different levels of difficulty could be applied. Nosepokes at non-rewarded dispensers were punished with a 1 s and 0.5 bar airpuff in certain conditions.

During their stays in the IC, different programs of various difficulty were applied to the L mice allowing and denying access to different corners or water dispensers (**Table 1**). As water was only offered in the cage corners, we assume that the mice were primarily entering the corners for fluid intake. Accordingly, attempts to visit corners and drink at certain water dispensers reflects the animals' foraging effort. Therefore, we use the term "foraging behavior" to describe the behavior of corner visiting for getting access to water. The IC tasks got more and more difficult over time including patterns in which the rewarded corner or side within a corner changed after each successful drinking event, like for example to the opposite corner or in a clock- or anticlockwise manner. The difficulty of the IC tasks was increased by allowing the L mice access to water in all four corners, but the L mice received an airpuff in addition to water in

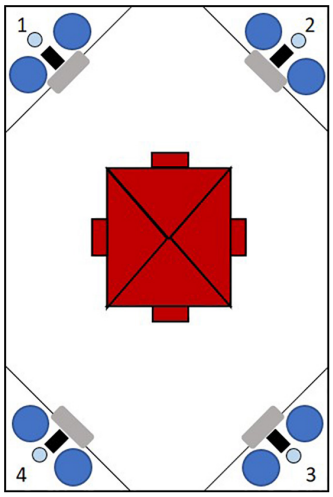

FIGURE 1 | Schematic figure of an IntelliCage from the top with four conditional corners marked by numbers. Each corner contains a radio frequency identification antenna (gray), a presence-sensor (black), an airpuff valve (light blue) and two water dispensers (dark blue). In the cage center four shelters (red) are arranged under the food rack.

TABLE 1 | Learning IntelliCage tasks assigned to learner mice within the IntelliCage.

IC Program	test duration (days)*	IC phase
Cornerlearning 1	14	1
Shuttling 1	12	1
Clockwise 1	11	1
Anticlockwise 1	14	1
Clockwise 2	9	1
Sidelearning	14	1
Cornerlearning 2	13	1
Shuttling 2	7	1
Clockwise 3	9	1
Anticlockwise 2	19	1
Clockwise 4	6	1
Anticlockwise 3	13	1
Clockwise 5	10	1
Complexclockwise 1	10	2
Complexshuttling	14	2
Complexanticlockwise 1	12	2
Clockwise 6	3	3
Complexclockwise 2	2	3
Complexanticlockwise 2	8	3

*See ***Supplementary Material*** for exact times in hh:mm:ss.*

three (incorrect) corners. Only in one (correct) corner, drinking was possible without receiving an airpuff. But drinking within the correct corner on the incorrect side leads to a corner change (for example clock- or anticlockwise). If the L mice drank on the correct side within the correct corner, the correct corner did not change (see Supplementary Material for full details on all IC cognition tasks). Some of the IC tasks were repeated, adding up to 51 learning tasks applied to the L mice in the IC. Both the NL and EC mice had access to water at all times but were punished by an airpuff if they stayed for longer than 15 s in a single corner to avoid a mouse occupying a corner and blocking the access for other individuals for too long. In addition, while the NL and EC mice had 15 s to drink, the L mice in the different IC tasks had 8–10 s to drink water within the correct corner or on the correct side. In order to drink again, all mice had first to leave the corner and re-enter it or visit another corner. In total, the mice were kept in the ICs in three phases of different lengths (**Table 1**, IC phase 1 = 32 tasks, IC phase 2 = 16 tasks, and IC phase 3 = 3 tasks). Even though the L mice had less time within the correct corner to drink and were not able to drink in all corners without punishment relative to the NL and EC mice, we compared the foraging behavior of the three treatment groups with each other. This was carried out by analyzing the number of corner visits that were performed (per week) by the mice. A visit was evaluated as soon as a mouse entered a corner, the RFID antenna registered the RFID transponder of the mouse and at the same time the presence sensor registered the presence of the mouse.

Standard Home-Cage

Mice were kept in type IV Macrolon cages (59 cm × 59 cm × 20 cm) per social group. Food and water were available *ad libitum*. Each home-cage contained two red houses, a plastic tube, and paper for nesting and bedding. No tests were applied directly in the home-cage, but mice were transferred to different tests over short time periods during their stays in the home-cage (for an overview of the timeline see **Figure 2** and **Supplementary Table 1**).

Barnes Maze Test

The Barnes maze test (BM) was carried out to test for spatial learning performance of L mice outside the IC. It uses the natural aversion of mice to open and exposed areas (Crawley, 1985) and measures how many errors they make to escape from such an area over the course of multiple trials. We used a round gray plastic platform of 100 cm diameter mounted about 120 cm above the ground as the exposed area. The platform contained 12 holes of 4 cm diameter evenly distributed along its edges, which could be closed via opaque black plexiglass lids. One hole could be connected via a 4 cm diameter PVC tube to a cage (Makrolon type III, filled with bedding transferred from the cage the mouse was housed in) placed under the platform. Visual cues (a bottle, a cloth, a metal stand, and a box) arranged around the platform served as spatial orientation cues. The test was performed during the light phase at 22 ± 1°C and with 54 ± 4% humidity in the experimental rooms. The illumination on the platform was 10–20 lux. Prior to each trial, the platform was cleaned with ethanol (70%).

For each trial, a single L mouse was placed in a start-cylinder (diameter: 10.5 cm, high: 21 cm) in the middle of the platform for 1 min. After this acclimatization phase, the start-cylinder was removed and the trial started. The number of attempts to enter covered holes to escape from the platform were evaluated from video recordings (Logitech HD webcam, Software: VirtualDub Version 1.9.11) to investigate learning performance. The BM test was performed twice at the age of 92 and 573 days. In the first run

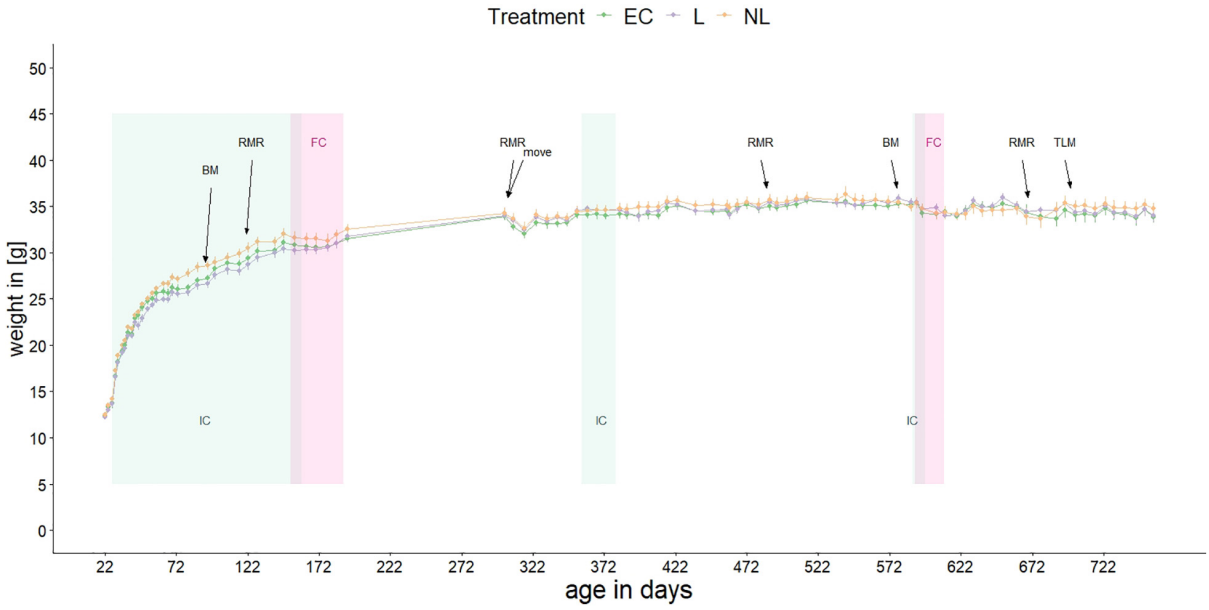

FIGURE 2 | Body mass of mice over their lifetime. Shown are the mean body masses of the three treatment groups. At 305 days of age, the mice moved to another animal facility. Repeated experiments were conducted: IC, IntelliCage; FC, female choice test; BM, Barnes maze test; RMR, resting metabolic rate; TLM, telomere length measurement. When not in the ICs, the mice were kept in standard home-cages within their social groups.

each trial lasted 5 min. Each mouse performed ten trials, two trials per day with an inter-trial interval of 30 min on five consecutive days. In the second run the test was performed in a different room and with different positions of the escape hole and the spatial cues surrounding the platform, but the same apparatus with the same illumination conditions was used. For the second run, each mouse performed six 3-min trials, two trials per day with an inter-trial interval of 30 min over three consecutive days. In both runs, if a mouse did not find the open hole within the given time, it was gently guided to the escape hole by the researcher's hand. To control for a possible handling-effect, each L mouse was assigned to an EC mouse. After the L mouse finished the BM test, the assigned EC mouse was placed on the platform of the maze for exactly the same time it took the L mouse to find the open hole. Contrary to the L mouse, the matched EC mouse could not escape from the platform, as all 12 holes were closed with the black opaque plexiglass.

For further cognitive stimulation outside the IC, a T-maze test, a tone conditioning test (carried out during the first IC phase) and a labyrinth experiment (after the third IC phase) were carried out. However, the results of these tests were not sufficiently conclusive to be included in this work due to methodological problems (e.g., we only know now why the T-Maze did not work: Habedank et al., 2021). Nevertheless, it is important to mention that the control group had other experiences outside the home-cage besides the BM test (for more information about the test procedure see the supplements).

Female Choice Tests

To test for possible differences of attractiveness to females between the three treatment groups of mice, the males were subjected to two female choice tests (FC), one at an age of 152 days (shortly after IC phase 1) and a second one at an age of 590 days (several months after IC phase 3). Females used in the first test were C57BL/6J mice naive to cognitive experiments (N = 16) and 152 days old. Females in the second test were the first generation offspring of C57BL/6J X BALB/C and subjected to cognitive tests including an IC phase themselves prior to the choice tests (N = 12, each female conducted a maximum of two choice tests). At the time of testing the females were 304 days old. The estrus status was not determined. Nevertheless, we must note that the estrus status may affect the female's behavior. In each test, one female was introduced into a type III Makrolon cage filled with fresh bedding for a 1 h habituation time. Accordingly, three males, one of each treatment group were randomly placed in three similar cages divided in half by a perforated clear plexiglass wall. The female's cage was connected to the three empty half compartments of the males' cages via PVC tubes. The location of the female and the time spent with each male over the course of 24 h was automatically recorded using light barriers installed in each tube.

Observations of Agonistic Behavior

To test whether the outcome of agonistic encounters between mice within each social group was influenced by the treatments, live observations on fighting behavior were conducted. Mice within the social groups in general lived rather peacefully together throughout their whole lives. At no time any mouse had to be removed from a group due to social incompatibility. The only times when agonistic behaviors occurred more frequently were during the weekly cage cleaning events, when mice were transferred to clean cages. Cage cleaning and accordingly live

observations of agonistic behavior took place once a week between 08:30 am and 11:30 am. The mice of one social group were placed in random order into an observation cage containing only fresh bedding, food, and water. One minute after all mice were placed in this cage, they were observed by an experimenter for 30 min. Within this time, the mice usually calmed down and agonistic behavior almost ceased to occur. The experimenter noted down the IDs of mice involved in fights as well as who won and lost each fight. The behavior *fight* began as soon as two mice began to circle one another with body contact. During a *fight*, the mice pushed one another with their paws and bodies or bit, sometimes even pushing the opponent with its back to the ground. *Fights* could be accompanied by vocalizations, too. A *fight* ended when one of the two opponents turned its head away from the other. The loser mouse was the one which turned away first. Fights between three or more mice were not included in the data set, as identification of individuals as well as the determination of winners and losers were less straightforward. Observations of agonistic behaviors were done when the mice were between 373 and 759 days old.

Resting Metabolic Rate

To test whether the different treatments of mice led to differences in their metabolism, respirometry measurements were repeatedly taken throughout their life. Respirometry is a method for determining the total energy turnover of an organism. The release of carbon dioxide is determined in relation to oxygen absorption. This allows determining the metabolic rate of an organism. The resting metabolic rate (RMR) measurement was carried out at the Department of Animal Behavior at Bielefeld University. The measurement was performed at 20 ± 1°C under low light conditions in an open flow system by measuring the oxygen consumption and carbon dioxide production with a continuous inflow (45 l/h, Mass Flow Meter FM-360, Tylan Corp., Torrance, CA, United States) of external fresh air. External air was transferred under ambient pressure to two transparent Plexiglas measuring chambers (14 cm × 20.5 cm × 14 cm), which were placed in a climate chamber (Rubarth Appaerate, Laatzen, Germany) and contained paper (for excretion absorption) but no water or food. Both chambers were located in a way that animals could not see, hear, or smell each other. For drying, the air was first pumped to two cooling devices (M&C Cooler, Ratingen, Germany) and then transferred to a molecular sieve. The oxygen consumption and carbon dioxide production were analyzed by an O_2 analyzer (Oxzillar FC, Sable Systems, Henderson, NV, United States) and CO_2 analyzer (Maihak AG, Hamburg, Germany). As a control, we compared dried outside air against the carbon dioxide and oxygen concentrations measured in the outflow of the metabolic chambers, in which the test animals rested. An initial control period of 10 min was used to assure the stability of the system. Over a period of 2.5 h, each animal's oxygen consumption and carbon dioxide production were measured across six periods of 10 min each. Between measurements, 1-min control intervals were interspersed to allow correction for system drifting if necessary. The specific RMR ($KJ/(d^*kg)^{-1}$) was calculated from the 3-min interval with the lowest, stable oxygen consumption throughout the measurement periods. Within the apparatus, two animals were measured simultaneously between 08:30 am and 05:00 pm. The evening before the measurement the mice were separated in type III Macrolon cages (with nesting and bedding, water and food from the home-cage) to habituate them to being separated during measurement. Directly before and after the measurement, body mass was measured. After RMR measurement, the mice were placed back into their home-cages. The RMR was measured four times, at an age of 138, 308, 482, and 665 days.

Telomere Lengths Measurement

To investigate whether repeated cognitive stimulation affects the length of telomeres, from all 44 mice still alive (15 L mice, 14 NL mice and 15 EC mice) at an age of 699 days, a blood sample was taken from the tail vein. Previous studies showed a link between telomere length and aging, age-related diseases or cognitive aging (Von Zglinicki, 2002; Brouilette et al., 2003; Benetos et al., 2004; Martin-Ruiz et al., 2006; Yaffe et al., 2011). By repeated cognitively stimulating the L mice, we assumed a difference in telomere length between the three treatments. Therefore, in our study telomere length should serve as a longevity marker (in addition to the classical survival analysis). DNA was extracted out of the leukocytes according to the instructions of the UltraClean® Blood DNA Sample Kit (Non-Spin, MO BIO Laboratorie, Inc.) and used in a telomere real time quantitative polymerase chain reaction (RT qPCR) performed with a BioRad PCR system consisting of the BioRad CFX96™ RealTime System as optical reaction module and BioRad C1000 Touch™ Thermal Cycler following Cawthon et al. (2003) and Callicott and Womack (2006). From each sample DNA and from a control DNA (single copy gene 36B4), three replicates were prepared. The dilutions of the sample DNA were: 100, 20, 4, 0.8, and 0.16 ng/5 µl. The dilutions of the control DNA were: 20, 4, 0.8, 0.16, and 0.032 ng/5 µl. In addition, a zero-control (Null Template Control-NTC) was prepared with three replicates. 10 µl GoTag®qPCR Master Mix, 1.8 µl telg-Primer, 1.8 µl telc-Primer and 1.4 µl PCR-water were added to the DNA samples. 10 µl GoTag®qPCR Master Mix, 1.8 µl forward-Primer, 1.8 µl reverse-Primer and 1.4 µl PCR-water were added to the control samples. The relative telomere length was then determined from the C_T value of the sample DNA and the C_T value of the control DNA.

Data Analysis

The following data analysis and visualization was performed with R (R Core Team, 2020; version 3.4.2) and Python (version 3.8), using the panda package (version 1.2.3).

Body Mass Measurement

For body mass analysis four linear mixed effects models (package nlme, Pinheiro et al., 2020) with weight (log transformed), housing condition (housed in the IC vs. housed in the home-cage) and age set as fixed effects were calculated for each of the three IC phases and in the home-cage. Animal ID served as a random effect. Model assumptions were checked visually by Q–Q plots and by plotting fitted versus residual values.

To test for effects between treatment groups within housing conditions, pairwise *post hoc* comparisons were conducted (package emmeans, Lenth, 2020).

IntelliCage

For each task in the IC, the performance of the L mice was determined by calculating the ratio of successful trials over all trials. For experiments where the task depended on a corner condition (i.e., choosing the correct corner), visit events were used as a measure for a trial, whereas for experiments where the task depended on a side condition (i.e., choosing the correct side within a corner), nosepoke events were used as a measure for a trial. A criterion of cognition was met when the proportion of successful trials within a task for a mouse was higher than 1.25 times chance level (e.g., 31.25% correct for cornerlearning), demonstrating that there was an understanding of the task.

The foraging behavior (number of corner visits) in the IC was analyzed in a Poisson GLMM (package lmerTest, Kuznetsova et al., 2017) including treatment group and phase as fixed effects and animal ID as random effect. To test for effects between treatment groups and between phases, a pairwise *post hoc* analysis with Tukey adjustment was conducted (package emmeans). Due to technical issues of the first part of the first IC phase, the changes in task objective during "cornerlearning 1" did not consistently change at the specified timestamp. Instead, for some L mice the correct drinking corner changed at a different moment than for other mice. Therefore, we decided to ignore the first 30 min of each "cornerlearning 1" task.

Barnes Maze Test

The two runs of the BM were evaluated separately, as they differed in execution (first run: 10 trials of 5 min each, second run: 6 trials of 3 min each). The number of errors was analyzed in a Poisson GLMM (package lme4, Bates et al., 2015) including errors and trial as fixed effects and animal ID as random effect. To test for effects between trials, a pairwise *post hoc* analysis with Tukey adjustment was conducted (package emmeans). Residuals of the model were visually inspected for homogeneity of variances and normal distribution by using QQ plots.

Female Choice Tests

Including only females which choose one of the males (i.e., which spent less than 50% of time alone in the starting cage and/or had less than 5% difference in the amount of time spent with their first versus second choice male) in the analysis, it comprised of 14 (out of 16) choices in the first test and 12 (out of 16) choices in the second test, respectively. For these females, we calculated the percentages of time spent with each specific male from the whole time a female spent in cages with males. These percentages were then analyzed in a GLMM (package lme4) including male treatment group, cage orientation within the test set-up and male mass rank (1–3) within the trio of males as fixed effects and female identity as random effect. In addition, we ran a model in which male mass rank was replaced by mean-centered male body mass to test for the effect of male body mass within the given set of males for each test and, more generally, across all tested males. When the model indicated a significant main effect of male treatment or cage, we conducted a *post hoc* comparison on the fixed effect using the false-discovery rate to adjust *p*-values (package emmeans).

Observation of Agonistic Behavior

The number of fights each mouse was involved in during the whole time period of life observations was calculated along with the number of fights this individual won. For each of the two variables, we ran a GLM (package lmerTest) to test the effect of treatment, cage and mean weight over the observation period.

Resting Metabolic Rate

For RMR analysis a linear mixed model (package lme4) was carried out with RMR, treatment and RMR test run as fixed effects and an interaction of treatment and RMR test run. Animal ID served as a random effect. If the model indicated a significant effect of treatment or test run, we conducted a pairwise *post hoc* analysis with Tukey adjustment (package emmeans).

Telomere Length

By the time of blood collection for telomere length measurement (699 days of age), 44 out of 48 mice were still alive (15 L mice, 14 NL mice, 15 EC). A linear model was conducted to analyze a possible effect of treatment on telomere length. Treatment was set as fixed effects (continuous effects). To obtain normally distributed data, the relative telomere length was transformed by logarithms.

Survival Analysis

For survival analysis, a Cox proportional-hazards model (Coxph) was calculated (package survival; Therneau and Grambsch, 2000). Treatment and telomere length were included as fixed effects in this model.

RESULTS

Description of Learning Outcomes

IntelliCage

The L mice were cognitively trained during three IC phases requiring them to perform various learning tasks to get access to water as reward (see **Supplementary Information**). To confirm that the mice were indeed cognitively stimulated by the tasks, their performance on each task was investigated. The mice performance was evaluated by calculating the percentage of successful trials over the entire task duration. To illustrate an example of the learning behavior of the group and the magnitude of inter-individual variability of the mice, the learning curve for the first learning task (cornerlearning 1) is shown in **Figure 3A**. A summary of task performance for the first IC phase per mouse is given in **Figure 3B**. The learning criterion was set to 1.25 × chance level. If a mouse did not reach this threshold, we assumed this mouse did not understand the underlying task's structure. Each mouse was able to learn at least 24 out of the 32 tasks, demonstrating that the mice were overall cognitively stimulated. The average performance of the mice per task is summarized in **Figure 3C** for the first IC phase.

During the second and third IC phase, more complex tasks were presented, which are shown in **Figure 4**. Again, the learning

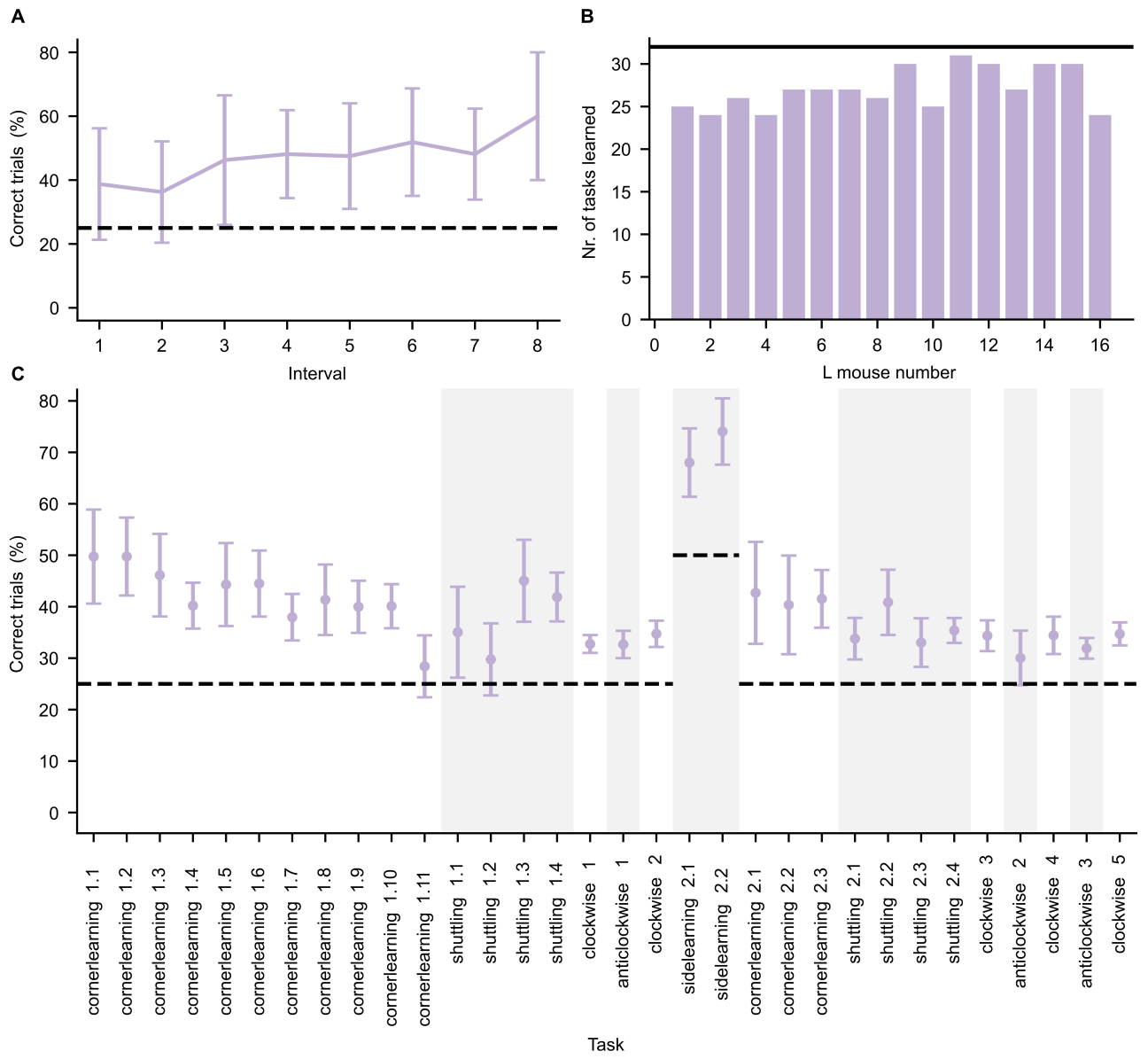

FIGURE 3 | Performance during the first IntelliCage phase. **(A)** Average performance (±standard deviation, $n = 16$) of all L mice for the first day of the cornerlearning task, divided in intervals of 10 trials per interval. Since the least active mouse performed less than 90 trials, only the first eight intervals are shown. The increase of performance over time reflects learning in the mice. **(B)** Number of tasks for which the L mice successfully achieved the learning criterion. The black horizontal line represents the total number of tasks. **(C)** Average performance (±standard deviation, $n = 16$) per learning task. Changes in task type are separated by alternating white and gray background color. The learning tasks are represented in chronological order. The dashed lines in panels **(A,C)** represent the chance level of correct performance in each task.

curve for the first learning task (complexclockwise 1) is shown in detail in **Figure 4A**. In contrast to the learning performances during the first IC phase, during the second and third IC phases the mice often did not achieve the learning criterion (**Figure 4B**). As illustrated in **Figure 4C**, the mice performed roughly at chance level for the tasks.

In **Figure 5** the daily foraging behavior of the three treatment groups in the IC are shown (see **Supplementary Figure 1** for the daily number of licks per treatment group). During the first phase, the foraging behavior of the L group was significantly higher (GLMM, $p < 0.001$) than the other groups. During the second and third phase the foraging behavior of the L and NL groups was almost identical (GLMM, $p > 0.05$), whereas the difference in foraging behavior between the EC and L group and EC and NL group were significantly different (GLMM, $p = 0.0025$ and $p < 0.0014$, respectively, for phase 2, $p = 0.03$ and $p = 0.029$, respectively, for phase 3). Furthermore, all treatment groups were significantly less active in subsequent phases (GLMM,

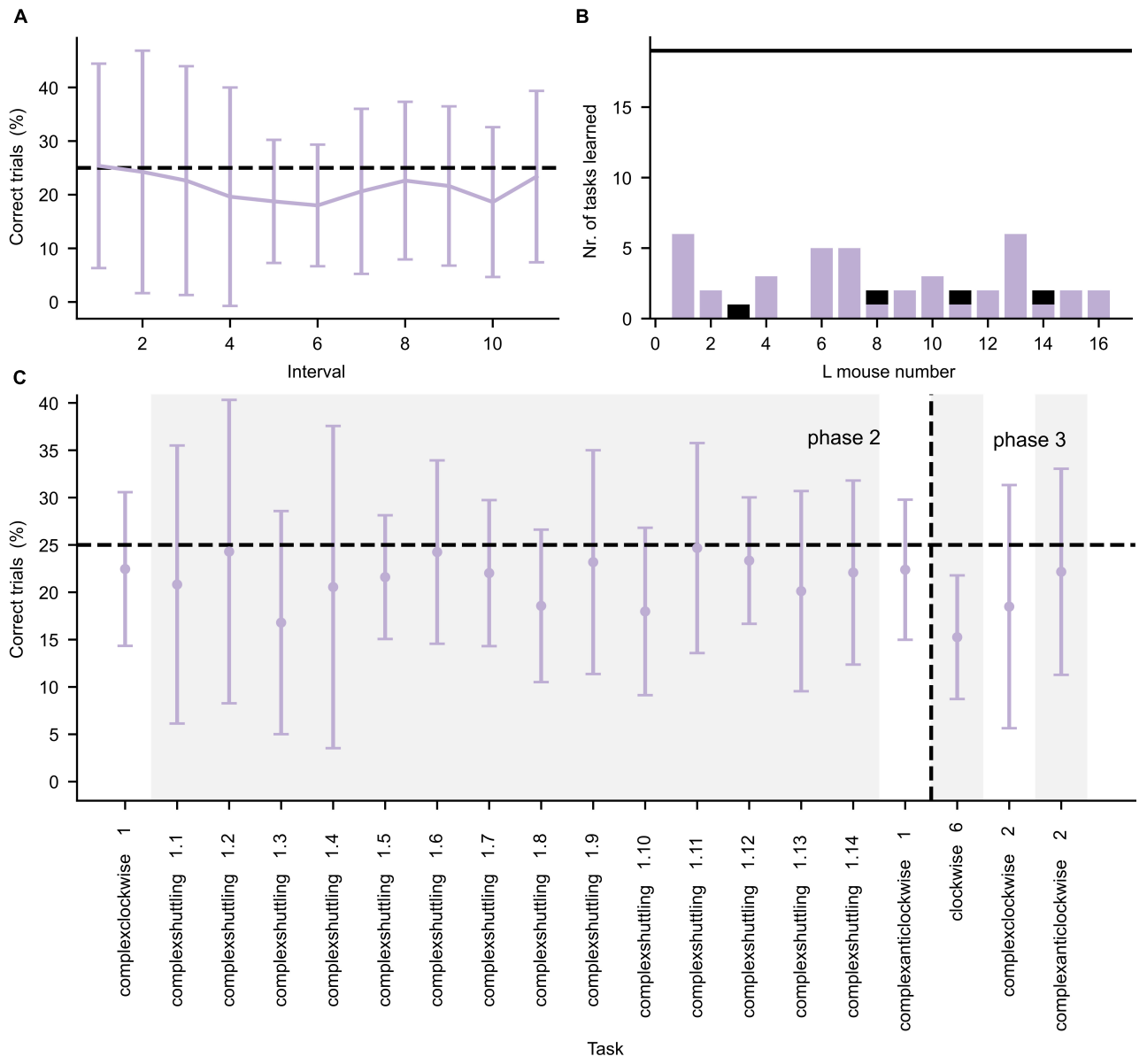

FIGURE 4 | Performance during the second and third IntelliCage phase. **(A)** Average performance (±standard deviation, $n = 16$) of all L mice for the first 550 trials of the Complexclockwise 1 learning task, divided in intervals of 50 trials per interval. **(B)** Number of tasks for which the L mice successfully achieved the learning criterion. The black horizontal line represents the total number of tasks. The purple and black bars represent successful tasks during phase 2 and 3, respectively. **(C)** Average performance (±standard deviation, $n = 16$) is given by the error bar data points. Changes in task are separated by alternating white and gray background color. The learning tasks are represented in chronological order. The horizontal dashed lines in both panels **(A,C)** represent the chance level of correct performance in each task. The vertical dashed line in panel **(C)** represents the separation between IC phase 2 and 3.

$p < 0.001$) except for the NL treatment group phase 2 and 3 ($p > 0.05$).

Barnes Maze Test

Spatial orientation and memory were investigated using the Barnes maze twice. L mice were able to learn and improve their performance in both runs (**Figure 6**). In the first run, the mice made significantly more mistakes in the first trial compared with all following trials (GLMM, $p < 0.001$, *post hoc* comparison).

The second run showed a similar picture for the number of errors made. The steepest decrease was observed between trial 1 and 2 (GLMM, $p < 0.001$).

Effects on Physiology/Body Maintenance

Body Mass

Body mass of the mice was measured between 21 and 757 days of age with a gap between day 192 and 302 (**Figure 2**). At the beginning of the study the mean body mass was 12.37 g, at the

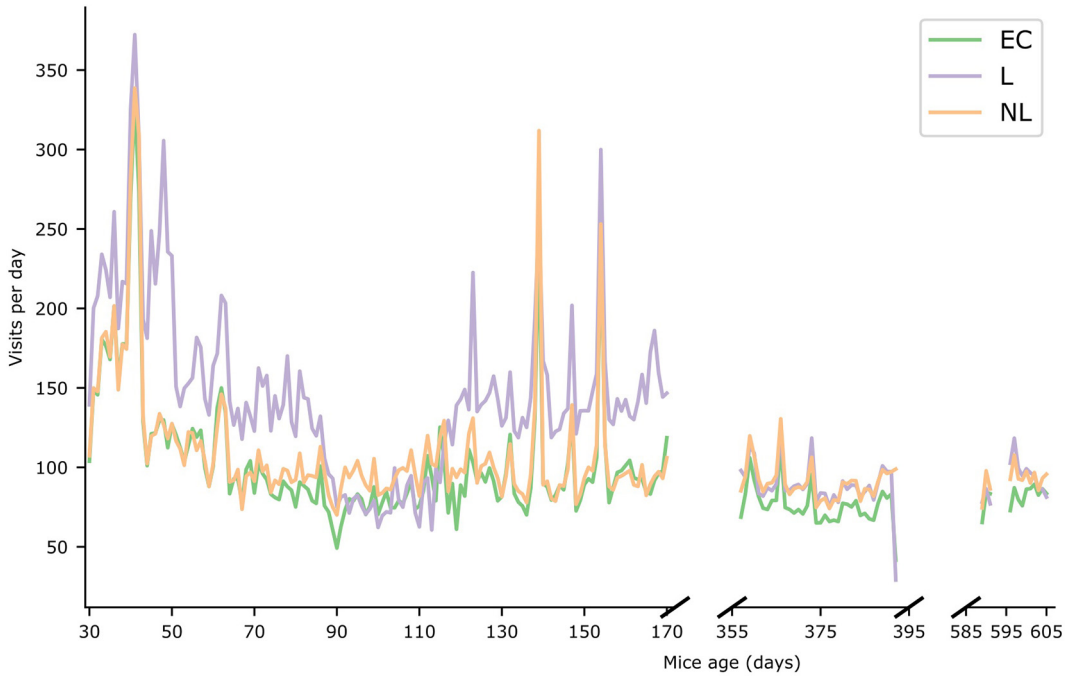

FIGURE 5 | Foraging behavior. Number of visits per day during the three IntelliCage phases. The X-axis represents the age of the mice (in days). The average value of the three different treatment groups are shown for the three IntelliCage phases.

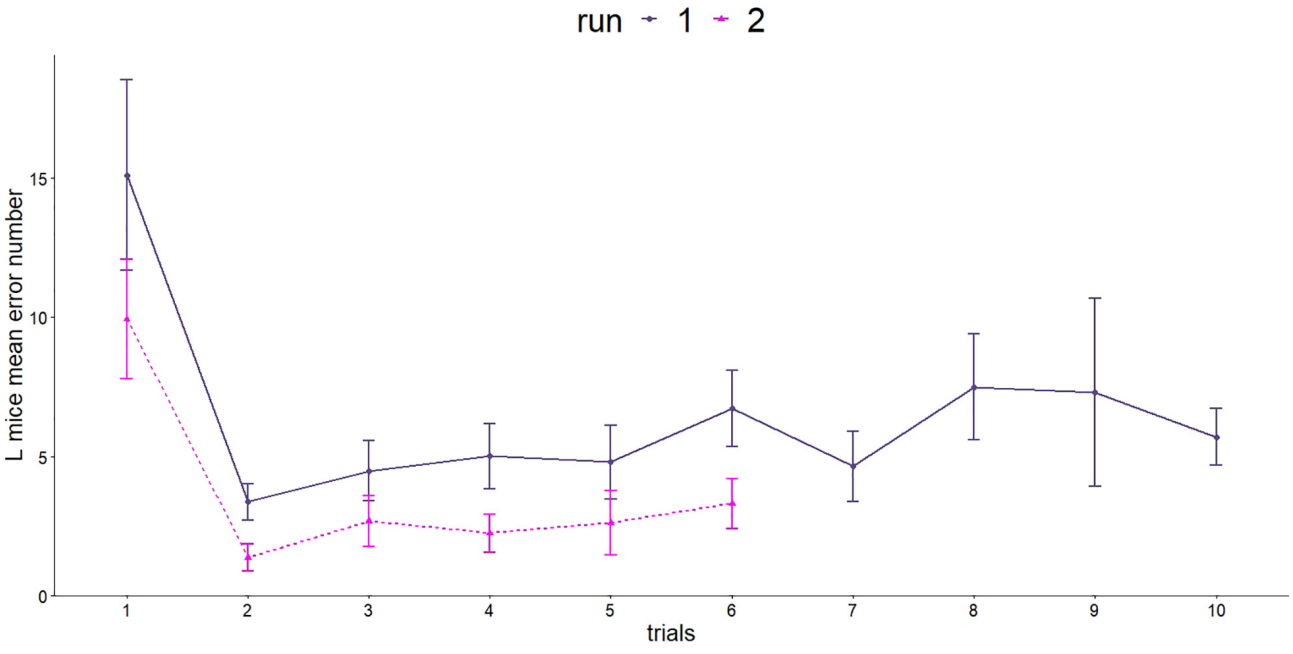

FIGURE 6 | Number of errors made in the Barnes maze test. Since the two runs were carried out differently (number of trials and length of trials), a separate statistical evaluation was carried out. First run, dark purple line: n = 16; age = 92 days; second run, purple line: n = 16, age = 573 days.

end of the measurement 34.12 g. When the mice were housed in the IC's for the first time, the body mass differed between L and NL mice (GLMM, *post hoc* comparison, $t = -2.71$, $p = 0.025$) while there was no difference between L and EC mice (GLMM, *post hoc* comparison, $t = 1.08$, $p = 0.531$) or between NL and EC mice (GLMM, *post hoc* comparison, $t = -1.63$, $p = 0.243$). When the mice were housed in the IC for the second and third time as well as in their home-cages, no treatment effect on weight

was detected (GLMM, *post hoc* comparison, IC phase 2: EC vs. L: $t = -0.88$, $p = 0.659$; EC vs. NL: $t = -0.70$, $p = 0.764$; L vs. NL: $t = -0.17$, $p = 0.984$; IC phase 3: EC vs. L: $t = -0.27$, $p = 0.960$; EC vs. NL: $t = -0.24$, $p = 0.969$; L vs. NL: $t = -0.03$, $p = 1$; home-cage: EC vs. L: $t = -0.03$, $p = 1$; EC vs. NL: $t = -0.91$, $p = 0.640$; L vs. NL: $t = -0.89$, $p = 0.657$).

Resting Metabolic Rate

A total of four measurements were performed to investigate the influence of learning treatment on RMR. When the mice were housed in the IC's, the RMR differed in the GLMM *post hoc* comparisons between L and EC mice (**Figure 7**, first RMR measurement: $t = 2.58$, $p = 0.04$), while there was no difference between L and NL ($t = -1.16$, $p = 0.25$) or EC and NL ($t = 1.43$, $p = 0.24$). When the mice were housed outside the IC's, no treatment effect was detected (GLMM, *post hoc* comparisons, second RMR measurement: L and EC: $t = -1.08$, $p = 0.61$; EC and NL: $t = -0.24$, $p = 0.81$; L and NL: $t = 0.84$, $p = 0.61$; third RMR measurement: EC and L: $t = -0.92$, $p = 0.61$; EC and NL: $t = -0.84$, $p = 0.61$; L and NL: $t = 0.07$, $p = 0.94$; fourth RMR measurement: EC and L: $t = 0.003$, $p = 0.8$, EC and NL: $t = -0.68$, $p = 0.75$, L and NL: $t = -0.68$, $p = 0.75$). While the RMR did not differ between the first two (GLMM, $t = 0.42$, $p = 0.68$) and last two measurements (GLMM, $t = 0.53$, $p = 0.68$), it increased significantly between the second and third measurement (GLMM, $t = 7.0$, $p < 0.001$).

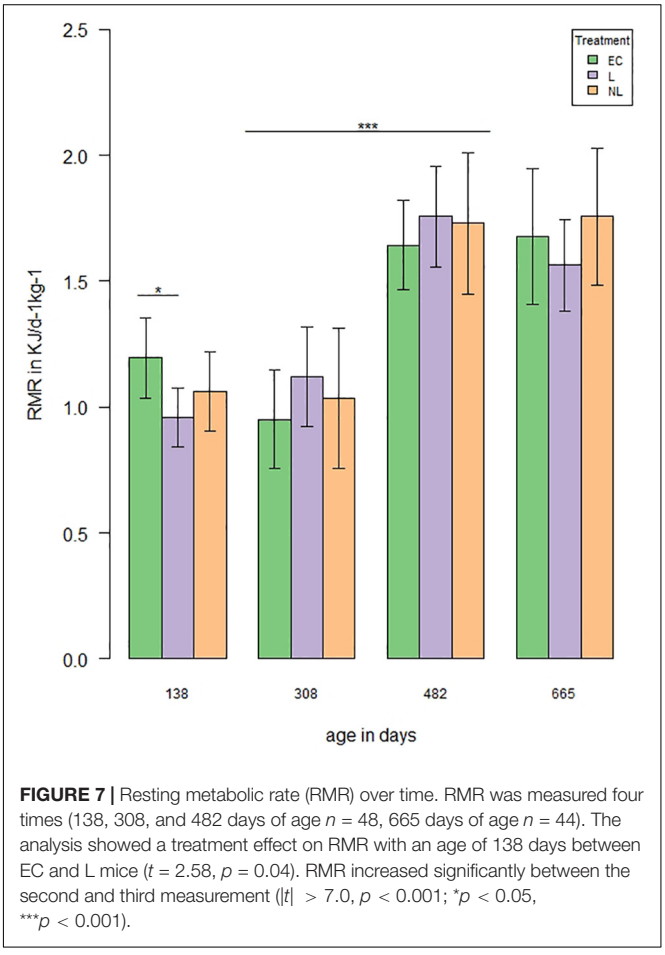

FIGURE 7 | Resting metabolic rate (RMR) over time. RMR was measured four times (138, 308, and 482 days of age $n = 48$, 665 days of age $n = 44$). The analysis showed a treatment effect on RMR with an age of 138 days between EC and L mice ($t = 2.58$, $p = 0.04$). RMR increased significantly between the second and third measurement ($|t| > 7.0$, $p < 0.001$; $^{*}p < 0.05$, $^{***}p < 0.001$).

Effects on Sexually Selected Traits

Female Choice Tests

In the first female choice test, statistical analyses did not indicate a cage bias (GLMM, $t = 1.03$, $p = 0.31$), or an effect of male weight calculated as either body mass (GLMM, $t = 0.40$, $p = 0.69$), measured directly before the test, or mass rank within the trio of males (GLMM, $t = -1.01$, $p = 0.32$). Female mice preferred NL males over L males (GLMM, *post hoc* comparison, $t = 2.47$, $p = 0.036$, **Table 2**) and EC males (GLMM, *post hoc* comparison, $t = 2.41$, $p = 0.036$) but did not differentiate between EC and L males (GLMM, *post hoc* comparison, $t = 0.14$, $p = 0.89$).

The second mate choice test likewise indicated no cage bias (GLMM, $t = 0.49$, $p = 0.62$) and no effect of body mass (GLMM, $t = 0.55$, $p = 0.60$) or mass rank within a trio (GLMM, $t = 0.45$, $p = 0.66$) on the choice of the females. In contrast to the first mate choice test, we found no indication for a preference of males of a specific treatment group (GLMM, $t = 0.55$, $p = 0.58$, **Table 2**).

Observations of Agonistic Behavior

Each cage of mice was observed for 26 h in which mice were involved in a mean number of 48 ± 4 fights. Mice of the L group were involved in fights significantly more often than the NL mice (GLM, *post hoc* comparison, $p = 0.011$) while no difference was found between NL and EC mice (**Table 2**). Additionally, heavier mice were more likely involved in fights (GLM, $z = 3.94$, $p < 0.001$). While heavier mice at the same time were more likely to win fights (GLM, $z = 0.02$, $p = 0.034$), no effect of treatment was found with regard to winning or losing fights (GLMM, *post hoc* comparisons, EC - L: $z = -1.23$, $p = 0.434$, EC-NL: $z = -0.49$, $p = 0.875$, L - NL: $z = 0.71$, $p = 0.756$).

Effects on Longevity

Telomere Length

For telomere length measurement 44 out of 48 mice were still alive. From these 44 mice 15 were L mice, 14 NL mice, and 15 EC mice. The treatment (**Figure 8**) did not influence the relative telomere length of the mice at an age of 699 days (linear model, adjusted $R^2 = 0.04$, $F = 1.84$, $p = 0.17$).

Survival Analysis

On average, the mice reached an age of 835 days. The shortest living mouse (EC) died at an age of 547 days, the oldest one reached an age of 1,218 (NL) days. Calculating a Coxph model, we investigated whether treatment or telomere length at the age of 699 days influenced longevity (**Figure 9**). All 44 mice which reached this age were included in the analysis. The model revealed no influence of the treatment on survival (EC vs. L: $z = -0.6$, $p = 0.548$; EC vs. NL: $z = -1.8$, $p = 0.072$; L vs. NL: $z = -1.2$, $p = 0.231$) and also no linkage between telomere length and survival (Coxph, $z = 0.3$, $p = 0.780$).

TABLE 2 | Effects of IntelliCage treatment on sexually selected traits.

Task	Behavioral response	EC	NL	L
Female choice I (age: 172 days)	% time spent with male	29.7 ± 5.0	**44.2 ± 5.2**	26.1 ± 5.1
Female choice II (age: 603 days)	% time spent with male	26.9 ± 8.7	35.7 ± 8.8	37.4 ± 8.7
Observations of agonistic behavior	# of fights involved	47.9 ± 5.8	42.8 ± 5.8	52.9 ± 5.8
	# of fights won	23.6 ± 4.2	22.7 ± 4.5	25.6 ± 2.9

Estimates in bold font indicate a significant difference from other treatment levels.

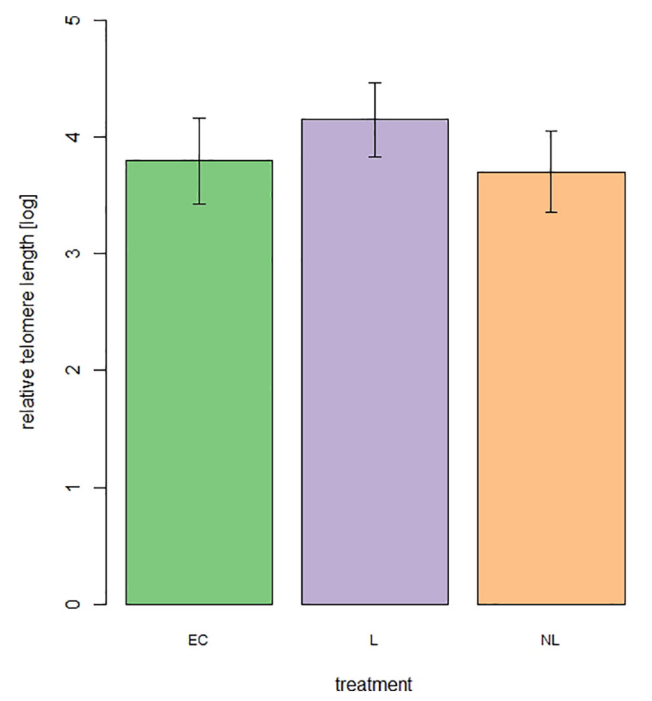

FIGURE 8 | Relative telomere length of male mice at the age of 699 days (n = 44 mice, 15 = L, 14 = NL, 15 = EC).

DISCUSSION

In this study we provide a unique observation of mice throughout their lives, from the juvenile stage to the natural death of the animals. By observing the mice during three consecutive phases, we obtained a lifetime data set. Our aim was to investigate if and how periods of cognitive stimulation affect animals on short term and over their lifetime. To produce cognitive stimulation, we used an automated IC system in which some mice in the social group were constrained by specific learning paradigms to where they could access water. Based on these environmental differences, we investigated a multitude of behavioral and physiological traits throughout the animals' lives as well as potential effects on longevity. Despite some problems with the automated testing system, the exposure of young mice to cognitive stimulation resulted in several immediate behavioral and physiological effects such as elevated foraging behavior, slower growth and a lower RMR. However, these effects did not manifest in long-term consequences. Re-exposing mice to new learning routines later in life did not result in similar learning success as in young mice. Accordingly, we also did not observe behavioral or physiological differences resurfacing during these later stages of cognitive stimulation. The explanation for no differences later in life could be that the second and third IC phases were much shorter compared to the first phase. Originally, it was planned to cognitively stimulate the L mice repeatedly, so as to induce a cognitively demanding life. Since the first phase resulted in a more obvious stimulation compared to the second and third phases, we assume that the first phase (while the mice were young) had a greater impact than the two shorter IC phases at later ages.

Several differences between the groups emerged during the first IC phase, when L mice were cognitively stimulated and most successful in learning. The L mice showed a higher foraging behavior in this phase, possibly resulting from searching for the correct drinking corner, while the NL and EC mice had access to water in all corners at all times. In addition, the L mice had less time to drink per visit than the EC and NL mice. The doors within the correct corners opened after a nosepoke for L mice for 8–10 s (depending on the IC program). For EC and NL mice the doors in all corners opened for 15 s after a nosepoke. To drink sufficiently, the L mice therefore had to visit the correct corners more often. The higher foraging behavior might have contributed to the slower weight gain compared with NL mice. However, the body mass of EC mice was comparably lower as for the L mice. At this time of life, the individual weights of mice are most variable (Eisen, 1976) and experiencing early life stress (Clutton-Brock et al., 1992; Lindström et al., 2005; Douhard et al., 2013) as well as the development of increased cognitive capacity (Kotrschal et al., 2015) have been shown to be accompanied by slower growth rates. L and EC mice were repeatedly placed on test set-ups by tail handling, which is known to increase stress and anxiety (Gouveia and Hurst, 2013; Ghosal et al., 2015). Since the NL mice were the least affected by experimental procedures, had unrestricted access to water, and were not cognitively stimulated, they likely had the most energy available to invest in growth, which resulted in a higher body mass gain. No difference in weight gain between the L and EC mice might indicate that the experienced stress in these two groups of mice might have been more influential as compared to a direct influence of the cognitive stimulation on the growth rate of the mice. The failure to find the same difference in older mice can be due to the much shorter duration of the IC phases 2 and 3.

By measuring the RMR, we examined a further physiological parameter and assumed that cognitively stimulated mice might have a higher energy turnover compared to non-stimulated

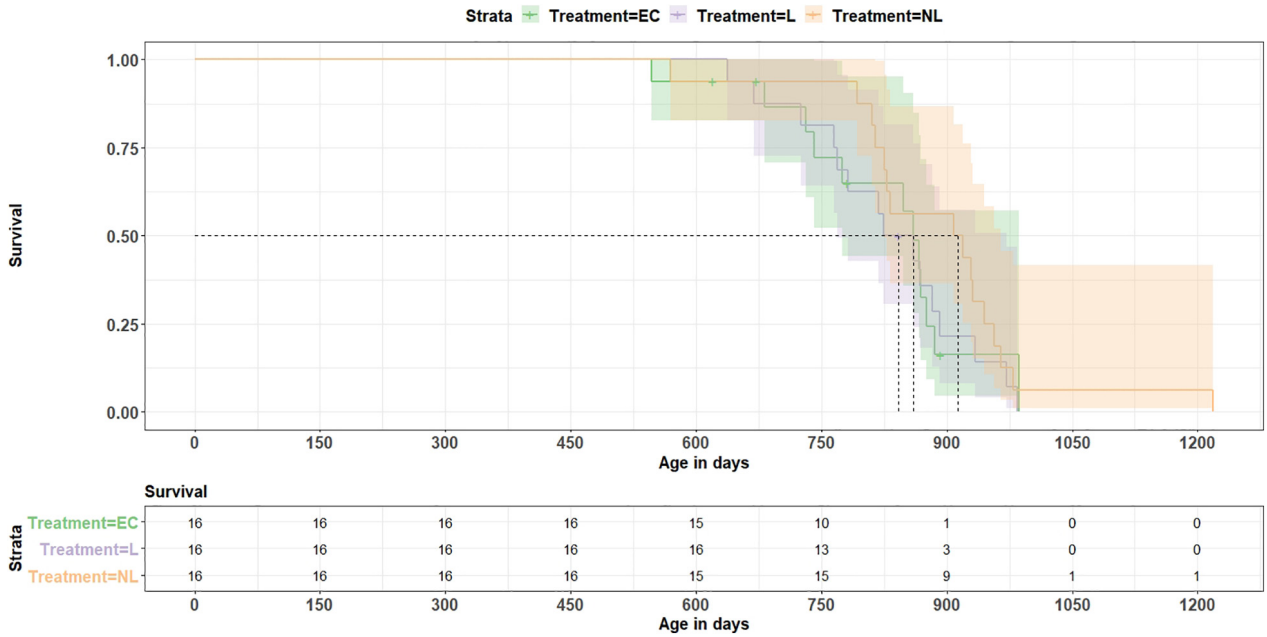

FIGURE 9 | Survival analysis. The data show no treatment effect (EC = equipment control group, L = learner, NL = non-learner) on survival ($p = 0.43$). Dashed lines indicate median length of survival for each treatment group.

mice. The RMR was measured at four different time points but differences between treatments were only found for the first measurement, when the mice were housed within the ICs. The L mice had a lower RMR compared to the EC mice, while EC and NL mice did not differ from each other. RMR is often referred to as the energetic cost of an organism's self-maintenance (McNab and Eisenberg, 1989; Speakman et al., 2004). Schubert et al. (2008) showed that increased foraging behavior resulted in reduced body mass and RMR. The L mice in our study also showed higher levels of foraging behavior and reduced body mass and RMR. The reduced RMR could be explained by an energy-saving strategy (Moe et al., 2007; Mathot et al., 2009), since a low RMR is commonly associated with demanding life-stages (e.g., under environmental stress or during reproduction). However, the energy-saving strategy did not seem to be necessary for a longer period of time, as the RMR differences were no longer present in the following measurements. This could be due to the fact that there were no more differences in foraging behavior or body mass. The underlying reason may be that the L mice were no longer successfully learning in the following IC phases. Interestingly, the RMR in all treatments increased with age.

Aging is characterized by declines in all physiological processes and concomitant changes in body composition. Age-related changes in body composition and physiological function are commonly reflected in a reduced metabolic rate in older individuals (Tzankoff and Norris, 1977; Piers et al., 1998). However, mixed patterns have been described depending on the species observed, the type of metabolic rate and the environment investigated (Elliott et al., 2015). Effects of the physical environment, the sex and the food availability have not been investigated systematically yet and the mechanism of metabolic aging is not well understood yet (Moe et al., 2007; Elliott et al., 2015). Here, we found an increase in metabolic rate with age, comparable to what has been found in rats from an age of 18 months on (McCarteer and Palmer, 1992). Nevertheless, the increase in RMR from 308 to 482 days of age is quite steep and we cannot exclude the possibility that other, external influences had an effect on the measurements. For example, while room and experimental conditions during the RMR measurements were kept constant, the animals had been relocated to another holding facility in between. We made sure that temperature, humidity and air pressure were comparable but it might have been that mice reacted to factors that escaped our perception.

In addition to the influence of cognitive stimulation on physiological processes, we investigated whether and how cognitive stimulation affects male attractiveness and dominance. Some studies showed that females of different species chose based on morphological features (Bischoff et al., 1985; Mateos and Carranza, 1995; Kodric-Brown and Nicoletto, 2001), scent marks (Ramm et al., 2008) or by cognitive performance [reviewed in Boogert et al. (2011)]. The question in our study was, whether or not male mice which were cognitively stimulated were also preferred by females. On the contrary, our results showed that female mice preferred NL mice, which were never cognitively stimulated like L mice or additionally stressed by tail handling like EC and L mice. Female mice did not differentiate between L and EC mice. Since females' choices were not influenced by the body mass of the males and they could not directly assess the males' cognitive abilities in the mate choice test, other factors have to be considered. Several volatile substances in

male mice urine have been shown to provide mate assessment signals (Stopka et al., 2012). Scent production is known to be associated with male social status, stress level, and other factors of the physical environment, thus allowing females to assess male quality (Novotny et al., 1985, 1990). The attractiveness of the scent mark is influenced by the quality and quantity of its compounds (Drickamer, 1992; Zala et al., 2004). Gosling et al. (2000) showed that high intensities of scent marks are costly in terms of body mass loss and loss of dominance status. As our L mice already had higher energy demands (higher activity, reduced body mass, and RMR), there may have been less energy available for the production of costly signals. The EC mice may have been stressed by the additional handling by being placed on test set-ups through tail handling without having the opportunity to escape this situation (like the L mice) which may also have had a negative effect on the chemical signals produced. Thus, the L mice and EC mice may have had a lower quality and/or quantity of scent marks and were therefore less attractive for female mice.

In our study the female mice only differentiated between the treatment groups during the first FC test. There was no female choice during the second FC. One explanation could be that there might not have been a detectable difference between treatments from the perspective of the females. This would be plausible as the L mice did not learn the tasks in the second and third IC phase which was accompanied by the fact that the treatments did not result in measurable physiological differences. Therefore, the females may not have been able to distinguish between the groups. Or the male mice were generally unattractive because of their age (592 days) regardless of the treatment. The post-reproductive phase appears to be strain-dependent and begins in wild mice by reduced fertilization from 570 days [reviewed in Brust et al. (2015)]. If this is also true for C57BL/6J mice, old age could possibly influence the female's choice during the second FC test. Finally, future experiments where females can directly assess the learning capabilities of male mice are needed to more clearly assess the validity of the influence of cognitive stimulation on female choice.

One further important factor in increasing attractiveness is the social rank of a male. But unfortunately, the social rank of the male mice in our study was not investigated during the first IC phase and the first FC test where the females differentiated between the groups. The observations recorded later showed that none of the three treatment groups differed with regard to winning or losing fights. However, the L mice were overall involved in more fights. Whether they also initiated them and thus have a greater potential for aggression cannot be determined with our data. However, it should be noted that fighting behavior was rare and only observed during cage cleaning. There was no need to remove individual mice from the groups at any time due to aggressive behavior and resulting injuries. We would argue that the relatively small environment and large number of males may simply not enable building territories and rank hierarchies.

In addition to physiological and fitness related costs of cognition we examined the influence of repeated cognitive stimulation on longevity. Therefore, we investigated telomere lengths in aged mice (699 days). It is known from the literature that the length of telomeres is associated with disease, loss of cognitive abilities, and longevity (Blasco et al., 1997; Lee et al., 1998; Rudolph et al., 1999; Cawthon et al., 2003; Benetos et al., 2004; Martin-Ruiz et al., 2006; Yaffe et al., 2011). Repeated cognitive stimulation might have led to higher energetic costs and compensation of these costs in terms of a reduced investment in other energetically costly physiological processes, which might reflect in shorter telomere lengths and lifespans. However, we did not find such an effect in our mice. Even though mice of the EC and L group both showed lower growth rates during the first IC phase, they caught up later in life, as it has been observed numerous times across many species under improving conditions (Metcalfe and Monaghan, 2001). Similarly, investment in other physiological processes may have been postponed rather than fully neglected if they became apparent due to the different environments experienced, especially during early life in the mice of this study. Still, both, early conditions experienced in life as well as compensation strategies (Metcalfe and Monaghan, 2001; Burton and Metcalfe, 2014) can hold long-term costs. The fact that we could not identify such costs may lie in the continuous availability of *ad libitum* food throughout the mice's life, which may have helped our animals to successfully catch up without suffering from long term consequences of the experienced early environment. In addition, one could argue that introducing cognitive tasks into the life of laboratory mice serves as cognitive enrichment, i.e., the possibility to use evolved cognitive skills to solve problems and control aspects of the environment (Clark, 2017). Cognitive enrichment is known to increase neuroplasticity properties and increase neural connectivity [reviewed in Petrosini et al. (2009)] and it possibly reduces boredom related abnormal behavior. This seemingly protects against the development of age-associated cognitive decline and functional impairments even in the presence of brain pathologies in laboratory rodents (Stern, 2002; Milgram et al., 2006).

As already mentioned earlier, during the first IC phase, the cognitive stimulation induced observable changes in behavior and physiology, while during the second and third IC phase, the L mice did not learn successfully within the IC. In addition, the L mice had to solve fewer cognitive tasks in the second and third IC phases compared to the first IC phase. Therefore, we might assume a cognitively demanding start in life rather than a cognitive demanding life in our study. This was reflected in our results. While differences in behavior and physiological processes were detected in young mice during the first IC phase, these differences were no longer present in the following measurements. Therefore, it is not surprising that no effects on longevity could be detected. Both the telomere length and the survival analysis showed no treatment differences.

Previous studies have already demonstrated that the IC is a useful tool for investigating learning behavior (Galsworthy et al., 2005; Mechan et al., 2009; Krackow et al., 2010; Endo et al., 2011; Voikar et al., 2018). As mentioned earlier, the L mice reached the learning criterion in the IC tasks only during the first phase. In the following phases, however, the cognitive stimulation was not successful. The tasks during the first IC phase were supposedly

easy to learn and similar to those of other IC studies. In contrast, the tasks during the second and third phase were chosen to be more complex and were conducted at an older age of the mice. We exclude that the L mice had severe age-related cognitive decline, as we were able to show that the L mice in the BM were also learning at the age of 573 days. This indicates that aged mice indeed are capable of showing reasonable performances which is also in accordance with literature data (Mechan et al., 2009). In addition, we retrospectively noticed that although the mice received an airpuff as punishment for visiting the wrong corner they still drank in these corners. This was unexpected as we assumed that they would avoid the airpuffs as is known from other experiments. So most likely they did not perceive the airpuff as a severe punishment, habituated after repeatedly being exposed to this stimulus, and accordingly developed a more relaxed attitude. Mice are known for notoriously using alternative strategies in tasks laid out by humans (Habedank et al., 2021). For example, in cognitive testing the use of semi-successful but often simpler strategies is common but often corrected for by the experimenter right away. Using an automated testing system prevented us from quickly noticing that the mice adopted an alternative strategy to solve some of the tasks and accordingly we could not adjust the experiment. Working with an automated system is advantageous in terms of not stressing the animals, for example by separating them from their social group or being in contact with a human experimenter on a regular basis. At the same time it leads to delayed feedback making it harder to detect flaws in the setup and execution of experiments, especially while these are running.

We demonstrated that the IC system, as an automated and home-cage based test system, is a useful method to keep mice with different treatments in one social group. Furthermore, the system worked extraordinarily well even for group housed male mice, which are often housed singly to prevent overt aggression. Hence, our setting supports calls of the current legislation (e.g., EU directive 2010/63) to house mice in social groups whenever possible. To keeping the mice within the ICs, they were well habituated to the test procedure and it was possible to observe the mice in their natural active phase. While the mice had to be handled for all tests outside the IC system, which is presumably associated with stress, the IC experiments themselves could be carried out without disturbance by the experimenter. We also showed that using the same group of mice repeatedly allowed us to perform experiments without treating laboratory mice as disposable goods like it is still common practice (Brust et al., 2015). This approach also reduces the overall number of experimental animals and is in accordance with the 3Rs (reduce, refine, and replace). In summary our data suggest that the IC system is a highly useful tool to conduct unique home-cage based long-term studies in social settings. In addition, our study showed that cognitive stimulation induces reversible short-term changes in behavior and physiology. We could not detect any long-term effects on behavior or physiology. Similarly, we also did not find persisting effects on sexually selected traits such as dominance or mate choice and no effects on longevity. To our best knowledge, our study provides the first unique long-term data set from male mice and we hope that this will guide future sustainable and responsible studies in laboratory animal science.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The animal study was reviewed and approved by LAVES, license 33.9-42502-05-14A430.

AUTHOR CONTRIBUTIONS

PK, AG, LL, and VB: conceptualization. PK, JT, and EH: conducting the experiments. PK, AG, MB, and VB: data analyzing and visualization. LL: project administration and supervision. PK, AG, MB, VB, and LL: article writing, review, and editing. All authors contributed to the article and approved the submitted version.

FUNDING

Parts of this work were funded by the Deutsche Forschungsgemeinschaft (DFG) as part of the research group "Severity Assessment for Animal-Based Research" (FOR 2591) as well as under Germany's Excellence Strategy—EXC 2002 "Science of Intelligence"—project number 390523135 (www.scienceofintelligence.de).

ACKNOWLEDGMENTS

Two of the IntelliCages used were courtesy of a collaborative project of LL together with Stefan Schlatt, Centre of Reproductive Medicine and Andrology, University of Muenster. The authors thank the animal caretakers, technicians, and colleagues, especially Phillipp M. Schindler, Christian Seebass, and Anja Ritz for their support in the animal husbandry and lab work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnbeh.2021.709775/full#supplementary-material

REFERENCES

Bates, D., Machler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48.

Benetos, A., Gardner, J. P., Zureik, M., Labat, C., Xiaobin, L., Adamopoulos, C., et al. (2004). Short telomeres are associated with increased carotid atherosclerosis in hypertensive subjects. *Hypertension* 43, 182–185. doi: 10.1161/01.HYP.0000113081.42868.f4

Bischoff, R. J., Gould, J. L., and Rubenstein, D. I. (1985). Tail size and female choice in the guppy (*Poecilia reticulata*). 17, 253–255.

Blasco, M. A., Lee, H. W., Hande, M. P., Samper, E., Lansdorp, P. M., DePinho, R. A., et al. (1997). Telomere shortening and tumor formation by mouse cells lacking telomerase RNA. *Cell* 91, 25–34. doi: 10.1016/S0092-8674(01)80006-4

Boogert, N. J., Fawcett, T. W., and Lefebvre, L. (2011). Mate choice for cognitive traits: a review of the evidence in nonhuman vertebrates. *Behav. Ecol.* 22, 447–459. doi: 10.1093/beheco/arq173

Boucherie, P. H., Loretto, M. C., Massen, J. J. M., and Bugnyar, T. (2019). What constitutes "social complexity" and "social intelligence" in birds? Lessons from ravens. *Behav. Ecol. Sociobiol* 73:12. doi: 10.1007/s00265-018-2607-2

Brandes, C. (1988). Estimation of heritability of learning behavior in honeybees (*Apis mellifera capensis*). *Behav. Genet.* 18, 119–132. doi: 10.1007/BF01067081

Brouilette, S., Singh, R. K., Thompson, J. R., Goodall, A. H., and Samani, N. J. (2003). White cell telomere length and risk of premature myocardial infarction. *Arterioscler. Thromb. Vasc. Biol.* 23, 842–846. doi: 10.1161/01.ATV.0000067426.96344.32

Brown, J. H. (2004). Toward a metabolic theory of ecology. *Ecology* 85, 1771–1789.

Brust, V., Schindler, P. M., and Lewejohann, L. (2015). Lifetime development of behavioural phenotype in the house mouse (*Mus musculus*). *Front. Zool.* 12:S17. doi: 10.1186/1742-9994-12-S1-S17

Buchanan, K. L., Grindstaff, J. L., and Pravosudov, V. V. (2013). Condition-dependence, developmental plasticity, and cognition: implications for ecology and evolution. *Trends Ecol. Evol.* 28, 290–296. doi: 10.1016/j.tree.2013.02.004. Condition-dependence

Burger, J. M. S., Kolss, M., Pont, J., and Kawecki, T. J. (2008). Learning ability and longevity: a symmetrical evolutionary trade-off in *Drosophila*. *Evolution* 62, 1294–1304. doi: 10.1111/j.1558-5646.2008.00376.x

Burton, T., and Metcalfe, N. B. (2014). Can environmental conditions experienced in early life influence future generations? *Proc. R. Soc. B Biol. Sci* 281:20140311. doi: 10.1098/rspb.2014.0311

Byrne, P. G., and Rice, W. R. (2006). Evidence for adaptive male mate choice in the fruit fly *Drosophila melanogaster*. *Proc. R. Soc. B Biol. Sci.* 273, 917–922. doi: 10.1098/rspb.2005.3372

Callicott, R. J., and Womack, J. E. (2006). Real-time PCR assay for measurement of mouse telomeres. *Comp. Med.* 56, 17–22.

Cawthon, R. M., Smith, K. R., O'Brien, E., Sivatchenko, A., and Kerber, R. A. (2003). Association between telomere length in blood and mortality in people aged 60 years or older. *Lancet* 361, 393–395. doi: 10.1016/S0140-6736(03)12384-7

Clark, F. (2017). Cognitive enrichment and welfare: current approaches and future directions. *Anim. Behav. Cogn.* 4, 52–71. doi: 10.12966/abc.05.02.2017

Clutton-Brock, T. H., Price, O. F., Albon, S. D., and Jewell, P. A. (1992). Early development and population fluctuations in soay sheep. *J. Anim. Ecol.* 61, 381–396.

Cole, E. F., Morand-Ferron, J., Hinks, A. E., and Quinn, J. L. (2012). Cognitive ability influences reproductive life history variation in the wild. *Curr. Biol.* 22, 1808–1812. doi: 10.1016/j.cub.2012.07.051

Crawley, J. N. (1985). Exploratory behavior models of anxiety in mice. *Neurosci. Biobehav. Rev.* 9, 37–44. doi: 10.1016/0149-7634(85)90030-2

Douhard, M., Gaillard, J. M., Delorme, D., Capron, G., Duncan, P., Klein, F., et al. (2013). Variation in adult body mass of roe deer: early environmental conditions influence early and late body growth of females. *Ecology* 94, 1805–1814. doi: 10.1890/13-0034.1

Drickamer, L. C. (1992). Oestrous female house mice discriminate dominant from subordinate males and sons of dominant from sons of sunordinate males by odour cues. *Anim. Behav.* 43, 868–870.

Dukas, R. (2004). Evolutionary biology of animal cognition. *Annu. Rev. Ecol. Evol. Syst.* 35, 347–374. doi: 10.1146/annurev.ecolsys.35.112202.130152

Dukas, R. (2006). Learning in the context of sexual behaviour in insects. *Anim. Biol.* 56, 125–141. doi: 10.1163/157075606777304258

Dunbar, R. I. M., and Shultz, S. (2007). Understanding primate brain evolution. *Philos. Trans. R. Soc. B Biol. Sci.* 362, 649–658. doi: 10.1098/rstb.2006.2001

Eisen, E. J. (1976). Results of growth curve analyses in mice and rats. *J. Anim. Sci.* 42, 1008–1023. doi: 10.2527/jas1976.4241008x

Elliott, K. H., Hare, J. F., Le Vaillant, M., Gaston, A. J., Ropert-Coudert, Y., and Anderson, W. G. (2015). Ageing gracefully: physiology but not behaviour declines with age in a diving seabird. *Funct. Ecol.* 29, 219–228. doi: 10.1111/1365-2435.12316

Endo, T., Maekawa, F., Võikar, V., Haijima, A., Uemura, Y., Zhang, Y., et al. (2011). Automated test of behavioral flexibility in mice using a behavioral sequencing task in intelliCage. *Behav. Brain Res.* 221, 172–181. doi: 10.1016/j.bbr.2011.02.037

Galsworthy, M. J., Amrein, I., Kuptsov, P. A., Poletaeva, I. I., Zinn, P., Rau, A., et al. (2005). A comparison of wild-caught wood mice and bank voles in the intellicage: assessing exploration, daily activity patterns and place learning paradigms. *Behav. Brain Res.* 157, 211–217. doi: 10.1016/j.bbr.2004.06.021

Ghosal, S., Nunley, A., Mahbod, P., Lewis, A. G., Smith, E. P., Tong, J., et al. (2015). Mouse handling limits the impact of stress on metabolic endpoints hhs public access author manuscript. *Physiol. Behav.* 150, 31–37. doi: 10.1016/j.physbeh.2015.06.021.Mouse

Gosling, L. M., Roberts, S. C., Thornton, E. A., and Andrew, M. J. (2000). Life history costs of olfactory status signalling in mice. *Behav. Ecol. Sociobiol.* 48, 328–332. doi: 10.1007/s002650000242

Gouveia, K., and Hurst, J. L. (2013). Reducing mouse anxiety during handling: effect of experience with handling tunnels. *PLoS One* 8:e66401. doi: 10.1371/journal.pone.0066401

Griffith, L. C., and Ejima, A. (2009). Courtship learning in *Drosophila melanogaster*: diverse plasticity of a reproductive behavior. *Learn. Mem.* 16, 743–750. doi: 10.1101/lm.956309

Gruszka, A., Matthews, G., and Szymura, B. (2010). *Handbook of Individual Differences in Cognition*. New York, NY: Springer-Verlag.

Habedank, A., Kahnau, P., and Lewejohann, L. (2021). Alternate without alternative: neither preference nor learning explains behaviour of C57BL/6J mice in the T-maze. *Behaviour* 158, 625–662. doi: 10.1163/1568539X-bja10085

Hollis, B., and Kawecki, T. J. (2014). Male cognitive performance declines in the absence of sexual selection. *Proc. R. Soc. B Biol. Sci* 281:20132873. doi: 10.1098/rspb.2013.2873

Huebner, F., Fichtel, C., and Kappeler, P. M. (2018). Linking cognition with fitness in a wild primate: Fitness correlates of problem-solving performance and spatial learning ability. *Philos. Trans. R. Soc. B Biol. Sci* 373:20170295. doi: 10.1098/rstb.2017.0295

Kodric-Brown, A., and Nicoletto, P. F. (2001). Female choice in the guppy (*Poecilia reticulata*): The interaction between male color and display. *Behav. Ecol. Sociobiol.* 50, 346–351. doi: 10.1007/s002650100374

Kotrschal, A., Corral-Lopez, A., Szidat, S., and Kolm, N. (2015). The effect of brain size evolution on feeding propensity, digestive efficiency, and juvenile growth. *Evolution* 69, 3013–3020. doi: 10.1111/evo.12784

Krackow, S., Vannoni, E., Codita, A., Mohammed, A. H., Cirulli, F., Branchi, I., et al. (2010). Consistent behavioral phenotype differences between inbred mouse strains in the intelliCage. *Genes Brain Behav.* 9, 722–731. doi: 10.1111/j.1601-183X.2010.00606.x

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* 82, 1–26. doi: 10.18637/jss.v082.i13

Laughlin, S. B., de Ruyter van Steveninck, R. R, Anderson, JC, et al. (1998). The metabolic cost of neural information. *Nat. Neurosci.* 1, 36–41.

Lee, C. P. (2003). "Innovation as a behavioural response to environmental challenges: a cost and benefit approach," in S. M. Reader, and K. N. Laland (eds.), *Animal Innovation*. Oxford: Oxford University Press. 261–277.

Lee, H. W., Blasco, M. A., Gottlieb, G. J., Horner, J. W., Greider, C. W., and DePinho, R. A. (1998). Essential role of mouse telomerase in highly proliferative organs. *Nature* 392, 569–574. doi: 10.1038/33345

Lenth, R. (2020). *Estimated Marginal Means, aka Least-Squares Means*. Available online at: https://cran.r-project.org/package=emmeans (accessed August 19, 2021).

Lindström, J., Metcalfe, N. B., and Royle, N. J. (2005). How are animals with ornaments predicted to compensate for a bad start in life? A dynamic optimization model approach. *Funct. Ecol.* 19, 421–428. doi: 10.1111/j.1365-2435.2005.00974.x

Loi, M., Mossink, J. C. L., Meerhoff, G. F., Den Blaauwen, J. L., Lucassen, P. J., and Joëls, M. (2017). Effects of early-life stress on cognitive function and

hippocampal structure in female rodents. *Neuroscience* 342, 101–119. doi: 10.1016/j.neuroscience.2015.08.024

Madden, J. R., Langley, E. J. G., Whiteside, M. A., Beardsworth, C. E., and Van Horik, J. O. (2018). The quick are the dead: Pheasants that are slow to reverse a learned association survive for longer in the wild. *Philos. Trans. R. Soc. B Biol. Sci* 373:20170297. doi: 10.1098/rstb.2017.0297

Martin-Ruiz, C., Dickinson, H. O., Keys, B., Rowan, E., Kenny, R. A., and Von Zglinicki, T. (2006). Telomere length predicts poststroke mortality, dementia, and cognitive decline. *Ann. Neurol.* 60, 174–180. doi: 10.1002/ana.20869

Mateos, C., and Carranza, J. (1995). Female choice for morphological features of male ring-necked pheasants. *Anim. Behav.* 49, 737–748. doi: 10.1016/0003-3472(95)80206-1

Mathot, K. J., Godde, S., Careau, V., Thomas, D. W., and Giraldeau, L.-A. (2009). Testing dynamic variance-sensitive foraging using individual differences in basal metabolic rates of zebra finches. *Oikos* 118, 545–552. doi: 10.1111/j.1600-0706.2008.17357.x

McCarteer, R. J., and Palmer, J. (1992). Energy metabolism and aging: a lifelong study of Fischer 344 rats. *Am. J. Physiol. - Endocrinol. Metab* 263, 448–452. doi: 10.1152/ajpendo.1992.263.3.E448

McEwen, B. S. (2007). Physiology and neurobiology of stress and adaptation: Central role of the brain. *Physiol. Rev.* 87, 873–904. doi: 10.1152/physrev.00041.2006

McNab, B. K., and Eisenberg, J. F. (1989). Brain size its relation to the rate of metabolism in mammals. *Am. Nat.* 133, 157–167.

Mechan, A. O., Wyss, A., Rieger, H., and Mohajeri, M. H. (2009). A comparison of learning and memory characteristics of young and middle-aged wild-type mice in the IntelliCage. *J. Neurosci. Methods* 180, 43–51. doi: 10.1016/j.jneumeth.2009.02.018

Mery, F., and Kawecki, T. J. (2002). Experimental evolution of learning ability in fruit flies. *Proc. Natl. Acad. Sci. U.S.A.* 99, 14274–14279. doi: 10.1073/pnas.222371199

Metcalfe, N. B., and Monaghan, P. (2001). Compensation for a bad start: grow now, pay later? *Trends Ecol. Evol.* 16, 254–260. doi: 10.1016/S0169-5347(01)02124-3

Milgram, N. W., Siwak-Tapp, C., Araujo, J., and Head, E. (2006). Neuroprotective effects of cognitive enrichment. *Aging Res. Rev.* 5, 354–369.

Moe, B., Angelier, F., Bech, C., and Chastel, O. (2007). Is basal metabolic rate influenced by age in a long-lived seabird, the snow petrel? *J. Exp. Biol.* 210, 3407–3414. doi: 10.1242/jeb.005090

Morand-Ferron, J. (2017). Why learn? The adaptive value of associative learning in wild populations. *Curr. Opin. Behav. Sci.* 16, 73–79. doi: 10.1016/j.cobeha.2017.03.008

Niven, J. E., and Laughlin, S. B. (2008). Energy limitation as a selective pressure on the evolution of sensory systems. *J. Exp. Biol.* 211, 1792–1804. doi: 10.1242/jeb.017574

Novotny, M., Harvey, S., Jemiolo, B., and Alberts, J. (1985). Synthetic pheromones that promote inter-male aggression in mice. *Proc. Natl. Acad. Sci. U.S.A.* 82, 2059–2061. doi: 10.1073/pnas.82.7.2059

Novotny, M., Harvey, S., and Jemiolo, B. (1990). *Chemistry of Male Dominance in the House Mouse, Mus Domesticus Department of Chemistry, Indiana University, Bloomington (Indiana 47405, USA) Received 2 January 1989*, Accepted 16 May 1989, Vol. 46. Bloomington, IN: Indiana University, 109–113.

Papini, M. R. (2002). Pattern and process in the evolution of learning. *Psychol. Rev.* 109, 186–201. doi: 10.1037/0033-295x.109.1.186

Petrosini, L., De Bartolo, P., Foti, F., Gelfo, F., Cutuli, D., Leggio, M. G., et al. (2009). On whether the environmental enrichment may provide cognitive and brain reserves. *Brain Res. Rev.* 61, 221–239. doi: 10.1016/j.brainresrev.2009.07.002

Piers, L. S., Soares, M. J., Mccormack, L. M., and O'Dea, K. (1998). Is there evidence for an age-related reduction in metabolic rate? *J. Appl. Physiol.* 85, 2196–2204. doi: 10.1152/jappl.1998.85.6.2196

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team (2020). *Linear and Nonlinear Mixed Effects Models*. Available online at: https://cran.r-project.org/package=nlme (accessed August 19, 2021).

Pravosudov, V. V., Lavenex, P., and Omanska, A. (2005). Nutritional deficits during early development affect hippocampal structure and spatial memory later in life. *Behav. Neurosci.* 119, 1368–1374. doi: 10.1037/0735-7044.119.5.1368

R Core Team (2020). *A language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Ramm, S. A., Cheetham, S. A., and Hurst, J. L. (2008). Encoding choosiness: female attraction requires prior physical contact with individual male scents in mice. *Proc. R. Soc. B Biol. Sci.* 275, 1727–1735. doi: 10.1098/rspb.2008.0302

Roelofs, S., Boleij, H., Nordquist, R. E., and van der Staay, F. J. (2016). Making decisions under ambiguity: Judgment bias tasks for assessing emotional state in animals. *Front. Behav. Neurosci.* 10:1–16. doi: 10.3389/fnbeh.2016.00119

Rudolph, K. L., Chang, S., Lee, H. W., Blasco, M., Gottlieb, G. J., Greider, C., et al. (1999). Longevity, stress response, and cancer in aging telomerase-deficient mice. *Cell* 96, 701–712. doi: 10.1016/S0092-8674(00)80580-2

Schubert, K. A., Vaanholt, L. M., Stavasius, F., Demas, G. E., Daan, S., and Visser, G. H. (2008). Female mice respond differently to costly foraging versus food restriction. *J. Exp. Biol.* 211, 2214–2223. doi: 10.1242/jeb.017525

Snell-Rood, E. C. (2013). An overview of the evolutionary causes and consequences of behavioural plasticity. *Anim. Behav.* 85, 1004–1011. doi: 10.1016/j.anbehav.2012.12.031

Sol, D., and Lefebvre, L. (2000). Behavioural flexibility predicts invasion success in birds introduced to New Zealand. *Oikos* 90, 599–605. doi: 10.1034/j.1600-0706.2000.900317.x

Speakman, J. R., Talbot, D. A., Selman, C., Snart, S., McLaren, J. S., Redman, P., et al. (2004). Uncoupled and surviving: Individual mice with high metabolism have greater mitochondrial uncoupling and live longer. *Aging Cell* 3, 87–95. doi: 10.1111/j.1474-9728.2004.00097.x

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *J. Int. Neuropsychol. Soc.* 8, 448–460. doi: 10.1017/S1355617702813248

Stopka, P., Stopkova, R., and Janotova, K. (2012). in *Mechanisms of Chemical Communication*, eds J. Pialek, M. Macholan, P. Munclinger, and S. J. Braid (Cambridge: Cambridge University Press), 191–220.

Therneau, T. M., and Grambsch, P. M. (2000). *A Package for Survival Analysis in R*.

Tryon, R. C. (1940). Studiesn in individual differences in maze ability. VII the specific components of maze ability, and a general theory of physiological components. *J. Comp. Psychol.* 30, 283–335.

Tzankoff, S. P., and Norris, A. H. (1977). Effect of muscle mass decrease on age-related BMR changes. *J. Appl. Physiol. Respir. Environ. Exerc. Physiol.* 43, 1001–1006. doi: 10.1152/jappl.1977.43.6.1001

Voikar, V., and Gaburro, S. (2020). Three pillars of automated home-cage phenotyping of mice: novel findings, refinement, and reproducibility based on literature and experience. *Front. Behav. Neurosci.* 14:575434. doi: 10.3389/fnbeh.2020.575434

Voikar, V., Krackow, S., Lipp, H. P., Rau, A., Colacicco, G., and Wolfer, D. P. (2018). Automated dissection of permanent effects of hippocampal or prefrontal lesions on performance at spatial, working memory and circadian timing tasks of C57BL/6 mice in IntelliCage. *Behav. Brain Res.* 352, 8–22. doi: 10.1016/j.bbr.2017.08.048

Von Zglinicki, T. (2002). Oxidative stress shortens telomeres. *Trends Biochem. Sci.* 27, 339–344. doi: 10.1016/S0968-0004(02)02110-2

Yaffe, K., Lindquist, K., Kluse, M., Cawthon, R., Harris, T., Hsueh, W.-C., et al. (2011). Telomere length and cognitive function in community-dwelling elders: Findings from the Health ABC Study. *Neurobiol. Aging* 23, 2055–2060. doi: 10.1016/j.neurobiolaging.2009.12.006.Telomere

Zala, S. M., Potts, W. K., and Penn, D. J. (2004). Scent-marking displays provide honest signals of health and infection. *Behav. Ecol.* 15, 338–344. doi: 10.1093/beheco/arh022

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

2.3 Publication 3

Determining the value of preferred goods based on consumer demand in a home-cage based test for mice

Pia Kahnau, Anne Jaap, Kai Diederich, Lorenz Gygax, Juliane Rudeck, Lars Lewejohann

Published in journal “Behavior Research Methods” 2022, 55, 751-766 (2023)
DOI: https://doi.org/10.3758/s13428-022-01813-8

The Supplementary Material for this article can be found online at:
https://link.springer.com/article/10.3758/s13428-022-01813-8#Sec12

Literature Review

Behavior Research Methods (2023) 55:751–766
https://doi.org/10.3758/s13428-022-01813-8

Determining the value of preferred goods based on consumer demand in a home-cage based test for mice

Pia Kahnau[1] · Anne Jaap[1] · Kai Diederich[1] · Lorenz Gygax[2] · Juliane Rudeck[1] · Lars Lewejohann[1,3]

Accepted: 14 February 2022 / Published online: 25 April 2022

Abstract
From the preference of one good over another, the strength of the preference cannot automatically be inferred. While money is the common denominator to assess the value of goods in humans, it appears difficult at first glance to put a price tag on the decisions of laboratory animals. Here we used consumer demand tests to measure how much work female mice expend to obtain access to different liquids. The mice could each choose between two liquids, one of which was free. The amount of work required to access the other liquid, by contrast, increased daily. In this way, the value of the liquid can be determined from a mouse's microeconomic perspective. The unique feature is that our test was carried out in a home-cage based setup. The mice lived in a group but could individually access the test-cage, which was connected to the home-cage via a gate. Thereby the mice were able to perform their task undisturbed by group members and on a self-chosen schedule with minimal influence by the experimenter. Our results show that the maximum number of nosepokes depends on the liquids presented. Mice worked incredibly hard for access to water while a bitter-tasting solution was offered for free whereas they made less nosepokes for sweetened liquids while water was offered for free. The results demonstrate that it is possible to perform automated and home-cage based consumer demand tests in order to ask the mice not only what they like best but also how strong their preference is.

Keywords Home-cage · IntelliCage · Group housing · Mice · Consumer demand · Preference test

Introduction

In economics, the principle of consumer demand is used to determine the best possible price of a product in order to achieve the highest possible profit. In contrast, the consumer demand test is used with animals as an operant task to assess the value of goods from the animal's point of view by examining the motivation to obtain or to avoid goods (Cooper, 2004; Lea, 1978). This is achieved by examining how much work animals are willing to perform to obtain goods or to avoid them. In this context, work performance can be equated with the paid price (Lea, 1978). By determining which price is paid for which goods by the animals, it is possible to determine the strength of the preference (Kirkden, Edwards, & Broom, 2003) with a higher price indicating a stronger preference. In addition, demand curves can be used to determine which goods are necessary or luxurious. Therefore, a consumer demand curve is plotted on logarithmic axes depicting the relation of the quantity consumed by the increase of price. Naturally, the amount of consumption is negatively influenced by the price, i.e., with increasing costs the consumption decreases (Dawkins, 1988; Lea, 1978). For necessary goods, which ensure survival or increase fitness, the slope is hardly influenced by the price, the so-called price elasticity is low. However, if the slope is strongly influenced by the price, this indicates that the goods are of little importance or even luxury goods (Cooper, 2004; Dawkins, 1988; Kirkden et al., 2003).

In past studies, animals had to press a lever (Ladewig et al., 2002; Lewejohann & Sachser, 2000) or a switch

✉ Pia Kahnau
pia.kahnau@bfr.bund.de

1 German Federal Institute for Risk Assessment (BfR), German Centre for the Protection of Laboratory Animals (Bf3R), Max-Dohrn Str. 8–10, 10589 Berlin, Germany

2 Animal Husbandry and Ethology, Albrecht Daniel Thaer-Institute of Agricultural and Horticultural Sciences, Faculty of Life Sciences, Humboldt-Universität zu Berlin, Berlin, Germany

3 Animal Behavior and Laboratory Animal Science, Institute of Animal Welfare, Freie Universität Berlin, Berlin, Germany

(Sherwin & Nicol, 1997) in order to receive a reward. The number of lever presses or the energy required to move the switch was used as the equivalent to price. Other obstacles such as a water-filled passageway (Sherwin & Nicol, 1996) or weighted one-way doors (Warburton & Mason, 2003) were also used to make access to the goods more costly.

Laboratory animals were often trained and tested individually in the consumer demand test. Therefore, the animals were either placed in an experimental setup for a few hours per day (Ladewig et al., 2002; Sørensen et al., 2004) or for several days consecutively, using the experimental setup as a home-cage (Manser et al., 1998; Sherwin, 1998; Sherwin & Nicol, 1996, 1997; Timberlake, 1984; Warburton & Nicol, 1998). However, by removing animals from their home-cages and keeping them individually during testing, the animal's well-being may be negatively affected (Krohn et al., 2006; Manouze et al., 2019). This in turn could have a negative effect on the motivation of the animals to work during the consumer demand test and thus affect the experimental data. Therefore, it seems advantageous to utilize a consumer demand test that allows testing animals that live in groups and in their home-cage with minimum influence of the experimenter. To the best of our knowledge, the first group-housed consumer demand test for mice was developed by Sherwin (Sherwin, 2003, 2004, 2007) who investigated the influence of cage mates on motivation for additional space. Mice were kept in groups, in which only one mouse was trained and thus had access to additional space. As the price increased, the trained mice continued to work for the access to additional space. However, the number of visits and time spent decreased as the price increased. The author argued that additional space seems to be an important resource regardless of the presence of cage mates (Sherwin, 2004).

To test all animals within a social group and to obtain individual data, radio frequency identification (RFID) technology can be used. Past studies showed that the IntelliCage (IC) is a valid home-cage based and automated test setup to analyze activity and learning behavior in mice (Endo et al., 2011; Galsworthy et al., 2005; Kahnau et al., 2021; Krackow et al., 2010; Mechan et al., 2009; Voikar et al., 2018). In addition, the IC allows determining the amount of consumption of liquids and identifying preferences if more than one liquid is presented at the same time.

Animals' relative preferences for goods have been tested using preference tests. They offer the opportunity to determine which goods are preferred, as the animals themselves can choose between different goods. Especially with regard to animal welfare, it is useful to determine the value of the goods used for improving the living conditions of animals (Dawkins, 1983, 1988, 1990). Preference tests have been widely used in mice, for example, to investigate which bedding and nesting material or enrichment items are preferred (Ago et al., 2002; Banjanin & Mrosovsky, 2000; Chmiel & Noonan, 1996; Freymann et al., 2017; Van Loo et al., 2004, 2005; Patterson-Kane, Harper, & Hunt, 2001; Van De Weerd et al., 1998). There are several different approaches to perform a preference test (Habedank et al., 2018), but usually a binary choice test is performed with two differing goods on offer. Whenever one of these goods is consumed more frequently, or more time is spent with it, it is considered as the preferred one. By combining multiple binary choice tests, it is possible to compare several goods against each other, resulting in a scaling with a defined order. In a previous preference test, we were able to determine a ranking of the liquids (first preference test: 0.2 mM sucrose solution > 10 mM NaCl solution = tap water > 0.4 mM sucrose solution > 10 mM HCl solution, second preference test: almond milk > apple juice > tap water > 10 HCl solution > 3 mM quinine solution) which were also used in this study (the data for this ranking is part of the R package simsalRbim https://talbotsr.com/simsalRbim/index.html). However, such a scaling is just an indicator of the preference under the assumption that all goods are equally accessible. A scaling cannot give information on how much the goods are needed, i.e., a scaling does not determine the strength of the demand for or against a certain good. In order to determine the strength of preference for different liquids, we carried out consumer demand tests using a home-cage based automated setup.

It has already been shown that mice and rats enter a test system, e.g., an automated radial eight-arm maze or a rodent virtual reality (VR) maze, independently from their home-cage through an RFID controlled gate system (Kaupert et al., 2017; Mei et al., 2020; Rivalan et al., 2017; Winter & Schaefers, 2011). In the present study, the setup consisted of a home-cage that was connected via a gate (AnimalGate) to the test-cage (the IC). The IC contained four computerized corners with two liquid dispensers each. Because of the gate, only one mouse was in the IC at a time. This was necessary to allow the individual mice to work undisturbed by group members when accessing the liquids. Otherwise, it would have been possible that the mice interfered with each other directly, for example by pushing each other from the corner of the IC. Given that the home-cage was connected to the test-cage, the mice were basically free to choose when to work for access to the liquids. Since only one mouse could enter the IC at a time, the remaining mice had to wait within the home-cage until the occupant of the IC had left it again. This made it possible to test the mice with minimal influence of the experimenter during their active phase, i.e., when they spontaneously decide to do so and a high level of motivation can be assumed accordingly.

The main objective of this study was to evaluate the feasibility of an automated consumer demand test in a home-cage using the IC system. With this system, we obtained individual data from all mice kept in one social group and

we were able to determine different strengths of preferences for different liquids. We expected that the ranking of the liquids would reflect that of the earlier study but provide a more detailed view on the strength of the preference of the tested liquids. Knowing how rewarding or how aversive certain liquids are perceived is a prerequisite for refinement of conditioning experiments. In addition, our group has suggested before that animal welfare can be improved, particularly outside of the actual experiment, by providing rewards (Lewejohann et al., 2020). Finally, taking the mouse's perspective in estimating the strength of preferences of goods will guide future experiments in refining housing and experimental conditions.

Animals and methods

Animals and housing conditions

The pre-test of this study with 11 mice was pre-registered in the Animal Study Registry (animalstudyregistry.org, doi:10.17590/asr.0000131). The implementation of the consumer test presented here was based on the experience of the pre-test and was not additionally pre-registered. For the present study, a total of 12 female C57BL/6J mice (Charles River, Sulzfeld, Germany) were used. To ensure maximum genetic and epigenetic independence between individuals, all mice had different mothers and foster mothers. The mice arrived at the institute at an age of 28 to 34 days. During the consumer demand test, the mice were 10 to 19 months old (from November 2019 until August 2020). At the time of testing, the mice were already familiar with the test setup of the consumer demand test because they had participated in the development of a home-cage based cognitive bias test (pre-registered as doi:10.17590/asr.0000121). All mice were handled by the tunnel handling method (Plexiglas, 17.5 cm in length, 4 cm in diameter, for a video tutorial on mouse handling see https://wiki.norecopa.no/index.php/Mouse_handling). Four mice had to be killed due to health issues unrelated to the experiment, and one mouse was found dead (Table 1). The mice were removed from the data analysis of the current run. Even before the experiment, all mice showed fur and whisker trimming behavior, which is commonly found in C57BL/6 mice (Sama et al., 2000).

The room temperature and the humidity of the housing/testing room was 22°C ± 3°C and 55% ± 15%, respectively. The dark/light cycle was set to 12/12 hours. Because of the switch from winter to summer time, the light switched on at 7:00 am (MET/CET) in winter months and at 8:00 am (MEST/CEST) in summer months. Half an hour before the light phase, a sunrise was simulated by a wake-up light (Philips HF 3510, 100–240 vac, 50–60 Hz, Philips Consumer Lifestyle B.V. Netherlands). Over 30 min, the light intensity gradually increased until it reached full intensity at 7:00 am or 8:00 am, respectively. The room lights switched on at 7:00 am or 8:00 am and the wake-up light switched off after 1.5 h. The wake-up light was positioned in one corner on the ground of the room with the light shining in the direction of the test setup. The daily visual inspection of the mice was performed between 8:00 and 10:00 am. Once a week, the mice were weighed, inspected for health, and tail-colored (Edding 700, colors: red, black, white, silver, yellow) for individual visual identification. On the same day, the experimental setup including the home-cage was cleaned. All nesting, bedding materials and other enrichment items were replaced, but a small handful of bedding was transferred from the old home-cage to the new home-cage.

For testing within the IC system, it is necessary to implant RFID transponders. Since there were some transponder losses after previous transponder implantations (see

Table 1 Experimental schedule. Four mice had to be killed due to health issues and one mouse was found dead

Run	Abb.	Working corner	Free corner	*n*	Age*	Duration*
1	WQ	Tap water	Quinine hydrochloride dihydrate, 1.3 mM	12	316	64
2	AW	Almond milk, Alnatura, Almond Drink, unsweetened, 1:3 dilution	Tap water	11	387	20
3	WN	Tap water	NaCl, 10 mM	11	427	9
4	S0.4W	Sucrose, 0.4 mM	Tap water	11	444	15
5	WH	Tap water	HCl, 10 mM	10	469	24
6	JW	Apple juice, Sachsenobst Apple juice clear, 1 :3 dilution	Tap water	10	510	24
7	WW	Tap water	Tap water	7	540	9
8	S0.2W	Sucrose, 0.2 mM	Tap water	7	561	10

n = number of mice present in the runs and included in data evaluation. Dilutions were made with tap water. *Abb.* abbreviations

* in days

supplements), we optimized our procedure. We assumed that the injection site was manipulated by the mice themselves or by group members in such a way that transponder loss occurred. In order to prevent this, the mice received an analgesic (meloxicam 1mg/kg, Meloxidyl by CEVA) the evening before instead of 2 h before the transponder implantation. The duration of the analgesic effect lasted until at least 3 h after implantation, but not until their active phase in the following evening/night. The mice received RFID transponders (Euro ID, FDX-B, ISO 11784/85) under isoflurane anesthesia (induction of anesthesia: 4 l/min 4%; maintenance of anesthesia: 1 l/min 1–2%) at an age of 35 to 41 days. No transponder was lost following this optimized procedure.

All 12 mice were housed as one social group in an automated and home-cage based test setup (Fig. 1). The test setup consisted of a home-cage connected to a test-cage (IntelliCage, TSE-Systems, Germany) via a gate (AnimalGate, TSE-Systems, Germany). This allowed the mice to be tested over a long period of time, in their active phase. The home-cage was equipped with 3–4 cm bedding (spruce/fir, 2.5–5 mm, JRS Lignocel FS, Germany), two red houses ("TheMouseHouse", Tecniplast, Italy), nesting material (eight paper tissues, six cotton rolls, six nesting paper stripes), four wooden bars to chew on, food *ad libitum* (LAS QCDiet, Rod 16, autoclavable, LASvendi, Germany), and one transparent handling tube (4 cm in diameter, 17.5 cm long). In order to gain access to water or the test liquids, the mice had to pass through the gate individually. The gate allowed only one mouse at a time to pass from the home-cage to the test-cage. This was made possible by three doors, one RFID antenna and eight infrared barriers within the gate. The doors remained closed until the mouse returned to the home-cage. The separation allowed the mice to be tested individually and undisturbed by group members. This also implies that the remaining mice in the cage had to wait until the one mouse left the IC again. The gate also contained a scale, which measured the weight of each mouse on each passage to the IC. Each corner of the IC had one RFID antenna and one presence sensor for individual mouse identification. The presence sensor detected changes in temperature. If there was a temperature change within the IC corner and a transponder was detected by the RFID antenna at the same time, this event was counted as a visit. Each corner also comprised two water dispensers. Each dispenser had one lickometer, which measured the number of licks. The access to the liquids could be denied or granted through doors for each dispenser. By performing a nosepoke on the nosepoke-sensors on each door, the doors could be opened by the mice. With the Designer software of the IntelliCage Plus software package the access permissions to certain corners within the IC could be defined for each mouse. In addition, the required number of nosepokes for the access to the liquids could also be defined using the Designer software.

Consumer demand test

For our consumer demand test, the strength of preference or aversion was tested for eight different liquids in eight sequential "runs". During single runs, one liquid was offered in both liquid dispensers of one IC corner for which the mice

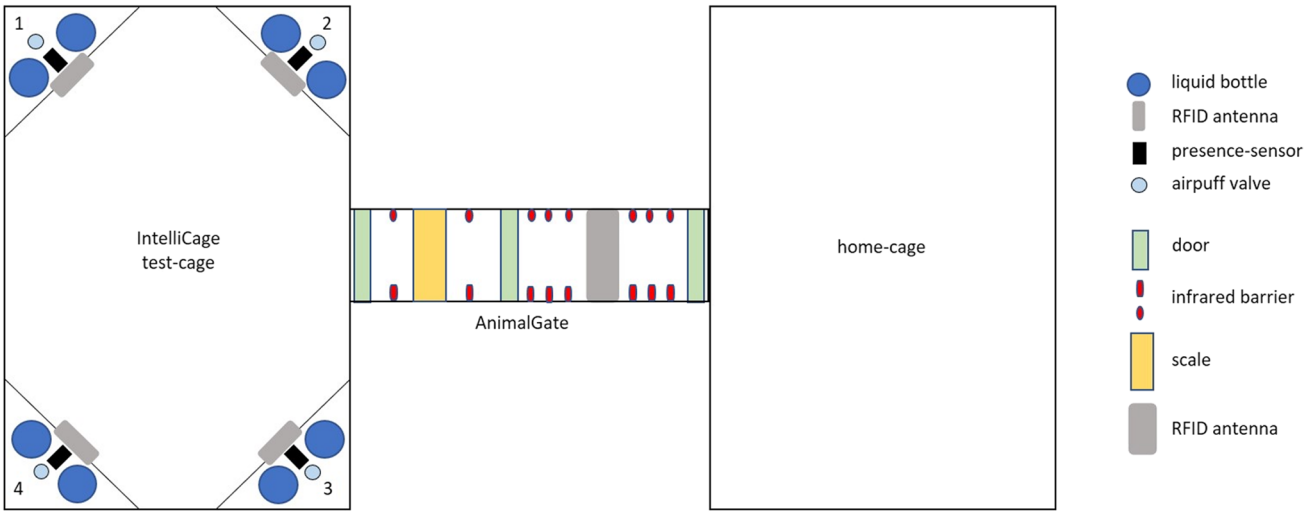

Fig. 1 Automated and home-cage-based test setup. The test setup consisted of a test-cage (IntelliCage), a gate (AnimalGate) and a home-cage. The IntelliCage contained bedding but no nesting or food. Within the IntelliCage each of the four corners was equipped with two water dispensers, one radio frequency identification (RFID) antenna, one presence-sensor, and one air-puff valve (air puffs were, however, not used during the consumer demand test). The AnimalGate contained three doors, eight infrared barriers, one scale and one RFID antenna. The home-cage contained bedding, nesting, two shelters, and food which was available *ad libitum*

had to make an increasing number of nosepokes every day (working corner). In both liquid dispensers of an adjacent corner the access to a second liquid was offered for the constant price of one nosepoke (free corner). This ensured that the mice did not suffer from thirst and had the possibility to drink at any time. To test the strength of preference, the mice had to work in four runs to gain access to supposedly positive tasting liquids (almond milk, apple juice, two sugar solutions) while at the same time water was offered in the free corner. To test aversion, the mice had to work in three runs to gain access to water while at the same time supposedly bad tasting liquids (bitter, sour, or salty-tasting solutions) were offered in the free corner.

In both working and free corners, the mice were able to drink for 10 s after making the required number of nosepokes. To drink again, the mice first had to leave the corner, re-enter it, and make the required number of nosepokes again. This ensured that while the price of access to the liquids changed, the quantity to be consumed per single access was constant. The working and free corner were the same for all mice but new positions were chosen after each run. This ensured that the new working corner for a new liquid was not used in the previous run.

The new free corner was again adjacent to it. In all runs, the two remaining corners were initially inactive. When a mouse did not execute the required number of nosepokes in the working corner for 2 days in a row, the additional two corners became active while the working and free corner became inactive for this mouse. The mice noticed such a change almost immediately. In the now active corners, the access to water was free (one nosepoke to open the door). This allowed excluding individual mice from the experimental conditions of a given run of the consumer demand test without having to remove them from their social group while the other mice could continue working for an increasing price.

Further on, the names of the single runs (eight runs in total) are abbreviated as follows: The first letter represents the liquid for which the mice had to work in the working corner (Table 1). The second letter represents the liquid that was available in the free corner. If the mice had to work in the working corner for access to, for example, almond milk while water was offered in the free corner, this run is abbreviated as AW. The A represents 3:1 dilution almond milk with tap water, the W represents tap water.

The sequence in which the paired liquids were presented was the same for all mice such that all mice experienced the same odors in the IC. First, the mice had to work for access to water while they had access to a bitter-tasting liquid in the free corner. This run served as training for the operant task (for more information on pre-tests see supplements) and provided data for the first pair of liquids at the same time. The number of required nosepokes to obtain access to the liquid in the working corner was increased daily by one, starting with one nosepoke at day one. For individual mice, each run ended as soon as they did not make the required nosepoke number on two consecutive days. One exception was the WQ run, which was stopped after 64 days, although ten of the 12 mice still made the required number of nosepokes. We decided to stop this run because participation with up to 64 nosepokes let us conclude that the aversion to quinine was very strong. From one run to the next, the mice had to work alternately for obtaining a positive liquid or avoiding a negative liquid (Table 1). Between each run and for 5–8 days on each occasion, all mice had access to water in all four corners by keeping all doors within the IC corners permanently open, therefore, the mice did not have to perform a nosepoke to open the doors. After the last run, the mice had to work one more time for access to water while access to quinine was free. Based on this run, we showed that all mice were still able to perform the operant task. Thus, the decline in motivation to work with rising prices across runs was not due to a nonspecific aging effect. In the last WQ control run, all seven mice that were still in the experiment made up to eight nosepokes for access to water. The run was then stopped (data not shown), because in six out of eight runs more than eight nosepokes were made (see Results section).

Data analysis

Data analysis and visualization was done with the open-source statistical software R, version 4.0.3 (R Core Team, 2020). Model assumptions were inspected visually by Q-Q plots and by visualizing variance homogeneity of the residuals versus the fitted values. The R package ggplot2 (Wickham, 2016) was used for data visualization.

The setup allowed the mice to enter the IC on their own and one at a time from the home-cage. All other mice had to wait until the IC was free again. The IC occupancy was analyzed based on the time duration during which each mouse was in the IC on each day. For this, the runs WW, WQ, and AW were considered. Runs WQ and AW were chosen to evaluate the influence on IC time of an aversive liquid (quinine) and a preferred liquid (almond milk). Run WW was chosen as a reference because water is a necessary good but should also be neutral compared to quinine and almond milk. The time spent in the IC was used as the outcome in a linear mixed-effects model (R package lme4; Bates et al., 2015). The experimental days were used as a continuous fixed effect (The data for days 53 and 54 of run WQ are missing due to technical problems with the AnimalGate.). The runs (factor reflected by sum-contrast with three levels: WW, WQ, AW) and the interaction of the runs and days were used as additional fixed effects. For the model, the variable day was

"centered". Day seven was chosen as the "middle" of all daily values for centering since observations for all three runs were still made on this day. The runs nested in animals were set as random effects. In addition, the individuality of the daily duration in the IC was evaluated. For this, we calculated the proportion of between-individual variance per total unexplained variance (between- plus within-individual variance) based on the estimated variance components in the model described so far. A confidence interval of this value was calculated using a parametric bootstrap approach with 1000 repetitions (R package lmerTest (Kuznetsova, Brockhoff, & Christensen, 2017) in combination with R package boot (Canty & Ripley, 2021; Davison & Hinkley, 1997)).

In addition to time spent in the IC, we analyzed the number of visits to the IC (IC entries). Since the run WQ ran the longest (with the highest price reached), this run was used for the evaluation. For each day, the sum of IC entries for both the light phase and the dark phase was determined for each mouse. Data were again missing for day 53 and 54 due to the technical problems with the AnimalGate. The logarithm of IC entries was used as the outcome in a linear mixed-effects model (R package nlme; Pinheiro et al., 2020). The experimental days (i.e., price) were used as a continuous fixed effect. The variable phase (factor with two levels: light and dark phase) and the interaction of day and phase were used as additional fixed effects. Sum-contrasts were used for the variable phase. To consider a possible effect of cleaning the setup that was suspected due to a waveform-shape in the number of visits, the variable day since cleaning was added as an additional continuous main effect. The variables day and day since cleaning were normalized for statistical analysis. It was added to the model as an additional fixed effect. The days nested in animals were set as the random effects. For further model assumption inspection, the homogeneity and shape of the residuals versus the variable day since cleaning were visually inspected.

The price paid for access to the liquids were nosepokes which the mice had to make inside the IC working corner. We assumed that as the number of nosepokes increased, the mice had to spend more time (visit duration) within the working corner. Run WQ was selected for analysis because in this run the mice made up to 64 nosepokes for access to water. Only visits in which the required nosepoke number and at least one lick was made, were considered. For the analysis, the visit duration was first determined for each price (required nosepoke number), each visit within the working corner, and each mouse. The logarithm of the visit duration was used as the outcome in a linear mixed-effects model, the price was used as a single fixed effect. Price (is equivalent to the individual test days) nested within the animal was used as the random effects. With this log-transformation, no serious deviations from the assumption could be detected.

The run WW can serve as a kind of control because in both, the working and the free corner, the same liquid was offered. Accordingly, we used the run WW as a reference for further evaluation and we compared the number of drinking events for water in the working corner and water in the free corner specifically in this run. The run WW ran for 9 days. Drinking events were defined as visits in which the mice made the required nosepoke number and drank. The number of these events were used as the outcome in a linear mixed-effects model (R package nlme). In this model, the nine experimental days were defined as days and used as a fixed effect (factor with nine levels). In addition, the type of corner (factor with two levels: working corner versus free corner) and the interaction of type of corner and day were used also as fixed effects. Again, sum-contrasts were used for day and type of corner. The test days nested in animals were set as the random effects.

We examined the maximum price paid by the mice for each liquid in the working corner of each liquid pair. For this, the maximum number of nosepokes they were willing to invest for gaining access was determined for each mouse and for each liquid within the working corner. A survival analysis was used to determine whether the maximum price paid depended on the liquids. This approach allowed for the correct handling of the censored data in the QW trial, i.e., the fact that the mice were still willing to continue working at even higher prices. We calculated this model with the R package survminer (Kassambara, Kosinski, & Biecek, 2020), which implements the cox proportional-hazard model (Coxph) and allowed to reflect the repeated measurement of the mice by defining animal as a "cluster". The maximum number of nosepokes was evaluated in dependence of the different runs (factor variable with eight levels) as a fixed effect.

Finally, we assessed the elasticity of demand. To do so, we analyzed the slopes of the consumer demand curves defined by the number of drinking events versus price for all liquids within the working corner. For this analysis, a linear mixed-effects model was used again. The log of the number of drinking events (plus 0.5 to allow the inclusion of zeros) was used as the outcome variable. The price (logarithm of the required number of nosepokes), the type of run, and their interaction was used as the fixed effects (using sum-contrasts) and run nested in animal as the random effects. With these log-transformations, we obtained normally distributed residuals. Based on this model, a single demand curve for the liquid in the working corners of each run could be estimated as follows: $Y = (I+R) + (SNN + IRNN) * NN$, where Y is the estimated (average logarithm of the) number of drinking events, I the intercept, R the main effect of the run, SNN the main effect of the price, IRNN the interaction of the run and the price and NN the price (number of nosepokes).

Results

IntelliCage occupancy

Mice spent the most time in the IC during the run WW (main effect run: $F_{2,12.01} = 16.73$, $p < 0.0001$; Fig. 2) and least during the run WQ. The IC time was not visibly influenced by day (and thus the price to be paid for a liquid; $F_{1,1011} = 0.03$; $p = 0.85$). Moreover, the interaction between day and run had no effect on the time spent in the IC ($F_{2,985.73} = 0.003$; $p = 1$). The proportion of between-animal variability was low at 13.92 % [3.84–25.64 CI] compared to the overall unexplained variability.

On average, the mice visited the IC around 11 times per day (Fig. 3). The mice entered the IC more often during the dark phase than during the light phase ($F_{1,11} = 34.58$; $p < 0.001$). The experimental days and the interaction of day and phase had no influence on the IC entries (day: $F_{1,1452} = 2.71$; $p = 0.1$; day:phase: $F_{1,1452} = 1.14$; $p = 0.29$). The wave-like pattern can be explained by the variable day since cleaning. According to this, the number of entries seemed to decrease after cleaning the setup, especially when comparing the cleaning day to the one that followed ($F_{1,1452} = 28.51$; $p < 0.001$, not shown).

With increasing price (required nosepoke number) the mice spent more time (visit duration) within the working corner to gain access (effect of price: $F_{63,686} = 110.49$; $p < 0.0001$; Fig. 4). From the figure it is seen that the mice spent around 16 s within the working corner for the price of one nosepoke. It was already around 38 s for the price of 32 nosepokes and around 64 s for the price of 64 nosepokes.

Comparison of drinking events for run WW

The mice made up to seven nosepokes to gain access to water in the working corner while water was available for the price of one nosepoke in the free corner (Fig. 5). On average, there were more drinking events in the free corner than in the working corner during the run WW (main effect corner: $F_{1,54} = 377.62$; $p < 0.0001$). Drinking events in the working corner decreased with increasing price whereas drinking events in the free corner increased (interaction: $F_{8,54} = 11.2$; $p < 0.0001$). In addition, drinking events appeared to decrease slightly with increasing days (main effect day: $F_{8,48} = 2.07$; $p = 0.058$).

Maximum price paid

The maximum price paid depended on the liquids (Coxph: $p < 0.0001$, Fig. 6). The mice paid the highest price (performed the highest number of required nosepokes) in run WQ. After 64 days, the run was stopped. Ten out of twelve

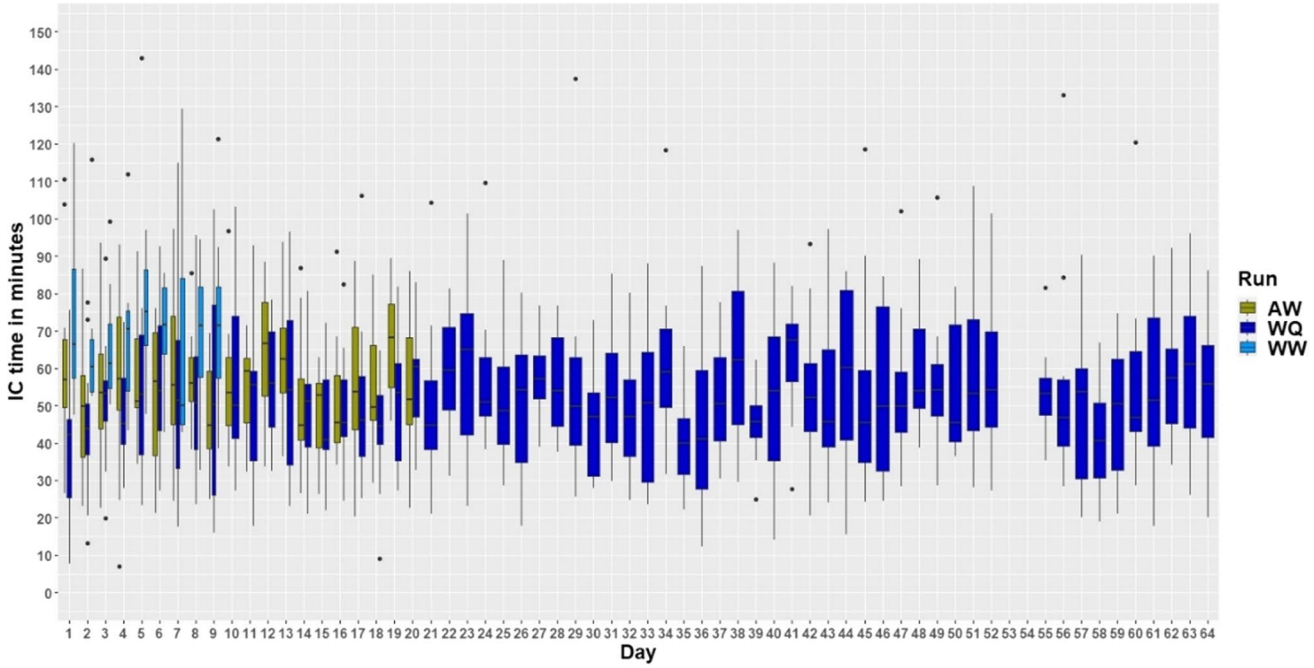

Fig. 2 Time the mice spent within the IntelliCage for the runs WW, WQ, and AW (W = water, Q = quinine, A = almond milk). The data for days 53 and 54 of run WQ are missing due to technical problems with the AnimalGate. On the *y*-axis, time spent in the IC by the mice is shown in minutes. The *x*-axis shows the experimental days, which can be equated with the price (number of nosepokes) for the liquids within the working corner

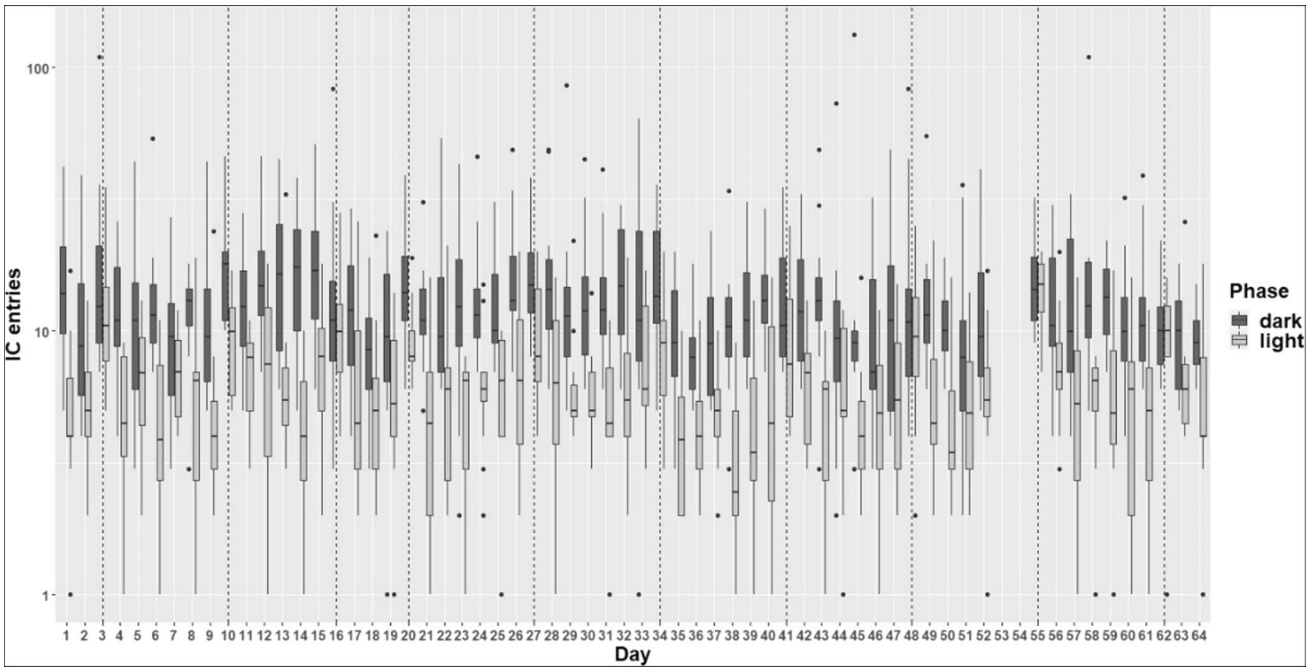

Fig. 3 Number of entries the mice made during the run WQ. The data for days 53 and 54 of run WQ are missing due to technical problems with the AnimalGate. On the *y*-axis, the IC entries are shown. The *x*-axis shows the experimental days, which can be equated with the price for the liquids within the working corner. The number of IC entries are shown on a logarithmic scale, while the labels are retained on the original scale. The *dashed lines* mark the days on which the setup was cleaned

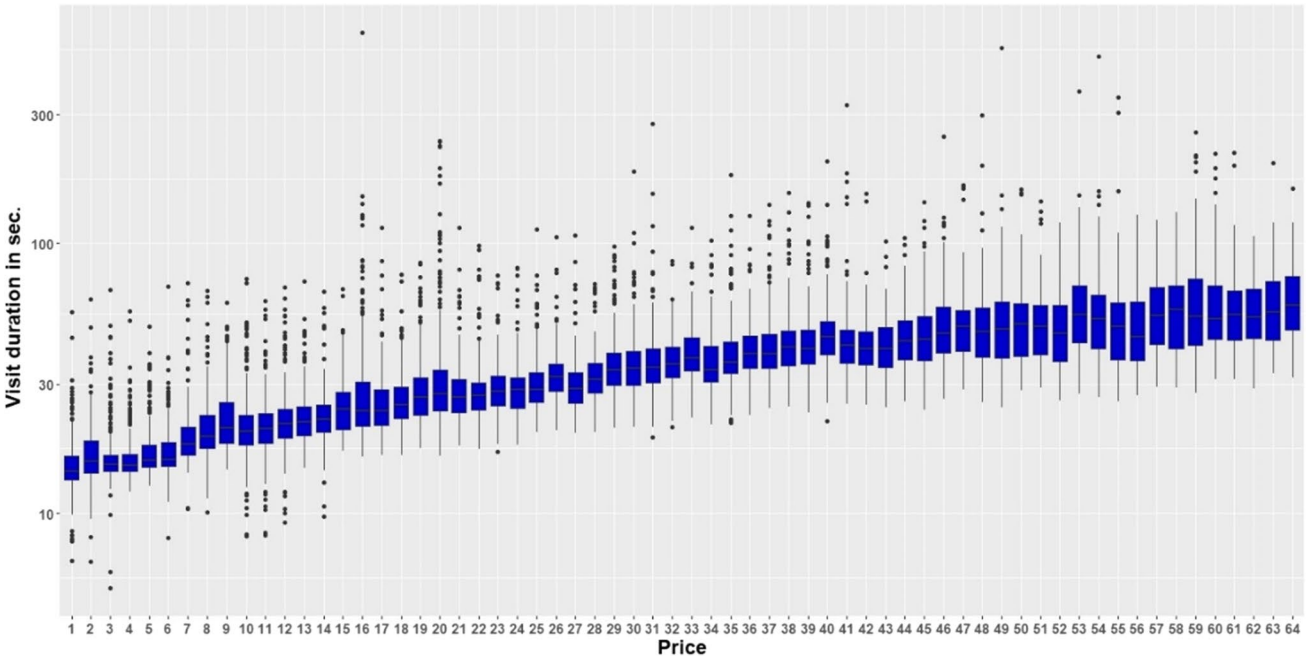

Fig. 4 Visit duration in working corner for run WQ (W = water, Q = quinine). On the *y*-axis, the time the mice spent within the working corner is shown. The *x*-axis shows the price the mice had to pay for the access to water. The price can be equated with the experimental days. The visit duration is shown on a logarithmic scale, while the original scale is retained for the axis labels

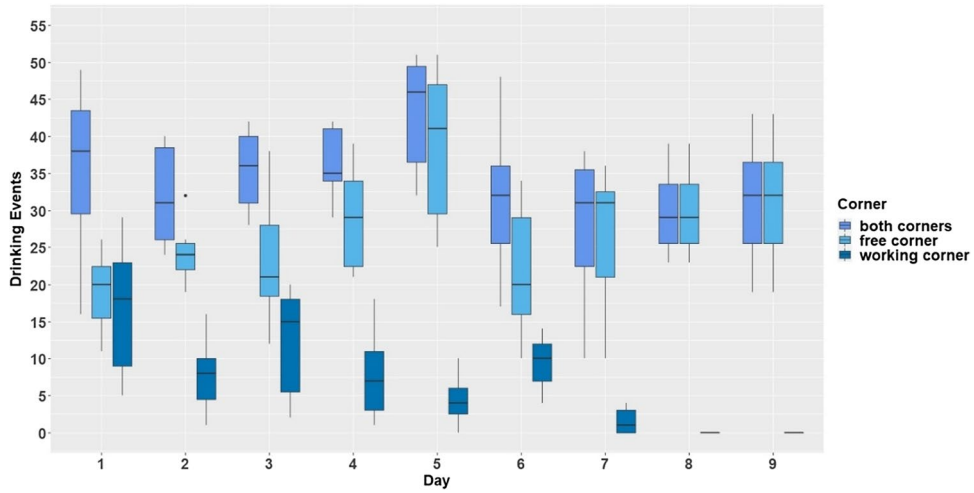

Fig. 5 Comparison of drinking events for water in run WW (W = water). The *y*-axis shows the drinking events which the mice made within the working corner and the free corner. The *x*-axis shows the experimental day. The day can be equated with the price the mice had to pay for access to water in the working corner while water within the free corner was available for the price of one nosepoke for all days

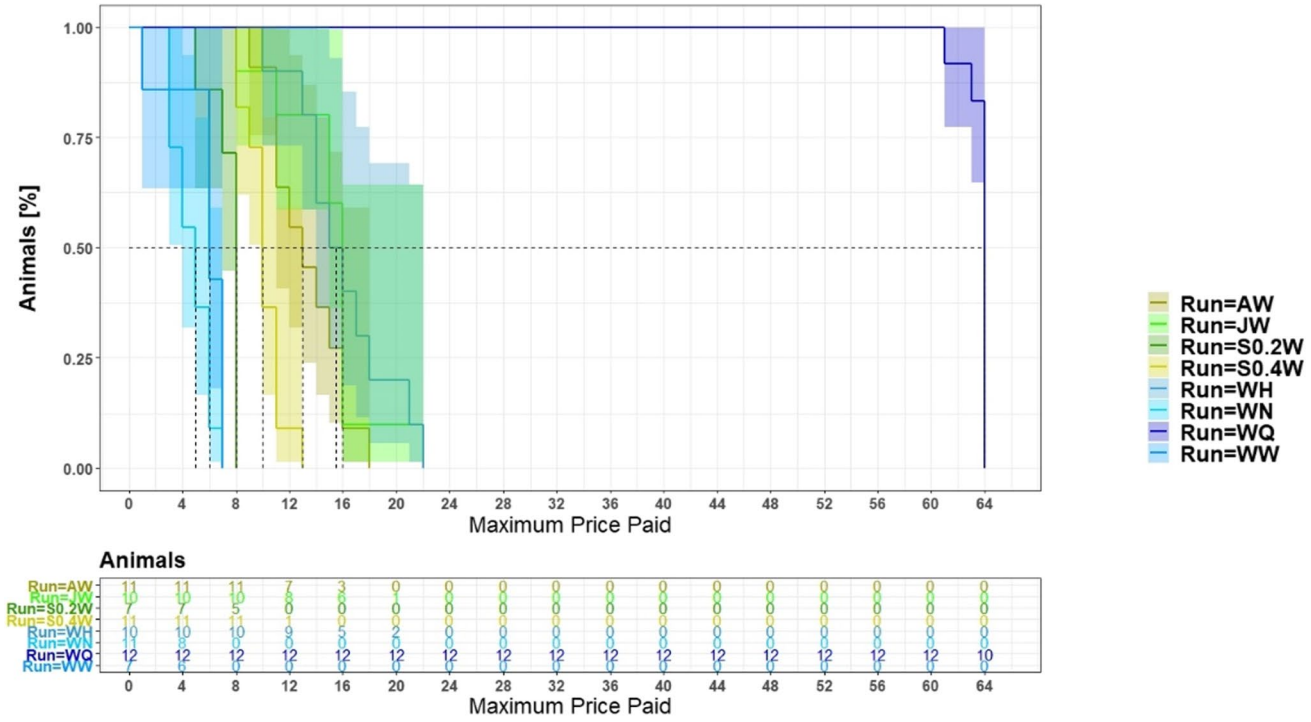

Fig. 6 Proportion of mice with specific maximum price paid in the different runs. The highlighted areas are the confidence intervals. The *y*-axis shows the animals which paid the required price. The *x*-axis shows the price the mice had to pay for the access to the liquids. The price is to be equated with the experimental days. W = water, A = almond milk, Q = quinine, N = NaCl, S = sucrose, H = HCl, J = apple juice. For order of runs and sample sizes, see Table 1

mice made up to 64 nosepokes to gain access to the tap water in the working corner when quinine water was provided in the free corner. To see if this overall influence was caused mainly by run WQ, the data of run WQ were removed for an additional analysis. The influence of the liquid combinations on participation could still be supported (Coxph: $p < 0.0001$).

In run WW and WN (W = water, N = NaCl), the mice paid the lowest maximum price with up to seven nosepokes to gain access to the liquid in the working corner. Mice paid

an equally low price for access to a 0.2 mM sucrose solution. For access to a higher concentrated sugar solution (0.4 mM), the mice made up to 13 nosepokes. Up to 18 nosepokes were paid for access to almond milk. In run JW and WH (J = apple juice, W = water, H = HCl), the mice made up to 22 nosepokes to gain access to the liquids in the working corner.

Consumer demand curve analysis

To investigate the willingness of the mice to work for different liquids the slopes of the demand curves were analyzed (Fig. 7, Table 2). The demand curves show the consumed amount (drinking events) on the *y*-axis in relation to the necessary price (required nosepoke number) on the *x*-axis. A more negative slope indicates a lower motivation of the mice to work for the access to the liquids. The comparison of the slopes of all demand curves showed differences compared to the slope of run WW except for run S0.2W. The demand curve of run WQ had the flattest slope. The slopes of run S0.4W, run WH, run JW, and run AW were steeper compared to run WW. The demand curve of run WN had the steepest slope.

In addition, the liquid amount consumed for the price of one single nosepoke can be analyzed. The comparison of the amount of drinking events for the price of one nosepoke showed that except for run WN, all runs differed from run WW (Fig. 7, Table 2: Intercept). The smallest amount was consumed in run WW and WN. The largest amount of drinking events was in run JW. The amounts of drinking events of the runs WQ, S0.2W, S0.4W, WH, and run AW were in between of run WW and run JW.

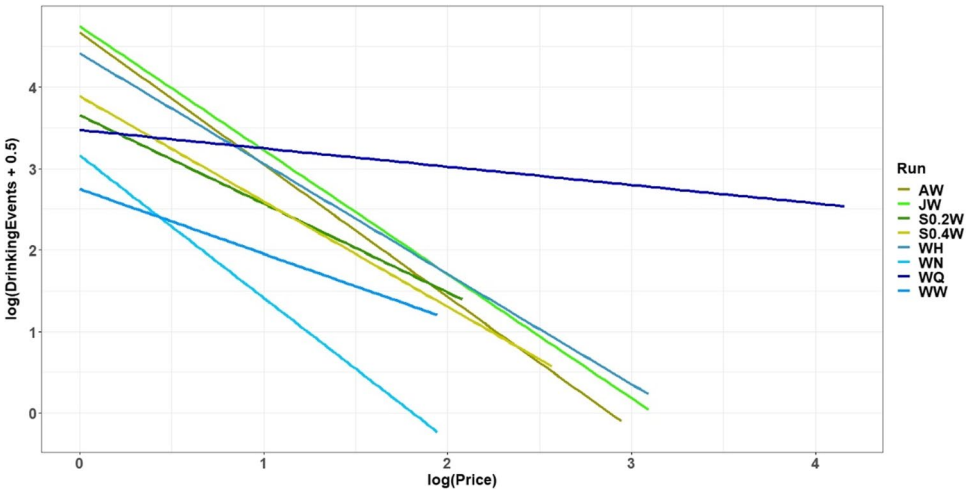

Fig. 7 Consumer demand curves. The data are plotted logarithmically. On the *x*-axis are the required nosepoke numbers for the access to the different liquids (price). On the *y*-axis are the values for the number of drinking events for each liquid. The curves end at the maximum number of nosepokes that was reached by any of the mice (A = almond milk, W = water, Q = quinine, N = NaCl, S = sucrose, H = HCl, J = apple juice)

Table 2 Results of the consumer demand analysis

Run	Slopes	Intercept	*p* slopes	*df*	t	*p* intercept	*df*	t
WN	– 1.75	2.78	0.00	1634	– 4.27	0.24	59	1.19
AW	– 1.62	4.67	0.00	1634	– 4.42	0.00	59	5.75
JW	– 1.52	4.77	< 0.001	1634	– 3.89	0.00	59	5.91
WH	– 1.35	4.43	< 0.01	1634	– 3.01	0.00	59	5.01
S0.4W	– 1.29	3.89	0.01	1634	– 2.56	< 0.01	59	3.36
S0.2W	– 1.09	3.65	0.21	1634	– 1.25	0.02	59	2.34
WW	– 0.79	2.74	reference	1634	– 4.55	reference	1634	9.66
WQ	– 0.23	3.47	< 0.01	1634	3.20	0.02	59	2.32

Run WW served as reference. W = water, Q = quinine, A = almond milk, N = NaCl, S = sucrose, H = HCl, J = apple juice, *df* = degrees of freedom

Discussion

The aim of the present study was to develop an automated and home-cage based test to determine the strength of preferences. For this purpose, we developed a test based on consumer demand theory and determined the strength of preference of mice for different liquids. Our test is using the RFID-based IC system, which makes it possible to test up to 12 mice in one social group over several months while obtaining individual data.

In our setup, the mice were able to independently enter the test-cage (IC) from the home-cage through a gate (AnimalGate). This allowed the mice to work undisturbed by other group members in order to gain access to the liquids. This was necessary because otherwise it would have been possible for group members to gain access to the corner by pulling, pushing, or biting the mouse that was "working". Anecdotally, we can report such behavior in experiments where multiple mice were housed within the IC. It is important to note, however, that the approach of individually channeling mice out of the cage can also result in "wait times" for the other mice in the home-cage. Other studies in which animals were allowed to enter the experimental cage individually have already examined how well individual entry worked, how long this entry lasted, how long habituation took, and how well animals performed in the actual test within the connected test-cage (Kaupert et al., 2017; Mei et al., 2020; Rivalan et al., 2017; Winter & Schaefers, 2011). Since the occupancy of the experimental cage could have an impact on the performance of the other animals, we are also interested in the occupancy of the test-cage. It was found that the time spent in the IC depended on the liquids offered. One might expect that the mice would spend more time in the IC if something positive, such as almond milk, was offered in addition to water. Interestingly, however, the mice spent more time in the IC when water was offered in both, the working and the free corner. The question arose whether some individuals occupied the IC more frequently than others, which would mean that access to the IC would be strongly influenced by these individuals. However, our analysis shows that this does not seem to be the case as we did not detect strong individual variation regarding the overall duration of IC time. This suggests that no single mouse consistently prevented other mice from accessing the liquids by primarily occupying the IC. With increasing required workload, the time spent in the working corner within the IC increased, however, the time spent in the IC per entry was not affected by the price. This suggests that above getting access to the liquids, the stay in the IC is perceived as an opportunity to explore additional space. This is consistent with the results of Sherwin (Sherwin, 2004), who showed that even with increasing price, mice continued to work for access to additional space (although the number of visits and time decreased with increasing price).

On average, the mice spent about 70 min in the IC during the run WW (W = water), which means that the IC is highly used when seven animals are present. Nevertheless, all mice were able to enter the IC and drink. Otherwise, we would have had to offer water separately to the mice that did not drink within 24 h as the IC system automatically warns if a mouse did not drink within 24 h. This was not the case during the entire consumer demand experiment. Through the IC entries, we were also able to show that the mice entered the IC primarily during the dark phase, which is the active phase of laboratory mice. This is in agreement with results of previous home-cage based experiments (Mei et al., 2020; Winter & Schaefers, 2011). However, our mice entered the test-cage more frequently on average (about 11 entries in our study compared to 5.5 entries per day in Winter & Schaefers, 2011). This may be due to the fact that in our study, any liquids were only offered in the test-cage and thereby forcing the mice to enter the test-cage whenever they felt thirsty.

Weekly cleaning of the cages affected the activity in terms of the number of entries made to the IC. The influence of cage changes on activity has already been shown using home-cage based activity measurement (Pernold et al., 2019). However, since all runs lasted for several weeks, we believe it is reasonable to assume that weekly cleaning of the cages did not influence the price the mice were willing to pay. Moreover, in all cases, the mice had two days to rejoin a run if they did not work for 1 day to access the offered liquid. Only after the mice had not worked for the access for two consecutive days, was the run ended for them. For differently structured experiments, however, the changes in daily activity related to cage cleaning may be of importance. Therefore, we recommend for home-cage based experimental designs to cautiously consider effects of intervention by the experimenter (i.e., cage cleaning, health inspections).

To assess the price paid by the mice, the amount of time the mice spent in the working corner was examined for the run WQ (W = water, Q = quinine). As the price increased, the mice also spent more time in the working corner. This shows that in addition to the movement expended to execute the nosepokes, work time can also be considered as another price component. To our knowledge, it had not been considered in recent consumer demand experiments how much time the animals had to spend on the work. In our experiment, this factor was of additional importance, as it possibly affects separating/singulating the mice into the test-cage, since only one mouse can be in the IC at a time. Nevertheless, as stated above, the overall occupation time of the IC was not affected by the increased amount of time spent in the corner.

Regarding home-cage based testing, it can be summarized that it is a well-functioning system for female mice to obtain individual data despite group housing. It should be noted, however, that male mice show much more conspicuous dominance behavior (Van Loo et al., 2004). However, we have recently shown that groups of 12 male mice of the strain C57Bl/6J can be housed without notable aggressive behavior in the IC for a very long time (Kahnau et al., 2021). To validate the suitability of our proposed home-cage based system for male mice, the same experimental design should be performed with males in a future study.

For the evaluation of motivation (the strength of preference) for or against a certain good, the maximum price the animals are willing to pay can be taken into account and compared for different goods (Kirkden et al., 2003). It has been stated that in order to compare the demand of different goods with each other, a benchmark value with a necessary good such as food should be generated (Cooper, 2004; Dawkins, 1983). We believe that our dataset indeed can be used as such a benchmark as it provides information how water as a necessity relates to different liquids either tasting better or worse. However, it should be noted that different wants for goods can interact with each other. Therefore, it is important to compare different motivations for wants in a meaningful way (Gygax, 2017) in order to obtain a suitable benchmark. Since water was offered in all runs in our study, the run WW, in which water was offered in both the working and free corner, was chosen as a reference. Although the drinking events in the free corner were higher than in the working corner, over 50% of the mice were willing to make seven nosepokes for getting access to water in the working corner. As the price exceeded seven nosepokes, all mice refused to work for water while they could have it for free in the other corner. The run WW was deliberately conducted at the end of the whole experiment because this enabled testing how willing the mice were to make nosepokes for water even after long experimental duration. However, the mice might have developed a habit to do nosepokes but this formed routine could not be related to the corner itself, as the position for the working and free corner within the test-cage were changed after each run. Therefore, we assume that the run WW can be used as a valid benchmark in our consumer demand experiment.

The phenomenon to perform an operant task in order to receive a reward in spite of the same reward being additionally available for free, is known as "contrafreeloading" (Jensen, 1963). Past studies showed that different species worked for access to food even while food was freely available. There seem to be individual differences as well as genetic influences (Jensen, Schütz, & Lindqvist, 2002; Lindqvist & Jensen, 2009). The willingness to work voluntarily despite not being obliged to do so, can be seen as an indication that work in itself has rewarding properties. This is especially true for laboratory animals, which usually live under conditions that limit their experience (Lewejohann et al., 2020). While wild mice spend time for foraging behavior, nest building or breeding, laboratory mice have a lot of time on their hands as there is not much else to do while they are “waiting" for the next experiment. Consequently, the determined boundary of seven nosepokes, which were performed as contrafreeloading, might serve as a benchmark in our artificial economy. This benchmark would indicate the maximum number of nosepokes mice are willing to perform due to their lack of alternative activities. This is also supported by the finding that the mice drank less water overall from the seventh day onwards during the run WW. If the "work" becomes too "expensive", a smaller amount of water is drunk, i.e., the water intake in the working corner is added to the basic requirement during contrafreeloading.

In the analysis of the maximum paid price, liquids for which the mice performed more than these seven nosepokes might be considered as having a higher priority than work in itself. In our study, this is true for all liquid combinations except water compared with NaCl, because the mice made only up to seven nosepokes in the run WN. The aversion to a NaCl concentration of 10 mM did not seem to be very strong, because the mice were not willing to work more not to drink this. The run S0.2W (0.2 mM sucrose concentration) also did only differ by one additional nosepoke with regard to the maximum paid price compared to working for water in both corners. Refusing to work more than eight nosepokes for a 0.2 mM sucrose concentration indicates a low strength of preference for mildly sweetened water.

To determine which goods are necessities or luxuries, a consumer demand curve can be plotted on logarithmic scales depicting the relation of the quantity consumed and the increase in price. A demand curve with low elasticity (the slope is hardly influenced by price) indicates necessary goods. However, a demand curve with high elasticity (slope strongly influenced by price) indicates luxury goods. It is important to note that if the quantity of goods that can be acquired per "purchase" is not constant, the price itself changes in terms of inflation. Therefore, Kirkden and Pajor note that the quantity of the good to be consumed should remain the same at any price to avoid other factors, such as time, influence the animal’s motivation (Kirkden & Pajor, 2006). Accordingly, the price of access to the liquids changed in our study, but the time the mice were able to drink remained constant (10 s). In our study, all consumer demand curves of the different runs were compared to the run in which water was offered in the working and free corner (run WW). The slope of the consumer demand curve of run S0.2W (S 0.2 = 0.2 mM sucrose) did not differ from the slope of the consumer demand curve of run WW. The slope of the consumer demand curve of run WN (N = NaCl) was even greater than the slope of run WW. This indicated

that the demand curve of WN was more influenced by the price (high price elasticity) and indicated a low motivation to work for access to water while access to 10 mM NaCl concentration was available for the price of only one nosepoke. In contrast to this, the motivation to work for water while access to quinine was available for the price of one nosepoke seemed to be very high. Accordingly, the price elasticity in the run WQ is the lowest.

In human microeconomics, consumer demand theory is based upon the amount of disposable income that can be spent on different goods in the market. In our experimental setup, however, there is only one good that a mouse can work for at a time. As a consequence, the value of the liquids might be overrated due to this methodological constraint. However, our approach allows us to directly relate the worth of the goods on the market to the workload the mice are willing to pay for access. Nevertheless, in our consumer demand test, all curves, except run WQ, seem to show a high price elasticity. This could be due to the fact that in all runs an alternative was offered in the free corner and thus the need to work was less strong. Therefore, it is appropriate to determine the motivation for getting different liquids, additionally by analyzing the maximum price paid.

To evaluate the motivation to obtain goods it is essential to ensure that the animals have indeed learned the operant task in order to exclude misinterpretation (Dawkins, 1990; Rutter & Duncan, 1992), which was also shown by our results of the pre-test (data shown in the supplements). In addition, the time point when the test is performed should be considered. Acosta and colleagues showed that for mice, which are nocturnal, the motivation to work for food is higher at night than during the day (Acosta et al., 2020). Also in our study, mice entered the test-cage more frequently in the dark phase than in the light phase. Thus, considering the time point of performance is crucial for avoiding misinterpretation of the demand curve or maximum paid price. Basically, the maximum amount of work the mice paid for the access to the different liquids tested can serve as a simple benchmark for future studies.

Home-cage based test setups have proven to be useful tools to overcome issues such as day/night rhythm or experimenter influence (reviewed in Richardson, 2012 and Voikar & Gaburro, 2020). In some experiments, it is necessary to keep the animals separately to obtain individual data. However, as mice are social animals, single-housing should be avoided if possible. So far, there are not many systems that allow testing mice in groups while obtaining individual data (some examples reviewed in Voikar & Gaburro, 2020). For the development of such automated and home-cage based systems, it is possible to use a gate to connect the home-cage to the test-cage (Winter & Schaefers, 2011). For example, Mei and colleagues used such a gate to connect a home-cage to an eight-arm radial maze (Mei et al., 2020). We also demonstrated that the mice were able to independently enter the IC several times a day to access the obtained liquids in the IC.

In our study, we developed an automated and home-cage based test setup by using the IC system, in which the mice were tested over several months, in their social group, familiar environment and during their active phase. As a result, influences such as the day/night rhythm or the experimenter could be minimized. In addition, each mouse could work undisturbed by cage mates for access to the liquids. This had the benefit that individual mice took the time they needed to pay the required price. Especially in run WQ, in which the mice made up to 64 nosepokes, the execution of the required nosepoke number took some time (as the duration increased with increasing price), the interruption-free environment will probably have facilitated the task. The mice were able to perform the operant task repeatedly, which reflected the motivation of the animal. By connecting the test-cage to the home-cage, the mice were free to choose if and when to do the required nosepoke number in the working corner (unless a cage mate was currently occupying the IC). Furthermore, by having the mice work again for access to water while they had free access to a bitter-tasting liquid in the free corner after the last run, we were able to show that even in old age the learning task was successfully performed by the mice (see supplement S1).

The consumer demand curves and number of animals that paid the corresponding price showed that the motivation was different and depended on the liquid. A previously conducted preference test already showed that almond milk and apple juice were preferred and the sour- and bitter-tasting liquids were least preferred. With the consumer demand test, these preferences were confirmed. Furthermore, we can show that the aversion to the bitter-tasting liquid is markedly stronger compared to the sour solution but also compared to the preference to almond milk or apple juice. The results may be used to select suitable stimuli for operant tasks in order to optimize learning behavior. We also found that operant conditioning was highly facilitated when the mice had to work for water to avoid drinking a bitter-tasting solution. The experience gained in that trial could then easily be transferred to working for rewarding liquids in consecutive trials. In addition, the data might be of interest for the husbandry of laboratory mice where acidified water is quite common due to the fact that acidification is used to keep water as pathogen-free as possible. Our data indicate a relatively strong aversion to a 10 mM HCl solution (pH = 2.5). This could serve as a reference for acidifying water to facilitate fluid intake by the animals.

With our test setup, it is currently only possible to examine different liquids. However, this setup is a proof of concept for future studies for example in order to optimize housing conditions. For example, when letting the animals choose

between different enrichment items, it would be important to also know how strong their preference is. Therefore, we are in the process of developing an automated and home-cage based test setup, combining the mouse positioning surveillance system (MoPSS, Habedank et al., 2021) with the knowledge gained from this study. In our view, one major lesson learned is to let the mice enter the test-cage independently and thus work undisturbed in it. We showed that this experiment could be carried out without a large amount of personnel time (approx. 30–40 min daily (checking the animals and their drinking behavior, preparing and changing liquids, checking the apparatus, approx. 1.5 h weekly cleaning of the setup). In addition, a certain basic technical understanding is advantageous as well as a daily control of the data to check whether the setup is running properly. This knowledge will give us the opportunity to integrate the animals' point of view in the husbandry as well as the experiments themselves and a more comprehensive understanding of the needs and wants of our laboratory mice.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.3758/s13428-022-01813-8. The raw data of the experiment can be found under: https://zenodo.org/record/6325238#.Yih7FXyZNPY.

Acknowledgements The authors thank the animal caretakers of the German Federal Institute of Risk Assessment, especially Carola Schwarck, for their support in the animal husbandry.

Funding Open Access funding enabled and organized by Projekt DEAL. This work was funded by the DFG (FOR 2591; LE 2356/5-1). The authors declare no competing interests.

Declarations

Ethical approval All experiments were approved by the Berlin state authority, Landesamt für Gesundheit und Soziales, under license No. G 0182/17 and were in accordance with the German Animal Protection Law (TierSchG, TierSchVersV).

References

Acosta, J., Bussi, I. L., Esquivel, M., Höcht, C., Golombek, D. A., & Agostino, P. V. (2020). Circadian modulation of motivation in mice. *Behavioural Brain Research, 382*(January), 112471. https://doi.org/10.1016/j.bbr.2020.112471

Ago, A., Gonda, T., Takechi, M., Takeuchi, T., & Kawakami, K. (2002). Preferences for paper bedding material of the laboratory mice. *Experimental Animals, 51*(2), 157–161.

Banjanin, S., & Mrosovsky, N. (2000). Preferences of mice, Mus Mus culus, for different types of running wheel. *Laboratory Animals, 34*(3), 313–318. https://doi.org/10.1258/002367700780384681

Bates, D., Machler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using {lme4}. *Journal of Statistical Soft ware, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Canty, A., & Ripley, B. D. (2021). boot: Bootstrap R (S-plus) function. R Package. Version 1.3–28.

Chmiel, D. J., & Noonan, M. (1996). Preference of laboratory rats for potentially enriching stimulus objects. *Laboratory Animals, 30*(2), 97–101. https://doi.org/10.1258/002367796780865790

Cooper, J. J. (2004). Consumer demand under commercial husbandry conditions: Practical advice on measuring Behavioural priorities in captive animals. *Animal Welfare* 13(SUPPL.):47–56.

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their applications*. Cambridge University Press.

Dawkins, M. S. (1983). Battery Hens Name Their Price: Consumer Demand Theory and the Measurement of Ethological 'Needs'. *Animal Behaviour, 31*(4), 1195–1205. https://doi.org/10.1016/S0003-3472(83)80026-8

Dawkins, M. S. (1988). Behavioural deprivation: A central problem in animal welfare. *Applied Animal Behaviour Science, 20*(3–4), 209–225. https://doi.org/10.1016/0168-1591(88)90047-0

Dawkins, M. S. (1990). From and animals point of view - motivation, fitness and animal welfare. *Behavioral and Brain Science, 13*, 1–61.

Endo, T., Maekawa, F., Võikar, V., Haijima, A., Uemura, Y., Zhang, Y., ..., Kakeyama, M. (2011). Automated test of behavioral flex ibility in mice using a behavioral sequencing task in IntelliCage. *Behavioural Brain Research, 221*(1), 172–181. https://doi.org/10.1016/j.bbr.2011.02.037

Freymann, J., Tsai, P. P., Stelzer, H., & Hackbarth, H. (2017). The impact of bedding volumes on laboratory mice. *Applied Animal Behaviour Science, 186*, 72–79. https://doi.org/10.1016/j.applanim.2016.11.004

Galsworthy, M. J., Amrein, I., Kuptsov, P. A., Poletaeva, I. I., Zinn, P., Rau, A., ..., Lipp, H. P. (2005). A comparison of wild-caught wood mice and Bank voles in the IntelliCage: Assessing explora tion, daily activity patterns and place learning paradigms. *Behav ioural Brain Research, 157*(2), 211–217. https://doi.org/10.1016/j.bbr.2004.06.021

Gygax, L. (2017). Wanting, liking and welfare: The role of affective states in proximate control of behaviour in vertebrates. *Ethology, 123*(10), 689–704. https://doi.org/10.1111/eth.12655

Habedank, A., Kahnau, P., Diederich, K., & Lewejohann, L. (2018). Severity assessment from an Animal's point of view. *Berliner Und Munchener Tierarztliche Wochenschrift, 131*(7–8), 304–320. https://doi.org/10.2376/0005-9366-18007

Habedank, A., Urmersbach, B., Kahnau, P., & Lewejohann, L. (2021). O mouse, where art thou? The mouse position surveillance system (MoPSS)—An RFID-based tracking system. *Behavior Research Methods.* https://doi.org/10.3758/s13428-021-01593-7

Jensen, G. D. (1963). Preference for Bar pressing over 'freeloading' as a function of number of rewarded presses. *Journal of Experimen tal Psychology, 65*(5), 451–454.

Jensen, P., Schütz, K., & Lindqvist, C. (2002). Red jungle fowl have more Contrafreeloading than white Leghorn layers: Effect of food deprivation and consequences for information gain. *Behaviour, 139*(9), 1195–1209.

Kahnau, P., Guenther, A., Boon, M. N., Terzenbach, J. D., Hanitzsch, E., Lewejohann, L., & Brust, V. (2021). Lifetime observation of cognition and physiological parameters in male mice. *Frontiers*

in Behavioral Neuroscience, 15(September). https://doi.org/10.3389/fnbeh.2021.709775

Kassambara, A., Kosinski, M., & Biecek, P. (2020). *Survminer: Drawing survival curves using 'Ggplot2'. R package version 0.4.8.*

Kaupert, U., Thurley, K., Frei, K., Bagorda, F., Schatz, A., Tocker, G., Rapoport, S., Derdikman, D., & Winter, Y. (2017). Spatial cognition in a virtual reality home-cage extension for freely moving rodents. *Journal of Neurophysiology, 117*(4), 1736–1748. https://doi.org/10.1152/jn.00630.2016

Kirkden, R. D., & Pajor, E. A. (2006). Using preference, motivation and aversion tests to ask scientific questions about animals' feelings. *Applied Animal Behaviour Science, 100*(1–2), 29–47. https://doi.org/10.1016/j.applanim.2006.04.009

Kirkden, R. D., Edwards, J. S. S., & Broom, D. M. (2003). A theoretical comparison of the consumer surplus and the elasticities of demand as measures of motivational strength. *Animal Behaviour, 65*(1), 157–178. https://doi.org/10.1006/anbe.2002.2035

Krackow, S., Vannoni, E., Codita, A., Mohammed, A. H., Cirulli, F., Branchi, I., ..., Lipp, H. P. (2010). Consistent behavioral phenotype differences between inbred mouse strains in the IntelliCage. *Genes, Brain and Behavior, 9*(7), 722–731. https://doi.org/10.1111/j.1601-183X.2010.00606.x

Krohn, T. C., Sørensen, D. B., Ottesen, J. L., & Hansen, A. K. (2006). The effects of individual housing on mice and rats: A review. *Animal Welfare, 15*(4), 343–352.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). ImerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Ladewig, J., Sørensen, D. B., Nielsen, P. P., & Matthews, L. R. (2002). The quantitative measurement of motivation: Generation of demand functions under open versus closed economies. *Applied Animal Behaviour Science, 79*(4), 325–331. https://doi.org/10.1016/S0168-1591(02)00156-9

Lea, S. E. (1978). The psychology and economics of demand. *Psychological Bulletin, 85*(3), 441–466. https://doi.org/10.1037/0033-2909.85.3.441

Lewejohann, L., & Sachser, N. (2000). Evaluation of Di Erent housing conditions for male laboratory mice by means of preference tests. *KTBL Schrift.*, 170–177.

Lewejohann, L., Schwabe, K., Häger, C., & Jirkof, P. (2020). Impulse for animal welfare outside the experiment. *Laboratory Animals, 54*(2), 150–158. https://doi.org/10.1177/0023677219891754

Lindqvist, C., & Jensen, P. (2009). Domestication and stress effects on Contrafreeloading and spatial learning performance in red jungle fowl (Gallus Gallus) and white Leghorn layers. *Behavioural Processes, 81*(1), 80–84. https://doi.org/10.1016/j.beproc.2009.02.005

Manouze, H., Ghestem, A., Poillerat, V., Bennis, M., Ba-M'hamed, S., Benoliel, J. J., ..., Bernard, C. (2019). Effects of single cage housing on stress, cognitive, and seizure parameters in the rat and mouse pilocarpine models of epilepsy. *ENeuro, 6*(4). https://doi.org/10.1523/ENEURO.0179-18.2019

Manser, C. E., Broom, D. M., Overend, P., & Morris, T. H. (1998). Investigations into the preferences of laboratory rats for Nestboxes and nesting materials. *Laboratory Animals, 32*(1), 23–35. https://doi.org/10.1258/002367798780559365

Mechan, A. O., Wyss, A., Rieger, H., & Hasan Mohajeri, M. (2009). A comparison of learning and memory characteristics of young and middle-aged wild-type mice in the IntelliCage. *Journal of Neuroscience Methods, 180*(1), 43–51. https://doi.org/10.1016/j.jneumeth.2009.02.018

Mei, J., Kohler, J., Winter, Y., Spies, C., Endres, M., Banneke, S., & Emmrich, J. V. (2020). Automated radial 8-arm maze: A voluntary and stress-free behavior test to assess spatial learning and memory in mice. *Behavioural Brain Research, 381*(November), 112352. https://doi.org/10.1016/j.bbr.2019.112352

Patterson-Kane, E. G., Harper, D. N., & Hunt, M. (2001). The cage preferences of laboratory rats. *Laboratory Animals, 35*(1), 74–79. https://doi.org/10.1258/0023677011911390

Pernold, K., Iannello, F., Low, B. E., Rigamonti, M., Rosati, G., Scavizzi, F., ..., Ulfhake, B. (2019). Towards large scale automated cage monitoring - diurnal rhythm and impact of interventions on in-cage activity of C57BL/6J mice recorded 24/7 with a non-disrupting capacitive-based technique. *PLoS One, 14*(2), 1–20. https://doi.org/10.1371/journal.pone.0211063

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2020). nlme: Linear and nonlinear mixed effects models. R package. Version 3.1–150. https://CRAN.R-project.org/package=nlme

R Core Team. (2020). R: *A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Richardson, C. A. (2012). Automated Homecage Behavioural analysis and the implementation of the three Rs in research involving mice. *Alternatives to Laboratory Animals : ATLA, 40*(5). https://doi.org/10.1177/026119291204000513

Rivalan, M., Munawar, H., Fuchs, A., & Winter, Y. (2017). An automated, experimenter-free method for the standardised, operant cognitive testing of rats. *PLoS One, 12*(1), 1–19. https://doi.org/10.1371/journal.pone.0169476

Rutter, S. M., & Duncan, I. J. H. (1992). Measuring aversion in domestic fowl using passive avoidance. *Applied Animal Behaviour Science, 33*(1), 53–61. https://doi.org/10.1016/S0168-1591(05)80084-X

Sarna, J. R., Dyck, R. H., & Whishaw, I. Q. (2000). The Dalila effect: C57BL6 mice barber whiskers by plucking. *Behavioural Brain Research, 108*(1), 39–45. https://doi.org/10.1016/S0166-4328(99)00137-0

Sherwin, C. M. (1998). The use and perceived importance of three resources which provide caged laboratory mice the opportunity for extended locomotion. *Applied Animal Behaviour Science, 55*(3–4), 353–367. https://doi.org/10.1016/S0168-1591(97)00049-X

Sherwin, C. M. (2003). Social context affects the motivation of laboratory mice, Mus Musculus, to gain access to resources. *Animal Behaviour, 66*(4), 649–655. https://doi.org/10.1006/anbe.2003.2239

Sherwin, C. M. (2004). The motivation of group-housed laboratory mice, Mus Musculus, for additional space. *Animal Behaviour, 67*(4), 711–717. https://doi.org/10.1016/j.anbehav.2003.08.018

Sherwin, C. M. (2007). The motivation of group-housed laboratory mice to leave an enriched laboratory cage. *Animal Behaviour, 73*(1), 29–35. https://doi.org/10.1016/j.anbehav.2006.04.008

Sherwin, C. M., & Nicol, C. J. (1996). Reorganization of behaviour in laboratory mice, Mus Musculus, with varying cost of access to resources. *Animal Behaviour, 51*(5), 1087–1093. https://doi.org/10.1006/anbe.1996.0110

Sherwin, C. M., & Nicol, C. J. (1997). Behavioural demand functions of caged laboratory mice for additional space. *Animal Behaviour, 53*(1), 67–74. https://doi.org/10.1006/anbe.1996.0278

Sørensen, D. B., Ladewig, J., Ersbøll, A. K., & Matthews, L. (2004). Using the cross point of demand functions to assess animal priorities. *Animal Behaviour, 68*(4), 949–955. https://doi.org/10.1016/j.anbehav.2003.12.017

Timberlake, W. (1984). A temporal limit on the effect of future food on current performance in an analogue of foraging and welfare. *Journal of the Experimental Analysis of Behavior, 41*(2), 117–124. https://doi.org/10.1901/jeab.1984.41-117

Van De Weerd, H. A., Van Loo, Van Zutphen, Koolhaas, J. M., & Baumans, V. (1998). Strength of preference for nesting material as environmental enrichment for laboratory mice. *Applied Animal*

Behaviour Science, 55(3–4), 369–382. https://doi.org/10.1016/S0168-1591(97)00043-9

Van Loo, P. L. P., Van De Weerd, Van Zutphen, & Baumans, V. (2004). Preference for social contact versus environmental enrichment in male laboratory mice. *Laboratory Animals, 38*(2), 178–188. https://doi.org/10.1258/002367704322968867

Van Loo, P. L. P., Blom, H. J. M., Meijer, M. K., & Baumans, V. (2005). Assessment of the use of two commercially available environmental enrichments by laboratory mice by preference testing. *Laboratory Animals, 39*(1), 58–67. https://doi.org/10.1258/0023677052886501

Voikar, V., & Gaburro, S. (2020). Three pillars of automated home-cage phenotyping of mice: Novel findings, refinement, and reproducibility based on literature and experience. *Frontiers in Behavioral Neuroscience, 14*(October), 1–15. https://doi.org/10.3389/fnbeh.2020.575434

Voikar, V., Krackow, S., Lipp, H. P., Rau, A., Colacicco, G., & Wolfer, D. P. (2018). Automated dissection of permanent effects of hippocampal or prefrontal lesions on performance at spatial, working memory and circadian timing tasks of C57BL/6 mice in IntelliCage. *Behavioural Brain Research, 352*, 8–22. https://doi.org/10.1016/j.bbr.2017.08.048

Warburton, H., & Mason, G. (2003). Is out of sight out of mind? The effects of resource cues on motivation in mink, Mustela Vison. *Animal Behaviour, 65*(4), 755–762. https://doi.org/10.1006/anbe.2003.2097

Warburton, H. J., & Nicol, C. J. (1998). Position of operant costs affects visits to resources by laboratory mice, Mus Musculus. *Animal Behaviour, 55*(5), 1325–1333. https://doi.org/10.1006/anbe.1997.0679

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag. https://ggplot2.tidyverse.org

Winter, Y., & Schaefers, A. T. U. (2011). A sorting system with automated gates permits individual operant experiments with mice from a social home cage. *Journal of Neuroscience Methods, 196*(2), 276–280. https://doi.org/10.1016/j.jneumeth.2011.01.017

2.4 Publication 4

Development of an IntelliCage-based cognitive bias test for mice

Pia Kahnau, Anne Jaap, Birk Urmersbach, Kai Diederich, Lars Lewejohann

Development of an IntelliCage-based cognitive bias test for mice
[version 1; peer review: 2 approved with reservations].

Published in journal “Open Research Europe” 2022, 2:128
DOI: https://doi.org/10.12688/openreseurope.15294.1

The Supplementary Material for this article can be found online at:
https://doi.org/10.5281/zenodo.7310180

Open Research Europe

Open Research Europe 2022, 2:128 Last updated: 12 SEP 2023

METHOD ARTICLE

Development of an IntelliCage-based cognitive bias test for mice [version 1; peer review: 2 approved with reservations]

Pia Kahnau[1], Anne Jaap[1], Birk Urmersbach[1], Kai Diederich[1], Lars Lewejohann[1,2]

[1]Laboratory Animal Science, German Federal Institute for Risk Assessment, Berlin, 10589, Germany
[2]Insitute of Animal Welfare, Animal Behavior and Laboratory Animal Science, Freie Universität Berlin, Berlin, 14163, Germany

v1 **First published:** 25 Nov 2022, **2**:128
https://doi.org/10.12688/openreseurope.15294.1
Latest published: 08 Aug 2023, **2**:128
https://doi.org/10.12688/openreseurope.15294.2

Abstract

The cognitive bias test is used to measure the emotional state of animals with regard to future expectations. Thus, the test offers a unique possibility to assess animal welfare with regard to housing and testing conditions of laboratory animals. So far, however, performing such a test is time-consuming and requires the presence of an experimenter. Therefore, we developed an automated and home-cage based cognitive bias test based on the IntelliCage system. We present several developmental steps to improve the experimental design leading to a successful measurement of cognitive bias in group-housed female C57BL/6J mice. The automated and home-cage based test design allows to obtain individual data from group-housed mice, to test the mice in their familiar environment, and during their active phase. By connecting the test-cage to the home-cage via a gating system, the mice participated in the test on a self-chosen schedule, indicating high motivation to actively participate in the experiment. We propose that this should have a positive effect on the animals themselves as well as on the data. Unexpectedly, the mice showed an optimistic cognitive bias after enrichment was removed and additional restraining. An optimistic expectation of the future as a consequence of worsening environmental conditions, however, can also be interpreted as an active coping strategy in which a potential profit is sought to be maximized through a higher willingness to take risks.

Keywords

cognitive bias, judgment bias, home-cage based, IntelliCage, conditioning, learning behavior, mice

This article is included in the Psychology gateway.

Open Peer Review

Approval Status ✓ ?

	1	2
version 2 (revision) 08 Aug 2023	✓ view	
version 1 25 Nov 2022	? view	? view

1. **David P. Wolfer**, University of Zurich, Zurich, Switzerland
ETH Zurich, Zurich, Switzerland

2. **Ewelina Knapska**, Nencki Institute of Experimental Biology, Warsaw, Poland

Any reports and responses or comments on the article can be found at the end of the article.

Open Research Europe

Open Research Europe 2022, 2:128 Last updated: 12 SEP 2023

This article is included in the Horizon 2020 gateway.

This article is included in the Neuroscience gateway.

This article is included in the COST Actions gateway.

Corresponding author: Pia Kahnau (pia.kahnau@bfr.bund.de)

Author roles: Kahnau P: Conceptualization, Formal Analysis, Investigation, Methodology, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing; **Jaap A**: Conceptualization, Writing – Review & Editing; **Urmersbach B**: Writing – Review & Editing; **Diederich K**: Project Administration, Writing – Review & Editing; **Lewejohann L**: Conceptualization, Funding Acquisition, Project Administration, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This research was financially supported by the European Union's Horizon 2020 research and innovation programme under the grant agreement No [861952](SGA3 COST – Maximising impact) through COST Action grant CA20135. This work was funded by the DFG (FOR 2591; LE 2356/5-1). The authors declare no competing interests.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Kahnau P, Jaap A, Urmersbach B *et al.* **Development of an IntelliCage-based cognitive bias test for mice [version 1; peer review: 2 approved with reservations]** Open Research Europe 2022, **2**:128 https://doi.org/10.12688/openreseurope.15294.1

First published: 25 Nov 2022, **2**:128 https://doi.org/10.12688/openreseurope.15294.1

Introduction

It has been shown that in both humans and animals, past experiences influence future expectations (Harding *et al.*, 2004; Mendl *et al.*, 2009; Paul *et al.*, 2005). Individuals with negative experiences or in a bad mood are more likely to be "pessimistic" about future events and, vice versa, individuals with positive experiences or in a good mood are more likely to be "optimistic". In the past, many tests for various species have been developed to investigate the emotional state of animals (Jirkof *et al.*, 2019). To examine the influence of emotional or affective states on expectations of future events, a number of cognitive bias tests have been developed (Boleij *et al.*, 2012; Harding *et al.*, 2004; Hintze *et al.*, 2018; Schlüns *et al.*, 2017; Verbeek *et al.*, 2014).

The common feature of these tests is the need for conditioning the subjects to scalable stimuli, *e.g.*, odors, tones, or spatial positions. The animals learn that they will receive a reward for the stimulus at one end of the scale and a punishment for the second stimulus at the other end of the scale. After successful conditioning, the actual test follows, in which ambiguous stimuli are presented to the animals. These ambiguous stimuli are located on the scale between the already known stimuli. The reaction towards these ambiguous stimuli is then measured and analyzed: It is assumed that if the response to the ambiguous stimulus is similar to the positively conditioned stimulus, the animals seem to expect a reward. In this case, they had a positive expectation of the future event, or in other words, they appear to be "optimistic". However, if the response resembles the response of the negatively conditioned stimulus, the animals seem to have a negative expectation or seem to be "pessimistic".

The first cognitive bias test was developed by Harding and colleagues in 2004 (Harding *et al.*, 2004). Rats were conditioned to press a lever in response to hearing the positively-associated tone-frequency to receive a reward or not to press a lever to avoid a punishment after hearing the negatively-associated tone-frequency. The cognitive bias test revealed that rats kept under unpredictable housing conditions were less likely to press the lever for a reward in response to ambiguous tone-frequencies than rats kept under normal housing conditions. It was thus concluded that the negative experience rendered them 'pessimistic'.

Although mice are the most commonly used experimental animals (Lewejohann *et al.*, 2020), it took eight years before the first results of a cognitive bias test for mice were published (Boleij *et al.*, 2012). Boleij and colleagues conditioned mice to various odor stimuli, which predicted either a palatable or an unpalatable food reward. First, it was shown that BALB/cJ mice were able to discriminate between odor stimuli, whereas 129P3/J mice were not. Second, it was shown that BALB/cJ mice tested under more aversive white light conditions had a higher latency in response to the ambiguous stimulus than mice tested under less aversive red-light conditions.

Further cognitive bias test methods followed in which mice were conditioned to spatial positions (Bailoo *et al.*, 2018; Kloke *et al.*, 2014; Novak *et al.*, 2015; Verjat *et al.*, 2021), to tactile stimuli (Novak *et al.*, 2016), to different tunnel lengths (Krakenberg *et al.*, 2019), to auditory stimuli (Jones *et al.*, 2017), to olfactory stimuli (Resasco *et al.*, 2021), or in an automated touchscreen-based set-up presenting different patterns on a screen (Krakenberg *et al.*, 2019). These studies showed that mice could be conditioned to the different stimuli and that the data plotted on the axis of stimuli increasing from negative to positive result in a sigmoidal curve (increasing s-shape slopes from the negative to the positive stimulus). These sigmoidal curves indicate that ambiguous stimuli are perceived differently compared to the conditioned stimuli, which is an important criterion for the validity of cognitive bias tests (Gygax, 2014; Hintze *et al.*, 2018; Krakenberg *et al.*, 2019).

So far, in all set-ups it is necessary for both the conditioning and the test itself to remove the mice from their home-cages and manually place them in the respective test set-ups. As a consequence, the animals have to be handled, taken out of their familiar environment, separated from their group members (if kept in groups) and forced to participate in the test irrespective of their current state of motivation. In fact, this may have a negative effect on the animals' state of mind during the conditioning phase and as a result the cognitive bias test might also be influenced. This implies that in order to minimize external influence on the cognitive bias, the best handling method has to be chosen (*e.g.*, known influence of tail handling compared to cup and tunnel handling on anxiety-like behavior (Hurst & West, 2010) and that the animals have to be very well-habituated to the test set-ups. Nevertheless, even with the best handling and habituation, a possibly negative influence of the separation from the home-cage and/or the group (Krohn *et al.*, 2006; Manouze *et al.*, 2019) as well as the experimenter's immediate influence on the mice, and thereby the test results, must be taken into account. To overcome this shortcoming, we have developed a home-cage based cognitive bias test for mice utilizing the IntelliCage system (TSE-Systems, Germany).

The IntelliCage is a home-cage based test system that allows automated data acquisition, which can improve the reproducibility of the data (reviewed in Voikar & Gaburro, 2020). Depending on size and weight of the animals, it is possible to keep up to 16 mice in the IntelliCage as one social group. Through radio frequency identification (RFID) technology and four conditioning corners, it is possible to study activity and learning behavior in social groups (Endo *et al.*, 2011; Kahnau *et al.*, 2021; Krackow *et al.*, 2010; Voikar *et al.*, 2018).

Our test set-up consisted of a home-cage, a gate (AnimalGate, TSE-Systems, Germany) and an IntelliCage (test-cage). Through the gate, it is possible to separate the mice and let them individually enter the test-cage. This is especially important to allow all individuals within the group to be conditioned and tested without disturbance by group members. Another advantage is that the mice can individually decide when to enter the test-cage and participate in the experiment, rather than being coerced by an experimenter-imposed schedule. As a result, the influence of the experimenter and the influence on the wake/sleep rhythm is reduced to a minimum, except for daily visual inspection and weekly cleaning of the set-up. It has

already been shown that rats and mice can independently transfer themselves from their home-cages to test-cages individually to perform tasks within test-cages (Kahnau *et al.*, 2022A; Kaupert *et al.*, 2017; Mei *et al.*, 2020; Rivalan *et al.*, 2017; Winter & Schaefers, 2011). A slight disadvantage is that since only one mouse can be within the test-cage at a time and other motivated mice have to wait until this mouse has left the test-cage. However, we could show in a recent experiment with a comparable set-up that no single mouse was constantly blocking others from getting access (Kahnau *et al.*, 2022A).

Within our automated and home-cage based test set-up, we conditioned female C57BL/6J mice to different tones. De Hoz and Nelken as well as Francis and colleagues already showed that mice were able to differentiate between different tones (De Hoz & Nelken, 2014; Francis & Kanold, 2017). Here, we present our different developmental steps and results of the cognitive bias tests. Our first hypothesis was that it is possible to condition mice within the IntelliCage based set-up and that the cognitive bias is influenced by the removal of enrichment and by repeated restraining. Here we present the individual developmental steps of our automated and home-cage based cognitive bias test, which were based on each other and the optimizations we implemented through previous experience. We show that it is possible to successfully condition mice in a relatively short time and measure the cognitive bias of mice, with minimal intervention and time investment by the experimenter.

Methods

Animals and housing conditions

In this study, three developmental steps with three different mouse groups (one developmental step per group) are presented in which different conditioning methods are described (Table 1). All three groups served as their own controls as before and after comparisons were made. All 36 female C57BL/6J mice were purchased from Charles River Sulzfeld, Germany. For each developmental step, the three groups consisted of 12 mice. This group size was chosen due to the size of the entire IC based set-up. All mice were four weeks old upon arrival but were bought at different time points. All efforts were undertaken to minimize animal suffering. No medical treatment was required at any time for the mice due to pain, suffering, or harm. Further details on the mouse groups are given at the respective developmental steps.

For the establishment of the home-cage based cognitive bias test, females were used exclusively since they can be kept in groups without complications due to little agonistic behavior. In addition, females do not show territorial behavior that excludes others (Mieske *et al.*, 2021) and at the beginning of the development of the set-up there was a concern that individual males could occupy the gate, and thus the test-cage.

We deliberately used an inbred strain to minimize genetic variability. However, despite all efforts of standardization, minimal genetic drift and varying epigenetic influences can occur during breeding. In order to randomize the factors that could not be controlled for, all mice in each experiment were born and raised by different mothers and foster mothers to ensure maximum genetic and epigenetic independence between individuals. Immediately after arrival a health inspection was performed and the mice were weighed and color-marked (edding 750, colors: black, white, red, yellow, silver) on the tail for visual identification. The mice were housed within the home-cage based set-up, and no data was recorded for the first two weeks. The day after arrival, tunnel handling training to reduce handling stress (Gouveia & Hurst, 2013; Hurst & West, 2010) was started and conducted for three weeks (see video tutorial).

One week after arrival, all mice received RFID transponders (Euro ID, FDX-B, ISO 11784/85). The evening before the transponder transplantation, an analgesic (meloxicam 1mg/kg, Meloxidyl by CEVA) was given orally by fixing the mice in the experimenter's hand, to reduce possible pain caused by implantation. The transponders were implanted under isoflurane anesthesia (induction of anesthesia: 4l/min 4%; maintenance of anesthesia: 1l/min 1-2%) subcutaneously in the neck region about 1cm behind the ears. Out of 36 mice, two mice lost their transponders by the morning after transponder implantation and the procedure had to be repeated. None of the 36 mice needed medical treatment after transponder implantation.

One week after transponder implantation, the mice moved to the housing room where also the home-cage based experiments

Table 1. Experimental procedure. IC = IntelliCage.

	Group one	Group two	Group three
Developmental step	1	2	3
Year	2019	2019	2020
Conditioning protocol	Gate: passing the gate	Corner: visiting the IC corner	Corner: visiting the IC corner
Tone	Sequences	Frequencies	Frequencies
Tone length	6.6 sec.	0.5 and 1 sec.	2 sec.
Airpuff length	1 sec.	1 sec.	2 sec.

were conducted. The room temperature and humidity were 22°C +/- 3°C and 55% +/- 15%. The light/dark cycle was set to 12/12 hours with light off at 7 pm in winter months and at 8 pm in summer because of the switch from winter to summer-time. The sunrise was simulated with a wake-up light (Philips HF 3510, 100-240 vac, 50-60 Hz, Philips Consumer Lifestyle B.V. Netherlands) half an hour before the room-light was switched on. The wake-up light was placed on the ground in a corner of the housing room with the light directed towards the animals. The light intensity increased gradually and reached the full intensity at 7/8 am (depending on season). The daily visual health inspection was performed between 7/8 am to 10 am (depending on season). The home-cage set-up was cleaned once a week. Bedding, nesting material, and enrichment items were replaced. A small handful of old bedding was transferred to the new home-cage. On the same day, the mice were weighed and re-color-marked.

Home-cage based set-up

In all developmental steps, the same home-cage based set-up was used. This set-up (Figure 1) consisted of three compartments: a home-cage, a gate (AnimalGate), and a test-cage (IntelliCage, IC). As the gate had doors, an RFID reader, and infrared barriers, it was possible to allow only one mouse at a time to pass through the gate from the home-cage into the IC. All other mice of the social group had to wait until the one mouse within the IC moved back through the gate into the home-cage.

The home-cage was a Makrolon type IV cage (floor space 2065 cm^2) with a filtertop equipped with 3-4 cm bedding (Poplar Granulate 2-3 mm, Altromin, Germany), two red triangle-shaped houses ("TheMouseHouse", Tecniplast, Italy), nesting material (eight papers, six paper nesting stripes and six cotton rolls), four wooden bars to chew on, and food *ad libitum* (autoclaved pellet diet, LAS QCDiet, Rod 16, Lasvendi, Germany). Within the home-cage was also an acrylic tube (4 cm diameter, 17.5 cm long), which was used for tunnel handling to reduce handling stress. Mouse group three additionally received nesting materials upon weekly changing: folded paper stripes, mid coarse wood wool and square hemp pads. Also, one resting platform and a running disk (InnoDome with InnoWheel, Bio-Serv) was placed within the home-cage and the mice received weekly changing toys filled with millet (organic peeled golden millet, Bohlsener Mühle) once per week.

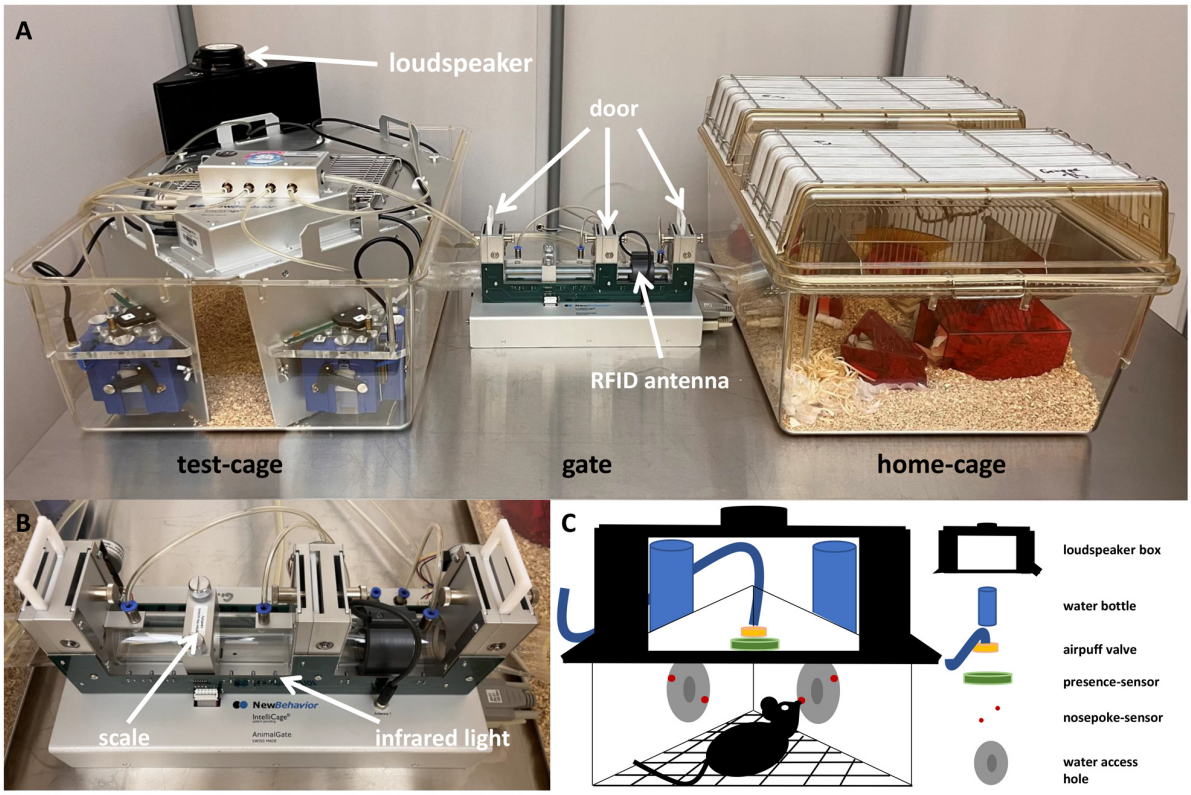

Figure 1. Home-cage based set-up based on the IntelliCage system. A: The set-up consisted of the IntelliCage used as the test-cage, which is connected through the AnimalGate to the home-cage. The IntelliCage was equipped with four conditioning corners and bedding. The home-cage was equipped with bedding, nesting, enrichment and food *ad libitum* (not shown here). The AnimalGate had three doors, one radio frequency identification (RFID) antenna. **B**: In addition, the AnimalGate had eight infrared barriers and one scale to measure the animal's weight during each gate passage. **C**: Within the IntelliCage corners, water could be provided. In addition, each corner had one radio frequency identification antenna, one presence-sensor, one airpuff-valve, two water dispensers and two doors.

The IC is a computer and RFID technology-based test system with four conditioning corners. Each corner contained an RFID antenna at the corner entrance, a presence sensor, which detected differences in temperature, two nosepoke infrared sensors, two doors through which the water access can be regulated, two water dispensers, and an airpuff valve for the possibility of a mild punishment (0.5 bar). Depending on the conditioning method, one or four IC corners were active, in which water was provided. The IC contained only bedding material.

In order to perform experiments within the IC system, it is necessary to habituate the mice to the system first. The mice had to learn how to pass through the AnimalGate and where to access water within the IC. For this purpose, the mice were habituated gradually to the AnimalGate and IC doors. Initially, all AnimalGate and IC doors were permanently open (phase: 'all doors open'). Thus, it was possible for all mice to move freely within the system. As a next step, the doors of the AnimalGate were closed, and opened only when a mouse entered the AnimalGate, which is similar to the next IC habituation step when the corner doors were closed and opened due to a visit (phase: 'visit open doors'). In the final phase of habituation, only one mouse could stay in the IC, and the IC doors could only be opened with a nosepoke.

Conditioning protocol

The basic requirement for performing a cognitive bias test is to condition the animals to scalable stimuli. In our study, the mice were conditioned to auditory stimuli. Three different conditioning protocols were performed with each of the different mouse groups. Common to all protocols was that the mice had to learn that for one presented tone (positive tone); they received water as a reward; if they made a nosepoke within the IC corner (correct behavior). For another tone (negative tone), they received an airpuff as a punishment, if they made a nosepoke (incorrect behavior). If the mice did not make a nosepoke after hearing the positive tone (incorrect behavior), they received no water. If the mice did not make a nosepoke after hearing the negative tone (correct behavior), they did not receive an airpuff (Table 2). All tones were created by using the online tool onlinetonegenerator.com and Audacity (AudacityCross-Platform Sound Editor).

Since the mice only had the opportunity to drink water in the IC, it was necessary to monitor whether all mice drank daily. If a mouse did not drink for 24 h, the mouse was offered water in a separate cage for 15 minutes. After these 15 minutes, they were placed back in the home-cage. If drinking did not occur in the IC for three consecutive days, these mice were taken out of the experiment by allowing them access to water within the IC corner without tones. These mice were no longer participating in the conditioning phase and cognitive bias test, but were still left in the group, leaving the social structure unchanged throughout the experiment.

Table 2. Description of the possible events during the conditioning within the IntelliCage corner.

Nosepoke	Positive Tone	Negative Tone
Yes	water	airpuff
No	nothing	nothing

For more clarity, the individual development steps are described individually below. The respective results and conclusions follow the method description of the individual development steps.

Analysis

Data analysis was done with the open-source statistical software R (version 4.0.3, RCoreTeam, 2020). For data visualization the R package ggplot2 (Wickham, 2016) was used. Model assumptions were inspected visually first by Q-Q plots, and secondly by visualizing variance homogeneity of the residuals versus fitted values. Individual animals served as the experimental unit, as only one mouse was in the test cage at a time. A total of 36 mice were used, which were divided into three groups (12 mice per group and developmental step). Since data were collected automatically, blinding was not necessary.

Analysis of data from gate conditioning protocol. For the gate conditioning protocol (detailed description below), the mice first had to learn which corner was the active corner. Therefore, the visit number of the active corner was compared to the visit number of the inactive corners for each mouse per day during the first 14 days (when only the positive tone was presented). A visit was recorded by the IC-system each time a mouse entered a corner, and both the RFID transponder number was detected and the presence-sensor was activated. The visit number was used as the outcome in a linear mixed-effects model (R package nlme [Pinheiro *et al.*, 2020]). The experimental days were used as a fixed effect (factor with 14 levels). The type of visit (factor with two levels: visits in active corners versus visits in inactive corners) and the interaction of type of visits and day were used also as fixed effects. The variable 'days nested in animals' (n = 12) were set as a random effect. Sum-contrasts were used for days and type of visits.

For the evaluation of the two gate conditioning runs (run 1 n = 11, run 2 n = 12), the frequency with which the mice passed the AnimalGate was first determined for each mouse for each day, *i.e.*, how often mice were presented with tone-sequences. The duration from entering to leaving the IC was defined as IC-session. From this, we determined how often the positive and negative tone-sequences were played (per animal, per day). Next, we determined how often the mice visited the active corner and made nosepokes on the nosepoke-sensor during the positive and negative tone-sequence IC-sessions. The number of nosepokes was used as the outcome in a linear mixed-effects model (R package nlme). In this model, the experimental days were defined as days and used as a fixed effect (factor with nine levels in AnimalGate conditioning run 1, factor with 14 levels in AnimalGate conditioning run 2). Within the statistical model, the type of tone-sequence (two-level factor: positive versus

negative tone-sequence) and the interaction of type of tone-sequence and day were also used as fixed effects. Sum-contrasts were used for day and type of tone-sequence. The variable 'experimental days nested in animals' was set as a random effect.

Analysis of data from corner conditioning protocol. For the evaluation of the corner conditioning protocol (detailed description below), the frequency with which the mice (group 2 n = 12, group 3 n = 12) visited the active corner within the IC was first determined for each mouse for each day, *i.e.*, how often mice were presented with tone-frequencies (inactive corners were blocked with a plug). From this, we determined how often the positive and negative tones were played (per animal, per day). Next, we determined how often the mice visited the active corner and made nosepokes at the nosepoke-sensor during the positive and negative tone. The number of nosepokes was used as the outcome in a linear mixed-effects model (R package nlme). In this model, the experimental days were defined as days and used as a fixed effect (factor with 48 levels). Within the statistical model, the type of tone-frequency (two level factor: positive tone-frequency versus negative tone-frequency) and the interaction of type of tone-frequency and day were also used as fixed effects. Sum-contrasts were used for day and type of tone-frequency. The variable 'experimental days nested in animals' was set as a random effect. To test for effects of interaction of day and tone-frequency, *post hoc* comparison was conducted (R package emmeans [Lenth, 2020]).

Learning success for visit conditioning. Descriptive statistics were used to assess individual learning success by observing correct nosepoke behavior. Correct nosepoke behavior at the positive tone was defined as a corner visit during which at least one nosepoke was made. Correct nosepoke behavior for the negative tone was defined as a corner visit without a nosepoke. For each mouse, we first determined how many positive tone trials and negative tone trials had occurred. Then, the numbers of positive tone trials with nosepokes and the number of negative tone trials without nosepokes were determined. Since the probabilities for the positive and negative tone trials were different, percentage values were calculated. From this, the corrected nosepoke behavior was plotted for each animal individually. The learning criterion was set as follows: First, we checked whether the values for the positive and negative tone were above the 50% chance level. Then, on 75% of the conditioning days, the correct nosepoke behavior had to be above the chance level in order to reach the learning criterion.

Cognitive bias test. All mice reaching the learning criterion were used in the cognitive bias test (test 1 and 2 n = 9). All other mice remained in the group, but no tones were presented when they entered the IC corner. For the cognitive bias test, the mice were presented with three additional (ambiguous) tones. First, for each mouse we determined how many nosepokes they made in response to the five different tones. The number of nosepokes was used as the outcome in a linear mixed-effects model (R package nlme). In this model, the tones (factor with five levels) and measurement (cognitive bias test 1: factor with three levels (baseline measurement 1, negative conditions and baseline measurement 2), cognitive bias test 2: factor with four levels (baseline measurement 1 and 2, negative conditions and baseline measurement 3) and the interaction were used as fixed effects. The variable 'treatment nested in animals' was set as a random effect. If the model indicated a significant effect of treatment or tone, we conducted a pairwise *post hoc* analysis (R package emmeans).

Body weight and IntelliCage behavior. For the evaluation of body weight, number of nosepokes and visits, the corresponding values were determined for each animal for each day (group 2 n = 12, group 3 n = 12). These three variables were used as the outcome in three different linear mixed-effects models (R package nlme). Treatment (group two: factor with eight levels (0%, 5%, 10%, 16%, 20%, 33% and 50% probability of negative tone and visit open doors), group three: factor with seven levels (0%, 20% and 50% probability of negative tone, nosepoke open doors, baseline measurement and negative conditions)), day (group two: factor with 75 levels, group three: factor with 100 levels) and the interaction of treatment and day was used as a fixed effect. The variable 'experimental days nested in animals' were set as a random effect.

Developmental Step 1

Methods

Animals. The 12 female mice of group one arrived at the institute in February 2019. At the start of the first developmental step, the mice were seven weeks old. After the experiment presented here, the mice were 18 weeks old and used in home-cage based learning tasks (data not published) and in a consumer demand test, which was also performed within the home-cage based set-up presented here (Kahnau *et al.*, 2022A). The mice started barbering behavior at the age of 18 weeks and immediately following the experiment presented here. Barbering behavior is commonly found in C57BL/6J mice (Kahnau *et al.*, 2022B; Sarna *et al.*, 2000). The reason for this behavior is not yet understood.

Gate conditioning protocol. The gate conditioning protocol was pre-registered in the Animal Study Registry (doi: 10.17590/asr.0000121). The mice were conditioned to tone-sequences. These sequences had a play time of 6.6 seconds at a frequency of 8 kHz and comprised either short tone-sequences with long breaks or long tone-sequences with short breaks (Figure 2).

Each mouse was randomly assigned one of two tone-sequences; thus six out of twelve mice had tone-sequence A and the other six had tone-sequence B as the positive tone stimulus. The other tone-sequence was consequently the negative stimulus. One loudspeaker was placed on top of the IC (on the grid) facing in the direction of the IC inside, allowing the mice to hear the tone-sequences. The tone-sequences were played when entering the IC after passing through the gate.

Within the IC, each mouse was randomly assigned one active corner (three mice per corner), in which the mice received either the water reward or an airpuff punishment depending

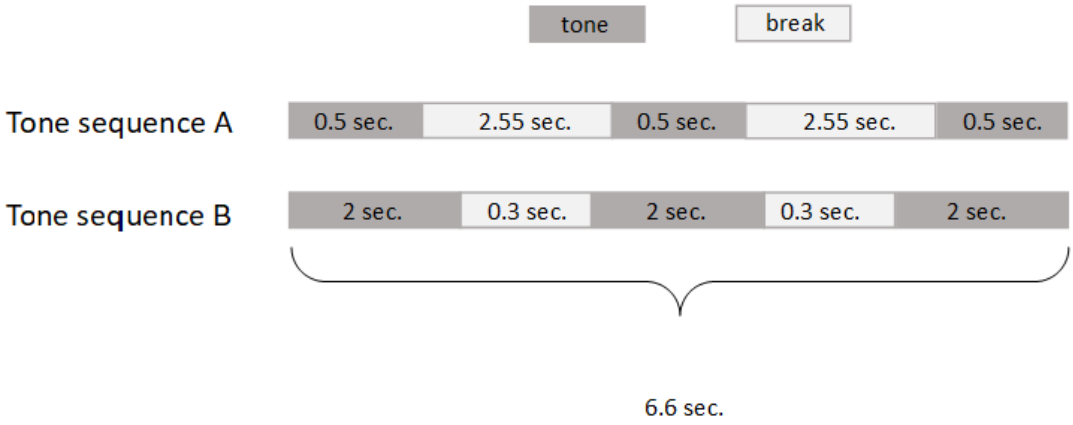

Figure 2. Tone-sequences used for AnimalGate conditioning.

on the tone-sequence. Visiting the other three corners had no consequences.

The mice had to learn first which corner their active corner was (one out of four) and second that a tone was played every time they entered the IC through the gate. This corner and positive tone conditioning ran for 14 days. When visiting the active corner and activating the nosepoke-sensor, the doors were opened for five seconds. To prevent the mice from staying too long inside the corner, an airpuff was released after another five seconds. To open the doors within the IC corner again, the IC had to be left through the gate (end of IC-Session). By re-entering the IC, a new trial was initiated.

After corner and positive tone conditioning, the negative tone-sequence was added. To prevent the mice from having too many negative experiences directly at the beginning of the conditioning phase, the probability of the negative tone being played was increased successively. Therefore, two runs were carried out. For gate conditioning run 1, the probability of playing the negative tone was 33%. For gate conditioning run 2, the probability of playing the negative tone was 50%. To initiate a new trial, the IC had to be re-entered through the gate, *i.e.*, mice that could not drink after a negative tone-sequence or did not drink after a positive tone-sequence had to leave and re-enter the IC for the next chance to drink.

Results

Corner and positive tone-sequence conditioning. The mice first had to learn which corner was the assigned active corner. Over a period of 14 days, the animals were successfully conditioned to the active corner (main effect visits: $F_{1,154} = 225.44$, $p < 0.0001$). The overall number of visits decreased over the experimental days (interaction: $F_{13,154} = 6.63$, $p < 0.0001$, Figure 3).

Gate conditioning protocol. The mice had to learn to make nosepokes after hearing positive tone-sequences and refrain from making nosepokes after hearing negative tone-sequences. In gate conditioning run 1 with 33% chance of hearing a negative tone sequence (Figure 4), the mice did not make more or less nosepokes after hearing positive or negative tone-sequences on average (main effect tone-sequence: $F_{1,90} = 0.22$; $p = 0.64$). The mice did not learn to differentiate between tone-sequences over time (interaction: $F_{8,90} = 0.82$; $p = 0.59$). However, the mice made fewer nosepokes regardless of tone-sequences over time (main effect day: $F_{8,80} = 4.58$; $p = 0.0001$).

In gate conditioning run 2 with the chance of hearing a negative tone sequence increase to 50% (Figure 5), the mice made, on average, more nosepokes for the positive tone-sequence (main effect tone-sequence: $F_{1,77} = 18.9$; $p < 0.0001$) but did not learn to differentiate between the tone-sequences (interaction: $F_{6,77} = 0.62$; $p = 0.71$). During run 2 the mice made more nosepokes over time regardless of tone-sequences (main effect day: $F_{6,66} = 2.45$; $p = 0.03$).

Discussion

The first developmental step was described as ‘gate conditioning protocol’, where tone-sequences were played whenever a mouse passed the gate and entered the IC. The initial idea of using tone-sequences was to easily create ambiguous sequences once the positive and negative sequences were successfully conditioned. Although it was possible to condition the mice to their respective randomly assigned IC corner, the mice were not able to distinguish between two tone-sequences. The mice were unable to associate a water reward with one tone-sequence and a mild airpuff punishment with another tone-sequence. The unsuccessful conditioning could have different reasons.

First, mouse-specific ultrasonic vocalization series can have a length of two seconds. They are variable in their sequence but are released at a more or less constant frequency. There are also short sequences (a few milliseconds long) that vary in both sequence and frequency (Ehret, 2018). Our artificially created,

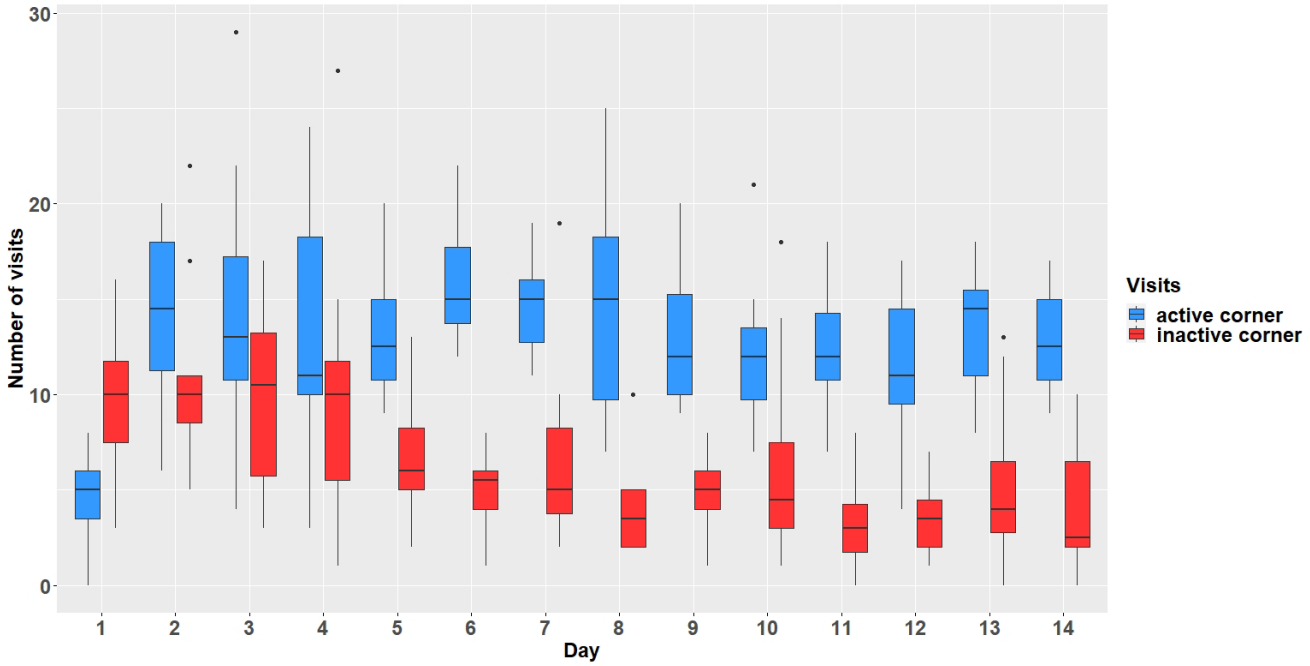

Figure 3. Comparison of visit numbers in active and inactive corners. The y-axis shows the number of visits which were made within the active and inactive corners. The x-axis shows the experimental days. n = 12.

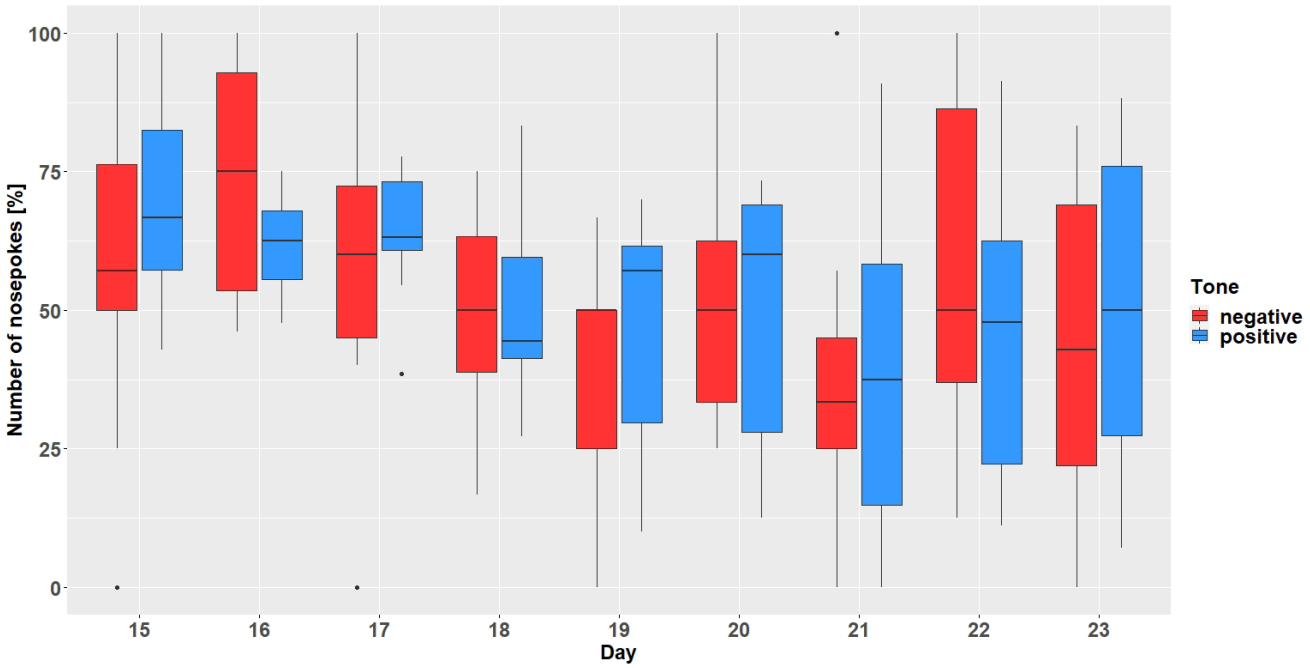

Figure 4. Gate conditioning run 1. The y-axis shows the number of nosepokes which were made in response to the presented tone-frequencies. Number of nosepokes are given in percent since the probability of the two tone-frequencies being played was different (positive = 67%, negative = 33%). After hearing a positive tone-frequency, a nosepoke had to be made, but not after hearing a negative tone-frequency. The x-axis shows the experimental days. n = 11.

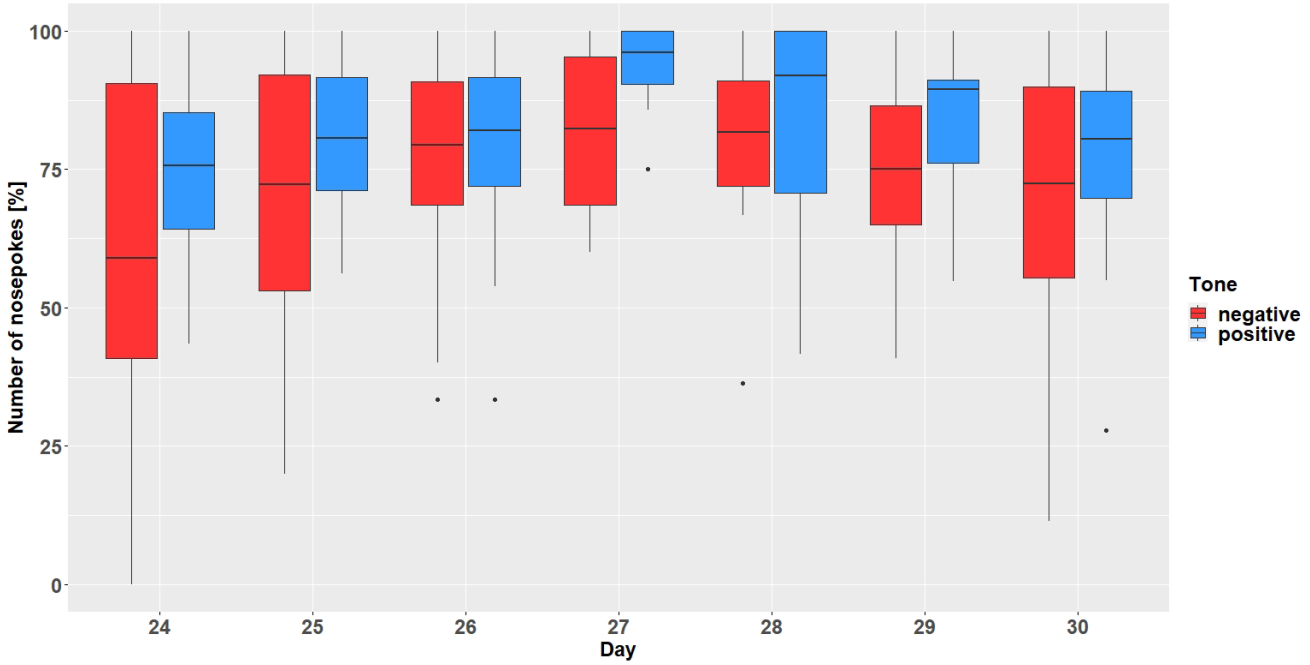

Figure 5. Gate conditioning run 2. The y-axis shows the number of nosepokes which were made in response to the presented tone-frequencies. The probability of the two tone-frequencies being played was 50:50. After hearing a positive tone-frequency, a nosepoke had to be made, but not after hearing a negative tone-frequency. The x-axis shows the experimental days. n = 12.

very static tone sequences at constant frequency had a length of 6.6 seconds, which may be too long to be perceived as relevant for the mice. The tone-sequences might have shown better results if shortened. To the best of our knowledge, there have been no experiments to condition mice to artificially created tone-sequences like the ones we used during developmental step 1. However, past studies showed the possibility to condition mice to tones, namely tone-frequencies (De Hoz & Nelken, 2014; Jones *et al.*, 2017). Therefore, we decided to use tone-frequencies instead for the next developmental step.

Second, the timing at which the tone-sequences during gate conditioning were presented was not optimal. Tones were initiated by each pass through the gate and played when the IC was entered. Whether the mouse then also directly visited the IC corner was probably dependent on how strong the motivation to drink was. Therefore, it might be possible that too much time passed between the tone and the actual corner visit, and thus, no association was established between these two events. The timing between stimulus presentation and event onset is important for successful conditioning, as shown, for example, by clicker training (Lattal, 2010).

Therefore, we decided to change the time point of tone presentation and relocated the conditioning completely to the IC corner. From then on, the sound was played when the mouse entered the IC corner. This improvement reduced the time span from the presentation of the stimulus to the corresponding nosepoke behavior to a minimum. To prevent a possible overlap effect of the unsuccessful conditioning on the next developmental step, we continued to work with a naïve mouse group.

Developmental Step 2

Methods

Animals. The twelve female mice of group two arrived at the institute in October 2019. At the start of the second developmental step presented here, the mice were 14 weeks old. The mice started barbering behavior at an age of 20 weeks, during the conditioning phase. At the end of the experiment, the mice were 26 weeks old and used in another experiment to develop a conditioned place preference test to assess severity of experimental procedures (publication in preparation).

Corner conditioning protocol. Since the gate conditioning protocol was not successful in group one, we improved the conditioning protocol and decided to no longer condition to tone-sequences but to tone-frequencies.

The hearing range of mice is between 2 kHz and 70 kHz (Heffner & Heffner, 2007). To find different frequencies with equal sound pressure levels (SPL) in the corner, a measuring microphone (miniDSP Umik-1 calibrated USB microphone) and the software Room EQ Wizard were used. In a study by de Hoz and Nelken, mice were successfully conditioned to tone-frequencies between 6 kHz and 13 kHz (De Hoz & Nelken, 2014). The same frequency range was used for our study. With a digital signal processor (miniDSP 2x4), the SPL of the played tone was optimized, to ensure that all tones were played at the

same volume within the corner. This was done to ensure that variations in SPL stemming from the speaker and confined space in which they are played were as small as possible. However, it must be emphasized that the perception of mice differs from that of humans and that there possibly are influences, which we are unable to detect and/or assess.

In addition to different tones, we decided to change the time when the tones were played. For the corner conditioning protocol, tone-frequencies (positive or negative tone) were played when entering an IC corner instead of when leaving the gate and entering the IC. One single corner within the IC was chosen as the active corner for all mice to set the focus of the mice to this corner and to ensure that the tone quality was the same for all mice. All other corners were made unreachable by 3D printed plugs made from gray polylactic acid (PLA). In order to initiate a new trial, the mice had to re-enter the active corner. During one IC session, multiple trials could be initiated by the mouse re-entering the active corner without having to leave the IC again (as it was the case for gate conditioning protocol). Within the active corner and after hearing the positive tone-frequency, the IC doors could be opened by a nosepoke for seven seconds.

To play the tone-frequencies, one loudspeaker was placed on top of the active corner directed towards the inside of the corner, so the mice were able to hear the tones. In order to be able to position the loudspeaker, it was integrated into a black 3D printed box (Figure 1C).

The tone-frequency at one end of the scale was 6.814 kHz at 70 decibel (dB), the other tone-frequency on the other end of the scale was 13.629 kHz at 70 dB. At the beginning of the conditioning phase, only the positive tone frequency was played during a visit in the active IC corner. The probability of the negative tone-frequency was increased progressively to avoid too many negative experiences at the beginning of the experiment (*Extended data* [Kahnau *et al.*, 2022C]).

The tone-frequencies had at first a length of 0.5 seconds. The tone length was extended to one second on experimental day 20. During the experimental phase, there were several technical problems and therefore, data for some days were lost. On several occasions, the body weight of the animals could not be recorded due to the AnimalGate being blocked by bedding material. Removing the bedding from the AnimalGate solved this problem. An unexpected failure of the control unit led to missing data recording on days 23, 99, 104, 105. The whole IC system had to be restarted to resolve these failures.

Results

Corner conditioning protocol. After visiting the active corner, one out of two tone-frequencies was randomly presented. In total (Figure 6), the mice made more nosepokes at the positive

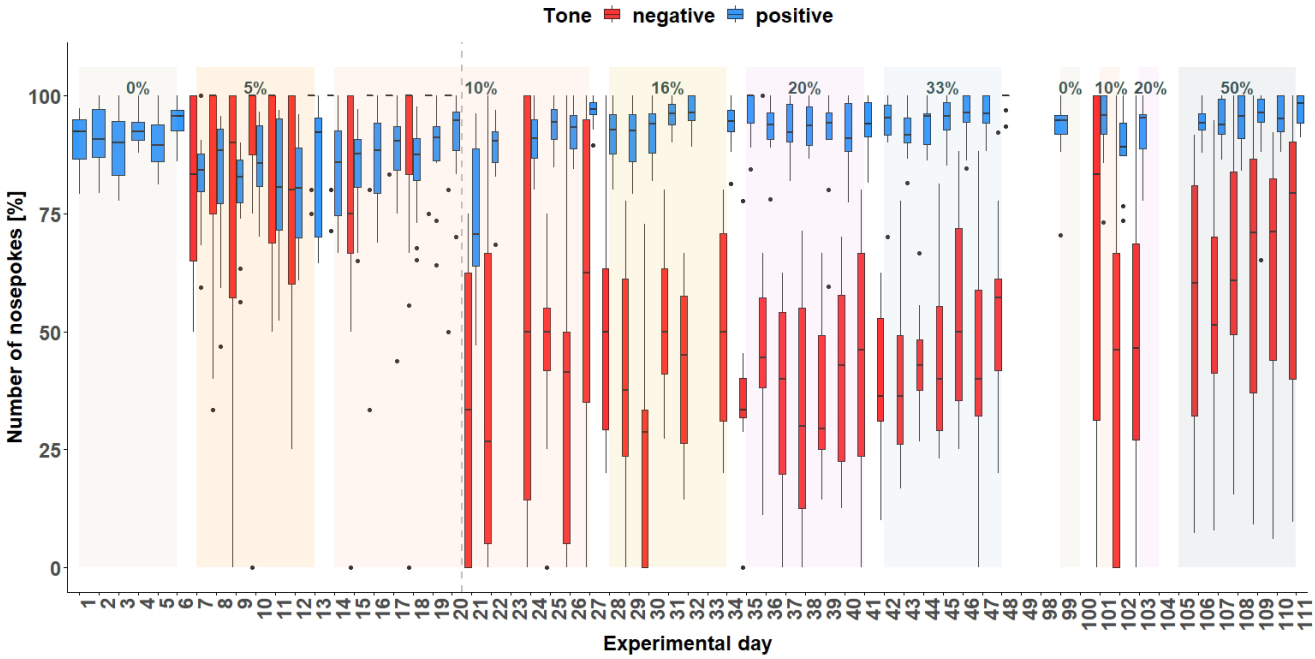

Figure 6. Corner conditioning group two. Number of nosepokes in percent made in response to two different tone-frequencies. The data for experimental day 23, 99, 104, 105 is missing due to technical problems with the IntelliCage system. No data from day 49 to 84 is available, because the mice were not in the home-cage based set-up as the set-up had to be maintained. From experimental day 85 the mice were kept in the set-up again. In order to habituate the mice to the set-up again, no sounds were played on days 85 to 98. On the y-axis, the number of nosepokes in percent is shown. The x-axis shows the experimental days. The dashed line marks the time point when the tone length was increased to one second. 0% = no negative tone, 5% = 5% negative tone probability, 10% = 10 percent negative tone probability, 16% = 16% negative tone probability, 20% = 20% negative tone probability, 33% = 33% negative tone probability, 50% = 50% negative tone probability.

tone compared to the number of nosepokes made at the negative tone-frequency (main effect tone: $F_{1,452} = 795$, $p < 0.0001$). The mice differentiated between the two-tone frequencies after the tone length was increased to 1 second on experimental day 20 (interaction: $F_{47,452} = 9.51$, $p < 0.0001$, table S3 *Extended data* [Kahnau *et al.*, 2022C]). In addition, the mice made less nosepokes in total after day 20 (main effect experimental day: $F_{47,473} = 5.39$, $p < 0.0001$).

Individual learning success. Since the results are considered for each mouse, the results are evaluated descriptively. The individual learning success was considered during the time period when the negative tone was played with a probability of 33% (Figure 7) and 50% (Figure 8). These were chosen because the negative tone was played enough times to allow a meaningful comparison of the nosepoke behavior.

At the time when the negative tone was played with a probability of 33%, seven mice (ro_si_2, ro_sw_2, sw_ge_2, sw_si_2, we_ro_2, we_si_2, and we_sw_2) out of 12 mice reached the learning criterion. Mouse ro_ge-2 stopped to drink before the negative tone was played with a probability of 33%.

Increasing the probability of the negative tone to 50% resulted in more incorrect nosepoke behavior in response to the negative tone. Only four mice (ro_sw_2, sw_si_2, we_ro_2 and we_si_2) out of 12 mice reached the learning criterion (75% of correct nosepoke behavior over 50%). The mice ro_ge_2 and ro_si_2 did not drink and were taken out of the experiment.

Body weight and IntelliCage behavior. Body weight, number of licks, and number of visits were recorded throughout the experimental period (Figure 9). Body weight was influenced by the treatment ($F_{7,803} = 2.33$, $p = 0.02$) as well as by the experimental day ($F_{1,803} = 211$, $p < 0.0001$). Also, the interaction treatment and day had an influence on body weight ($F_{7,803} = 2,36$, $p = 0.02$). In addition, the number of licks over time was influenced by treatment ($F_{7,813} = 20.71$, $p < 0.0001$) as well as experimental day ($F_{1,813} = 14,3$, $p > 0.0001$). This influence seems to be particularly strong on individual experimental days (interaction: $F_{17,813} = 7,99$, $p < 0.0001$), which is also reflected in the number of visits (interaction: $F_{7,813} = 22.31$, $p < 0.0001$). These were also influenced by the treatment ($F_{7,814} = 47$, $p < 0.0001$) but not influenced by the experimental day ($F_{1,815} = 0.05$, $p = 0.83$).

Discussion

The second developmental step was described as 'corner conditioning protocol', where tone-sequences were played whenever a mouse visited the active IC corner. With this protocol, it was possible for the first time for single mice to distinguish between two different tone-frequencies within the set-up presented here. Two mice ceased drinking in the IC during the conditioning phase. Therefore, these mice were excluded from the experiment, *i.e.*, for them the tone presentation was turned off and they were able to open the doors by a nosepoke at each visit. Dropouts also occurred in other studies, where individual animals did not reach the learning criterion and thus the actual test phase (*e.g.*, Bračić *et al.*, 2022; Hintze *et al.*, 2018; Kloke *et al.*, 2014; Krakenberg *et al.*, 2019).

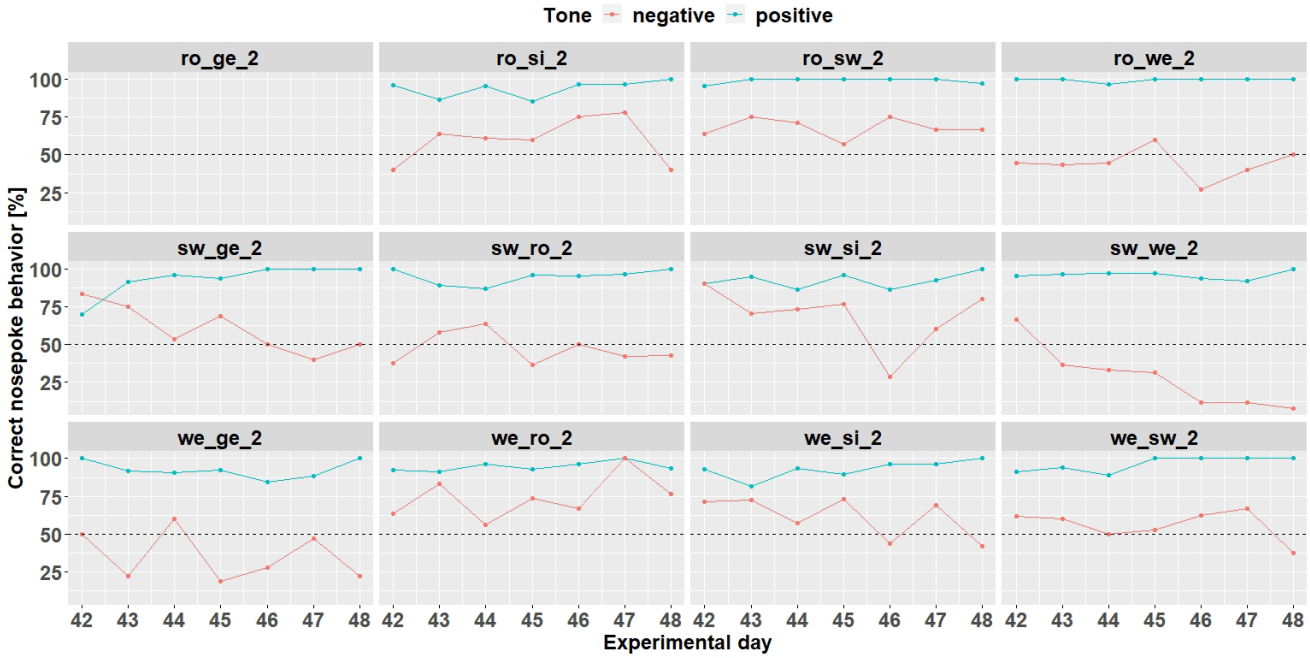

Figure 7. Individual learning success during conditioning when the negative tone-frequency was presented with a probability of 33%. Mouse ro_ge_2 was taken out of the experiment. Learning criterion 75% of correct nosepoke behavior over 50%.

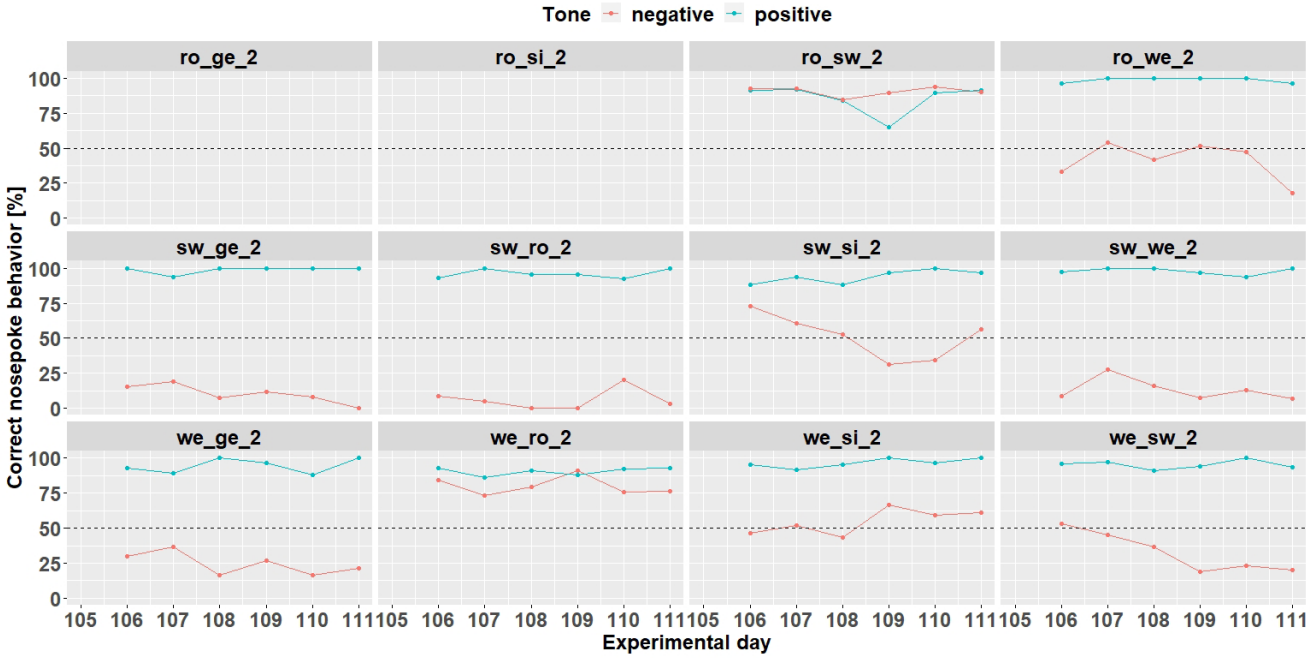

Figure 8. Individual learning success during conditioning when the negative tone-frequency was presented with a probability of 50%. The data of day 105 is missing due to technical problems with the IntelliCage system. Ro_ge_2 and ro_si_2 did not participate any longer in the experiment. Learning criterion 9 trials over 50% out of 12.

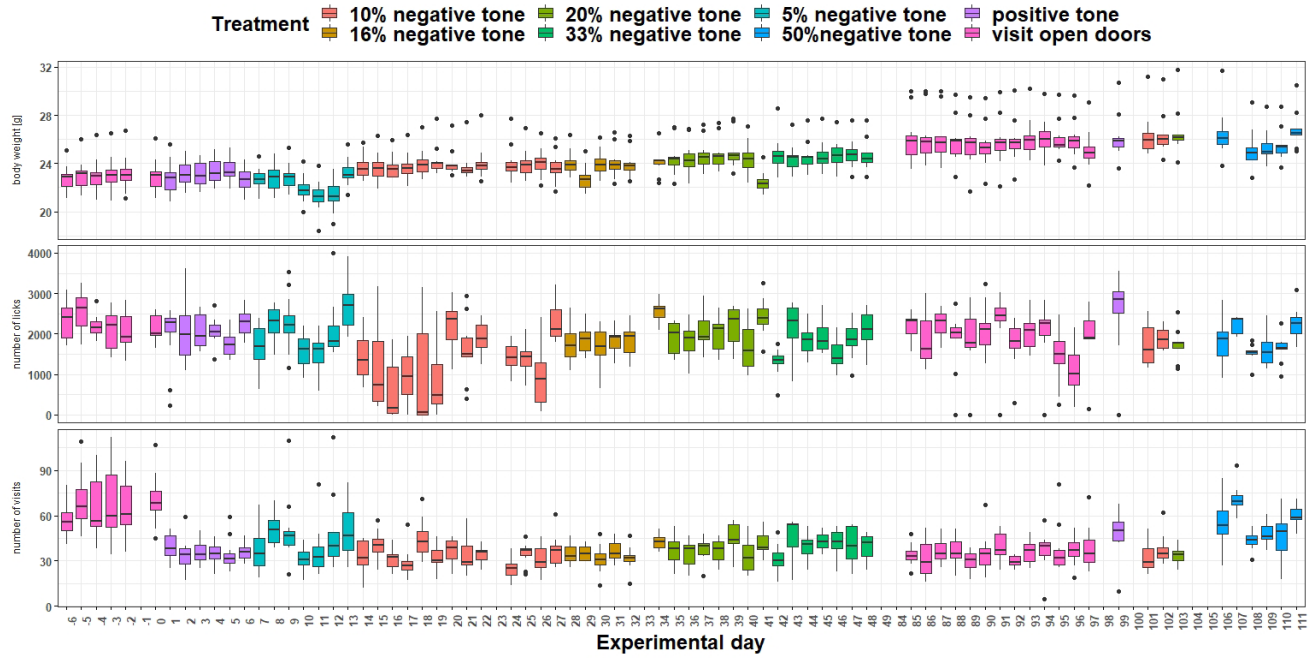

Figure 9. Measurement of body weight, IntelliCage corner visits and lick number over time. The x-axis shows the experimental days. On the y-axis first the body weight, second the lick number and third the visit number is shown. Different tones with different playback probabilities were presented throughout the experimental period (treatment). The data of experimental day -1, 23, 33, 98, 100, 104 and 105 are missing, due to technical issues. During experimental days 49 to 84 no tones were played.

The other ten mice of the group continued to drink within the IC but they did not initially distinguish between the two different tones. After changing the tone length (from half a second to one second), significant differences in the nosepoke behavior depending on the tone could be detected. The mice did more nosepokes in response to the positive tone compared to the nosepoke number for the negative tone. However, the number of nosepokes for the negative tone increased when the probability of it occurring was increased (up to 50%).

This was particularly evident in the examination of individual learning performance, when nosepoking was barely supressed by the negative tone. Overall, mice made many correct responses for the positive tone, but markedly fewer correct responses for the negative tone. Accordingly, the mice seemed to have a high motivation to perform nosepokes regardless of the outcome.

There was also an increase in the number of visits over the course of the experiment. However, the number of licks per day hardly changed. The explanation might be that the possibility to drink was reduced by increasing the number of trials with the negative tone. Thus, to get the same amount of liquid, more visits had to be made. It may be that the motivation to interact with the nosepoke sensor was so strong that the risk of punishment was accepted. This would be in line with literature data showing that mice continue to operate a lever although it was associated with a stimulation of 'aversive brain regions' (Cazala, 1986).

By giving many incorrect responses to the negative tone, the mice also received a correspondingly high number of airpuffs, which in turn could have led to habituation to the airpuff. The punishment would therefore no longer be perceived as a valid punishment (Kahnau *et al.*, 2021). Another explanation could be that the permanent presentation of the tones caused them to no longer be perceived as relevant but rather as a kind of background noise, and nosepokes were made independently of the tones.

In conventional tests, mice were placed in a designed test apparatus for a defined test period and were exposed to the stimuli for that defined time (*e.g.*, Bailoo *et al.*, 2018; Boleij *et al.*, 2012; Kloke *et al.*, 2014; Krakenberg *et al.*, 2019; Richter *et al.*, 2012). After the test phase, the mice were transferred back to their home-cages, where they spent their time undisturbed until the next test phase. On the contrary, in our system, which also served as the home-cage, no such breaks occurred. Thus, the stimulus might have had none or little relevance and the focus might be on opening the doors, driven by the motivation to drink.

Our results suggest that rest periods should be included in order to maintain the concentration and/or motivation of the mice. Therefore, for the next developmental step, we decided to schedule breaks, while the mice had access to the water without presentation of the tones, between the individual conditioning and testing phases. To exclude possible influences from previous conditioning phases, we again worked with another naïve mouse group in the next developmental step.

Developmental Step 3

Methods

Animals. The twelve female mice of group three arrived at the institute in September 2020. At the start of the third developmental step, the mice were six weeks old. At the end of this experiment, the mice were 21 weeks and used in various cognitive experiments (data not published) and in an experiment to develop a home-cage based consumer demand test based on the mouse positioning surveillance system (data not published yet). The mice started barbering behavior at the age of 31 weeks, 10 weeks after the experiment presented here.

Corner conditioning protocol. In order to successfully condition the mice of group three to tone-frequencies, further modifications were made to the corner conditioning protocol described earlier. This experiment was pre-registered in the Animal Study Registry (doi: 10.17590/asr.0000228). In the active corner and after hearing the positive tone-frequency, the IC doors could be opened by a nosepoke for ten seconds. In addition, the tone length as well as the airpuff length was extended to two seconds. The tone-frequencies for the first conditioning phase of group three were the same as for group two (6.814 kHz at 70 dB and 13.629 kHz at 70 dB). For the second conditioning phase, tone-frequencies between 6.814 kHz at 70 dB and 9.636 kHz at 70 dB were used. Also, for group three, the probability of the negative tone was increased step by step (*Extended data*).

Cognitive bias test. After the conditioning phase, the cognitive bias test followed. This was done by adding ambiguous tone-frequencies, which were calibrated between the positive and negative tone-frequencies (first cognitive bias test: 8.103 kHz, 9.636 kHz, 11.459 Hz, second cognitive bias test: 7.431 kHz, 8.103 kHz, 8.836 kHz). For the determination of these ambiguous tones, the geometric mean, which is the perceived middle between two tones, was used. To determine the geometric mean (GM), the square of the product of the two chosen tone frequencies is calculated.

$$GM = \sqrt{f_1 \cdot f_2}$$

The tritone of the original low and high frequency is then used as the respective high and low frequency to calculate two additional tritones, generating a scale of five tones, each perceptibly equidistant to their neighbors. The SPL was checked with a measuring microphone and the Room Acoustics Software.

The probability for each of the three ambiguous tone frequencies to be played was 5%. By entering the active corner, one of the five different tone-frequencies was randomly presented. The mice received water by performing a nosepoke at the positive tone, and received an airpuff by performing a nosepoke at the negative tone. The mice received neither a reward nor a punishment for the ambiguous tones. For data evaluation, the

nosepoke behavior toward the ambiguous tone-frequencies was measured.

During baseline measurement, the housing conditions were as described in section "Home-cage based set-up". To manipulate the cognitive bias, the housing conditions were changed. The mice had less bedding (2cm high), less nesting material (four papers), less housing (one mouse house), no running disk, one handling tube, two wooden gnawing sticks, no active enrichment and no resting platform. For further treatment effect, the mice were additionally restrained. For this purpose, the mice were handled by tail and placed in a tube. In the tube, the mice were unable to move and had to remain in the tube for three minutes. This procedure was performed on four consecutive days at 08:00 to 9:30 o'clock during the cognitive bias measurement. The order in which the mice were restrained was randomized for each day using the R statistical program.

Results

Corner conditioning protocol. From day 57 (Figure 10), the tone-frequencies were changed. In total, the mice made more nosepokes in response to the positive tone compared to the number of nosepokes made in response to the negative tone-frequency ($F_{1,429}$ = 3578, p < 0.0001). The experimental days also seem to have an influence on the nosepoke number (main effect experimental day: $F_{48,418}$ = 4.77, p < 0.0001) as well as the interaction of day and tone ($F_{48,429}$ = 6.13, p < 0.0001).

Individual learning success conditioning phase 1. Conditioning phase 1 run for 11 days (Figure 11). Experimental day 16 was quite noticeable, where all mice performed worse. It was found that a technical problem occurred during the tone playback. Therefore, for learning success evaluation only 10 days were used.

Nine mice (ro_ge_3, ro_sw_3, ro_we_3, sw_ge_3,sw_si_3, sw_we_3, we_ge_3 we_ro_3 and we_si_3) out of 12 mice reached the learning criterion (75% of correct nosepoke behavior over 50%). The mice ro_si_3, sw_ro_3 and we_sw_3 stopped to drink and were taken out of the experiment at day 18.

Cognitive bias test 1. During the first CB test (Figure 12), the tone-frequencies influenced the number of nosepokes, which were made after hearing the tone-frequencies ($F_{4,96}$ = 28.55, p < 0.0001). A *post hoc* comparison showed that, except for the negative and near-negative tone (tone-frequency which is close to the negative tone-frequency), the mice discriminated between the different frequencies (Table 3). Also, the treatment (baseline measurement and negative treatment (less bedding and nesting, no enrichment and daily restraining) had an influence on the nosepoke behavior of the mice ($F_{2,16}$ = 5.08, p = 0.02). A *post hoc* comparison showed that the mice made less nosepokes during baseline 1 measurement compared to baseline 2 measurement and negative treatment (Table 4). The interaction of tone-frequency and treatment had no influence on the nosepoke behavior ($F_{8,96}$= 1.05, p = 0.4).

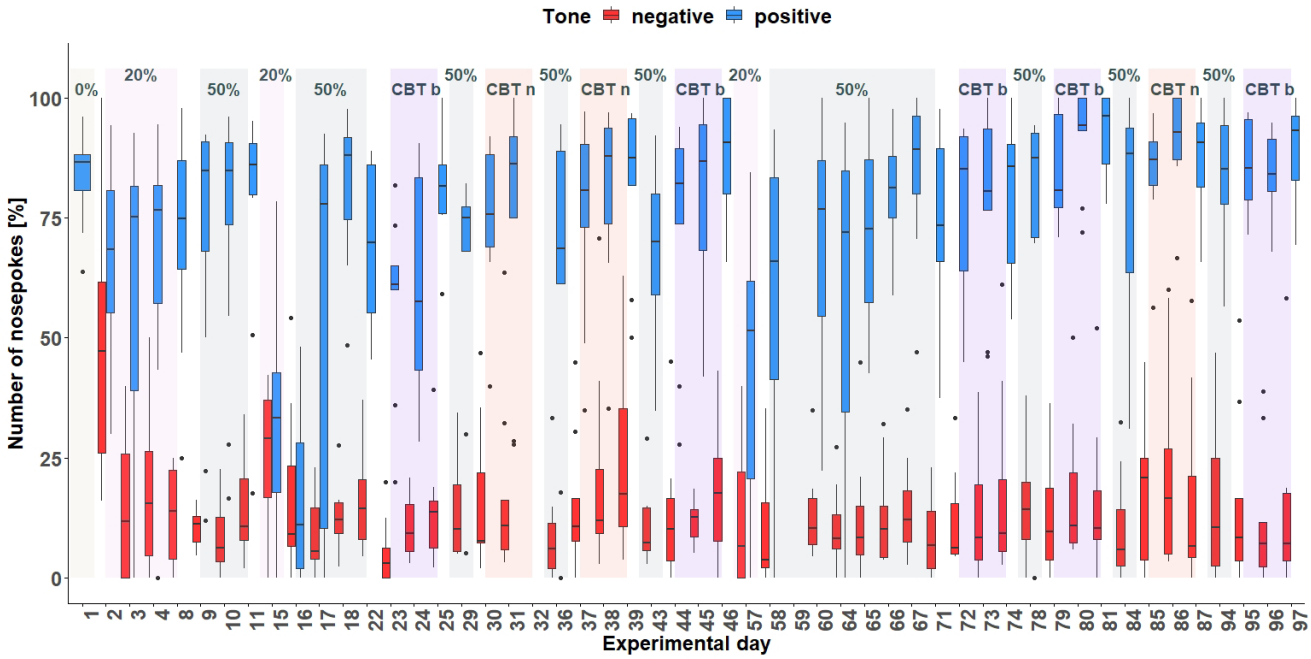

Figure 10. Corner conditioning group three. Number of nosepokes in percent made in response to two tone-frequencies. The data for experimental day 32 and 59 are missing due to technical problems with the IntelliCage system. There was an experimental break between day 47 and 56. After each treatment, no tones were presented. On the y-axis, the number of nosepokes in percent are shown. The x-axis shows the experimental days. During the experimental time period, the tones were presented with different probabilities. 0% = no negative tone, 20% = 20% negative tone probability, 50% = 50% negative tone probability, CBT b = cognitive bias measurement baseline, CBT n = cognitive bias measurement under negative conditions with less bedding and nesting, no enrichment and daily restraining.

Open Research Europe 2022, 2:128 Last updated: 12 SEP 2023

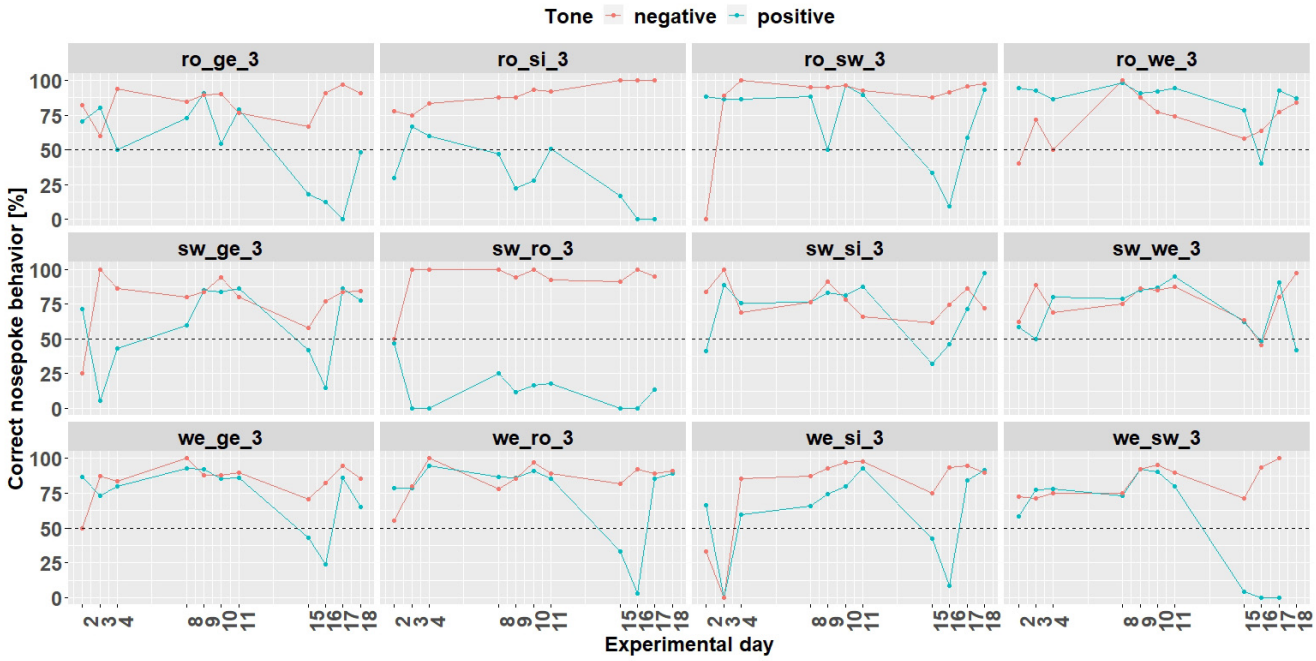

Figure 11. Individual learning success during conditioning phase 1 of mouse group three. On day 16, due to technical problems, the tones were not played correctly The mice ro_si_3, sw_ro_3 and we_sw_3 were excluded from the experiment from day 18 onwards. Learning criterion: 75% of correct nosepoke behavior over 50%.

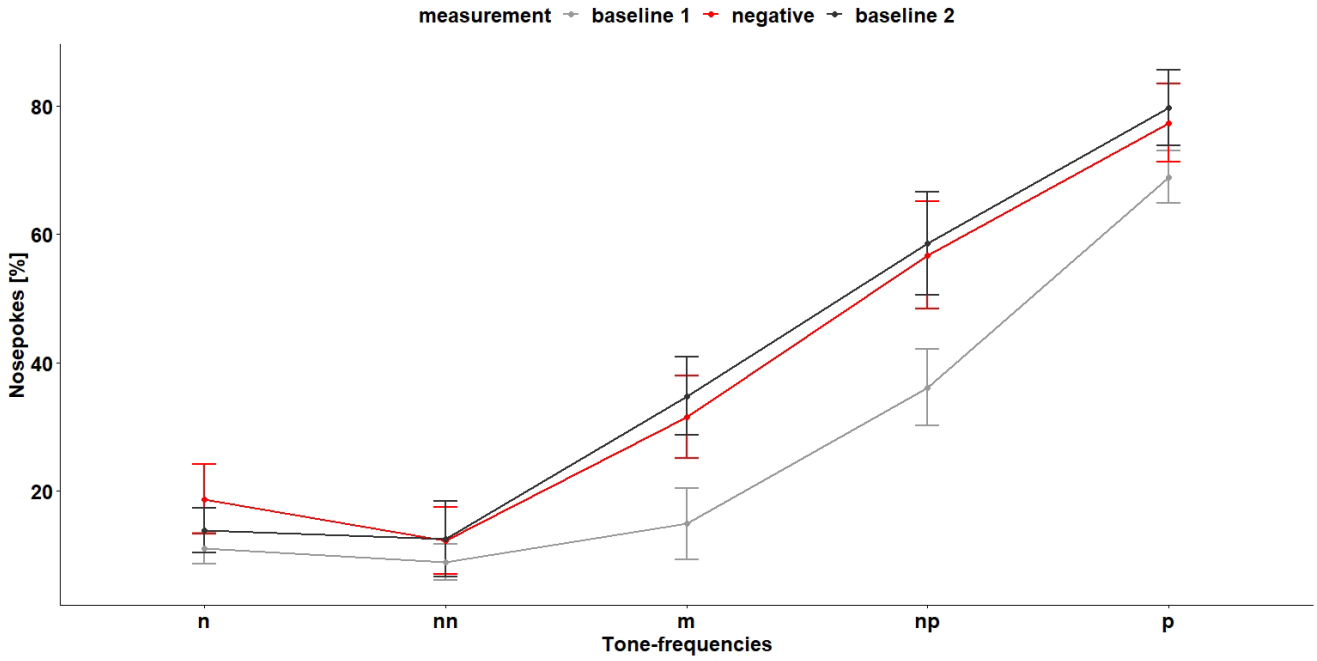

Figure 12. Cognitive bias test 1. The x-axis shows the tone-frequencies with n = negative tone, nn = near-negative tone, m = middle tone, np = near-positive tone and p = positive tone. The y-axis shows the number of nosepokes in percent made in response to the tone-frequencies. During negative measurement the housing conditions were changed compared (less bedding and nesting and no enrichment) to baseline measurement and the mice were restrained daily. n = 9.

Table 3. Results of the *post hoc* comparison of the performed nosepokes in response to the tone-frequencies for the first cognitive bias test. n = negative tone, nn = near-negative tone, m = middle tone, np = near-positive tone and p = positive tone.

Comparison	Estimate	SE	df	t.Ratio	p-Value
m – n	12.52	3.87	96	3.24	<0.001
m – nn	15.79	3.87	96	4.08	<0.001
m – np	-23.41	3.87	96	-6.06	<0.0001
m – p	-48.27	3.87	96	-12.48	<0.0001
n – nn	3.27	3.87	96	0.85	0.4
n – np	-35.93	3.87	96	-9.29	<0.0001
n – p	-60.79	3.87	96	-15.72	<0.0001
nn – np	-39.2	3.87	96	-10.14	<0.0001
nn – p	-64.06	3.87	96	-16.57	<0.0001
np – p	-24.86	3.87	96	-6.43	<0.0001

Table 4. Results of the *post hoc* comparison of the performed nosepokes during baseline measurement and negative treatment in response to the tone-frequencies. b = baseline, n = negative treatment (less bedding and nesting, no enrichment and daily restraining)

Comparison	Estimate	SE	df	t.Ratio	p-Value
b1 – b2	11.917	3	16	-3.979	<0.01
b1 – n	-11.345	3	16	-3.788	<0.01
b2 – n	0.573	3	16	0.191	0.85

Individual learning success conditioning phase 2. Due to technical issues, the data of day 59 are missing and was excluded for learning success evaluation. During conditioning phase 2 (Figure 13) 8 (ro_sw_3, ro_we_3, sw_ge_3, sw_si_3, sw_we_3, we_ge_3, we_ro_3 and we_si_3) out of 12 mice reached the learning criterion. The mice ro_si_3 and we_sw_3 stopped to drink and were taken out of the experiment at day 59. The mouse sw_ro_3 stopped to drink, too, and was taken out of the experiment at day 65.

Also, during the second CB test (Figure 14), the tone-frequencies influenced the number of nosepokes ($F_{4,112}$ = 27.27, p < 0.0001). Again, the mice did not differentiate between the negative and near-negative tone but between all other tone-frequencies (Table 5). The measurement and the interaction of tone-frequency and treatment had no influence on the nosepoke number (main effect treatment: $F_{3,21}$ = 1.67, p = 0.2, interaction: $F_{13,112}$ = 0.62, p = 0.8).

Body weight and IntelliCage behavior. Body weight (Figure 15) was influenced by the treatment ($F_{5,963}$ = 17.4, p < 0.0001) as well as by the experimental day ($F_{1,963}$ = 196, p < 0.0001). Over time, body weight increased continuously. Also, the interaction of experimental day and treatment influenced body weight ($F_{5,963}$ = 12.52, p < 0.0001). The number of licks (Figure 15) over time were influenced by treatment ($F_{5,963}$ = 30.79, p < 0.0001) but not by experimental day ($F_{1,963}$ = 0.03, p = 0.9) or the interaction of experimental day and treatment ($F_{5,963}$= 1.8, p = 0.1). The number of visits (Figure 15) were influenced by treatment ($F_{5,963}$ = 50.29, p < 0.0001). The analysis showed a tendency towards influence of the experimental day on the visit numbers ($F_{1,963}$ = 3.5, p = 0.06). However, the interaction of experimental day and treatment had an influence on the visit number ($F_{5,963}$ = 6.8, p < 0.0001).

Discussion

The third developmental step was also described as 'corner conditioning protocol', where tone-frequencies were played whenever a mouse visited the active IC corner. The tone length was changed again (from one to two seconds) compared to developmental step two. The assumption was that this change would allow the mice to discriminate the tone-frequencies more easily. In the study by de Hoz and Nelken, the tone-frequencies were played throughout the complete time of a corner visit. The playing of the tone was stopped only after the mouse left the corner and was re-initiated by a new corner visit (De Hoz & Nelken, 2014). This extreme adjustment of playback length was not considered for our experiment, since it is not known how the individual tone presentation length influences the nosepoke behavior, and thus, the cognitive bias of the mice. There was a potential for individual visit durations to have an influence on the individual mouse assessment of ambiguous tone, making the results difficult to interpret and thus reducing the validity of the data.

Like in developmental step two, some mice in group three could not be conditioned to the tone-frequencies. However, the remaining mice learned effectively and made more nosepokes in response to the positive tone compared to the negative tone. In addition, the mice seemed to be more hesitant in nosepoke behavior compared to the mice in developmental step two. This becomes evident when examining individual learning performance: There were slightly fewer correct responses for the positive tone and more correct responses for the negative tone. This implies that they performed less nosepokes overall, which has a positive effect on the number of correct answers for the negative tone but a negative effect on the answers for the positive tone. The airpuff seems to be perceived as negative. However, since some mice had to be excluded in this and in the previous developmental step because they stopped drinking, it should be considered whether the airpuff of 0.5 bar is too intense and might be reduced which could reduce the drop-out rate.

In other studies, punishment is not used at all (Graulich *et al.*, 2016; Hintze *et al.*, 2018; Novak *et al.*, 2016; Verjat *et al.*, 2021), as it is discussed that punishment during conditioning and in the

Open Research Europe 2022, 2:128 Last updated: 12 SEP 2023

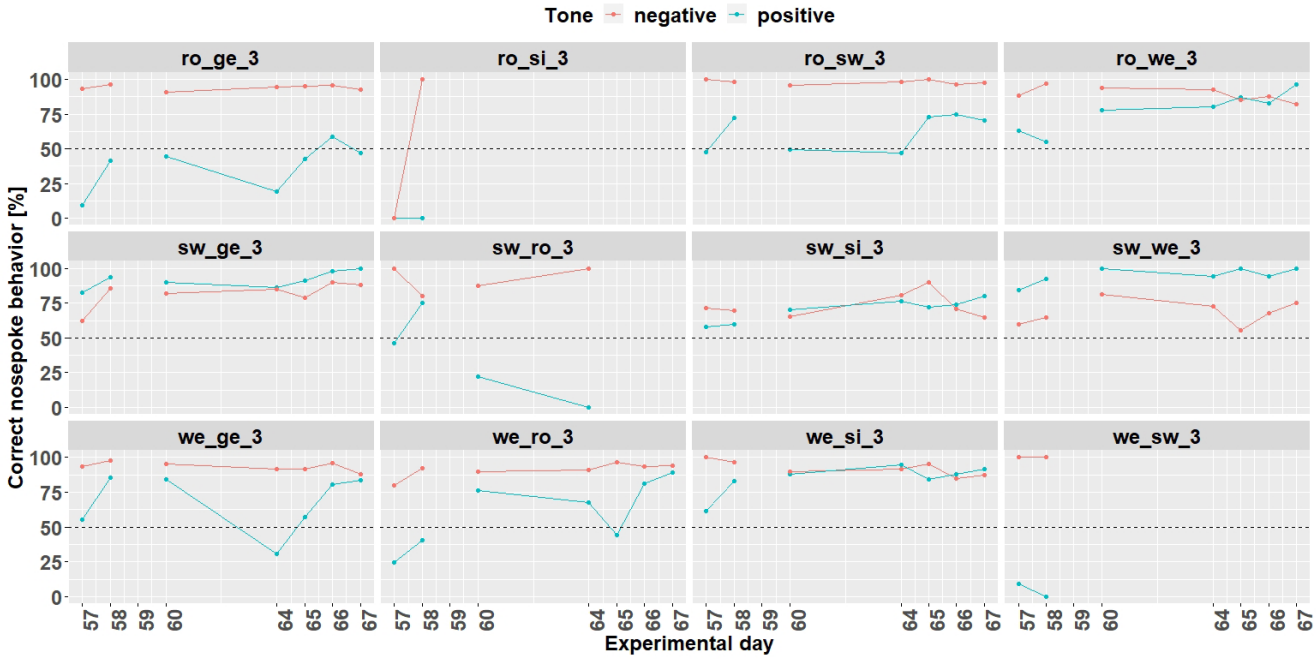

Figure 13. Individual learning success during conditioning phase 2 of mouse group three. Data of day 59 is missing due to technical problems.

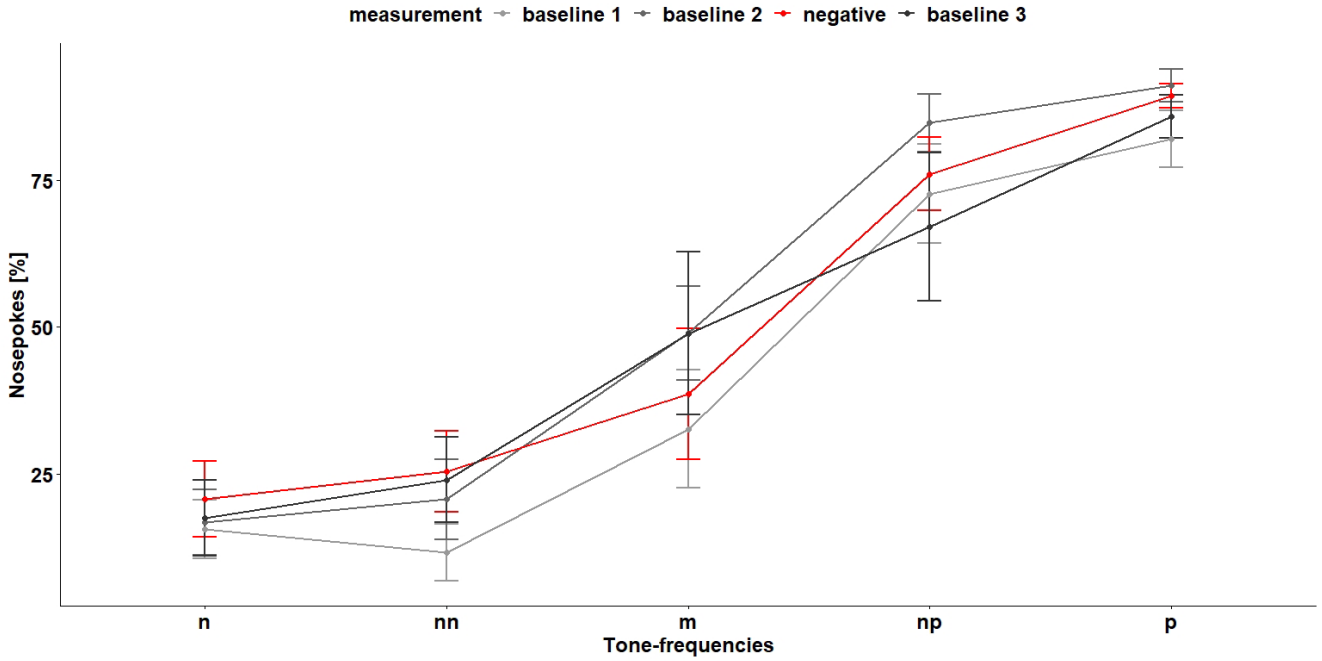

Figure 14. Cognitive bias test 2. The x-axis shows the tone-frequencies with n = negative tone, nn = near-negative tone, m = middle tone, np = near-positive tone and p = positive tone. The y-axis shows the number of nosepokes in percent made in response to the tone-frequencies. During negative treatment the housing conditions were changed compared to baseline measurement and the mice were restrained daily. n = 9

Table 5. Results of the *post hoc* comparison of the performed nosepokes in response to the tone-frequencies for the first cognitive bias test. n = negative tone, nn = near-negative tone, m = middle tone, np = near-positive tone and p = positive tone.

Comparison	Estimate	SE	df	t.Ratio	p-Value
m – n	24.72	4.42	112	5.59	<0.0001
m – nn	21.92	4.42	112	5.0	<0.0001
m – np	-32.89	4.42	112	-7.44	<0.0001
m – p	-44.87	4.42	112	-10.15	<0.0001
n – nn	-2.79	4.42	112	-0.63	0.53
n – np	-57.61	4.42	112	-13.03	<0.0001
n – p	-69.59	4.42	112	-15.74	<0.0001
nn – np	-54.82	4.42	112	-12.4	<0.0001
nn – p	-66.80	4.42	112	-15.1	<0.0001
np – p	-11.98	4.42	112	-2.71	<0.001

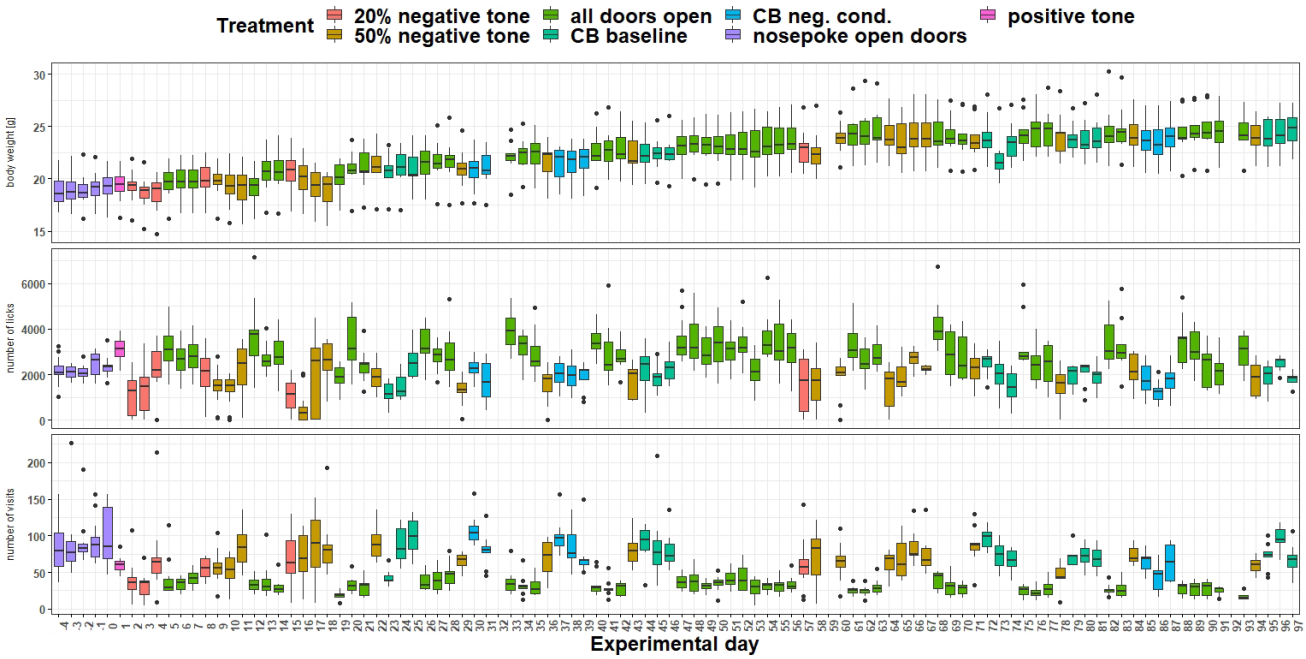

Figure 15. Measurement of body weight, IntelliCage corner visits and lick number over time. The x-axis shows the experimental days. On the y-axis, first the body weight, second the lick number and third the visit number is shown. Different tones with different playback probabilities were presented throughout the experimental period (treatment). The data of experimental day 32, 59 and 92 is missing, due to technical issues.

test itself may already have an influence on the cognitive bias (Roelofs *et al.*, 2016). However, conditioning with punishment seems to be easier to learn and thus seems to succeed faster (Lagisz *et al.*, 2020). In our system, we chose to use a punishment because the behavior of the mice can be interpreted clearly. The mice want to avoid the airpuff and therefore do not poke

when the negative tone is presented. We come to this conclusion based on our experience of mice immediately performing nosepokes upon entering the IC corners if no airpuffs are included in an experimental design.

Lagisz and colleagues identified in their systematic review and meta-analysis that a go/go active choice paradigm (go to receive a reward and go to avoid a punishment) leads to the most sensitive set-up (Lagisz *et al.*, 2020). It is discussed wheather in a go/no-go paradigm the no-go behavior could be related to reduced activity or motivation and not to negative expectation of the future event (Enkel *et al.*, 2010; Matheson *et al.*, 2008). Nevertheless, we chose in our system a go/no-go paradigm. The mice had to nosepoke (go) to receive the reward (water) and not to nosepoke (no-go) to avoid the punishment (airpuff). In addition, the mice had to leave the IC corner and re-enter it (go) to initiate a new trial. We chose a go/no-go paradigm for the same reason that we used the airpuff as a punishment. The behavior in response to the tones is more easily distinguished and interpreted. In addition, by self-initiating the trial, there are no waiting times and the mice have the possibility to complete the trial in a self-determined manner (Hintze *et al.*, 2018; Krakenberg *et al.*, 2019). This choice of experimental design allows us to assume that the mice are highly motivated and facilitates the derivation of a conclusive interpretation of the mice's behavior.

In the third development step we also analyzed the visit and lick behavior. Both seem to be influenced by the treatment (breaks, conditioning or cognitive bias measurement). By starting conditioning, fewer visits and licks were made. It can be assumed that the lick number is also influenced by the circumstance that the IC doors were permanently open during the breaks. This allowed the mice to drink more per visit during the breaks, which consequently reduced the number of visits and increased the number of licks. The data suggest that the mice need more time to drink, as weight was also affected by the treatment. It would therefore be reasonable to increase IC open-door-time. However, the open-door-time should not be so long that the number of visits is reduced because more licks might be made per visit and thus fewer visits are needed and made overall. This in turn would lead to a reduced number of trials for evaluation.

All three groups of mice showed barbering behavior over their lifespan. This behavior occurred at different ages in the respective groups. Group one showed barbering behavior immediately after the experiments presented here, group two during the experiments and group three a few weeks after the experiments presented here. However, it is likely that the behavior was present earlier, as it was only visible through fur lesions. Barbering is a common behavior in female C57 mice (Garner, 2005; Kahnau *et al.*, 2022B). The reasons for the occurrence of this behavior are still unknown. To gain a better understanding of the behavioral course of barbering, we have developed a score sheet (Kahnau *et al.*, 2022B). Whether and what influence barbering has on the mice and thus on the experimental data is unclear. We assume that the influence on the data presented here is rather low, as we were able to condition mice and measure the cognitive bias. Nevertheless, it is necessary to investigate this behavior further and to report it if it occurs.

Because we assumed successful conditioning in developmental step three, the cognitive bias test followed. With the automated and home-cage based set-up presented here, it was possible to measure the cognitive bias of female C57BL/6J mice. Our data showed a sigmoidal curve of data points decreasing from positive tone-frequency to negative tone-frequency. Our result suggests that the ambiguous tone-frequencies are perceived and interpreted differently with respect to the previously conditioned tone-frequencies, which is a basic requirement of a valid cognitive bias test (Gygax, 2014; Hintze *et al.*, 2018; Krakenberg *et al.*, 2019).

We hypothesized that mice living in enriched housing conditions (from 28 days of age) would be affected in their emotional state by removal of enrichment and additional restraining. In fact, we were able to detect a change in the cognitive bias. The mice showed more nosepoke behavior while kept under negative conditions compared to the time of the first baseline measurement, indicating a positive, optimistic cognitive bias. This increased nosepoke behavior was still evident during the second baseline measurement, when the negative conditions had been eliminated. This result is surprising because studies in rats showed that rats housed under negative housing conditions showed a negative cognitive bias (Burman *et al.*, 2009; Harding *et al.*, 2004) and a transfer from standard to enriched housing conditions led to a shift from pessimistic to optimistic cognitive bias (Brydges *et al.*, 2011; Richter *et al.*, 2012). So far, only Resasco and colleagues were able to measure an influence of housing conditions on cognitive bias in mice. Unlike to our study, enriched housed mice seemed to have a positive expectancy related to the ambiguous stimulus compared to standard housed mice (Resasco *et al.*, 2021).

The question arises why the mice in our experiment seem to have a more optimistic cognitive bias after removing enrichment and with restraining. One explanation might be that the mice experienced boredom due to the removal of enrichment, since most stimulating objects had been removed. According to optimal arousal theory, individuals strive for an optimal arousal state. If an individual does not have this arousal state and/or experiences boredom, it would seek something arousing/stimulating. However, if the arousal state is too strong, the individual would seek less arousing stimuli (Mitchell *et al.*, 1984).

In our experiment, this could indicate that the mice did not have an optimal arousal state due to the removal of enrichment and that this is targeted by an increased willingness to take risks to receive an airpuff. However, the mice were also additionally restrained. Thus, it is not possible to identify which factor (removal of enrichment or restraining) or both factors had an influence on the cognitive bias. The influence also seems to be so strong that an increased nosepoke behavior (compared to the first baseline measurement) could also be detected for the second baseline measurement. This raised the question of whether the

mice really had a more optimistic cognitive bias or whether the tones were too "easy" to distinguish. Therefore, we decided to reduce the tone scalar.

The mice also learned to discriminate between tones which were closer to each other, and learned when they received water and when they received an airpuff. Therefore, another cognitive bias test was performed.

During the second test phase, we again observed a sigmoidal curve in the data, but no change in cognitive bias. This result is consistent with other studies (Bailoo *et al.*, 2018; Bračić *et al.*, 2022). It should be noted that during the second test phase, the period of negative conditions was significantly lower. It is possible that one week has no influence on the cognitive bias of mice or that the experiences already made have led to a kind of habituation. It is also possible that the test systems developed so far, including the system presented here, were not sensitive enough. In addition, the possible change in cognitive bias might not last long enough to be measured or is covered by positive stimulation due to cognition training (Krakenberg *et al.*, 2019). Another reason why we could not measure a change could be that a group of mice serving as their own control is not informative enough, as we cannot rule out a temporal carrying over effect for the second baseline. However, the study of Bracic and colleagues showed that the cognitive bias was repeatable over multiple measurements (Bračić *et al.*, 2022). Further experiments are necessary to better interpret the results presented here. For example, it is necessary to test whether a mouse group can serve as its own control group.

It should be noted that a too-frequent repetition of presenting the ambiguous stimuli could also lead to mice learning that neither reward nor punishment occurs with ambiguous stimuli (Roelofs *et al.*, 2016). Therefore, it is necessary to ensure that the ambiguous stimuli are distributed in an appropriately high trial number of positive and negative stimuli (Krakenberg *et al.*, 2019).

In some mouse test systems, the trial number per session is relatively low, *i.e.*1-32 trials (Boleij *et al.*, 2012; Kloke *et al.*, 2014; Novak *et al.*, 2016). In contrast, in the set-up of Hintze and colleagues and in the automated touch-screen system of Krakenberg and colleagues, up to 54 trials per day could be performed. However, it was always necessary to remove the mice from their familiar environment, thus separating them from their group members (with the exception of individual housing) and determining the time of the test, which could have an influence on the motivation to participate in the test.

In our set-up, the number of trials varied depending on how frequently the IC corners were visited (group three: 4 - 214 visits = trials), but were distributed over the entire day. The animals decided independently from the experimenter when to enter the IC (if the IC was not already occupied by another mouse), which makes a high motivation to participate in the test plausible. Even though only one mouse could be in the IC at a time, it was possible for all mice to enter the IC several times a day, and thus, initiate trials in the IC itself. This was also shown in an automated and home-cage based consumer demand test, for which a similar test setup was used as described here (Kahnau *et al.*, 2022A). It is also not necessary to manipulate the night/day rhythm (as *e.g.*, in Krakenberg *et al.*, 2019), as in experiments in which the presence of an experimenter during data acquisition is required. This, in turn, drastically reduces the time required (daily control of animals and set-up of about 30 minutes).

Conclusions

The cognitive bias test seems to be a suitable test method to measure the affective state of animals (Lagisz *et al.*, 2020). So far, however, these tests are very labor intensive and require animals to be tested outside of their home cages, which has implications for the animals and thus the data. (*e.g.*, Bračić *et al.*, 2022; Hintze *et al.*, 2018; Kloke *et al.*, 2014; Krakenberg *et al.*, 2019). Therefore, we aimed to develop an automated and home-cage based cognitive bias test for mice.

In the study presented here, we describe the developmental steps for such a test concluding in a method that allows measuring the cognitive bias in mice. By presenting the various stages of development, we intended to provide a better understanding of the structure of the test method. We also contribute to providing comprehensive information to the scientific community that can be used to develop further automated and home-cage based systems.

Automation and home-cage based testing offers the advantages of testing the mice in their familiar environment and during their active phase. The influence of the animals on each other is reduced, as only one mouse can be in the test-cage at a time. Also, the influence of the experimenter is reduced to a minimum. The fact that the mice can choose the time of the experiment and initiate trials themselves gives them control over what occurs and suggests that the mice are highly motivated. All this, in turn, might have a positive impact on the validity of the data.

We were able to measure and manipulate the cognitive bias of the mice although further research is needed for a better understanding of the mice's cognitive bias. We will continue to develop our test system and use it to assess the burden of commonly used behavioral tests such as the Water Maze Test, and to include the perspective of the mouse in this assessment.

Ethical approval

All experiments were approved by the Berlin state authority, Landesamt für Gesundheit und Soziales (LAGeSo), under license No. G 0182/17 and were in accordance with the German Animal Protection Law (TierSchG, TierSchVersV).

Data availability

Underlying data

Zenodo: Data and supplementary material for the paper "Development of an IntelliCage based Cognitive Bias Test for Mice.", https://doi.org/10.5281/zenodo.7310180 (Kahnau *et al.*, 2022C)

Extended data

Zenodo: Data and supplementary material for the paper "Development of an IntelliCage based Cognitive Bias Test for Mice.", https://doi.org/10.5281/zenodo.7310180 (Kahnau *et al.*, 2022C)

One .pdf file contains supplementary material, the other .pdf file the ARRIVE checklist. In addition, there is one .xlsx file for each developmental step that contains all the data for that step. Each of these three .xlsx files contains a ReadMe sheet describing all variables. The different .txt files contain the R scripts that were used to analyze the data.

Authors' contribution

Conceptualization: Pia Kahnau, Lars Lewejohann, Anne Jaap

Investigation/Methodology: Pia Kahnau

Technical support: Birk Urmersbach

Formal Analysis: Pia Kahnau

Project Administration: Lars Lewejohann, Kai Diederich, Pia Kahnau

Funding acquisition: Lars Lewejohann

Supervision: Lars Lewejohann

Writing – original draft preparation: Pia Kahnau

Writing – review and editing: Pia Kahnau, Anne Jaap, Birk Urmersbach, Kai Diederich, Lars Lewejohann

Acknowledgements

The authors thank the animal caretakers of the German Federal Institute for Risk Assessment, especially Carola Schwarck, Lisa Gordijenko and Heiderose Camus, for their support in the animal husbandry.

Lars Lewejohann and Pia Kahnau are members of the COST Action CA20135 – Improving biomedical research by automated behaviour monitoring in the animal home-cage (TEATIME).

References

Bailoo JD, Murphy E, Boada-Saña M, *et al.*: **Effects of cage enrichment on behavior, welfare and outcome variability in female mice.** *Front Behav Neurosci.* 2018; **12**: 232.
PubMed Abstract | Publisher Full Text | Free Full Text

Boleij H, van't Klooster J, Lavrijsen M, *et al.*: **A test to identify judgement bias in mice.** *Behav Brain Res.* 2012; **233**(1): 45–54.
PubMed Abstract | Publisher Full Text

Bračić M, Bohn L, Siewert V, *et al.*: **Once an optimist, always an optimist? Studying cognitive judgment bias in mice.** *Behav Ecol.* 2022; **33**(4): 775–788.
PubMed Abstract | Publisher Full Text | Free Full Text

Brydges NM, Leach M, Nicol K, *et al.*: **Environmental enrichment induces optimistic cognitive bias in rats.** *Anim Behav.* 2011; **81**(1): 169–175.
Publisher Full Text

Burman OHP, Parker RMA, Paul ES, *et al.*: **Anxiety-induced cognitive bias in non-human animals.** *Physiol Behav.* 2009; **98**(3): 345–350.
PubMed Abstract | Publisher Full Text

Cazala P: **Self-stimulation behavior can be elicited from various 'aversive' brain structures.** *Behav Brain Res.* 1986; **22**(2): 163–171.
PubMed Abstract | Publisher Full Text

de Hoz L, Nelken I: **Frequency tuning in the behaving mouse: Different bandwidths for discrimination and generalization.** *PLoS One.* 2014; **9**(3): e91676.
PubMed Abstract | Publisher Full Text | Free Full Text

Ehret G: **Characteristics of Vocalization in Adult Mice.** In: *Handbook of Behavioral Neuroscience.* 2018; **25**: 187–195.
Publisher Full Text

Endo T, Maekawa F, Võikar V, *et al.*: **Automated test of behavioral flexibility in mice using a behavioral sequencing task in IntelliCage.** *Behav Brain Res.* 2011; **221**(1): 172–181.
PubMed Abstract | Publisher Full Text

Enkel T, Gholizadeh D, Von Bohlen Und Halbach O, *et al.*: **Ambiguous-cue interpretation is biased under stress- and depression-like states in rats.** *Neuropsychopharmacology.* 2010; **35**(4): 1008–1015.
PubMed Abstract | Publisher Full Text | Free Full Text

Francis NA, Kanold PO: **Automated operant conditioning in the mouse home cage.** *Front Neural Circuits.* 2017; **11**: 10.
PubMed Abstract | Publisher Full Text | Free Full Text

Garner JP: **Stereotypies and other abnormal repetitive behaviors: Potential impact on validity, reliability, and replicability of scientific outcomes.** *ILAR J.* 2005; **46**(2): 106–117.
PubMed Abstract | Publisher Full Text

Gouveia K, Hurst JL: **Reducing Mouse Anxiety during Handling: Effect of Experience with Handling Tunnels.** *PLoS One.* 2013; **8**(6): e66401.
PubMed Abstract | Publisher Full Text | Free Full Text

Graulich DM, Kaiser S, Sachser N, *et al.*: **Looking on the bright side of bias—Validation of an affective bias test for laboratory mice.** *Appl Anim Behav Sci.* 2016; **181**: 173–181.
Publisher Full Text

Gygax L: **The A to Z of statistics for testing cognitive judgement bias.** *Anim Behav.* 2014; **95**: 59–69.
Publisher Full Text

Harding EJ, Paul ES, Mendl M: **Animal behaviour: cognitive bias and affective state.** *Nature.* 2004; **427**(6972): 312.
PubMed Abstract | Publisher Full Text

Heffner HE, Heffner RS: **Hearing ranges of laboratory animals.** *J Am Assoc Lab Anim Sci.* 2007; **46**(1): 20–22.
PubMed Abstract

Hintze S, Melotti L, Colosio S, *et al.*: **A cross-species judgement bias task: Integrating active trial initiation into a spatial Go/No-go task.** *Sci Rep.* 2018; **8**(1): 5104.
PubMed Abstract | Publisher Full Text | Free Full Text

Hurst JL, West RS: **Taming anxiety in laboratory mice.** *Nat Methods.* 2010; **7**(10): 825–826.
PubMed Abstract | Publisher Full Text

Jirkof P, Rudeck J, Lewejohann L: **Assessing Affective State in Laboratory Rodents to Promote Animal Welfare—What Is the Progress in Applied Refinement Research?** *Animals.* 2019; **9**(12): 1026.
Publisher Full Text

Jones S, Paul ES, Dayan P, *et al.*: **Pavlovian influences on learning differ between rats and mice in a counter-balanced Go/NoGo judgement bias task.** *Behav Brain Res.* 2017; **331**: 214–224.
PubMed Abstract | Publisher Full Text | Free Full Text

Kahnau P, Guenther A, Boon MN, *et al.*: **Lifetime Observation of Cognition and Physiological Parameters in Male Mice.** *Front Behav Neurosci.* 2021; **15**: 709775.
PubMed Abstract | Publisher Full Text | Free Full Text

Kahnau P, Jaap A, Diederich K, *et al.*: **Determining the value of preferred goods based on consumer demand in a home-cage based test for mice.** *Behav Res Methods.* 2022A; 0123456789.
PubMed Abstract | **Publisher Full Text**

Kahnau P, Jaap A, Hobbiesiefken U, *et al.*: **A preliminary survey on the occurrence of barbering in laboratory mice in Germany.** *Animal Welfare.* 2022B; **31**(4): 433–436.
Publisher Full Text

Kahnau P, Jaap A, Urmersbach B, *et al.*: **Data and supplementary material for the paper "Development of an IntelliCage based Cognitive Bias Test for Mice".** [Data set]. Zenodo. 2022C.
http://www.doi.org/10.5281/zenodo.7310180

Kaupert U, Thurley K, Frei K, *et al.*: **Spatial cognition in a virtual reality home-cage extension for freely moving rodents.** *J Neurophysiol.* 2017; **117**(4): 1736–1748.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Kloke V, Schreiber RS, Bodden C, *et al.*: **Hope for the best or prepare for the worst? Towards a spatial cognitive bias test for mice.** *PLoS One.* 2014; **9**(8): e105431.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Krackow S, Vannoni E, Codita A, *et al.*: **Consistent behavioral phenotype differences between inbred mouse strains in the IntelliCage.** *Genes Brain Behav.* 2010; **9**(7): 722–731.
PubMed Abstract | **Publisher Full Text**

Krakenberg V, Woigk I, Garcia Rodriguez L, *et al.*: **Technology or ecology? New tools to assess cognitive judgement bias in mice.** *Behav Brain Res.* 2019; **362**: 279–287.
PubMed Abstract | **Publisher Full Text**

Krohn TC, Sørensen DB, Ottesen JL, *et al.*: **The effects of individual housing on mice and rats: A review.** *Animal Welfare.* 2006; **15**(4): 343–352.
Reference Source

Lagisz M, Zidar J, Nakagawa S, *et al.*: **Optimism, pessimism and judgement bias in animals: A systematic review and meta-analysis.** *Neurosci Biobehav Rev.* 2020; **118**: 3–17.
PubMed Abstract | **Publisher Full Text**

Lattal KA: **Delayed Reinforcement of Operant Behavior.** *J Exp Anal Behav.* 2010; **93**(1): 129–139.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Lenth R: **Estimated Marginal Means, aka Least-Squares Means.** 2020.
Reference Source

Lewejohann L, Schwabe K, Häger C, *et al.*: **Impulse for animal welfare outside the experiment.** *Lab Anim.* 2020; **54**(2): 150–158.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Manouze H, Ghestem A, Poillerat V, *et al.*: **Effects of single cage housing on stress, cognitive, and seizure parameters in the rat and mouse pilocarpine models of epilepsy.** *eNeuro.* 2019; **6**(4): ENEURO.0179-18.2019.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Matheson SM, Asher L, Bateson M: **Larger, enriched cages are associated with 'optimistic' response biases in captive European starlings (*Sturnus vulgaris*).** *Appl Anim Behav Sci.* 2008; **109**(2–4): 374–383.
Publisher Full Text

Mei J, Kohler J, Winter Y, *et al.*: **Automated radial 8-arm maze: A voluntary and stress-free behavior test to assess spatial learning and memory in mice.** *Behav Brain Res.* 2020; **381**: 112352.
PubMed Abstract | **Publisher Full Text**

Mendl M, Burman OHP, Parker RMA, *et al.*: **Cognitive bias as an indicator of animal emotion and welfare: Emerging evidence and underlying mechanisms.** *Appl Anim Behav Sci.* 2009; **118**(3–4): 161–181.
Publisher Full Text

Mieske P, Diederich K, Lewejohann L: **Roaming in a land of milk and honey: Life trajectories and metabolic rate of female inbred mice living in a semi naturalistic environment.** *Animals.* 2021; **11**(10): 3002.
Publisher Full Text

Mitchell D, Koleszar AS, Scopatz RA: **Arousal and T-maze choice behavior in mice: A convergent paradigm for neophobia constructs and optimal arousal theory.** *Learn Motiv.* 1984; **15**(3): 287–301.
Publisher Full Text

Novak J, Bailoo JD, Melotti L, *et al.*: **An exploration based cognitive bias test for mice: Effects of handling method and stereotypic behaviour.** *PLoS One.* 2015; **10**(7): e0130718.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Novak J, Stojanovski K, Melotti L, *et al.*: **Corrigendum to 'Effects of stereotypic behaviour and chronic mild stress on judgement bias in laboratory mice' [Appl. Anim. Behav. Sci. 174 (2016) 162–172].** *Appl Anim Behav Sci.* 2016; **179**: 104.
Publisher Full Text

Paul ES, Harding EJ, Mendl M: **Measuring emotional processes in animals: The utility of a cognitive approach.** *Neurosci Biobehav Rev.* 2005; **29**(3): 469–491.
PubMed Abstract | **Publisher Full Text**

Pinheiro J, Bates D, DebRoy S, *et al.*: **Linear and Nonlinear Mixed Effects Models**. 2020.
Reference Source

RCoreTeam: **A language and environment for statistical computing**. 2020.
Reference Source

Resasco A, MacLellan A, Ayala MA, *et al.*: **Cancer blues? A promising judgment bias task indicates pessimism in nude mice with tumors.** *Physiol Behav.* 2021; **238**: 113465.
PubMed Abstract | **Publisher Full Text**

Richter SH, Schick A, Hoyer C, *et al.*: **A glass full of optimism: Enrichment effects on cognitive bias in a rat model of depression.** *Cogn Affect Behav Neurosci.* 2012; **12**(3): 527–542.
PubMed Abstract | **Publisher Full Text**

Rivalan M, Munawar H, Fuchs A, *et al.*: **An automated, experimenter-free method for the standardised, operant cognitive testing of rats.** *PLoS One.* 2017; **12**(1): e0169476.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Roelofs S, Boleij H, Nordquist RE, *et al.*: **Making decisions under ambiguity: Judgment bias tasks for assessing emotional state in animals.** *Front Behav Neurosci.* 2016; **10**: 119.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Sarna JR, Dyck RH, Whishaw IQ: **The Dalila effect: C57BL6 mice barber whiskers by plucking.** *Behav Brain Res.* 2000; **108**(1): 39–45.
PubMed Abstract | **Publisher Full Text**

Schlüns H, Welling H, Federici JR, *et al.*: **The glass is not yet half empty: agitation but not Varroa treatment causes cognitive bias in honey bees.** *Anim Cogn.* 2017; **20**(2): 233–241.
PubMed Abstract | **Publisher Full Text**

Verbeek E, Ferguson D, Lee C: **Are hungry sheep more pessimistic? The effects of food restriction on cognitive bias and the involvement of ghrelin in its regulation.** *Physiol Behav.* 2014; **123**: 67–75.
PubMed Abstract | **Publisher Full Text**

Verjat A, Devienne P, Rödel HG, *et al.*: **More exploratory house mice judge an ambiguous situation more negatively.** *Anim Cogn.* 2021; **24**(1): 53–64.
PubMed Abstract | **Publisher Full Text**

Voikar V, Gaburro S: **Three Pillars of Automated Home-Cage Phenotyping of Mice: Novel Findings, Refinement, and Reproducibility Based on Literature and Experience.** *Front Behav Neurosci.* 2020; **14**: 575434.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Voikar V, Krackow S, Lipp HP, *et al.*: **Automated dissection of permanent effects of hippocampal or prefrontal lesions on performance at spatial, working memory and circadian timing tasks of C57BL/6 mice in IntelliCage.** *Behav Brain Res.* 2018; **352**: 8–22.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Wickham H: **ggplot2: Elegant Graphics for Data Analysis**. Springer-Verlag New York. 2016.
Reference Source

Winter Y, Schaefers ATU: **A sorting system with automated gates permits individual operant experiments with mice from a social home cage.** *J Neurosci Methods.* 2011; **196**(2): 276–280.
PubMed Abstract | **Publisher Full Text**

Chapter 3

Discussion

The objective of my dissertation was to develop and perform automated and home-cage based experiments for mice. The used set-up is based on the IntelliCage (IC) system. The IC was initially used in its conventional manner for repetitive cognitive stimulation, with the IC itself being the home-cage. For following experiments, additional components were added to the IC. Thus, the IC was part of the home-cage based set-up and served as the test-cage. By using a gate (AnimalGate, AG), another cage was connected, which served as the home-cage. This extended set-up was used to establish and conduct a Consumer Demand (CD) and a Cognitive Bias (CB) test. The studies presented here demonstrate the use of automated and home-cage based experiments with its multiple advantages but also indicate aspects that need to be considered. With the development of home-cage based experiments, the experimental conditions for laboratory mice could be improved and the experiments themselves allow to study the mouse's point of view. Thus, it is possible to optimize the experimental conditions in order to minimize the burden on the animals. This in turn will provide valid data that will ultimately be more transferable to humans.

3.1 Development challenges

The first challenge was to reduce the very high transponder loss rate of 50% to 60% within 24 hours after injection. However, the transponders are mandatory when using RFID-based systems and the transponder injection had to be repeated if transponders were lost. Repeated anesthesia may affect the well-being of the mice (Hohlbaum et al. 2017) and an optimization of the procedure was necessary (mentioned in the supplements of chapter 4). The main change was the time point of analgesia administration. The analgesia was no longer administered two hours before implantation, but the evening before. This change resulted in a transponder loss rate reduction to 6%. Transponder loss rates are often not reported in scientific publications, possibly because none are lost. However, it is more likely that this is considered to be of little relevance and is therefore not mentioned. By not only mentioning the loss rate but also describing the improvement of the method, it is possible for other scientists to adapt their transponder injection procedure so that the welfare of the animals is not compromised by repeated transponder implantations.

After injecting all mice with transponders, the habituation to the set-up began. While habituation to the IC when used as a home-cage (chapter 3) lasted only a few days, habituation to the

expanded set-up required significantly more time. But first, a technical problem with the AG had to be solved. Immediately at the start of the AG use, it turned out that the door management was not properly adjusted. Individual doors opened and closed too quickly, and it was not possible for the mice to pass through the gate without physical contact with the doors. The AG settings had to be adjusted, which could not be done by myself and required an update from the supplier.

Second, the protocol of habituation to the set-up had to be established. At the beginning, the mice were kept in a type IV cage during the two-week institute quarantine/acclimatization phase. Later, the mice were transferred to the automated and home-cage based set-up. The mice had to be habituated to the AG and the IC. Initially, all doors were open and all mice were able to explore the set-up without restriction and learn that water is exclusively offered in the IC. Step by step, the AG doors and IC doors within the IC corners were closed. Thus, the mice were adapted to the movements of the doors as well as to the separation within the IC. At the beginning, the set-up habituation took four weeks. By keeping mice groups within the IC based set-up immediately after arrival at the institute (with all doors open) and experiences with the functioning of the AG as well as the behavior of the mice, the set-up habituation was reduced to 16 days.

After set-up habituation only one (out of 12 mice) mouse could stay in the IC at a time. At the beginning there was concern that not all mice would be able to enter the IC with sufficient frequency within 24 hours. Since water was only offered in the IC, it could have resulted in not all mice having access to water sufficiently frequently. However, it turned out that all mice entered the IC several times and especially during the dark phase (chapter 4). Nevertheless, it occurred sometimes that individual mice entered the IC several times within 24 hours but even without drinking. The reason for not drinking could be no learning but was not always apparent, as it happened that the mice entered the IC corners and performed nosepokes but did not drink. However, if mice did not drink, they were given water for 15 minutes in a separate type III cage. On repeated failure to drink, other corners were released or the tones in the CB test were turned off (described in chapters 3 and 4). Not drinking occurred relatively rarely and we were able to show that when water was offered only in the IC, the mice drank. It has also been shown that laboratory animals enter a test-cage independently from the home-cage (Mei et al. 2020; Kaupert et al. 2017; Rivalan et al. 2017; Winter & Schaefers 2011). The unique feature in the studies presented here, however, is that not only 4-6 mice were housed as one group, but 12. Our results show that even in large group sizes, all mice were able to enter the test-cage several times per 24 hours.

Another more extensive issue was that individual ICs turned off for no identifiable reason. This led to data losses on individual days during the experiments (chapter 4 and 5). Again, I had to rely on the supplier's support.

These technical issues highlight a disadvantage of using commercial systems: It relies on the supplier's support. Even though the communication with the supplier was positive, it took some time until the appropriate updates could be delivered. In self-developed systems it is possible to react faster to technical problems, but the development of own systems requires also a lot of time and especially expertise (Habedank et al. 2022).

An important factor for the development and establishment of experiments is to determine the optimal methodology. In the experiments presented here, the mice had to complete different learning tasks. During the long-term study (chapter 3), it was shown that the mice reached the learning criterion only during the first IC phase. As discussed in the publication, the later learning tasks might have been too complex. The mice chose a different strategy to access water than we expected. They simply endured the airpuff or it was not aversive enough and did not provide an obstacle to find the correct corner to drink without receiving an airpuff. Therefore, it is important to be able to monitor the success of an experiment while it is running in order to identify the use of alternative strategies by the animals.

Even during the pre-test of the CD experiment, it became rapidly obvious that it was not feasible to determine whether the mice had learned the operant task or whether the motivation to work for access to almond milk was low (discussed in chapter 4). Only the adaptation of the method to work not for almond milk but for water instead of drinking a bitter tasting liquid allowed a conclusion about the learning success of the mice.

The development of a suitable conditioning method for the CB test was particularly complex and time-consuming. As described in chapter 5, different methods were tested. At each step of the development important knowledge was gained, which finally led to a successful conditioning method. In my opinion, the description of individual development steps and pre-tests are very valuable. This offers the possibility to get an understanding of the applied method. It also offers the possibility to use this knowledge to develop other experiments. All the failed experiments and the finally successful experiment help to get a better understanding of the behavior of the experimental mice.

3.2 Advantages of the automated and home-cage based system

The first presented study (chapter 3) demonstrated the usability of the IC for cognitively stimulation over a long period of time. Even though the IC can be used as a home-cage based system, there is a possibility that individual mice influence each other. It was observed that individual mice pulled each other out of the IC corners. This influence was eliminated by extending the IC with a gate and another cage, allowing only one mouse to be within the IC to solve the task undisturbed.

Expanding the set-up, the test-cage became part of the home-cage. This has the advantage that the mice can be tested in their familiar environment. For the experiments presented here, the mice were not removed from their home-cage and not actively separated from their group members. Past studies showed that the handling method, the separation and testing within an unknown environment may influence the well-being of the animals (Manouze et al. 2019; Gouveia & Hurst 2013; Hurst & West 2010; Krohn et al. 2006; Chesler et al. 2002; Crabbe et al. 1999). To reduce stress, the mice were well habituated to the home-cage based set-up and a tunnel handling method was used. In order to promote this method, we (members of the working group Laboratory Animal Science at the German Federal Institute for Risk Assessment) established our own tube-handling protocol with a short video (https://wiki.norecopa.no/index.php/Mouse_handling).

The attached test-cage allowed the mice to determine when to enter the test-cage and thus determine the timing of the experiment. Moreover, the mice had an influence on their daily routine and were not externally determined by the experimenter at which time of day/night the experiment is carried out. Also, the self-initiation of the single trials (by entering the IC corners independently and repeatedly) contributes to more self-determination which may lead to a reduction of frustration (Krakenberg et al. 2019; Hintze et al. 2018). Through self-determined participation, a high motivation to solve the tasks may be assumed and contribute to the production of valid data.

Laboratory mice are nocturnal animals, nevertheless certain experiments are performed during the light phase (Habedank et al. 2021) or the day/night rhythm is reversed (Krakenberg et al. 2019), which in turn requires time and a habituation phase. In the experiments presented here none of this was necessary. The mice were tested during their active phase with no influence on their natural diurnal rhythm. Therefore, it is likely that the motivation to participate in the experiment is high.

By automatizing the experiments, data were recorded without the presence of the experimenter, eliminating observer bias and increasing reproducibility (Voikar & Gaburro 2020; Pernold et al. 2019; Krackow et al. 2010). This advantage is particularly important when multiple experimenters are involved, as in the long-term study presented in chapter 3.

A further major advantage of the automated and home-cage based set-up is the significantly reduced daily time requirement. While conventional methods require the experimenter to be permanently present during the experiment, the daily time required was about 20 to 60 minutes, depending on the experiment (and without technical problems).

3.3 Limitations

In 2020, mice were used most frequently for scientific research in Germany (Bundesinstitut für Risikobewertung 2021). Therefore, improvements in husbandry and experimental conditions could have an impact on a large number of laboratory animals. In the studies presented here, mice of strain C57BL/6J were used which is the most commonly used strain (van de Lagemaat et al. 2017). However, it should be noted that there are differences in behavior between different strains (Pitzer et al. 2021; König et al. 2020; O'Leary et al. 2013; Podhorna & Brown 2002; Connolly & Lynch 1981). Therefore, the results presented here cannot have general validity. Consequently, the experiments should be repeated with mice from other strains as well. However, it should be considered whether it is possible to transfer the results of strains that behave very similarly in order to avoid unnecessary animal experiments. Through systematic analyses, it is possible to compare the results of different experiments. Some results also suggest that similar interventions are even likely to lead to comparable results in different species (Mieske et al., 2022). However, to improve husbandry and experimental conditions, a comprehensive understanding of the behavior of our laboratory mice is required.

In the studies presented here, both male and female mice were used, but only one sex per study. In the long-term study, four social groups with 12 males were kept and, as discussed earlier, no agonistic behaviors requiring intervention occurred. Nevertheless, it would be useful to repeat the study with females, as there are also sex-specific behavioral differences (Tucker & McCabe 2017; Van Den Berg et al. 2015). For example, males take higher risks than females (Gomes et al. 2022). This could mean that females would have reached the learning criterion during IC learning phase two or three because they would have avoided the airpuff. This in turn could have been reflected on physiological parameters such as resting metabolic rate or body weight.

For the other two studies, only females were used. As described above, there was initial concern that not all mice would have adequate access to water, as water was only offered in the IC. We assume that individual males claim access to the gate, and thus to the IC and water, through increased territorial behavior (Kappel et al. 2017). As a result, water would have to be offered separately to the mice frequently and frequent intervention would have a negative influence on the mice and thus on the data.

Nevertheless, it is necessary to test the automated and home-cage based set-up also with male mice. Only in this way we can find out which preferences or aversions the males have or in which emotional state they are. And only in this way can the laboratory conditions also for male mice be improved.

A major limitation of the experiments presented here is that only liquids could be used in the IC system. Nevertheless, the results are important because they show that water can be used as a reward without limiting the access to water in time. The results of the CD experiment may also serve in the choice of appropriate rewards or punishments in future experiments. Nevertheless, it is necessary to be able to test the preference or aversion for other goods such as bedding or nesting material and enrichment items. This allows to study the wants and needs of the mice and to improve the laboratory conditions.

The original aim was to use the experiments presented here to evaluate the burden of commonly performed behavioral tests with the inclusion of the mice's perspective. Due to the long development time of the CB Test in particular, the burden of the behavioral tests could not be performed. Nevertheless, we now have experiments that allow us to measure the wants and needs as well as the emotional state of the mice in an automated and home-cage based manner.

3.4 Conclusion and outlook

The aim of this dissertation was to improve the test conditions for mice in experiments and to reduce the burden. To achieve these aims, the IC system was used to develop and perform automated and home cage-based experiments. In the experiments presented here, mice successfully completed different learning tasks. It was possible to better understand the influence of cognitive stimulation, to investigate the motivation to access certain liquids for work, and to measure the emotional state of the mice. Furthermore, an experimental set-up was designed that allows individual data collection without being disturbed/influenced by group members while solving different tasks.

Through this dissertation, it was demonstrated that the development of automated and home-cage based experiments requires a great amount of time and patience, as well as flexibility. Nevertheless, the development of such experiments and set-ups is important and working systems offer many advantages with them.

A follow-up experiment will use the CB test presented here to evaluate the burden of a Water Maze experiment from the mice's perspective. For this purpose, mice will experience a water maze experiment and the CB will be measured. The CB will then be compared to the CB of mice that will not have had this experience. This should make it possible to elucidate the burden in a Water Maze test.

In a further project, the experience gained here will be used to advance the Mouse Position Surveillance System (MoPSS) we have already developed (Habedank et al., 2022). The aim is to develop a modular system in which a home-cage is connected to a test-cage by a gate.

Further test-boxes with nosepoke-sensors will be connected to the test-cage. Thus, it will be possible to let mice work in a CD experiment, beside liquids, also for e.g. nest material or enrichment items.

As a result, it will be possible to better understand the wants and needs of mice and to design their environment in such a way to accommodate as many natural behaviors as possible. This will lead to the generation of valid data, which in turn will have a positive effect on the transfer-ability to humans.

Especially because animal testing cannot be completely eliminated in the near future, it is our responsibility to study and understand our laboratory animals as best we can. Only in this way are we able to make everyday laboratory life as comfortable as possible for the laboratory mice entrusted to us.

Chapter 4

Summary

The use of an automated and home-cage based test system to improve behavioral experiments for group housed mice

One possibility to improve laboratory conditions for animals is to conduct the experiments in a home-cages manner. For the experiments presented here, the RFID (radio frequency identification)-based IntelliCage (IC) served as an automated and home-cage-based test system.

Chapter 2 describes behavioral methods that are available to assess the burden on laboratory animals in animal studies. It describes the importance of examining animal behavior in order to conclude about burden, for example, by assessing activity. By using preference tests, the perspective of the animals can be included, since the animals are (indirectly) asked what they want or do not want. The Consumer Demand test then offers the possibility to determine the strength of the preference or aversion. Through the Cognitive Bias test, it is also possible to measure the emotional state of the animals.

In chapter 3, the IC was used in a long-term study to cognitively stimulate mice in the IC repeatedly, while at the same time another subset of mice was never cognitively stimulated. However, the study showed treatment effects only in the early life phase of the mice. Young mice that were cognitively stimulated showed higher activity, lower growth and a lower resting metabolic rate. They were also less attractive to female mice. However, these results were not evident at later life stages. Furthermore, no effect of cognitive stimulation on dominance or longevity could be detected.

In chapter 4, an extended IC based set-up (IC and home-cage connected by a gate) was used to investigate the strength of preference/aversion for different liquids in a Consumer Demand test. For this, the mice had to make more nosepokes daily to gain access to different liquids. The data show that the number of nosepokes was dependent on the liquids offered. The mice made many nosepokes to avoid drinking a bitter-tasting liquid, while they made significantly fewer nosepokes to gain access to sweet-tasting liquids.

In chapter 5, the extended IC based setup was used to develop a home-cage based Cognitive Bias test. In this study, the focus was on three developmental steps leading to a functional testing protocol. We showed successful conditioning, as well as measurement of cognitive bias in female mice.

In chapter 6, the main developmental challenges and advantages of using the home-cage based system but also limitations are summarized. For instance, the method of RFID transponder implantation had to be improved, or various technical problems had to be solved. Chapter 6 also describes the advantages of using home-cage based systems. We were able to successfully record the behavior of mice over several months. We obtained individual data, the mice were tested in their familiar environment and during their natural active phase. The influence of the experimenter was reduced to a minimum, as well as the influence of the mice on each other, when only one mouse could stay in the IC at a time and thus solved the tasks undisturbed by group members.

In conclusion, performing home-cage based experiments offered the possibility to refine the laboratory conditions for laboratory animals. This will again help to obtain valid data, which will be beneficial for humans.

Chapter 5

Zusammenfassung

Die Verwendung eines automatisierten und heimatkäfigbasierten Testsystems zur Verbesserung von Verhaltensexperimenten für Mäuse

Eine Möglichkeit, die Laborbedingungen für Tiere zu verbessern, ist die Versuche innerhalb des Heimatkäfigs durchzuführen. Bei den hier vorgestellten Experimenten diente das RFID (*Radio Frequency Identification*) -basierte *IntelliCage* System (IC) als automatisiertes und heimatkäfigbasierendes Testsystem.

Kapitel 2 beschreibt die Verhaltensmethoden, die zur Bewertung der Belastung von Labortieren in Tierversuchen zur Verfügung stehen. Es wird beschrieben, wie wichtig es ist, das Verhalten der Tiere zu untersuchen, um Rückschlüsse auf die Belastung zu ziehen, zum Beispiel durch die Bewertung der Aktivität. Durch die Verwendung von Präferenztests kann die Perspektive der Tiere einbezogen werden, da die Tiere (indirekt) gefragt werden, was sie wollen oder nicht wollen. Der *Consumer Demand* Test bietet dann die Möglichkeit, die Stärke der Präferenz oder Abneigung zu bestimmen. Durch den *Cognitive Bias* Test ist es außerdem möglich, den emotionalen Zustand der Tiere zu messen.

In Kapitel 3 wurde der IC in einer Langzeitstudie eingesetzt, um Mäuse im IC wiederholt kognitiv zu stimulieren, während gleichzeitig eine weitere Gruppe von Mäusen nie kognitiv stimuliert wurde. Die Studie zeigte jedoch nur in der frühen Lebensphase der Mäuse Unterschiede zwischen den Gruppen. Junge Mäuse, die kognitiv stimuliert wurden, zeigten eine höhere Aktivität, ein geringere Körpergewichtsentwicklung und einen niedrigeren Ruhestoffwechsel. Sie waren auch weniger attraktiv für weibliche Mäuse. Diese Unterschiede waren jedoch nicht mehr in späteren Lebensphasen auszumachen. Es konnten außerdem keine Auswirkung der kognitiven Stimulation auf Dominanz oder Langlebigkeit festgestellt werden.

In Kapitel 4 wurde ein erweiterter IC-basierter Aufbau (IC und Heimkäfig, verbunden durch eine Schleuse) verwendet, um die Stärke der Präferenz/Abneigung für verschiedene Flüssigkeiten in einem *Consumer Demand* Versuch zu untersuchen. Für den Zugang zu unterschiedlichen Flüssigkeiten, mussten die Mäuse täglich mehr *Nosepokes* machen. Die Daten zeigen, dass die Anzahl der *Nosepokes* von den angebotenen Flüssigkeiten abhing. Die Mäuse machten viele *Nosepokes*, um eine bitter schmeckende Flüssigkeit nicht trinken zu müssen, während sie deutlich weniger *Nosepokes* machten, um Zugang zu süß schmeckenden Flüssigkeiten zu erhalten.

In Kapitel 5 wurde die erweiterte Testapparatur eingesetzt, um einen heimatkäfigbasierten *Cognitive Bias* Test zu entwickeln. In dieser Studie lag der Schwerpunkt auf den drei Entwicklungsschritten, die zu einem funktionalen Testprotokoll führten. Wir zeigten die erfolgreiche Konditionierung sowie die Messung der Erwartungsvalenz (*Cognitive Bias*) bei weiblichen Mäusen.

In Kapitel 6 wurden die wichtigsten Herausforderungen bei der Entwicklung und die Vorteile der Verwendung des Heimkäfigsystems, aber auch die Einschränkungen zusammengefasst. So musste beispielsweise die Methode der RFID-Transponder-Implantation verbessert werden, oder es mussten verschiedene technische Probleme gelöst werden. In Kapitel 6 wurden außerdem die Vorteile der Verwendung von Heimkäfigsystemen beschrieben. Wir konnten das Verhalten von Mäusen über mehrere Monate hinweg erfolgreich aufzeichnen. Wir erhielten individuelle Daten, die Mäuse wurden in ihrer gewohnten Umgebung und während ihrer natürlichen aktiven Phase getestet. Der Einfluss der Experimentatorin wurde auf ein Minimum reduziert, ebenso wie der Einfluss der Mäuse untereinander, da sich jeweils nur eine Maus im IC aufhalten konnte und somit die Versuche ungestört von Gruppenmitgliedern absolvieren konnte.

Zusammenfassend lässt sich sagen, dass die Durchführung von Experimenten in Heimkäfigen die Möglichkeit bietet, die Laborbedingungen für Labortiere zu verbessern. Dies wird wiederum dazu beitragen, valide Daten zu erhalten, die für den Menschen von Nutzen sein werden.

References

Bundesinstitut für Risikobewertung (2021):
Zahlen zu den im Jahr 2020 verwendeten Versuchstieren.
Abgerufen am: 24.10.2022 um 13:52 Uhr, von: https://www.bf3r.de/cm/343/zahlen-zu-den-im-jahr-2020-verwendeten-versuchstieren.pdf

Chesler E J, Wilson S G, Lariviere W R, Rodriguez-Zas S L, Mogil J S (2002):
Influences of laboratory environment on behavior [1].
Nat Neurosci., *5*(11), 1101–1102. https://doi.org/10.1038/nn1102-1101

Connolly M S, Lynch C B (1981):
Circadian variation of strain differences in body temperature and activity in mice.
Physiol Behav., *27*(6), 1045–1049. https://doi.org/10.1016/0031-9384(81)90368-1

Crabbe J C, Wahlsten D, Dudek B C (1999):
Genetics of mouse behavior: Interactions with laboratory environment.
Science, 284(5420), 1670–1672. https://doi.org/10.1126/science.284.5420.1670

Endo T, Maekawa F, Võikar V, Haijima A, Uemura Y, Zhang Y, Miyazaki W, Suyama S, Shimazaki K, Wolfer D P, Yada T, Tohyama C, Lipp H P, Kakeyama M (2011):
Automated test of behavioral flexibility in mice using a behavioral sequencing task in IntelliCage.
Behav Brain Res., 221(1), 172–181. https://doi.org/10.1016/j.bbr.2011.02.037

Freund J, Brandmaier A M, Lewejohann L, Kirste I, Kritzler M, Krüger A, Sachser N, Lindenberger U, Kempermann G (2013):
Emergence of Individuality in Genetically Identical Mice.
Science, 340(May), 756–760.

Gomes M G S, Tractenberg S G, Orso R, Viola T W, Grassi-Oliveira R (2022):
Sex differences in risk behavior parameters in adolescent mice: Relationship with brain-derived neurotrophic factor in the medial prefrontal cortex.
Neurosci Lett., 766, 136339. https://doi.org/10.1016/j.neulet.2021.136339

Gouveia K, Hurst J L (2013):
Reducing Mouse Anxiety during Handling: Effect of Experience with Handling Tunnels.
PLoS ONE, 8(6). https://doi.org/10.1371/journal.pone.0066401

Habedank A, Kahnau P, Lewejohann L (2021):
Alternate without alternative: neither preference nor learning explains behaviour of C57BL/6J mice in the T-maze.
Behaviour. https://doi.org/DOI 10.1163/1568539X-bja10085

Habedank A, Urmersbach B, Kahnau P, Lewejohann L (2022):
O mouse, where art thou? The Mouse Position Surveillance System (MoPSS)—an RFID-based tracking system.
Behav Res Methods, 54(2), 676–689. https://doi.org/10.3758/s13428-021-01593-7

Hintze S, Melotti L, Colosio S, Bailoo J D, Boada-Saña M, Würbel H, Murphy E (2018):
A cross-species judgement bias task: Integrating active trial initiation into a spatial Go/No-go task.
Sci Rep., 8(1), 1–13. https://doi.org/10.1038/s41598-018-23459-3

Hobbiesiefken U, Urmersbach B, Jaap A, Diederich K, Lewejohann L (2021):
Rating enrichment items by group-housed laboratory mice in multiple binary choice tests using an RFID-based tracking system.
bioRxiv preprint, https://doi.org/10.1101/2021.10.20.465117

Hohlbaum K, Bert B, Dietze S, Palme R, Fink H, Thöne-Reineke C (2017):
Severity classification of repeated isoflurane anesthesia in C57BL/6JRj mice - Assessing the degree of distress.
PLoS ONE, 12(6), 1–21. https://doi.org/10.1371/journal.pone.0179588

Hurst J L, West R S (2010):
Taming anxiety in laboratory mice.
Nat Methods, 7(10), 825–826. https://doi.org/10.1038/nmeth.1500

Jirkof P, Fleischmann T, Cesarovic N, Rettich A, Vogel J, Arras M (2013):
Assessment of postsurgical distress and pain in laboratory mice by nest complexity scoring.
Lab Anim., 47(3), 153–161. https://doi.org/10.1177/0023677213475603

Kahnau P, Guenther A, Boon M N, Terzenbach J D, Hanitzsch E, Lewejohann L, Brust V (2021):
Lifetime Observation of Cognition and Physiological Parameters in Male Mice.
Front Behav Neurosci., 15(September). https://doi.org/10.3389/fnbeh.2021.709775

Kappel S, Hawkins P, Mendl M T (2017):
To group or not to group? Good practice for housing male laboratory mice.
Animals, 7(12), 1–25. https://doi.org/10.3390/ani7120088

Kaupert U, Thurley K, Frei K, Bagorda F, Schatz A, Tocker G, Rapoport S, Derdikman D, Winter Y (2017):
Spatial cognition in a virtual reality home-cage extension for freely moving rodents.
Journal of Neurophysiology, 117(4), 1736–1748. https://doi.org/10.1152/jn.00630.2016

König C, Plank A C, Kapp A, Timotius I K, von Hörsten S, Zimmermann K (2020):
Thirty Mouse Strain Survey of Voluntary Physical Activity and Energy Expenditure: Influence of Strain, Sex and Day–Night Variation.
Front Neurosci., 14(July). https://doi.org/10.3389/fnins.2020.00531

Krackow S, Vannoni E, Codita A, Mohammed A H, Cirulli F, Branchi I, Alleva E, Reichelt A, Willuweit A, Voikar V, Colacicco G, Wolfer D P, Buschmann J U F, Safi K, Lipp H P (2010):
Consistent behavioral phenotype differences between inbred mouse strains in the IntelliCage.
Genes Brain Behav., 9(7), 722–731. https://doi.org/10.1111/j.1601-183X.2010.00606.x

Krakenberg V, Woigk I, Garcia Rodriguez L, Kästner N, Kaiser S, Sachser N, Richter S H (2019):
Technology or ecology? New tools to assess cognitive judgement bias in mice.
Behav Brain Res., 362, 279–287. https://doi.org/10.1016/j.bbr.2019.01.021

Kramer K, Van De Weerd H, Mulder A, Van Heijningen C, Baumans V, Remie R, Voss H P, Van Zutphen B F M (2004):
Effect of conditioning on the increase of heart rate and body temperature provoked by handling in the mouse.
Altern Lab Anim., 32(SUPPL. 1A), 177–181.
https://doi.org/10.1177/026119290403201s29

Krohn T C, Sørensen D B, Ottesen J L, Hansen A K (2006):
The effects of individual housing on mice and rats: A review.
Anim Welf., 15(4), 343–352.

Lewejohann L, Hoppmann A M, Kegel P, Kritzler M, Krüger A, Sachser N (2009):
Behavioral phenotyping of a murine model of Alzheimer's disease in a seminaturalistic environment using RFID tracking.
Behav Res Methods, 41(3), 850–856. https://doi.org/10.3758/BRM.41.3.850

Manouze H, Ghestem A, Poillerat V, Bennis M, Ba-M'hamed S, Benoliel J J, Becker C, Bernard C (2019):
Effects of single cage housing on stress, cognitive, and seizure parameters in the rat and mouse pilocarpine models of epilepsy.

ENeuro, 6(4). https://doi.org/10.1523/ENEURO.0179-18.2019

Mei J, Kohler J, Winter Y, Spies C, Endres M, Banneke S, Emmrich J V (2020):
Automated radial 8-arm maze: A voluntary and stress-free behavior test to assess spatial learning and memory in mice.
Behav Brain Res., 381(November), 112352. https://doi.org/10.1016/j.bbr.2019.112352

Mieske P, Diederich K, Lewejohann L (2021):
Roaming in a land of milk and honey: Life trajectories and metabolic rate of female inbred mice living in a semi naturalistic environment.
Animals, 11(10). https://doi.org/10.3390/ani11103002

Mieske P, Hobbiefiefken U, Fischer-Tenhagen C, Heinl C, Hohlbaum K, Kahnau P, Meier J, Wilzopolski J, Butzke D, Rudeck J, Lewejohann L, Diederich K (2021):
Bored at home? A systematic review on the effect of environmental enrichment on the welfare of laboratory rats and mice.
Front Vet Sci., 9. https://doi.org/10.3389/fvets.2022.899219

Miller A L, Flecknell P A, Leach M C, Roughan J V (2011):
A comparison of a manual and an automated behavioural analysis method for assessing post-operative pain in mice.
Appl Anim Behav Sci., 131(3–4), 138–144. https://doi.org/10.1016/j.applanim.2011.02.007

O'Leary T P, Gunn R K, Brown R E (2013):
What are we measuring when we test strain differences in anxiety in mice?
Behav Genet., 43(1), 34–50. https://doi.org/10.1007/s10519-012-9572-8

Park J, Jung M S, Moon E, Lim H J, Oh C E, Lee J H (2021):
Prediction of locomotor activity by infrared motion detector on sleep-wake state in mice.
Clin Psychopharmacology and neuroscience, 19(2), 303–312.
https://doi.org/10.9758/cpn.2021.19.2.303

Pernold K, Iannello F, Low B E, Rigamonti M, Rosati G, Scavizzi F, Wang J, Raspa M, Wiles M V, Ulfhake B (2019):
Towards large scale automated cage monitoring - Diurnal rhythm and impact of interventions on in-cage activity of C57BL/6J mice recorded 24/7 with a non-disrupting capacitive-based technique.
PLoS ONE, 14(2), 1–20. https://doi.org/10.1371/journal.pone.0211063

Pitzer C, Kurpiers B, Eltokhi A (2021):
Gait performance of adolescent mice assessed by the CatWalk XT depends on age, strain and sex and correlates with speed and body weight.

Scientific Reports, 11(1), 1–18. https://doi.org/10.1038/s41598-021-00625-8

Podhorna J, Brown R E (2002):
Strain differences in activity and emotionality do not account for differences in learning and memory performance between C57BL/6 and DBA/2 mice.
Genes, Brain Behav., 1(2), 96–110. https://doi.org/10.1034/j.1601-183X.2002.10205.x

Rivalan M, Munawar H, Fuchs A, Winter Y (2017):
An automated, experimenter-free method for the standardised, operant cognitive testing of rats.
PLoS ONE, 12(1), 1–19. https://doi.org/10.1371/journal.pone.0169476

Russel W M S, Burch R L (1952):
The Prinziple of Humane Experimental Technique.
Special Edition. Universities Federation for Animal Wlfare (UFAW),
ISBN: 0-900767-78-2

Späni D, Arras M, König B, Rülicke T (2003):
Higher heart rate of laboratory mice housed individually vs in pairs.
Lab Anim., 37(1), 54–62. https://doi.org/10.1258/002367703762226692

Steele, A D, Jackson W S, King O D, Lindquist S (2007):
The power of automated high-resolution behavior analysis revealed by its application to mouse models of Huntington's and prion diseases.
Proceedings of the National Academy of Sciences of the United States of America, 104(6), 1983–1988. https://doi.org/10.1073/pnas.0610779104

Ticher A, Ashkenazi I E (1995):
Compound structure of rodents activity rhythm.
Physiology and Behavior, 57(1), 37–40. https://doi.org/10.1016/0031-9384(94)E0108-G

Tucker L B, McCabe J T (2017):
Behavior of male and female C57Bl/6J mice is more consistent with repeated trials in the elevated zero maze than in the elevated plus maze.
Front Behav Neurosci., 11(January), 1–8. https://doi.org/10.3389/fnbeh.2017.00013

van de Lagemaat L N, Stanford L E, Pettit C M, Strathdee D J, Strathdee K E, Elsegood K A, Fricker D G, Croning M D R, Komiyama N H, Grant S G N (2017):
Standardized experiments in mutant mice reveal behavioural similarity on 129S5 and C57BL/6J backgrounds.
Genes Brain Behav., 16(4), 409–418. https://doi.org/10.1111/gbb.12364

Van Den Berg W E, Lamballais S, Kushner S A (2015):
Sex-Specific Mechanism of Social Hierarchy in Mice.

Neuropsychopharmacology, 40(6), 1364–1372. https://doi.org/10.1038/npp.2014.319

Voikar V, Gaburro S (2020):
Three Pillars of Automated Home-Cage Phenotyping of Mice: Novel Findings, Refinement, and Reproducibility Based on Literature and Experience.
Front Behav Neurosci, 14(October), 1–15. https://doi.org/10.3389/fnbeh.2020.575434

Voikar V, Krackow S, Lipp H P, Rau A, Colacicco G, Wolfer D P (2018):
Automated dissection of permanent effects of hippocampal or prefrontal lesions on performance at spatial, working memory and circadian timing tasks of C57BL/6 mice in IntelliCage.
Behav Brain Res., 352(August 2017), 8–22. https://doi.org/10.1016/j.bbr.2017.08.048

Weegh N, Füner J, Janke O, Winter Y, Jung C, Struve B, Wassermann L, Lewejohann L, Bleich A, Häger C (2020):
Wheel running behaviour in group-housed female mice indicates disturbed wellbeing due to DSS colitis.
Lab Anim., 54(1), 63–72. https://doi.org/10.1177/0023677219879455

Winter Y, Schaefers A T U (2011):
A sorting system with automated gates permits individual operant experiments with mice from a social home cage.
J Neurosci Methods, 196(2), 276–280. https://doi.org/10.1016/j.jneumeth.2011.01.017

Appendix

Author's contribution

Publication: Behavioral methods for severity assessment

Pia Kahnau: article conceptualization, literature research, writing of the original draft, figure preparation, review and editing, correspondence with reviewers

Co-authors: cooperation in article conceptualization, figure preparation, cooperation in writing, review and editing, cooperation with correspondence with reviewers, funding acquisition, supervision

Publication: Lifetime observation of cognition and physiological parameters in male mice

Pia Kahnau: experimental planning and conducting (IC phase 2), animal care taking (IC phase 2), use, care and cleaning of the IntelliCage system (IC phase 2), article conceptualization, literature research, writing of the original draft, data analysis and visualization, figure preparation, correspondence with reviewers

Co-authors: experimental planning and conducting, animal care taking, cooperation in article conceptualization, literature research, data analysis and visualization, writing, review and editing, cooperation with correspondence with reviewers, funding acquisition, supervision

Publication: Determining the value of preferred goods based on consumer demand in a home-cage based test for mice

Pia Kahnau: development of study, experimental planning and conducting, animal care taking, interchange with animal care takers, configuration, use, maintenance and cleaning of IntelliCage system, correspondence with supplier TSE, literature research, writing of the original draft, data analysis and visualization, figure preparation, correspondence with reviewers

Co-authors: cooperation in experimental planning, cooperation in article conceptualization, literature research, data analysis, writing, review and editing, cooperation with correspondence with reviewers, funding acquisition, supervision

Publication: Development of an IntelliCage based cognitive bias test for mice

Pia Kahnau: development of study, experimental planning and conducting, animal care taking, interchange with animal care takers, configuration, use, maintenance and cleaning of IntelliCage system, correspondence with supplier TSE, literature research, writing of the original draft, data analysis and visualization, figure preparation

Co-authors: cooperation in experimental planning, technical support, cooperation in article conceptualization, writing, review and editing, funding acquisition, supervision

Publication Index

Original Articles

Habedank A, Urmersbach B, Kahnau P, Lewejohann L:

O mouse, where art thou? The Mouse Position Surveillance System (MoPSS) – an RFID based tracking system.

Accepted for publication by Behavior Research Methods, 2021

Published 2022, https://doi.org/10.3758/s13428-021-01593-7

Habedank A, Kahnau P, Lewejohann L:

Alternate without alternative: Neither preference nor simple learning behavior shown by C57BL/6J mice in the T-maze.

Published in Behavior, 2021, https://brill.com/view/journals/beh/158/7/article-p625_4.xml

Kahnau P, Guenther A, Boon M N, Terzenbach J D, Hanitzsch E, Lewejohann L, Brust V:

Lifetime Observation of Cognition and Physiological Parameters in Male Mice.

Published in Behavioral Neuroscience, 2021, https://www.frontiersin.org/articles/10.3389/fnbeh.2021.709775/full

Kahnau P, Jaap A, Diederich K, Gygax L, Rudeck J, Lewejohann L:

Determining the value of preferred goods based on consumer demand in a home-cage based test for mice.

Published in Behavior Research Methods, 2021, https://doi.org/10.3758/s13428-022-01813-8

Kahnau P, Jaap A, Urmersbach B, Diederich K, Lewejohann L:

Development of an IntelliCage-based cognitive bias test for mice.

Published in Open Research Europe, 2022 https://doi.org/10.12688/openreseurope.15294.1

Reviews

Habedank A*, Kahnau P*, Diederich K, Lewejohann L:

Severity assessment from an animal's point of view.

Berliner Münchener tierärztlichen Wochenschrift, 2018

doi: 10.2376/0005-9366-18007

* These authors contributed equally

Kahnau P, Habedank A, Diederich K, Lewejohann L:

Behavioral Methods for Severity Assessment.

Published in Animals 2020, https://www.mdpi.com/2076-2615/10/7/1136

Survey

Kahnau P, Jaap A, Hobbiesiefken U, Mieske P, Diederich K, Thöne-Reineke C, Lewejohann L, Hohlbaum K:

A preliminary survey on the occurrence of barbering in laboratory mice in Germany.

In press in Animal Welfare, 2022, https://doi.org/10.7120/09627286.31.4.009

Talks

Kahnau P (2018):

Von optimistischen und pessimistischen Mäusen: Der Cognitive Bias Test

BB3R Spring School, Freie Universität Berlin, Berlin

Kahnau P (2018):

The Cognitive Bias Test as a means for severity assessment from an animal's point of view.

DRS PhD Symposium, Freie Universität Berlin, Berlin

Kahnau P (2019):

Der Cognitive Bias Test als Methode zur Beurteilung der Belastung in Tierversuchen aus Sicht der Tiere.

Klausurwoche Wohlergehen von Tieren – ethische, wissenschaftstheoretische, und biologische Perspektiven, Münster

Kahnau P (2019):

The Cognitive Bias Test as a means for severity assessment from an animal's point of view.

PreDocSymposium, German Federal Institute for Risk Assessment (BfR), Berlin

Kahnau P (2020):
The Cognitive Bias Test as a means for severity assessment from an animal's point of view.
BB3R-Symposium, Freie Universität Berlin, Berlin

Kahnau P (2022):
Home-cage based set-ups for mice.
Berlin-Oxford Open Science Summer School, Berlin

Kahnau P (2022):
Home-Cage based cognitive bias Test for mice.
BB3R Autumn School, Freie Universität Berlin, Berlin

Posters

Kahnau P (2018):
Von optimistischen und pessimistischen Mäusen: Der Cognitive Bias Test
BB3R Spring School, Freie Universität Berlin, Berlin

Kahnau P, Habedank A, Lewejohann L (2018):
The cognitive Bias Test as a means for severity assessment from an animal's point of view.
BB3R-3R Symposium, Freie Universität Berlin, Berlin

Kahnau P, Habedank A, Lewejohann L (2019):
The cognitive Bias Test as a means for severity assessment from an animal's point of view.
PreDocSymposium, German Federal Institute for Risk Assessment (BfR), Berlin

Kahnau P, Habedank A, Lewejohann L (2019):
The cognitive Bias Test as a means for severity assessment from an animal's point of view.
Tagung der Ethologischen Gesellschaft, Hannover
Abstract in Program Book

Kahnau P, Habedank A, Lewejohann L (2020):
Home cage based consumer demand test to investigate the strength of preferences in mice.
Tagung der Ethologischen Gesellschaft, Tübingen
Abstract in Program Book

Danksagung

Bereits im Bachelorstudium nach der ersten Vorlesung Verhaltensbiologie von Prof. Norbert Sachser war mir klar: DAS will ich machen. Mich faszinierte die Frage, warum sich Tiere verhalten, wie sie sich verhalten. Die Möglichkeit während meiner Promotion Methoden zu entwickeln, Mäuse indirekt zu fragen was sie wollen oder in welchem emotionalen Status sie sich befinden, war daher für mich eine spannende Herausforderung und vor allem eine große Freude. Ohne Unterstützung auf unterschiedlichste Weise wäre es mir aber nicht möglich gewesen diese Arbeit zu vollenden.

Zuerst möchte ich den wichtigsten Akteuren dieser Dissertation danken: den Mäusen. Die Forschung habe ich mit ihnen aber eben auch für sie gemacht. Ziel ist es sich für die einzusetzen, die es für sich selbst nicht können und die Laborbedingungen aller Mäuse zu verbessern.

Neben den Mäusen möchte ich mich aber auch bei all den Menschen bedanken, die mich auf diesem Weg begleitet, unterrichtet und unterstützt haben. Allen voran meinem Mentor Prof. Lars Lewejohann. Seine offene und optimistische Art erlaubte es mir eigene Ideen zu entwickeln, die zu den Lösungen der vielen kleinen und großen Probleme beitrugen. Zu jedem Zeitpunkt war er erreichbar und gab mir Selbstvertrauen die Expertin meiner Arbeit zu sein.

Dem Mentoren-Team gehörten auch Dr. Kai Diederich und Prof. Christa Thöne-Reineke an, denen ich ebenfalls für ihre Unterstützung danken möchte. Die ruhige, verständnisvolle und vor allem diplomatische Art von Kai half dabei die richtigen Antworten unter anderem auf Reviewer-Kommentare zu finden. Durch Gespräche mit ihm wurden mir viele bürokratische und wissenschaftliche Prozesse verständlich und halfen dabei Anträge und Versuchsabläufe zu optimieren. Bei Christa möchte ich mich für ihre motivierende und positive Art bedanken. Durch ihre Anerkennung meines Ideenreichtums und Durchhaltevermögens bin ich immer mit einem optimistischen Gefühl aus unseren Meetings gegangen.

Besonders möchte ich mich bei Anne Jaap bedanken, die mit mir zusammen im selben Projekt geforscht hat. Ihre Fähigkeit, alles von allen Seiten zu betrachten, hat mir geholfen meinen Horizont zu erweitern und Lösungen für Probleme in den unterschiedlichen Versuchen zu finden. Unsere Unterhaltungen über so viele verschiedene Themen haben mir immer sehr gefallen. Auch waren durch ihre emotionale Unterstützung die unterschiedlichen Fehlschläge viel einfacher zu bewältigen.

Auch für die offene, lustige und dadurch sehr angenehme Atmosphäre innerhalb der Fachgruppe 92 möchte ich mich bei meinen (auch ehemaligen) Kollegen Birk, Paul, Katharina, Benjamin, Ute und Juliane bedanken. Die Unterhaltungen über Mäuse, Verhalten, Wissen-

schaft aber auch Themen wie Filme, Musik und das Leben selbst waren immer sehr unterhaltsam und boten Möglichkeiten kleiner Arbeitspausen. Besonders die Corona-Zeit machte deutlich, wie wichtig ein guter kollegialer Zusammenhalt ist.

Bedanken möchte ich mich auch bei Lorenz Gygax, Vera Brust und Anja Guenther, die mich bei der Datenauswertung unterstützt haben. Ihre Bereitschaft und Geduld alle meine Fragen zu verstehen (was im Bereich der Statistik gar nicht so leicht ist) und ausführlich zu beantworten, führte dazu, dass ich immer selbständiger meine Daten auswerten konnte und auch fremde Arbeiten viel besser einschätzen kann. Dies gibt mir die wichtige Sicherheit im Austausch mit anderen Wissenschaftlern.

In der Tierpflege wurde ich von den Tierpflegerinnen des Bundesinstituts für Risikobewertung unterstützt. Vor allem an den Wochenenden und Tagen, an denen ich nicht selbst vor Ort sein konnte, übernahmen sie die wichtige Aufgabe der Kontrolle der Mäuse und der Versuche. Es war sicher nicht immer leicht die komplizierten und am Anfang noch nicht ausgereiften Versuchskonzepte nachzuvollziehen. Ich bedanke mich sehr für diese Unterstützung. Besonderer Dank gilt Carola Schwarck, Lisa Gordijenko und Heidi Camus, die sich vor allem intensiv in das nicht einfache IntelliCage System eingearbeitet haben.

Unterstützt bei der Pflege der Mäuse wurde ich auch von Dr. Annalena Riedasch. Als Tierärztin hat sie mir beigebracht die Anzeichen richtig zu deuten, um den Gesundheitszustand der Mäuse richtig einschätzen zu können. Durch sie habe ich sicheren Umgang gerade mit alten Mäusen gelernt. Vielen Dank für die vielen lehrreichen Unterhaltungen.

Ich möchte mich auch bei meiner Familie bedanken, bei meiner Mutter und meinen Geschwistern Thorsten, Miriam und Timo Kahnau, die mich nach Berlin hat ziehen lassen, um hier zu promovieren. Dadurch habe ich einige schöne und lustige Events verpasst und habe eindeutig zu wenig Zeit mit ihnen verbringen können. Es war sicher nicht immer leicht nachzuvollziehen, warum ich so viel und oft an eigentlich freien Tagen gearbeitet habe. Danke für das Verständnis. Auch meinen Schwiegereltern Renate und Peter Grundmann möchte ich für ihr immer großes Interesse an mir und meinem Lebensweg danken. Durch ihren eigenen Lebensweg können sie sehr gut nachvollziehen, wie anstrengend es sein kann, sich nach oben zu arbeiten und dafür auch mal den ein oder anderen Umweg zu gehen.

Dieses Verständnis hat mir vor allem mein Ehemann Felix Grundmann entgegengebracht. Er hat mich bei so vielen so wichtigen Etappen begleitet und unterstützt. Durch seine liebevolle, einfühlsame und verständnisvolle Art habe ich mich immer weiterentwickeln können. Ich danke ihm dafür, dass er mich besser kennt als ich mich selbst. Wahrscheinlich kann er am besten einschätzen, was es für mich bedeutet, diese Dissertation nach langer, schöner, intensiver aber auch anstrengender Zeit beenden zu können.

Finanzierungsquellen

Die Arbeiten wurden finanziell unterstützt durch DFG (FOR 2591; LE 2356/5-1)

Interessenskonflikte

Im Rahmen dieser Arbeit bestehen keine Interessenskonflikte durch Zuwendungen Dritter.

Selbstständigkeitserklärung

Hiermit bestätige ich, dass ich die vorliegende Arbeit selbstständig angefertigt habe.
Ich versichere, dass ich ausschließlich die angegebenen Quellen und Hilfen in Anspruch genommen habe.

Berlin, den 03.05.2023

Pia Kahnau

mbvberlin mensch und buch verlag

49,90 Euro | ISBN: 978-3-96729-224-4