



# Effects of Cage Enrichment on Behavior, Welfare and Outcome Variability in Female Mice

Jeremy D. Bailoo<sup>1\*</sup>, Eimear Murphy<sup>1</sup>, Maria Boada-Saña<sup>1</sup>, Justin A. Varholick<sup>1</sup>, Sara Hintze<sup>2</sup>, Caroline Baussière<sup>1</sup>, Kerstin C. Hahn<sup>3</sup>, Christine Göpfert<sup>3</sup>, Rupert Palme<sup>4</sup>, Bernhard Voelkl<sup>1</sup> and Hanno Würbel<sup>1</sup>

<sup>1</sup>Division of Animal Welfare, Veterinary Public Health Institute, University of Bern, Bern, Switzerland, <sup>2</sup>Division of Livestock Sciences, Department of Sustainable Agricultural Systems, University of Natural Resources and Life Sciences Vienna (BOKU), Vienna, Austria, <sup>3</sup>Institute for Animal Pathology, University of Bern, Bern, Switzerland, <sup>4</sup>Department of Biomedical Sciences, University of Veterinary Medicine Vienna, Vienna, Austria

The manner in which laboratory rodents are housed is driven by economics (minimal use of space and resources), ergonomics (ease of handling and visibility of animals), hygiene, and standardization (reduction of variation). This has resulted in housing conditions that lack sensory and motor stimulation and restrict the expression of species-typical behavior. In mice, such housing conditions have been associated with indicators of impaired welfare, including abnormal repetitive behavior (stereotypies, compulsive behavior), enhanced anxiety and stress reactivity, and thermal stress. However, due to concerns that more complex environmental conditions might increase variation in experimental results, there has been considerable resistance to the implementation of environmental enrichment beyond the provision of nesting material. Here, using 96 C57BL/6 and SWISS female mice, respectively, we systematically varied environmental enrichment across four levels spanning the range of common enrichment strategies: (1) bedding alone; (2) bedding + nesting material; (3) deeper bedding + nesting material + shelter + increased vertical space; and (4) semi-naturalistic conditions, including weekly changes of enrichment items. We studied how these different forms of environmental enrichment affected measures of animal welfare, including home-cage behavior (time-budget and stereotypic behavior), anxiety (open field behavior, elevated plus-maze behavior), growth (food and water intake, body mass), stress physiology (glucocorticoid metabolites in fecal boluses and adrenal mass), brain function (recurrent perseveration in a two-choice guessing task) and emotional valence (judgment bias). Our results highlight the difficulty in making general recommendations across common strains of mice and for selecting enrichment strategies within specific strains. Overall, the greatest benefit was observed in animals housed with the greatest degree of enrichment. Thus, in the super-enriched housing condition, stereotypic behavior, behavioral measures of anxiety, growth and stress physiology varied in a manner consistent with improved animal welfare compared to the other housing conditions with less enrichment. Similar to other studies, we found no evidence, in the measures assessed here, that environmental enrichment increased variation in experimental results.

**Keywords:** environmental enrichment, variation, animal welfare, mice, behavioral phenotypes

## OPEN ACCESS

### Edited by:

S. Tiffany Donaldson,  
University of Massachusetts Boston,  
United States

### Reviewed by:

Valerie J. Bolivar,  
Wadsworth Center, United States  
Chantal Mathis,  
Centre National de la Recherche  
Scientifique (CNRS), France

### \*Correspondence:

Jeremy D. Bailoo  
jbailoo@gmail.com

**Received:** 20 July 2018

**Accepted:** 14 September 2018

**Published:** 26 October 2018

### Citation:

Bailoo JD, Murphy E, Boada-Saña M, Varholick JA, Hintze S, Baussière C, Hahn KC, Göpfert C, Palme R, Voelkl B and Würbel H (2018) Effects of Cage Enrichment on Behavior, Welfare and Outcome Variability in Female Mice.  
*Front. Behav. Neurosci.* 12:232.  
doi: 10.3389/fnbeh.2018.00232

## INTRODUCTION

The mouse, *Mus musculus*, is the most widely used animal in research, and yet surprisingly little is known about how its behavioral biology relates to the social and physical aspects of current laboratory housing conditions (Latham and Mason, 2004). Consequently, the Committee updating the US Guide for the Care and Use of Laboratory Animals identified a critical lack of empirical evidence on “space and housing needs” and “enrichment, exercise and human contact” (Garber, 2011; National Research Council, 2011). We have recently reported on a systematic assessment of the effects of space allowance on measures of animal welfare in laboratory mice (Bailoo et al., 2018); here we report on the effects of cage enrichment on behavior and measures of welfare in female laboratory mice.

Housing conditions for laboratory mice have been shaped primarily by economics (minimal use of space, equipment and labor), ergonomics (ease of handling, visibility of animals), hygiene (easy to sanitize) and standardization (minimization of variation; see Olsson and Dahlborn, 2002; Baumans and Van Loo, 2013, for further details). Typically, mice have been kept in transparent “shoe-box” cages with bedding, food and water. Such housing conditions lack sensory and motor stimulation and may prevent mice from performing species-typical behaviors, such as nest building (Würbel, 2001; Latham and Mason, 2004). Consequently, such housing conditions are associated with signs of impaired welfare, including abnormal repetitive behavior (Garner et al., 2004a,b, 2011; Garner, 2005; Würbel, 2006; Gross et al., 2012) and anxiety (Chapillon et al., 1999; van Praag et al., 2000; Würbel, 2001). Based on such evidence, Switzerland has declared environmental enrichment (i.e., nesting material) mandatory (The Swiss Federal Council, 2008), while the EU Directive 2010/63/EU (European Parliament Council, 2010) and the US Guide for the Care and Use of Laboratory Animals (2011) only recommend enrichment of rodent cages. Thus, an increasing number of researchers are using nesting material, shelters, gnawing sticks and other enrichment items with the goal of improving the welfare of laboratory rodents (The Swiss Federal Council, 2008; European Parliament Council, 2010; Baumans and Van Loo, 2013).

Environmental enrichment is used to increase sensory and motor stimulation, to facilitate species-typical behavior, and to provide the animals with some degree of control over their environment (Dawkins, 1988; Newberry, 1995; Olsson and Dahlborn, 2002; Nithianantharajah and Hannan, 2006; Gross et al., 2011b; Bennett et al., 2018). Current evidence indicates that even the addition of simple forms of enrichment to standard laboratory cages can improve the welfare of laboratory mice, as shown by reduced abnormal repetitive behavior (DeLuca, 1997; Würbel et al., 1998; Nevison et al., 1999a; Latham and Mason, 2010; Tilly et al., 2010; Bechard et al., 2011; Gross et al., 2011b, 2012) and reduced measures of anxiety (e.g., Chapillon et al., 1999; Roy et al., 2001; Benaroya-Milshtein et al., 2004; Binder et al., 2004; Görtz et al., 2008; Sztainberg and Chen, 2010), although not all studies have confirmed these findings (van de Weerd et al., 1994; Nevison et al., 1999a; Zhu et al., 2006).

In most studies, cage enrichment adds several resources, environmental complexity and sometimes novelty to the standard housing conditions. However, one study has found that within standard laboratory cages neither complexity nor novelty of enrichments had beneficial effects beyond those of nesting material alone (Gross et al., 2011b). Arguably, nesting material is the only enrichment that has consistently been found to be beneficial for mouse welfare (for review, see Olsson and Dahlborn, 2002).

Nesting material is an important resource for laboratory mice as it allows for the expression of species-typical nest building behavior, facilitates thermoregulation and provides shelter, if provided in sufficient quantity (Bult and Lynch, 1997; Sherwin, 1997; van de Weerd et al., 1997; Olsson and Dahlborn, 2002; Smith and Corrow, 2005; Gaskill et al., 2009, 2011, 2012; Gross et al., 2011b). Mice are highly motivated to construct nests (Nicol et al., 2008), and prefer nesting material to nest boxes (van de Weerd et al., 1998). Furthermore, for resting and maintenance behavior, mice prefer much higher ambient temperatures (30–32°C) than are common in animal facilities (20–26°C)—without nesting material, laboratory mice are at a higher risk for experiencing chronic cold stress (e.g., Johnson et al., 2017).

Other forms of enrichment, by contrast, have produced inconsistent effects (for reviews, see Jennings et al., 1998; Olsson and Dahlborn, 2002; Benefiel et al., 2005; Smith and Corrow, 2005; Whittaker et al., 2012; Baumans and Van Loo, 2013). For example, in some studies the provision of shelters has been associated with increased levels of aggression in male mice (McGregor and Ayling, 1990; Haemisch et al., 1994; Howerton et al., 2008), but this seems to depend strongly on the strain of mouse (van de Weerd et al., 1994; Chapillon et al., 1999; Nevison et al., 1999a).

The aim of the present study was to replicate and extend the results of previous studies on the effects of cage enrichment on mouse behavior and measures of welfare. We varied cage enrichment across four levels, including cages: (i) without any enrichment (Barren, B); (ii) with nesting material (Nesting, N); (iii) with deep bedding, shelters and additional vertical space besides nesting material (Enriched, E); and (iv) large pet cages attached to a laboratory cage, offering multiple resources and different items each week for active engagement (Super-Enriched, SE). To increase the generality of our findings, we studied an inbred and an outbred strain of mouse. However, similar to other proof-of-concept studies on environmental enrichment in mice (Van Loo et al., 2003; Wolfer et al., 2004), we only studied female mice as some forms of enrichment are associated with escalating aggression in male mice. Our primary outcome measures of animal welfare were stereotypy performance in the home-cage and measures of anxiety in behavioral tests, as these two measures were most consistently found to be improved by environmental enrichment. Additionally, a range of secondary outcome measures covering different domains of animal welfare were also included: home-cage behavior, measures of growth, endocrine stress responses, brain function, and emotional state. We tested the hypothesis that the welfare of mice increases with increasing degrees of

cage enrichment across our four treatment groups, following the prediction:  $B < N < E < SE$ .

Since environmental enrichment renders the animals' environment more complex, concerns have been raised that cage enrichment might lead to higher variability in experimental results. Although empirical evidence does not support these concerns (Augustsson et al., 2003; Wolfer et al., 2004; Baumans et al., 2010; van de Weerd et al., 2010) they seem to persist (e.g., Toth et al., 2011). Therefore, we additionally assessed how cage enrichment affected variability in all measured outcomes.

## MATERIALS AND METHODS

### Experimental Design, Animals and Housing Conditions

We used a 4 (cage enrichment)  $\times$  2 (mouse strain) factorial design conducted in two consecutive batches of equal size; the second batch began 3 weeks after the end of the first batch. Subjects were 192 female mice, 96 each of the inbred strain C57BL/6J (C57) and the outbred strain RjOrl:SWISS (SWISS) from Janvier Labs, France. Each batch was comprised of 48 newly weaned non-sibling mice (21–25 days old at delivery) per strain. All mice were ear-tattooed for identification by the same two experimenters (JB and EM) on the day following arrival at the laboratory.

A review of the literature, comparing variation in behavioral outcomes between barren cages and cages with nesting material alone, yielded an effect size range of 0.85–1.0 (Cohen's  $d$ ) resulting in a sample size of  $n = 12$  per treatment group. Expecting some of our behavioral measures, such as the behavior in the elevated plus-maze, to be more variable and thus yield a smaller effect size, and to accommodate for possible attrition, we adjusted our sample size to  $n = 24$  per experimental condition per strain.

Mice were randomly allocated to the four housing conditions, B, N, E and SE, described in further detail in **Supplementary Table S1**. They were housed in groups of three per cage, with four replicate cages per strain and treatment in each batch (see **Supplementary Table S2**). Each batch of mice was delivered in four boxes, two per strain, containing 24 animals each. Each box of mice was allocated to cages and treatment groups sequentially to minimize individual differences in behavior between cages. Animals were housed in two housing rooms, located on either side of the test room (**Supplementary Figure S1**), with half of the animals per strain per batch in each room. Cage height on the rack was counterbalanced by strain and room between batches.

The B, N and E housing conditions consisted of a Makrolon<sup>®</sup> Type three cage, which besides food (Kliba Nafag #3430, Switzerland) and tap water *ad libitum*, contained either bedding only (B; Lignocel<sup>®</sup> select, see **Supplementary Table S1** for depth), bedding and nesting material (N), or deeper bedding, nesting material, a tunnel, a shelter and increased vertical space (E). To increase the generality of our findings, we used two different types of nesting material (10 g of Sizzle

Pet<sup>®</sup> or three paper tissues), tunnels (rat tunnel Plexx EU #13104 or rat retreat Plexx EU #13154) and shelters (arch Plexx EU #13244 or hut Plexx EU #13169), counterbalanced across cages.

The SE housing condition consisted of a Makrolon<sup>®</sup> Type three cage connected to a Savic<sup>™</sup> Mickey XL pet cage, by a clear polycarbonate tunnel 6 cm in diameter. The layout of the SE system and the enrichments used are displayed in **Supplementary Figures S2, S3**. Briefly, the Type three cage contained a paper tunnel (Plexx EU #14152), a paper shelter (Plexx EU #13244), and nesting material (10 g Sizzle Pet<sup>®</sup> and 3 $\times$  paper tissues). The Mickey XL cage contained an elevated platform made of polycarbonate (410 cm<sup>2</sup> in floor area, with a 5 cm wall), with a wooden ladder leading up to it and a wooden shelter on top of it. A plastic hammock was attached to the cage lid above the elevated platform where a secondary source of food and water was also accessible. Two wooden coconuts attached to the cage lid and joined by a suspended bridge made of rope and wood, a paper shelter, and a paper tunnel were also provided. These items in the Mickey XL cage were present throughout the study. To stimulate activity and exploration further, and to increase behavioral diversity, additional items were rotated in the SE cage on a weekly basis (see **Supplementary Table S3**).

### Husbandry Procedures

Animals were kept on a reversed 12:12 light/dark cycle with lights on at 19:00 h. A red light emitting diode (LED) remained on throughout the entire cycle. Temperature was maintained at  $22 \pm 1^\circ\text{C}$ , with an average humidity of 40%. Husbandry procedures were conducted weekly (see **Supplementary Methods**).

Because some outcome measures involved long periods of testing and manipulation by the experimenter, all mice were habituated to being handled beginning upon arrival at the laboratory (see **Supplementary Methods**).

### Outcome Variables

Outcome variables covered a range of measures related to animal welfare. Our primary outcome measures were: (i) stereotypy performance in the home-cage; and (ii) measures of anxiety in two behavioral tests (elevated plus-maze test and open field test). In addition, we assessed a range of secondary outcome measures covering different domains of animal welfare, including measures of: (i) home-cage behavior (time budget, use of enrichments); (ii) growth (food and water intake, body weight); (iii) endocrine stress responses (glucocorticoid metabolites in fecal boluses, adrenal weight); (iv) brain function (inhibitory control of behavior as measured by recurrent perseveration in a two-choice guessing task); and (v) emotional state (judgment bias in a spatial Go/No-Go task; **Figure 1**).

For assessing recurrent perseveration and judgment bias, one focal animal from three of the four cages per treatment, strain and batch were randomly selected ( $n = 48$  in total). Cage, treatment and strain were counterbalanced across three experimenters (EM, MB-S, JB) such that: (1) no two

experimenters tested animals from the same treatment at the same time; (2) no experimenter tested animals from the same treatment consecutively; and (3) no experimenter tested animals from the same strain consecutively. For judgment bias testing, one experimenter was replaced (MB-S by SH) towards the end of the training period due to an emergency leave of absence in batch 1, and for batch 2, mice from one experimenter (EM) were shared for two sessions by the other two experimenters (JB and MB-S) due to illness.

### Stereotypy Performance in the Home-Cage and Other Home-Cage Behavior

Cage enrichment has previously been reported to reduce abnormal repetitive behavior (DeLuca, 1997; Würbel et al., 1998; Nevison et al., 1999b; Tilly et al., 2010; Bechard et al., 2011), and one study has found that nesting material is associated with a reduction of stereotypic behavior in particular (Gross et al., 2012). To evaluate the effects of increasing degrees of enrichment on the incidence of stereotypic behavior, all cages were recorded for 24 h each, prior to the end of the experiment.

Video-recordings were scored separately for stereotypic behavior and other home-cage behavior using Noldus Observer XT (version 10.5) by EM, CB and JB. From each cage all animals were observed and their behavior scored, although, to evaluate the relation between stereotypic behavior and recurrent perseveration, data from focal animals were used (the same ones tested in the guessing task). Videos of two E cages (all intervals) and one SE cage (one interval) were unavailable for scoring due to equipment failure (see below for a description of the intervals).

Stereotypic behavior was scored using a previously validated ethogram (Novak et al., 2016; see also **Supplementary Table S4**). Mice were observed for 15 min, within four 1-h time windows, distributed across the dark phase (07:30–08:30, 09:30–10:30, 12:30–13:30 and 15:30–16:30). These time windows had been determined by pilot observations to represent two time intervals of high activity and two of low activity. Stereotypic behavior and general activity were sampled using one-zero sampling with 15 s intervals, yielding 240 data points per mouse. Stereotypy performance is reported here as a proportion of active time. Ten percent of all videos were rescored for assessment of intra- and inter-rater reliabilities, which were high throughout (Jansen et al., 2003),  $\kappa = 0.94$  and  $0.96$ , respectively.

In addition to stereotypic behavior, we also assessed the extent to which the mice engaged with the cage environment, including engagement with enrichments, and how the provided resources within the cage affected general daily activities and facilitated species-typical behavior (**Supplementary Table S5**; category, environmental manipulation). Of note, in terms of engagement with the cage environment, we evaluated manipulation of bedding (e.g., digging) in the B cage, while we additionally evaluated engagement with the provided enrichment items in the N, E and SE cages. Thus, engagement with the cage environment was quantified even in the barren cage, where no enrichments were present. We did not categorize time spent on the lid as engagement with

the environment, although these data were analyzed and are described in “Anxiety in the Elevated Plus-Maze Test and Open Field Test” section in our analysis of time budget.

To evaluate engagement with the cage environment, we used the ethogram described under **Supplementary Table S5**. Mice were observed for 30 min, within six 1-h time windows, four in the dark phase (the same as above) and two in the light phase (20:30–21:30, 03:00–04:00), using instantaneous sampling at 1-min intervals and yielding 180 data points per mouse. Ten percent of all videos were rescored for assessment of intra- and inter-rater reliabilities, which were high throughout (Jansen et al., 2003),  $\kappa = 0.85$  and  $0.90$ , respectively.

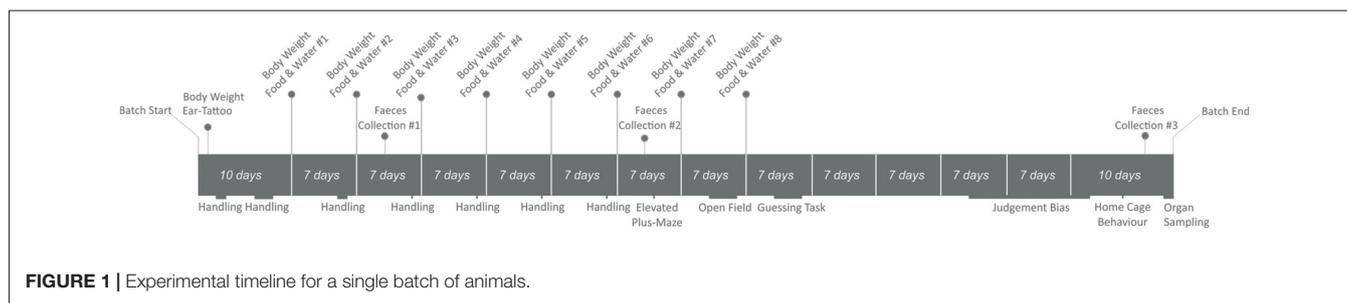
### Anxiety in the Elevated Plus-Maze Test and Open Field Test

Besides reduced stereotypy performance, cage enrichment in mice has most consistently been associated with reduced measures of anxiety in behavioral tests such as the elevated plus-maze test or the open field test. Both elevated plus-maze behavior (Pellow et al., 1985; Carola et al., 2002) and open field behavior (Denenberg, 1969; Carola et al., 2002) have been validated for differences in anxiety in both rats and mice. On the elevated plus-maze, less time spent on, and fewer entries into, the open arms, as well as reduced locomotion reflect higher levels of anxiety (Walf and Frye, 2007). In the open field, longitudinal assessment of the pattern of locomotor behavior in the open field across repeated exposures has been demonstrated to provide information about how animals behaviorally cope with a stressor and how the hypothalamus-pituitary-adrenal (HPA) axis differentially operates between groups of animals (Whimbeay and Denenberg, 1967; Baillou et al., 2013). Specifically, reduced levels of habituation (i.e., reduced exploration and/or time in the center across days of testing) and/or increased levels of sensitization (i.e., greater exploration and/or time in the center across days of testing), between groups of animals, are associated with increased anxiety and vice versa.

In the present experiment, we used three elevated plus-mazes and three open field arenas. Elevated plus-mazes were made of polycarbonate with infrared (850 nm) backlit floors. Each maze consisted of four arms, each 30 cm in length and 6 cm wide, and a center square measuring  $6 \times 6$  cm. Two arms opposite to each other were open, with a small lip around the perimeter 0.5 cm high, while the remaining two arms were enclosed, with walls 15 cm high. The open field arenas were made of polycarbonate with dimensions  $45 \times 45 \times 45$  cm<sup>3</sup> with infrared (850 nm) backlit floors.

For both tests, mice from each cage were randomly assigned to one of three experimenters (EM, MB-S and JB). The test order of cages was counterbalanced by treatment and strain in blocks of four across eight blocks, whereby all four treatments were represented in each block, with two treatments per strain. For consecutive blocks, the strain associated with treatment in the previous block was alternated.

Behavior on the elevated plus-maze was assessed in a single session while behavior in the open field was assessed across four consecutive days; all tests were conducted between 10:00 h



**FIGURE 1** | Experimental timeline for a single batch of animals.

and 13:00 h. The test duration for both tests was 5 min. The outcome variables of interest in the elevated plus-maze test were: (1) distance traveled; and (2) time spent in the open arms. In the open field test changes in: (1) distance traveled; (2) time in the center; and (3) time in the corners, across 4 days of testing was evaluated.

For each test session, the cage to be tested was brought to the test room and the overhead lights (120 lux) were turned on. The three animals were then removed from their cage and placed into the apparatus by the assigned experimenter. At the end of the test, the animals were replaced into the home-cage, and the cage returned to the housing room. Between test sessions, the arenas were cleaned with 70% isopropanol. Outcome measures were scored live by Noldus EthoVision XT (version 11.5). Accuracy of video tracking was subsequently evaluated by JB from video recordings ensuring issues associated with automated tracking were eliminated (Bailoo et al., 2010). The detection settings for tracking were selected so that both the percentage of samples in which the subject was not found and the percentage of samples skipped were less than 1% per trial.

### Growth: Body Weight and Food and Water Intake

Longitudinal assessment of food and water intake and body weight can provide information about the different types of experienced stress, since acute stress is associated with a loss of and chronic stress with a gain in body weight (Klok et al., 2007; McEwen, 2007; Torres and Nowson, 2007; Lutter et al., 2008). Food and water intake are coupled and thus expected to be positively correlated.

Body weight of each mouse and food and water intake at the cage level, were recorded at weekly cage changes. Food and water intake were assessed for each cage by subtracting the weight of the remaining food and water from that of food and water provided at the last cage change. Thus, we measured food disappearance rather than food intake, as particulates of food dropped in the bedding and spilled water remained unaccounted for. For analysis, food and water intake were corrected for the number of animals per cage when there were fewer than three animals in the cage (see “Attrition” section, for further details).

### Endocrine Stress Responses: Fecal Glucocorticoid Metabolites and Adrenal Weight

Non-invasive methods of quantifying circulating levels of glucocorticoids, a primary product of the activation of the

HPA stress system, are preferable to invasive methods such as blood sampling, because they do not elicit a stress response. In mammals, glucocorticoids are metabolized by the liver and are excreted in both urine and feces. A validated method of analyzing glucocorticoid metabolites in mice (Touma and Palme, 2005) was used as a measure of endocrine stress responses prompted by the different housing conditions.

Feces were collected, with a minimum of six boluses per mouse per cage, at three time points in the dark phase under red light approximately 24 h after cage changes. Because the transit time of by-products of corticosterone secretion in feces is between 8 h and 10 h (Touma et al., 2003; Touma and Palme, 2005), it is important to note here that we are measuring, in part, arousal/stress as a consequence of the husbandry procedures associated with cage change (Balcombe et al., 2004). Samples were immediately frozen at  $-20^{\circ}\text{C}$  and later processed (blinded to experimental treatment, JV and RP) to assess the concentration of  $5\alpha\text{-}3\beta$ ,  $11\beta$ -corticosterone metabolites (ng/0.05 g feces) as described by Touma et al. (2003, 2004). In total 563 samples were processed; four samples were missing due to attrition (see “Attrition” section), and nine samples were not processed because of insufficient sample material.

At the end of the experiment, animals were killed by anesthesia with 5% isoflurane followed by asphyxiation with  $\text{CO}_2$ , performed by JB. Within 2 min, the animals were transported to the dissection laboratory. Dissections were performed by KH and CG. Animals were first weighed, then the adrenals were extracted and weighed using a precision scale to the nearest 10,000th of a gram (Mettler AE 160). Chronic exposure to stress has been associated with higher levels of circulating glucocorticoids and heavier adrenal glands (van de Weerd et al., 1994; McEwen, 2007). Other organs (brain, heart, kidney, liver, spleen) were also extracted and weighed using a precision scale to the nearest thousandth of a gram (Mettler Toledo ME 802) to obtain additional outcome measures to assess treatment-dependent variability in the data. Organ weights were corrected by body weight.

### Brain Function: Recurrent Perseveration in a Two-Choice Guessing Task

The expression of stereotypic behavior has been found to correlate with recurrent perseveration, a form of impaired inhibitory control of behavior, both in humans (autistic children)

and in captive mammals and birds (Turner, 1997; Garner et al., 2003, 2011; Gross et al., 2012). A positive correlation between stereotypy levels and recurrent perseveration has also been observed in mice, notably in C57 mice (Garner et al., 2011), but other studies have yielded mixed results (Latham and Mason, 2010; Gross et al., 2011a, 2012; Novak et al., 2016).

To assess recurrent perseveration, we used a slightly modified apparatus and procedure to the one described previously (Novak et al., 2016; Bailoo et al., 2018). Briefly, all focal mice ( $n = 48$ ) were trained and tested under red light in three virtually identical apparatuses. Each apparatus consisted of a rectangular arena with two goal-holes at one end and a trapezoid-shaped start-box at the opposite end (see **Supplementary Figure S4**). Animals were first habituated to the apparatus across 3 days and then shaped to retrieve rewards (BioServ chocolate pellets, 20 mg) from both goal holes. Animals were then tested across 2 days, 60 trials/day, where one of the two goals was rewarded with a probability equaling the proportion of responses to the other side in the previous 20 trials (see Novak et al., 2016; see Bailoo et al., 2018 and **Supplementary Figure S5** for further details). Perseveration score (logit [P]), as well as the frequencies of pure repetitions (LLLL, RRRR) and pure alternations (RLRL, LRLR) relative to all possible tetragrams of consecutive choices ( $n = 16$ ) were calculated for each individual and analyzed.

### Emotional State: Judgment Bias in a Spatial Go/No-Go Task

One approach to evaluate the valence (i.e., the positivity or negativity) of emotions in animals is to investigate how decisions in ambiguous situations are biased by the underlying emotional states of the animals—as assessed by cognitive judgment bias tasks (Harding et al., 2004). Previous research evaluating judgment biases in rats indicates that the transfer from standard housing conditions to enriched environments is associated with a relative shift from “pessimistic” to “optimistic” judgments based on the expectancy of non-reward and reward, respectively (e.g., Brydges et al., 2011; Richter et al., 2012). However, no study has investigated the effect of environmental enrichment on judgment biases of mice.

To assess judgment biases, we developed a task which integrates active trial initiation into a spatial Go/No Go task (see Hintze et al., 2018 for further details). Briefly, all focal mice ( $n = 48$ ), were trained and tested under red light in three virtually identical apparatuses (see **Supplementary Figure S6**). Each apparatus consisted of a rectangular arena with a row of nine goal-holes at one end (five of which were used for this task), and a trapezoid-shaped area containing a nose-poke at the opposite end. Mice were first trained to initiate each trial by nose-poking, and in several subsequent steps, were trained to discriminate that the location of an open goal-hole (at the extreme ends) signaled either reward (Go trial) or non-reward (No-go trial). Once animals had learned this discrimination, ambiguous test trials with open goal-holes at three equidistant intermediate locations between the extreme ends, were interspersed among Go and No-Go trials, across six

test sessions. The Go: No-go response ratio to all five goal-holes (positive, three ambiguous, negative) was used as a measure of judgment bias.

### Attrition

In batch 1, one C57 mouse housed in the B condition was euthanized prior to the end of the experiment because the animal was favoring the right side of its head (lopsided tilt). Behavioral symptoms indicated mild distress, although post-mortem pathology yielded no diagnosis. In batch 2, one C57 mouse allocated to E condition was found dead at delivery.

### Ethical Statement

This study was carried out in accordance with the guidelines of the Swiss Animal Welfare Ordinance (TSchV 455.1). It was approved by the Cantonal Veterinary Office in Bern, Switzerland (permit number: BE16/16).

### Statistical Analyses

All statistical analyses were performed using IBM<sup>TM</sup> SPSS Statistics (version 23), except for the judgment bias task data, which were analyzed using R (version 3.3.2). For parametric models run in SPSS, assumptions of normally distributed errors and homogeneity of variance were examined graphically and, based on these inspections, no transformations of data were needed. Strains of mice were analyzed separately. In SPSS, batch was included in all analyses as a fixed effect—no significant differences were observed and this predictor is not discussed further. *P*-values less than 0.05 were considered statistically significant for all analyses and are presented as actual values rounded to three decimal places. Raw data for all outcome measures will be made available upon request.

Home-cage and stereotypic behavior was analyzed using Kruskal-Wallis independent sample tests due to the high degree of skewness observed in the data. Data was summed across all observation intervals and expressed as proportions. Significant effects were probed using Kruskal-Wallis pairwise comparisons with a Dunn-Bonferroni correction.

Measures of anxiety—elevated plus-maze and open field behavior—were analyzed using the MIXED procedure. Housing condition and, for the open field only, day of testing, were treated as categorical fixed effects. Mouse nested in cage was added as a random effect. Significant effects were probed with Bonferroni corrected pairwise comparisons.

Food and water intake and body weight were analyzed using the MIXED procedure, with housing condition treated as a categorical fixed effect, and week as a continuous covariate. For body weight, mouse nested in cage was added as a random effect. Significant effects were probed with Bonferroni corrected linear contrasts.

One of our measures of stress physiology—fecal glucocorticoid metabolite concentrations—was analyzed using the MIXED procedure. Housing condition and time point of assessment were treated as categorical fixed effects. Mouse nested in cage was added as a random effect. Significant effects were probed with Bonferroni corrected pairwise comparisons. Adrenal mass between housing conditions was analyzed using

Kruskal-Wallis independent sample tests due to the high degree of skewness observed in the data. Significant effects were probed using Kruskal-Wallis pairwise comparison with a Dunn-Bonferroni correction.

Measures of brain function—recurrent perseveration and patterned responding in a two-choice guessing task—were analyzed using the MIXED procedure. Housing condition was treated as a categorical fixed effect. Significant effects were probed with Bonferroni corrected pairwise comparisons.

Our measure of emotional valence—responding in a judgment bias task—was analyzed using the function `glmer` of the package `lme4` (“family:” binomial, including the “logit” link function). Housing condition, trial type, and their two-way interaction were fixed effects, while trial type nested in session per test day nested in test day, nested in animal ID nested in batch was used as a random effect.

To assess variation between housing conditions, the coefficient of variation (CV), the ratio of the standard deviation to the mean, for each outcome was calculated from the raw data with the exception of home-cage and stereotypic behavior, guessing task and judgment bias.

## RESULTS

### Engagement With the Housing Environment

Engagement with the housing environment was defined as the mouse being either in contact with an enrichment item (inside, on, or under) or actively manipulating an enrichment item (N, E, SE) or bedding (all groups). The degree of cage enrichment strongly affected engagement with the housing environment in both strains (C57:  $H = 116.20$ ,  $p < 0.001$ ; SWISS:  $H = 114.80$ ,  $p < 0.001$ , **Figure 2**). *Post hoc* comparisons indicated that in both strains of mice, animals housed in the barren condition were less often observed to be engaged with the housing environment compared to all other groups. Additionally, in both strains, animals housed in the super enriched condition were more often observed engaged with the housing environment compared to animals housed in the nesting condition, and in SWISS mice only, compared to the enriched condition. In the SE housing condition, the mice were engaged with enrichments on average more than 85% of the observed time.

### Stereotypic Behavior

The degree of cage enrichment had a strong effect on the expression of stereotypic behavior in the home-cage in both strains (C57:  $H = 34.21$ ,  $p < 0.001$ ; SWISS:  $H = 21.77$ ,  $p < 0.001$ ; **Figure 3**). This was mainly due to greatly reduced stereotypy levels in mice in SE cages compared to mice from all other treatment groups, while there were no consistent differences between mice in B, N and E cages.

The degree of cage enrichment also had effects on specific forms of stereotypies, but these varied with the strain of mouse. The expression of bar-mouthing was affected by cage enrichment in both strains (C57:  $H = 21.82$ ,  $p < 0.001$ ; SWISS:  $H = 16.48$ ,  $p < 0.001$ ; **Supplementary Figure S7**). However, circling on the

cage lid was affected by cage enrichment in C57 mice only (C57: *circling*:  $H = 20.05$ ,  $p < 0.001$ ), while back-flipping, route-tracing on the cage lid and twirling were affected by cage enrichment in SWISS mice only (SWISS: *back-flipping*:  $H = 20.68$ ;  $p < 0.001$ ; *twirling*:  $H = 13.81$ ;  $p < 0.001$ ; *route-tracing on the lid*:  $H = 11.00$ ;  $p = 0.012$ ; **Supplementary Figure S7**). These effects were mostly due to higher levels of stereotypies observed in B and N cages compared to E and SE cages, although bar-mouthing in C57 mice was higher in E cages compared to all other cages. However, there were large individual differences in both the form and levels of stereotypy performance, and many 0 values for most forms of stereotypies, which precluded further analysis of treatment effects on specific forms of stereotypies.

### Other Home-Cage Behavior

In both strains, variation in home-cage behavior depending on the degree of cage enrichment was mostly determined by the extent of unseen behavior (C57; *active*:  $H = 43.28$ ,  $p < 0.001$ ; *inactive*:  $H = 53.05$ ;  $p = 0.001$ ; *unseen*:  $H = 98.14$ ,  $p = 0.001$ ; SWISS; *active*:  $H = 30.48$ ,  $p < 0.001$ ; *inactive*:  $H = 75.42$ ;  $p < 0.001$ ; *unseen*:  $H = 94.78$ ,  $p < 0.001$ , **Figure 4**). Behavior was recorded as unseen when enrichments obscured the view of the mice such that the specific pattern of behavior could not be identified unambiguously. That both the relative amount of seen active and inactive behavior decreased with increasing enrichment indicates that engagement with enrichment affected both active and inactive components of behavior. That seen inactive behavior decreased more than seen active behavior further indicates that mice have a stronger preference for shelter when inactive.

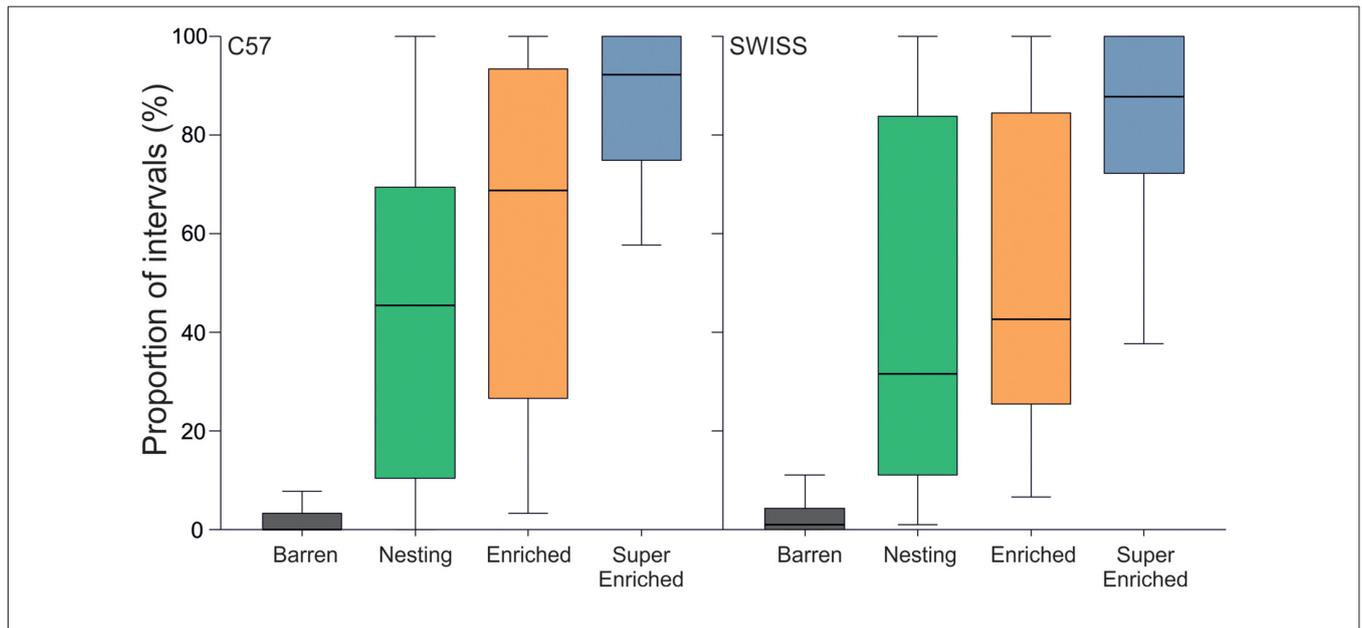
Comparing variation among specific patterns of seen active behavior indicated that engagement with enrichments affected specific patterns of active behavior differently. For example, the proportion of grooming decreased with the degree of enrichment across all treatment groups, while the proportion of active behavior on the lid was greatly reduced, and the proportion of active behavior on the floor increased, in SE mice compared to mice from all other groups (**Figure 5**).

### Anxiety in the Elevated Plus-Maze Test and Open Field Test

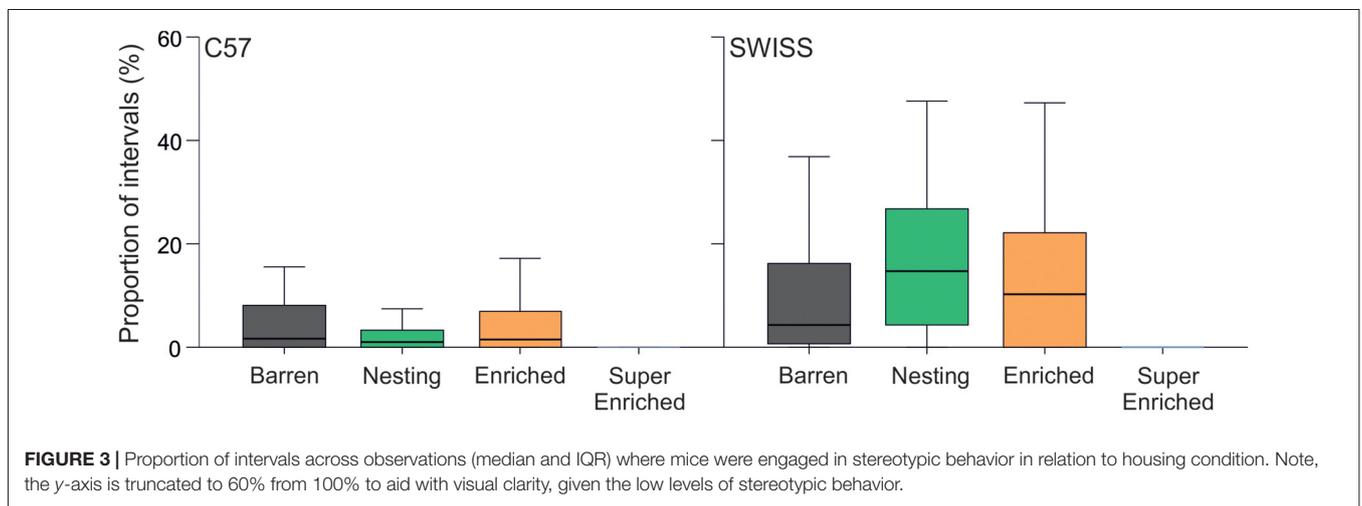
Behavior in the elevated plus-maze test varied with the type of housing condition in C57 but not SWISS mice. However, this effect was only observed with respect to the total distance traveled but not time-in-open arms (*distance traveled*: C57:  $F_{(3,28)} = 6.53$ ,  $p = 0.002$ ; SWISS:  $F_{(3,28)} = 0.09$ ,  $p = 0.967$ ; *time-in-open arms* C57:  $F_{(3,28)} = 0.95$ ,  $p = 0.080$ ; SWISS:  $F_{(3,28)} = 2.33$ ,  $p = 0.420$ ; **Figure 6**). Closer inspection of the data indicated that C57 mice housed in E and SE cages traveled longer distances than those housed in B and N cages, while there were no consistent effects with respect to time-in-open arms.

A secondary *post hoc* analysis evaluating time spent in the center of the maze yielded no differences between our treatment groups (C57:  $F_{3,91} = 2.29$ ,  $p = 0.083$ ; SWISS:  $F_{3,92} = 0.46$ ,  $p = 0.709$ ).

Behavior in the open field varied with the type of housing condition in C57 but not SWISS mice. This effect was observed



**FIGURE 2** | Proportion of intervals across observations (median and IQR) where mice were engaged with the housing environment.



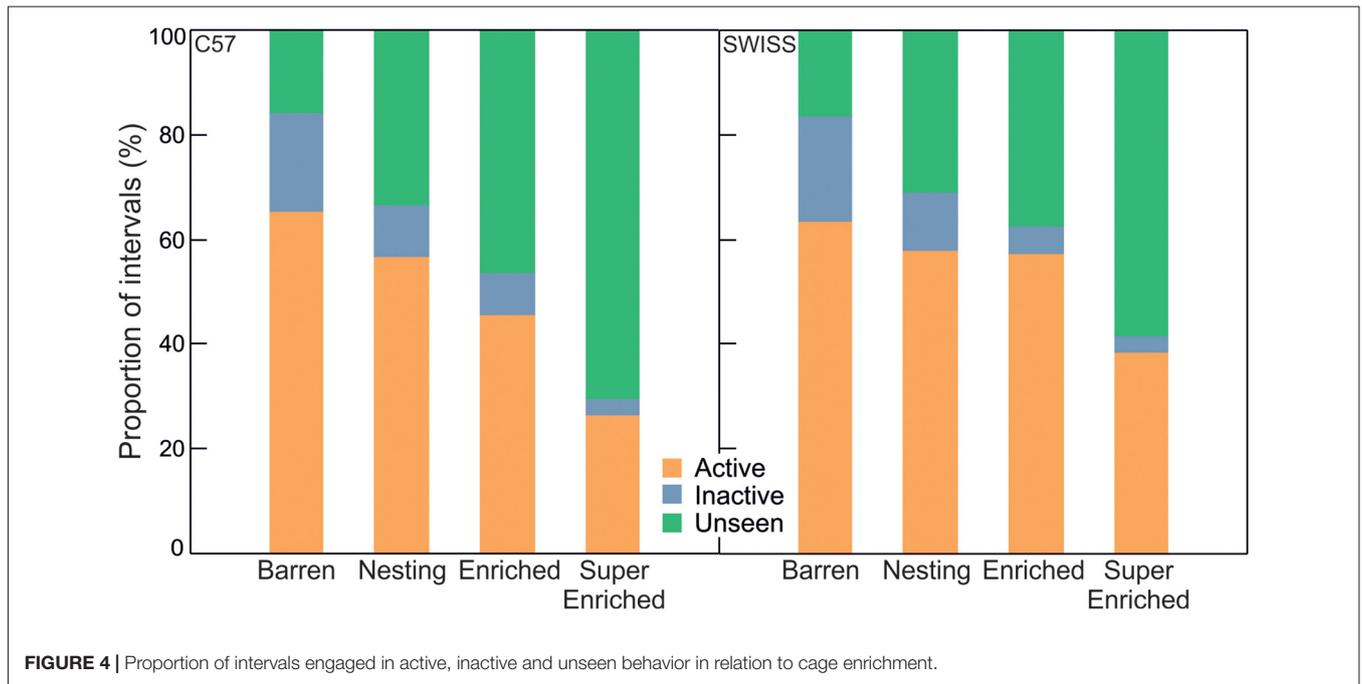
**FIGURE 3** | Proportion of intervals across observations (median and IQR) where mice were engaged in stereotypic behavior in relation to housing condition. Note, the y-axis is truncated to 60% from 100% to aid with visual clarity, given the low levels of stereotypic behavior.

with respect to both total distance traveled and time in corners, but not time in the center (*distance traveled* C57:  $F_{(3,90)} = 8.18, p < 0.001$ ; SWISS:  $F_{(3,90)} = 2.106, p = 0.105$ ; *time-in-center* C57:  $F_{(3,89)} = 0.11, p = 0.234$ ; SWISS:  $F_{(3,88)} = 1.45, p = 0.957$ ; *time-in-corners* C57:  $F_{(3,89)} = 3.43, p = 0.021$ ; SWISS:  $F_{(3,89)} = 2.25, p = 0.088$ ; **Figure 7**). In both strains, behavior in the open field also varied across days of testing (*distance traveled* C57:  $F_{(3,144)} = 100.32, p < 0.001$ ; SWISS:  $F_{(3,171)} = 9.44, p < 0.001$ ; *time-in-center* C57:  $F_{(3,169)} = 28.65, p < 0.001$ ; SWISS:  $F_{(3,173)} = 7.84, p < 0.001$ ; *time-in-corners* C57:  $F_{(3,173)} = 68.04, p < 0.001$ ; SWISS:  $F_{(3,174)} = 21.95, p < 0.001$ ; **Figure 7**). However, there was no interaction between the type of housing condition and day of testing (*distance traveled* C57:  $F_{(9,144)} = 1.64, p = 0.109$ ; SWISS:  $F_{(9,171)} = 1.84, p = 0.064$ ; *time-in-center* C57:  $F_{(9,169)} = 0.63, p = 0.773$ ; SWISS:

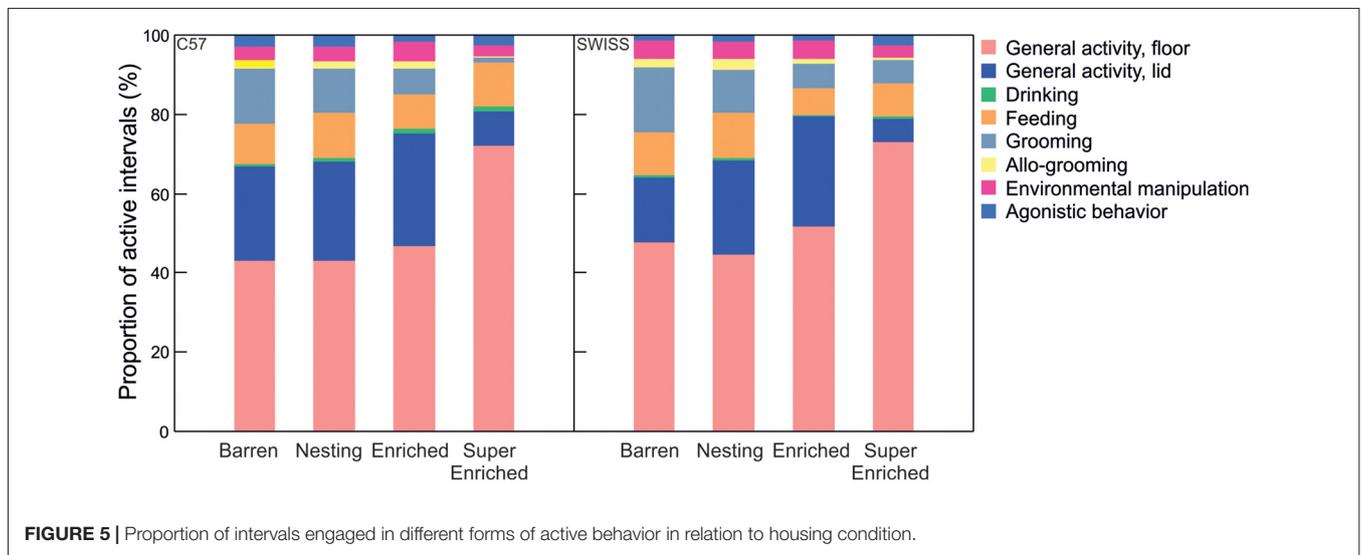
$F_{(9,173)} = 0.72, p = 0.687$ ; *time-in-corners* C57:  $F_{(9,173)} = 1.62, p = 0.114$ ; SWISS:  $F_{(9,174)} = 1.59, p = 0.121$ ; **Figure 7**). Closer inspection of the data indicated that both distance traveled and time in the center decreased while time in the corner increased across the 4 days of testing. Furthermore, C57 mice housed in SE cages traveled shorter distances compared to all other groups.

### Growth: Body Weight and Food and Water Intake Per Mouse

In both strains and across all housing conditions, food intake was positively correlated with both water intake and body weight (averaged at the cage level)—except for C57 mice housed in E and SE groups (**Supplementary Table S6**).



**FIGURE 4 |** Proportion of intervals engaged in active, inactive and unseen behavior in relation to cage enrichment.



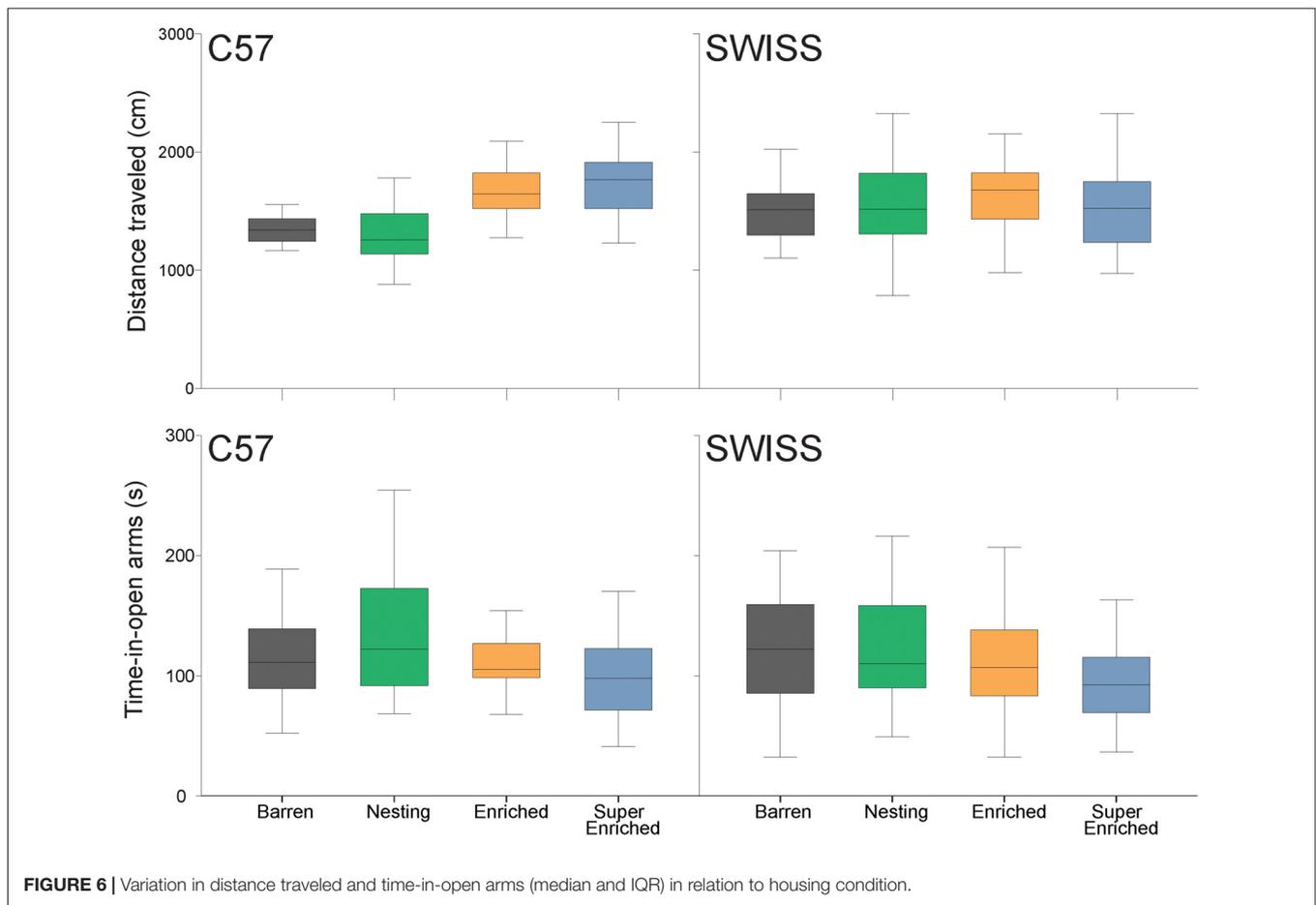
**FIGURE 5 |** Proportion of intervals engaged in different forms of active behavior in relation to housing condition.

In C57 mice only, water intake was positively correlated with body weight (averaged at the cage level) across all housing conditions—a similar relationship was observed only in SWISS mice housed in the SE condition (**Supplementary Table S6**).

Body weight at arrival did not vary between housing conditions in both strains (C57:  $F_{(3,87)} = 0.79, p = 0.504$ ; SWISS:  $F_{(3,88)} = 1.55, p = 0.208$ ). Body weight increased with age in both strains (C57:  $F_{(8,333)} = 725.88, p < 0.001$ ; SWISS:  $F_{(8,348)} = 495.52, p < 0.001$ ; **Figure 8**). Furthermore, body weight varied with the type of housing condition in C57, but not SWISS mice (C57:  $F_{(3,99)} = 13.04, p < 0.001$ ; SWISS:

$F_{(3,88)} = 0.70, p = 0.543$ ), and there was an interaction between the type of housing condition and week in C57, but not SWISS mice (C57:  $F_{(24,333)} = 3.77, p < 0.000$ ; SWISS:  $F_{(24,348)} = 1.51, p = 0.060$ ; **Figure 8**). *Post hoc* comparisons indicated that in C57 mice only, mice housed in the SE condition were heavier than mice of all other housing conditions. This difference emerged 1 week after arrival and persisted for the duration of the experiment.

In both strains, there was no main effect of the type of housing condition on food intake (C57:  $F_{(3,28)} = 1.19, p = 0.332$ ; SWISS:  $F_{(3,28)} = 2.88, p = 0.054$ ), but food intake varied by week (C57:  $F_{(7,66)} = 11.16, p < 0.001$ ; SWISS:  $F_{(7,107)} = 9.27,$



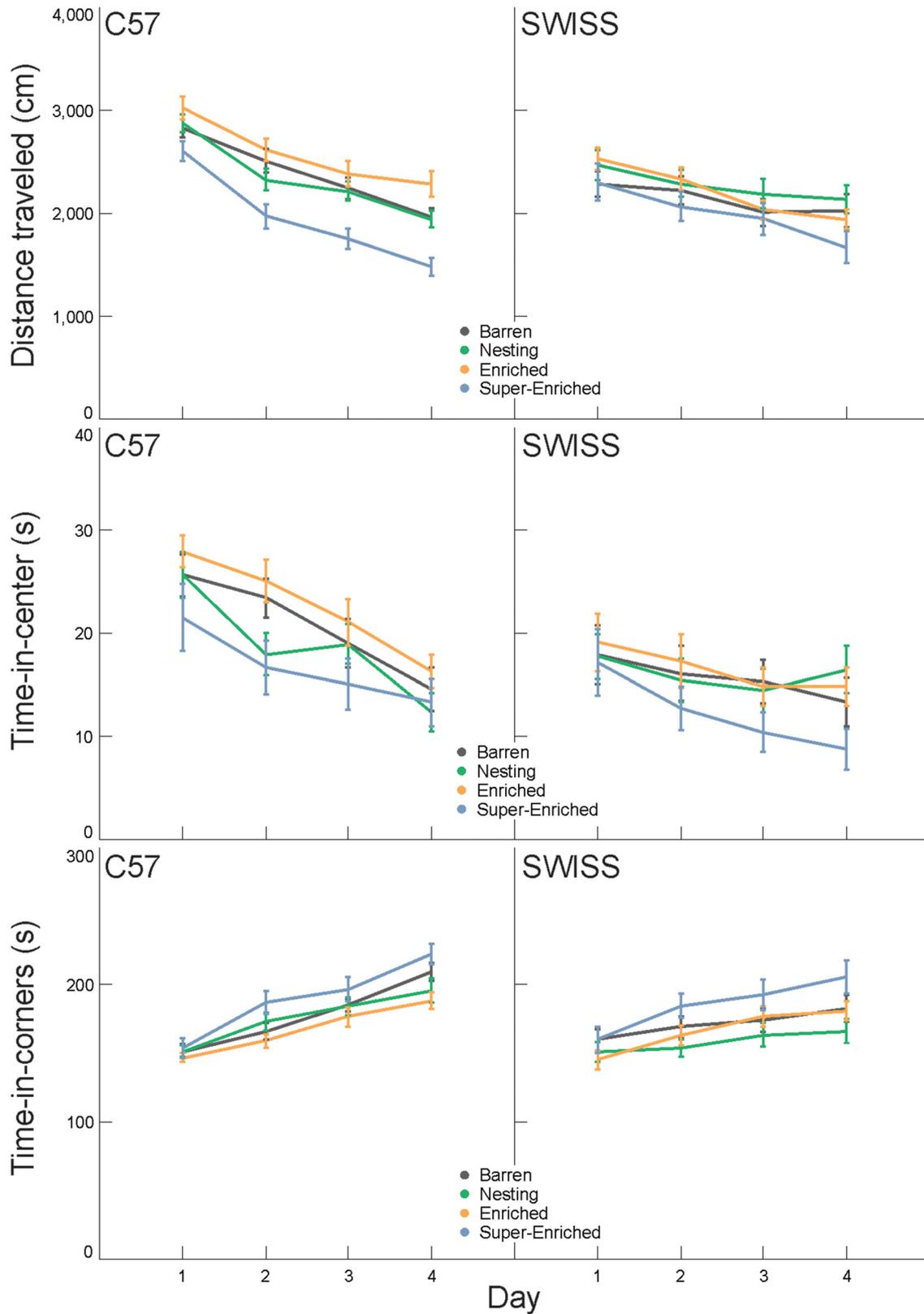
$p < 0.001$ ). Furthermore, there was an interaction between the type of housing condition and week in both strains (C57:  $F_{(21,66)} = 3.76$ ,  $p < 0.001$ ; SWISS:  $F_{(21,107)} = 4.27$ ,  $p < 0.001$ ; **Figure 9**). Closer inspection of the data indicated that food intake initially increased in both strains and all treatment groups, but flattened or decreased with age, whereby these changes in the time course of food intake differed between treatment groups. Thus, there was a relative decrease in food intake with increasing degree of enrichment at later time points.

Water intake increased with age (C57:  $F_{(7,118)} = 5.2$ ,  $p < 0.001$ ; SWISS:  $F_{(7,120)} = 8.66$ ,  $p < 0.001$ ) and there was a main effect of the type of housing condition on water intake (C57:  $F_{(3,28)} = 15.37$ ,  $p < 0.001$ ; SWISS:  $F_{(3,28)} = 3.01$ ,  $p = 0.047$ ). Furthermore, in SWISS, but not C57 mice, water intake also varied depending on the interaction between the type of housing condition and week (C57:  $F_{(21,118)} = 0.880$ ,  $p = 0.606$ ; SWISS:  $F_{(21,120)} = 2.14$ ,  $p = 0.006$ ; **Figure 9**). *Post hoc* comparisons indicated, in both strains of mice, that average water intake was higher in the SE cages. In C57 mice, water intake was higher in the SE group from week 3 compared to all other groups. In SWISS mice, water intake was higher in the SE group on week 6 compared to B and E groups.

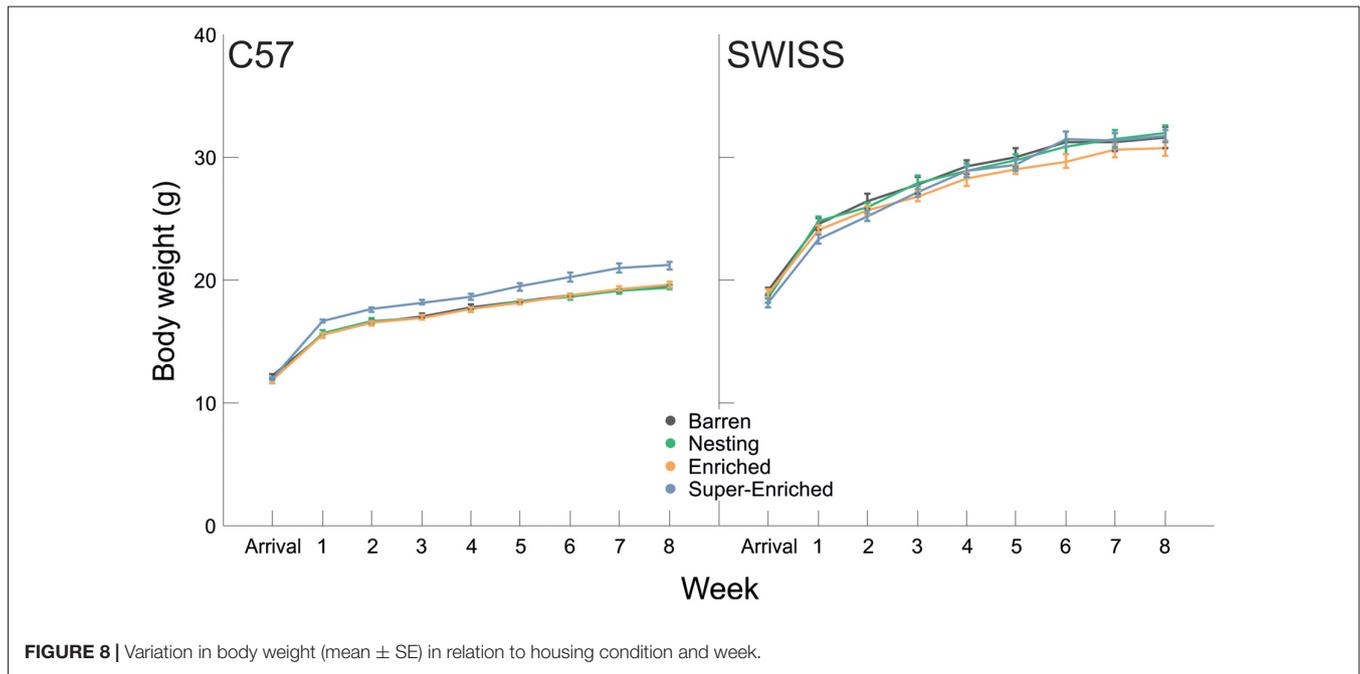
## Endocrine Stress Responses: Fecal Glucocorticoid Metabolites and Adrenal Weight

Glucocorticoid metabolite concentration increased across the three time points in SWISS mice, but not in C57 mice (C57:  $F_{(2,125)} = 0.52$ ,  $p = 0.596$ ; SWISS:  $F_{(2,142)} = 20.67$ ,  $p < 0.000$ ; **Figure 10**). Furthermore, there was an effect of the type of housing condition in SWISS, but not C57 mice (C57:  $F_{(3,91)} = 0.18$ ,  $p = 0.910$ ; SWISS:  $F_{(3,89)} = 4.38$ ,  $p = 0.006$ ). *Post hoc* comparisons indicated that glucocorticoid metabolite concentration was, on average, higher in mice housed in B and N cages compared to mice housed in E and SE cages across all three time-points; but only in SWISS mice. There was also an interaction between type of housing condition and time point in C57 mice, but not SWISS mice (C57:  $F_{(6,126)} = 2.38$ ,  $p = 0.032$ ; SWISS:  $F_{(6,142)} = 0.64$ ,  $p = 0.687$ ), but *post hoc* analyses yielded no consistent effects.

Variation in adrenal weight partly reflected the effects found in glucocorticoid metabolite concentrations. Thus, the type of housing condition affected adrenal weight in SWISS mice, but not C57 mice (C57:  $H = 2.95$ ,  $p = 0.400$ ; SWISS:  $H = 13.12$ ,  $p = 0.004$ ; **Figure 11**). However, *post hoc* comparisons indicated a difference in adrenal weight only between mice from N cages compared to mice from E cages.



**FIGURE 7** | Variation in distance traveled, time-in-center and time-in-corners (mean ± SE) in relation to housing condition and day of testing.



### Brain Function: Recurrent Perseveration in a Two-Choice Guessing Task

Recurrent perseveration was not affected by housing condition, stereotypy level, or by the interaction between the type of housing condition and stereotypy level (C57; *housing condition*:  $F_{(3,16)} = 0.57$ ,  $p = 0.645$ ; *stereotypy level*:  $F_{(1,16)} = 0.38$ ,  $p = 0.850$ ; *housing condition*  $\times$  *stereotypy level*:  $F_{(3,16)} = 0.43$ ,  $p = 0.732$ ; SWISS; *housing condition*:  $F_{(3,16)} = 1.09$ ,  $p = 0.384$ ; *stereotypy level*:  $F_{(1,16)} = 0.38$ ,  $p = 0.549$ ; *housing condition*  $\times$  *stereotypy level*:  $F_{(3,16)} = 0.71$ ,  $p = 0.558$ ; **Figure 12**).

Response patterns in the guessing task varied with the type of housing condition in C57 mice, but not SWISS mice (C57;  $F_{(9,80)} = 4.31$ ,  $p < 0.001$ ; SWISS;  $F_{(3,16)} = 1.22$ ,  $p = 0.295$ ; **Figure 13**). Closer inspection of the data indicated that C57 mice housed in N cages made more pure repetitions compared to all other groups, albeit to one side only (LLLL).

### Emotional State: Judgment Bias in a Spatial Go/No-Go Task

All C57 mice reached the learning criterion for testing, whereas two SWISS mice were excluded during the Go/No-go Discrimination stage (see Hintze et al., 2018 for further methodological details). Total training duration for all stages (Habituation, Shaping for Trial Initiation, Left-Right Discrimination, Go/No-go Discrimination) was  $14.20 \pm 1.63$  sessions for C57 mice and  $15.50 \pm 1.79$  sessions for those SWISS mice that reached the test criterion.

During testing, the animals' decision as to whether or not to go when confronted with the different trial types varied as an interaction between the type of housing condition and trial type for both C57 and SWISS mice (C57:  $\chi^2_3 = 13.46$ ,  $p = 0.004$ ; SWISS:  $\chi^2_3 = 22.97$ ,  $p < 0.001$ ; **Figure 14**).

Closer inspection of the data revealed no consistent differences in response patterns depending on the type of housing condition. In both strains, the strongest treatment effect was observed in near-negative (NN) trials, as indicated by fewer go responses in mice from B cages in C57 mice, and in mice from SE cages in SWISS mice, compared to mice from all other treatment groups. However, NN trials were also associated with the largest variability in the number of go responses.

### Variability in the Measured Outcome Variables

To assess potential effects of the type of housing condition on the variability of experimental results, we calculated the CV for all measured outcome variables, with the exception of home-cage and stereotypic behavior, and for judgment bias and guessing task data where only a subset of animals were tested ( $n = 6$  per strain per housing condition). The CV is the ratio of the standard deviation to the mean, which yields a dimensionless, standardized measure of dispersion. It thus allowed us to compare variation estimates directly between the different outcomes measured in this experiment.

CVs varied greatly depending on the measured outcome variable (**Figure 15, Supplementary Table S7**). Most CVs were relatively small and there were no consistent relationships between cage enrichment and CV across the range of outcome variables assessed in this study.

## DISCUSSION

To study the effects of environmental enrichment on measures of animal welfare in mice, we systematically varied housing conditions across four levels of enrichment and measured animal welfare in a multi-faceted way. Overall, the greatest benefit

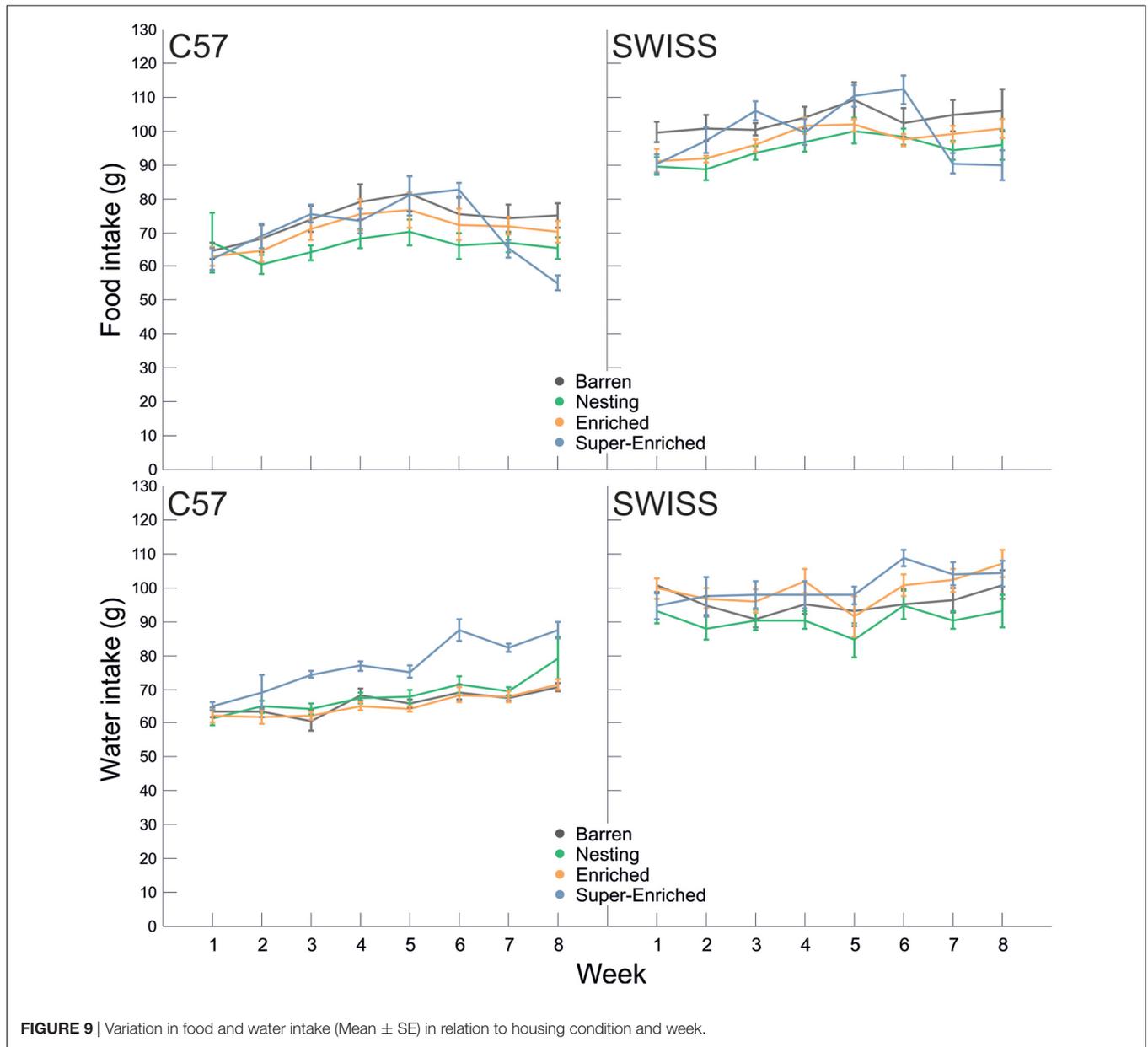
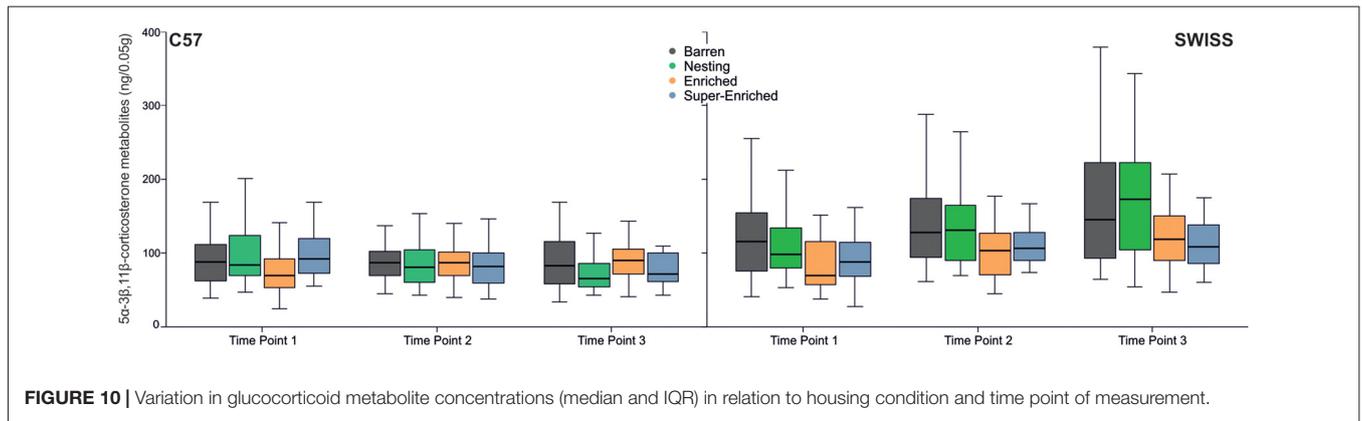


FIGURE 9 | Variation in food and water intake (Mean ± SE) in relation to housing condition and week.

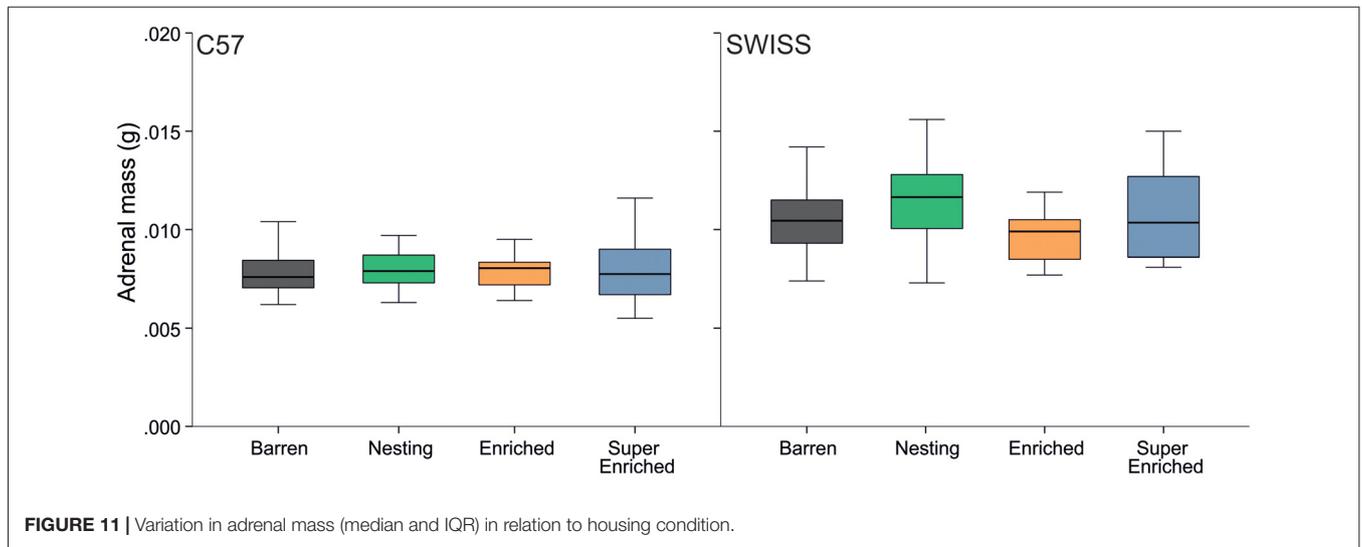
to welfare was observed in animals housed with the most extensive enrichment, the SE housing condition—stereotypic behavior, some behavioral measures of anxiety, growth and stress physiology, all varied in a manner consistent with improved animal welfare. We also assessed the effect of the different housing conditions on variation in outcome measures. Similarly to other studies (Augustsson et al., 2003; Wolfer et al., 2004; Baumans et al., 2010; van de Weerd et al., 2010), we found no evidence that environmental enrichment increased variation in experimental results, for the outcome measures assessed here.

When evaluating whether and how animals engaged with the enrichment items and evaluating the relation to welfare, substantial differences were noted—animals were more often observed to be in, on or under items, and manipulated items

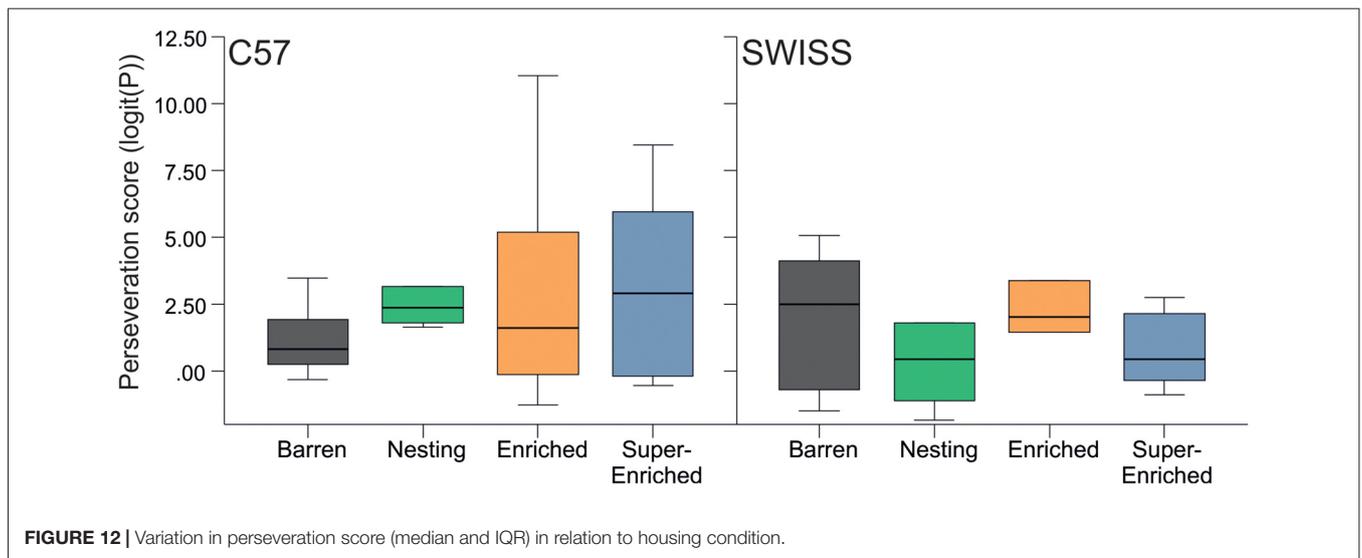
as the degree of enrichment increased. One possible explanation for this difference may be that there is a higher probability for contact with different items, simply by chance, as the number of items increased and the “free” floor area decreased. If this explanation were true, we would predict that despite higher levels of engagement with the housing environment, animal welfare would not be improved. This explanation may apply to N and E groups—even though engagement with enrichment increased with the degree of enrichment, no consistent improvements in measures of welfare were observed in comparison to the B housing condition. For example, the expression of stereotypic behavior did not vary in a systematic way between B, N and E housing groups even though engagement with enrichments increased with increasing degrees of enrichment in these groups.



**FIGURE 10** | Variation in glucocorticoid metabolite concentrations (median and IQR) in relation to housing condition and time point of measurement.



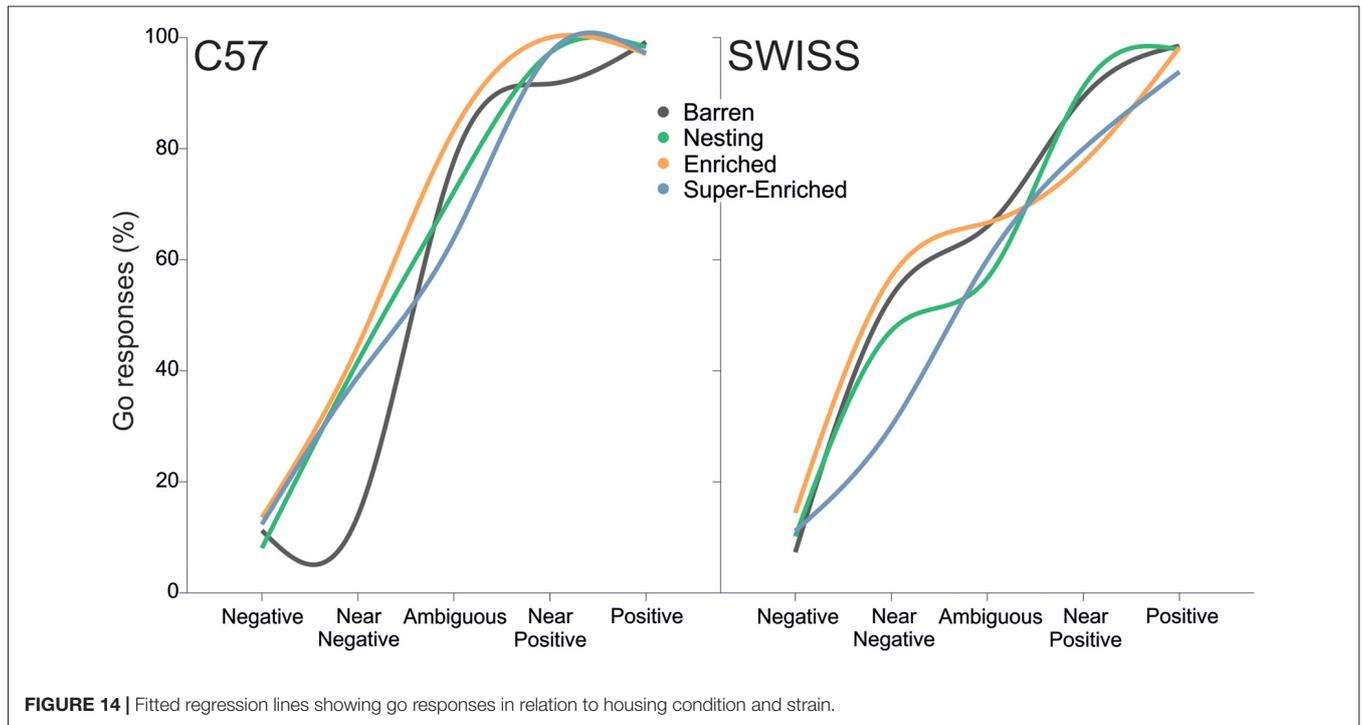
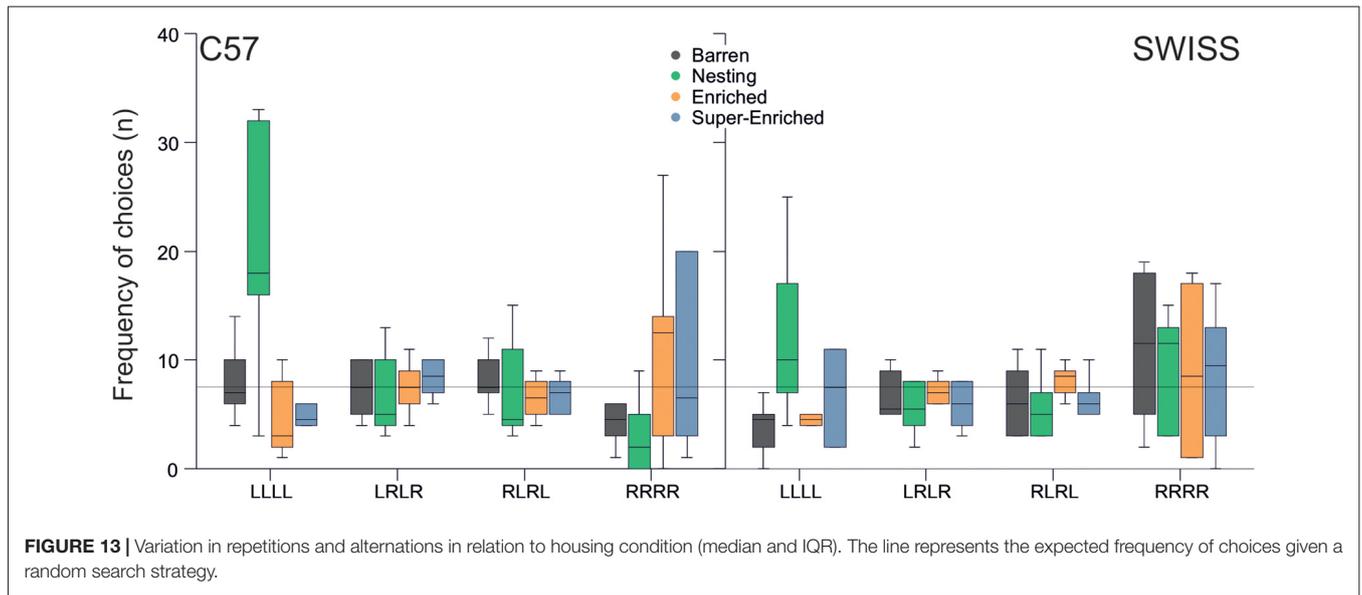
**FIGURE 11** | Variation in adrenal mass (median and IQR) in relation to housing condition.



**FIGURE 12** | Variation in perseveration score (median and IQR) in relation to housing condition.

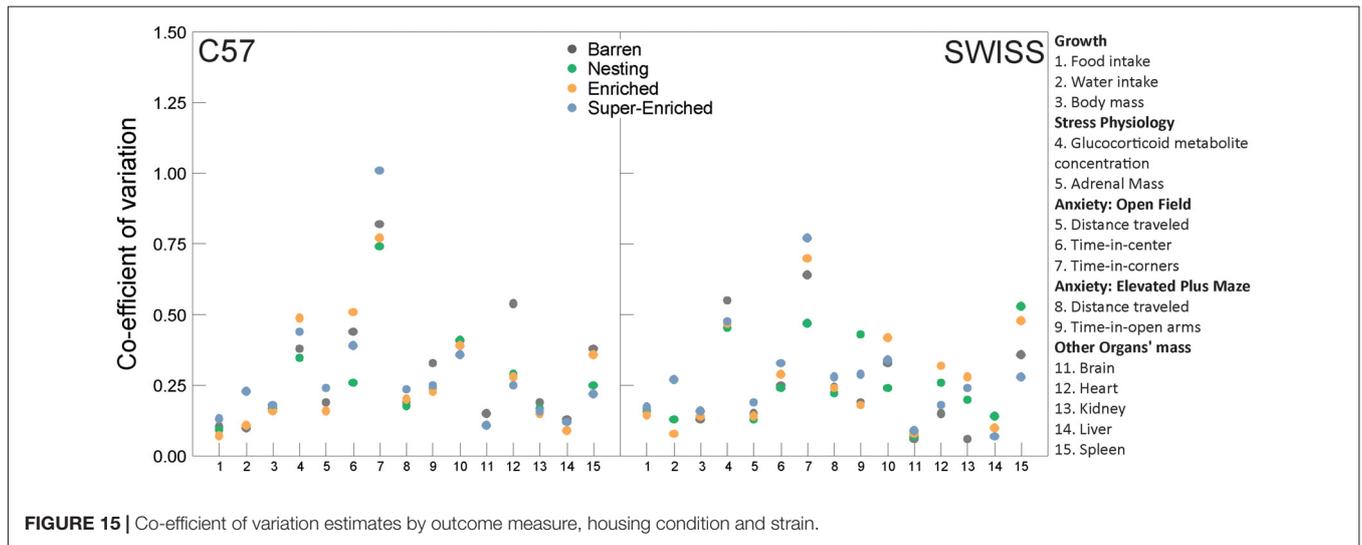
However, this does not apply to the SE housing condition. Although the SE condition was the most enriched, it also

offered the most “free” floor and lid space. Furthermore, in the SE housing condition, virtually no stereotypic behavior was



observed—suggesting improved welfare in this group (Mason and Latham, 2004). It is possible that the prevalence of stereotypic behavior was underestimated in the SE group, given that the highest levels of unseen behavior were also observed in this group. We find this explanation unlikely given that bar-mouthing, the most commonly reported form of stereotypy in laboratory mice (Würbel, 2006), was observable and recorded since the bars of the cage were not covered by enrichment items. Animals in the SE group displayed extremely low levels of bar-mouthing (on average <1%). It is similarly unlikely that

mice were performing any of the other forms of stereotypy categorized here when unseen, given the amount of space that is necessary for the performance of these behaviors; most of the unseen behavior was recorded when mice were under or in enrichment items. Therefore, this difference in stereotypy performance, with respect to housing condition, likely reflects the true incidence of behavioral expression. Underestimation of stereotypic behavior, when animals were unseen, in B, N and E cages is also unlikely given that unseen behavior in these cages was primarily scored when animals were either in the nest or in a



**FIGURE 15 |** Co-efficient of variation estimates by outcome measure, housing condition and strain.

huddle and where one mouse visibly blocked another. Neither of these two scenarios is likely to have masked the display of stereotypic behavior.

When assessing differences in time budget, some differences were observed—for example, the proportion of grooming and of active behavior on the lid decreased with the degree of enrichment across all treatment groups. In contrast to stereotypic behavior, however, other components of the animals’ time budget categorized here could occur when an animal was in or under enrichment items. For example, we frequently observed nesting material shaking in a rhythmic pattern, indicative of grooming or allo-grooming behavior; this behavior was coded as unseen, given that the animals were not visible. Therefore, it is quite likely that our time–budget assessment underestimates the prevalence of specific behaviors and these data should therefore be interpreted with caution. On the other hand, as our data indicated that grooming was often performed out of sight, it may also suggest that mice prefer to perform grooming in a shelter, if a shelter is available.

Differences in engagement with enrichment items were also associated with improvements in animal welfare, albeit in a strain dependent way. C57 mice housed with a greater degree of enrichment, i.e., E and SE groups, were more exploratory, and therefore less anxious, in the elevated plus–maze. In the open field, C57 mice housed in the SE groups displayed a classic habituation response to the novelty of the arena across days of testing—on average lower levels of exploration and more time spent inactive in the corner of the arena. Thus, in general, C57 mice housed in SE cages displayed an anxiolytic profile in comparison to the other housing conditions.

In terms of our secondary outcome measures, some concordant differences were observed. For example, when comparing differences in stress physiology, a consistent difference was observed, but only in SWISS mice—animals housed in E and SE groups had lower levels of glucocorticoid metabolites in feces across all three time points of measurement

compared to B and N groups. This result suggested that SWISS mice housed in E and SE conditions experienced lower levels of chronic stress. Our other measure of stress physiology, adrenal weight, varied differently with housing condition—SWISS mice housed in E groups had smaller adrenals compared to those housed in the N groups. Both adrenal weight and glucocorticoid metabolites in feces have been used previously as indicators of experienced chronic stress in mice (Tsai et al., 2002; Akre et al., 2011; Gurfein et al., 2014; Bailoo et al., 2018), albeit with different degrees of sensitivity relative to housing conditions. For example, one study has found that animals housed in enriched environments tended to have smaller adrenals; although this difference was not statistically significant (Tsai et al., 2002). In contrast, studies that measured corticosterone secretion in feces have consistently found that enriched housing conditions are associated with decreased levels of circulating glucocorticoids (Akre et al., 2011; Gurfein et al., 2014). Therefore, the discrepancy in the pattern of differences between fecal glucocorticoids and adrenal weight with respect to housing condition may simply reflect differences in measurement sensitivity.

Across time, food intake initially did not vary by housing condition—it did however vary by time. Food intake initially increased, peaking at around 7–8 weeks of age (puberty/early adulthood) and then either flattened or decreased with age. In contrast, water intake increased across time in the SE condition in both strains, but did not vary between the other housing conditions. In C57 mice only, animals housed in the SE condition weighed more than animals in all other groups. Notably, food intake and body weight (averaged at the cage level) were uncorrelated in C57 mice housed in E and SE groups. Taken together, these patterns of differences with respect to growth, most likely reflect variation in metabolic need. Importantly, as ambient temperatures in mouse facilities are kept below the thermoneutral zone of mice, increasing the risk of cold stress (Gordon, 1985, 1993; Johnson et al., 2017), increased opportunity for structural complexity, and in particular for the construction

of elaborate nests, may act as a buffer to cold stress—resulting in heavier animals. However, we were unable to examine this hypothesis further, as we did not measure temperatures inside the cages (micro-climate). Food and water intake are tightly linked to metabolism and are, in general, positively correlated (Gordon, 1993; Gordon et al., 1998). However, in relative terms, water intake in mice is an inelastic requirement—in the absence of water, mice quickly dehydrate (Harkness et al., 2013). Increased water intake in the SE housing condition is therefore most likely reflective of a difference in metabolic need prompted by increased activity and possibly reflecting increased engagement with the environment.

Animals housed in the different housing conditions did not vary in our measures of brain function, recurrent perseveration and patterned responding in a two-choice guessing task; at least not in the predicted way. Specifically, animals housed in barren environments have been reported to display a higher incidence of stereotypic behavior and, in turn, higher levels of perseverative behavior and patterned responding (Garner et al., 2011). In this experiment, no association was found between housing condition, stereotypy level and recurrent perseveration. This lack of association is, however, in line with more recent studies using this experimental paradigm (Latham and Mason, 2010; Gross et al., 2011a, 2012; Novak et al., 2016; Bailoo et al., 2018); the levels of stereotypy observed here were similar to those in the study reporting a relationship (Garner et al., 2011). These findings therefore indicate that either this test paradigm does not measure recurrent perseveration reliably or that cage stereotypies in these mice do not reflect behavioral disinhibition as measured by recurrent perseveration. A difference in patterned responding was observed with respect to housing condition, in C57 mice only—animals housed in the nesting condition made more repetitive choices, compared to all other groups; albeit to one side (LLLL). Previous studies have indicated that C57 mice tend to show more alternations (LRLR or RLRL) than repetitions (LLLL or RRRR; Bailoo et al., 2018) and that enriched housing reduces the number of repetitive choices compared to barren housing (Gross et al., 2011a). The series of paradoxical results across studies suggests that these may be study specific idiosyncrasies of unknown etiology.

Studies investigating the effect of transfer from standard housing conditions to enriched environments on judgment biases in rats have reported a shift in judgment biases from more “pessimistic” to more “optimistic” responses (Brydges et al., 2011; Richter et al., 2012). In the present experiment, the overall response pattern, a monotonic graded response, was observed in mice from both strains and across all four housing conditions—confirming that the internal consistency criteria of this test were met (for further details, see Hintze et al., 2018). However, the only variations depending on housing conditions were that C57 mice housed in the B group and SWISS mice kept in the SE group showed a lower number of go responses when confronted with near negative trials compared to the other housing conditions. These findings may indicate either that this test was not sensitive enough to detect variation in animal welfare or that the observed differences in animal welfare were not

associated with differences in emotional valence. However, judgment bias was assessed at the end of a series of other tests and after extensive training, with daily removal from the home cage and handling for testing, which may have masked treatment effects. Without further study, this explanation remains essentially speculative. Therefore, these results—the first report of judgment biases in mice as a consequence of environmental supplementation—require replication and further study.

When comparing variation across our outcome measures in relation to housing conditions, no systematic patterns of differences were found. These results add to a growing body of evidence against the often heard, yet unsubstantiated, concerns that environmental enrichment would increase variation in outcome measures, thereby inducing a need for larger sample sizes to detect treatment effects (Augustsson et al., 2003; Wolfer et al., 2004; Baumans et al., 2010; van de Weerd et al., 2010).

Overall, we found that the mice readily used the enrichments when available, and that engagement with enrichment increased with increasing degree of enrichment—indicating that the enrichments offered the animals choices, which they integrated in the expression of their behavior. However, we also found that with the exception of extensive enrichment in the SE condition, variation in enrichment did not produce consistent variation in our measures of welfare. According to a large body of literature, there is no doubt that nesting material is beneficial for laboratory mice in many ways (Bult and Lynch, 1997; Sherwin, 1997; van de Weerd et al., 1997; Olsson and Dahlborn, 2002; Smith and Corrow, 2005; Gaskill et al., 2009, 2011, 2012; Gross et al., 2011b). Previous studies also found that nesting material alone attenuates stereotypic behavior (Gross et al., 2012). However, the lack of consistent differences between mice housed in B, N and E conditions together with the substantial differences between these mice and those housed in SE conditions suggested that considerably more extensive enrichment strategies may be needed to achieve substantial improvements in welfare. From the present results, we do not know whether it was the larger space, the frequency of bedding changes, the more complex environment, the types of enrichments, or the novelty that contributed most to the beneficial effect of the SE conditions as all of these factors were confounded. We therefore suggest that further studies under more extensive conditions are needed to inform decisions on minimal requirements for the housing of laboratory mice. Such an approach may allow for the identification of relevant resources or combinations of resources necessary for improving the welfare of laboratory housed rodents. It may also lend insight into relevant mental or cognitive capacities underlying species-typical behaviors (e.g., learning, memory, spatial navigation)—permitting for the development of innovative housing solutions that recapitulate the use of these capacities in the laboratory (for review, see Bennett et al., 2018; and for examples of evaluation and application, Dutton et al., 2018). In the 70s and 80s, such an approach was pursued in a range of farm animals (e.g., laying hens: Fölsch et al., 1983; rabbits: Stauffacher, 1992; pigs: Stolba and Wood-Gush, 1984), resulting in welfare-friendly prototypes

of housing conditions, from which practicable solutions were further developed for specific contexts of animal use (e.g., laboratory rabbits: Stauffacher, 1994).

The present results also highlight the difficulty in making general recommendations for improving the housing environment of laboratory mice. For example, even though the mice engaged more with the environment as the degree of supplementation increased, specific improvements in measures of welfare varied in a strain specific way. Furthermore, we only studied females and the needs of male mice may differ from those of female mice due to their higher propensity of escalating aggression (Kappel et al., 2017). Thus, we need to take into consideration that there may be no single solution to meet the needs of all strains of mice and both sexes, and enrichment strategies may always have to be adjusted to the specific strain and sex being used.

## AUTHOR CONTRIBUTIONS

HW: funding acquisition and resources. JB, EM and HW: conceptualization. JDB and EM: methodology. JB: project

## REFERENCES

- Akre, A. K., Bakken, M., Hovland, A. L., Palme, R., and Mason, G. (2011). Clustered environmental enrichments induce more aggression and stereotypic behaviour than do dispersed enrichments in female mice. *Appl. Anim. Behav. Sci.* 131, 145–152. doi: 10.1016/j.applanim.2011.01.010
- Augustsson, H., van de Weerd, H. A., Kruitwagen, C. L. J. J., and Baumans, V. (2003). Effect of enrichment on variation and results in the light/dark test. *Lab. Anim.* 37, 328–340. doi: 10.1258/00236770322389898
- Bailoo, J. D., Bohlen, M. O., and Wahlsten, D. L. (2010). The precision of video and photocell tracking systems and the elimination of tracking errors with infrared backlighting. *J. Neurosci. Methods* 188, 45–52. doi: 10.1016/j.jneumeth.2010.01.035
- Bailoo, J. D., Jordan, R. L., Garza, X. J., and Tyler, A. N. (2013). Brief and long periods of maternal separation affect maternal behavior and offspring behavioral development in C57BL/6 mice. *Dev. Psychobiol.* 56, 674–685. doi: 10.1002/dev.21135
- Bailoo, J. D., Murphy, E., Varholick, J. A., Novak, J., Palme, R., and Würbel, H. (2018). Evaluation of the effects of space allowance on measures of animal welfare in laboratory mice. *Sci. Rep.* 8:713. doi: 10.1038/s41598-017-18493-6
- Balcombe, J. P., Barnard, N. D., and Sandusky, C. (2004). Laboratory routines cause animal stress. *Contemp. Top. Lab. Anim. Sci.* 43, 42–51.
- Baumans, V., and Van Loo, P. L. P. (2013). How to improve housing conditions of laboratory animals: the possibilities of environmental refinement. *Vet. J.* 195, 24–32. doi: 10.1016/j.tvjl.2012.09.023
- Baumans, V., van Loo, P. L. P., and Pham, T. M. (2010). Standardisation of environmental enrichment for laboratory mice and rats: utilisation, practicality and variation in experimental results. *Scand. J. Lab. Anim. Sci.* 37, 101–114.
- Bechard, A., Meagher, R., and Mason, G. (2011). Environmental enrichment reduces the likelihood of alopecia in adult C57BL/6J mice. *J. Am. Assoc. Lab. Anim. Sci.* 50, 171–174.
- Benaroya-Milshtein, N., Hollander, N., Apter, A., Kukulansky, T., Raz, N., Wilf, A., et al. (2004). Environmental enrichment in mice decreases anxiety, attenuates stress responses and enhances natural killer cell activity. *Eur. J. Neurosci.* 20, 1341–1347. doi: 10.1111/j.1460-9568.2004.03587.x
- Benefiel, A. C., Dong, W. K., and Greenough, W. T. (2005). Mandatory “enriched” housing of laboratory animals: the need for evidence-based evaluation. *ILAR J.* 46, 95–105. doi: 10.1093/ilar.46.2.95
- Bennett, A. J., Bailoo, J. D., Dutton, M., Michel, G. F., and Pierre, P. J. (2018). Psychological science applied to improve captive animal care: a model for administration, formal analysis and writing—original draft. JB, EM, MB-S, JV, SH, CB, KH, CG and RP: investigation. JB EM MB-S, JV, SH, KH, CG, RP, BV and HW: writing—review and editing.

## FUNDING

This study was funded by the ERC Advanced Grant (No. 322576 “REFINE”) to HW.

## ACKNOWLEDGMENTS

We would like to thank Zeljko Kragic for building the test apparatus used in this study and Dr. Janja Novak for technical assistance related to the behavioral coding.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnbeh.2018.00232/full#supplementary-material>

- development of a systematic, evidence-based assessment of environmental enrichment for nonhuman primates. *PsyArXiv* doi: 10.31234/osf.io/79xky
- Binder, E., Droste, S., Ohl, F., and Reul, J. M. H. M. (2004). Regular voluntary exercise reduces anxiety-related behaviour and impulsiveness in mice. *Behav. Brain Res.* 155, 197–206. doi: 10.1016/j.bbr.2004.04.017
- Brydges, N. M., Leach, M., Nicol, K., Wright, R., and Bateson, M. (2011). Environmental enrichment induces optimistic cognitive bias in rats. *Anim. Behav.* 81, 169–175. doi: 10.1016/j.anbehav.2010.09.030
- Bult, A., and Lynch, C. B. (1997). Nesting and fitness: lifetime reproductive success in house mice bidirectionally selected for thermoregulatory nest-building behavior. *Behav. Genet.* 27, 231–240. doi: 10.1023/A:1025610130282
- Carola, V., D'Olimpio, F., Brunamonti, E., Mangia, F., and Renzi, P. (2002). Evaluation of the elevated plus-maze and open-field tests for the assessment of anxiety-related behaviour in inbred mice. *Behav. Brain Res.* 134, 49–57. doi: 10.1016/s0166-4328(01)00452-1
- Chapillon, P., Manneché, C., Belzung, C., and Caston, J. (1999). Rearing environmental enrichment in two inbred strains of mice: 1. Effects on emotional reactivity. *Behav. Genet.* 29, 41–46. doi: 10.1023/A:1021437905913
- Dawkins, M. S. (1988). Behavioural deprivation: a central problem in animal welfare. *Appl. Anim. Behav. Sci.* 20, 209–225. doi: 10.1016/0168-1591(88)90047-0
- DeLuca, A. M. (1997). Environmental enrichment: does it reduce barbering in mice? *Anim. Welf. Inf. Cent. Newsl.* 8, 7–8.
- Denenberg, V. H. (1969). Open-field behavior in the rat: what does it mean? *Ann. N Y Acad. Sci.* 159, 852–859. doi: 10.1111/j.1749-6632.1969.tb12983.x
- Dutton, M. B., Pierre, P. J., Bailoo, J. D., Warkins, E., Michel, G. F., and Bennett, A. J. (2018). A model quantitative assessment tool for nonhuman primate environmental enrichment plans. *BioRxiv:341206 [Preprint]*. doi: 10.1101/341206
- European Parliament Council. (2010). 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. *J. European Union.* 33–79 [Accessed on 11 October 2018].
- Fölsch, D. W., Rist, M., Munz, G., and Teygeler, H. (1983). Entwicklung eines tiergerechten legehennen-haltungssystem: die volierenhaltung. *Landtechnik* 6, 255–257.
- Garber, J. C. (2011). On the care and use of US lab animals. *Nature* 476:152. doi: 10.1038/476152a

- Garner, J. P. (2005). Stereotypies and other abnormal repetitive behaviors: potential impact on validity, reliability, and replicability of scientific outcomes. *ILAR J.* 46, 106–117. doi: 10.1093/ilar.46.2.106
- Garner, J. P., Dufour, B., Gregg, L. E., Weisker, S. M., and Mench, J. A. (2004a). Social and husbandry factors affecting the prevalence and severity of barbering ('whisker trimming') by laboratory mice. *Appl. Anim. Behav. Sci.* 89, 263–282. doi: 10.1016/j.applanim.2004.07.004
- Garner, J. P., Weisker, S. M., Dufour, B., and Mench, J. A. (2004b). Barbering (Fur and Whisker Trimming) by laboratory mice as a model of human trichotillomania and obsessive-compulsive spectrum disorders. *Comp. Med.* 54, 216–224.
- Garner, J. P., Mason, G. J., and Smith, R. (2003). Stereotypic route-tracing in experimentally caged songbirds correlates with general behavioural disinhibition. *Anim. Behav.* 66, 711–727. doi: 10.1006/anbe.2002.2254
- Garner, J. P., Thogerson, C. M., Dufour, B. D., Würbel, H., Murray, J. D., and Mench, J. A. (2011). Reverse-translational biomarker validation of Abnormal Repetitive Behaviors in mice: an illustration of the 4P's modeling approach. *Behav. Brain Res.* 219, 189–196. doi: 10.1016/j.bbr.2011.01.002
- Gaskill, B. N., Gordon, C. J., Pajor, E. A., Lucas, J. R., Davis, J. K., and Garner, J. P. (2012). Heat or insulation: behavioral titration of mouse preference for warmth or access to a nest. *PLoS One* 7:e32799. doi: 10.1371/journal.pone.0032799
- Gaskill, B. N., Rohr, S. A., Pajor, E. A., Lucas, J. R., and Garner, J. P. (2011). Working with what you've got: changes in thermal preference and behavior in mice with or without nesting material. *J. Therm. Biol.* 36, 193–199. doi: 10.1016/j.jtherbio.2011.02.004
- Gaskill, B. N., Rohr, S. A., Pajor, E. A., Lucas, J. R., Garner, J. P., Pajor, E. A., et al. (2009). Some like it hot: mouse temperature preferences in laboratory housing. *Appl. Anim. Behav. Sci.* 116, 279–285. doi: 10.1016/j.applanim.2008.10.002
- Gordon, C. J. (1985). Relationship between autonomic and behavioral thermoregulation in the mouse. *Physiol. Behav.* 34, 687–690. doi: 10.1016/0031-9384(85)90365-8
- Gordon, C. J. (1993). *Temperature Regulation in Laboratory Rodents*. New York, NY: Cambridge University Press.
- Gordon, C. J., Becker, P., and Ali, J. S. (1998). Behavioral thermoregulatory responses of single- and group-housed mice. *Physiol. Behav.* 65, 255–262. doi: 10.1016/s0031-9384(98)00148-6
- Görtz, N., Lewejohann, L., Tomm, M., Ambrée, O., Keyvani, K., Paulus, W., et al. (2008). Effects of environmental enrichment on exploration, anxiety, and memory in female TgCRND8 Alzheimer mice. *Behav. Brain Res.* 191, 43–48. doi: 10.1016/j.bbr.2008.03.006
- Gross, A. N., Engel, A. K. J., Richter, S. H., Garner, J. P., and Würbel, H. (2011a). Cage-induced stereotypies in female ICR CD-1 mice do not correlate with recurrent perseveration. *Behav. Brain Res.* 216, 613–620. doi: 10.1016/j.bbr.2010.09.003
- Gross, A. N., Engel, A. K. J., and Würbel, H. (2011b). Simply a nest? Effects of different enrichments on stereotypic and anxiety-related behaviour in mice. *Appl. Anim. Behav. Sci.* 134, 239–245. doi: 10.1016/j.applanim.2011.06.020
- Gross, A. N., Richter, S. H., Engel, A. K. J., and Würbel, H. (2012). Cage-induced stereotypies, perseveration and the effects of environmental enrichment in laboratory mice. *Behav. Brain Res.* 234, 61–68. doi: 10.1016/j.bbr.2012.06.007
- Gurfein, B., Davidenko, O., Premenko-Lanier, M., Milush, J. M., Acree, M., Dallman, M. F., et al. (2014). Environmental enrichment alters splenic immune cell composition and enhances secondary influenza vaccine responses in mice. *Mol. Med.* 20, 179–190. doi: 10.2119/molmed.2013.00158
- Haemisch, A., Voss, T., and Gärtner, K. (1994). Effects of environmental enrichment on aggressive behavior, dominance hierarchies, and endocrine states in male DBA/2J mice. *Physiol. Behav.* 56, 1041–1048. doi: 10.1016/0031-9384(94)90341-7
- Harding, E. S., Paul, E. J., and Mendl, M. (2004). Animal behaviour: cognitive bias and affective state. *Nature* 427:6972. doi: 10.1038/427312a
- Harkness, J. E., Turner, P. V., VandeWoude, S., and Wheler, C. L. (2013). *Harkness and Wagner's Biology and Medicine of Rabbits and Rodents*. 5th Edn. Ames, IA: Wiley-Blackwell.
- Hintze, S., Melotti, L., Colosio, S., Bailoo, J. D., Boada-Saña, M., Würbel, H., et al. (2018). A cross-species judgement bias task: integrating active trial initiation into a spatial Go/No-go task. *Sci. Rep.* 8:5104. doi: 10.1038/s41598-018-23459-3
- Howerton, C. L., Garner, J. P., and Mench, J. A. (2008). Effects of a running wheel-igloo enrichment on aggression, hierarchy linearity and stereotypy in group-housed male CD-1 (ICR) mice. *Appl. Anim. Behav. Sci.* 115, 90–103. doi: 10.1016/j.applanim.2008.05.004
- Jansen, R. G., Wiertz, L., Meyer, E. S., and Noldus, L. P. J. J. (2003). Reliability analysis of observational data: problems, solutions, and software implementation. *Behav. Res. Methods Instrum. Comput.* 35, 391–399. doi: 10.3758/bf03195516
- Jennings, M., Batchelor, G. R., Brain, P. F., Dick, A., Elliott, H., Francis, R. J., et al. (1998). Refining rodent husbandry: the mouse. Report of the Rodent Refinement Working Party. *Lab. Anim.* 32, 233–259. doi: 10.1258/002367798780559301
- Johnson, J. S., Taylor, D. J., Green, A. R., and Gaskill, B. N. (2017). Effects of nesting material on energy homeostasis in BALB/cAnNCrl, C57BL/6NCrl, and Crl:CD1(ICR) mice housed at 20°C. *J. Am. Assoc. Lab. Anim. Sci.* 56, 254–259.
- Kappel, S., Hawkins, P., and Mendl, M. T. (2017). To group or not to group? Good practice for housing male laboratory mice. *Animals* 7:E88. doi: 10.3390/ani7120088
- Klok, M. S. D., Jakobsdottir, S., and Drent, M. L. (2007). The role of leptin and ghrelin in the regulation of food intake and body weight in humans: a review. *Obes. Rev.* 8, 21–34. doi: 10.1111/j.1467-789x.2006.00270.x
- Latham, N., and Mason, G. (2004). From house mouse to mouse house: the behavioural biology of free-living *Mus musculus* and its implications in the laboratory. *Appl. Anim. Behav. Sci.* 86, 261–289. doi: 10.1016/j.applanim.2004.02.006
- Latham, N., and Mason, G. (2010). Frustration and perseveration in stereotypic captive animals: is a taste of enrichment worse than none at all? *Behav. Brain Res.* 211, 96–104. doi: 10.1016/j.bbr.2010.03.018
- Lutter, M., Sakata, I., Osborne-Lawrence, S., Rovinsky, S. A., Anderson, J. G., Jung, S., et al. (2008). The orexigenic hormone ghrelin defends against depressive symptoms of chronic stress. *Nat. Neurosci.* 11, 752–753. doi: 10.1038/nn.2139
- Mason, G. J., and Latham, N. R. (2004). Can't stop, won't stop: is stereotypy a reliable animal welfare indicator? *Anim. Welf.* 13, 57–69.
- McEwen, B. S. (2007). Physiology and neurobiology of stress and adaptation: central role of the brain. *Physiol. Rev.* 87, 879–904. doi: 10.1152/physrev.00041.2006
- McGregor, P. K., and Ayling, S. J. (1990). Varied cages result in more aggression in male CFLP mice. *Appl. Anim. Behav. Sci.* 26, 277–281. doi: 10.1016/0168-1591(90)90143-2
- National Research Council. (2011). *Guide for the Care and Use of Laboratory Animals*. 8th Edn. Washington, DC: The National Academies Press.
- Nevison, C. M., Hurst, J. L., and Barnard, C. J. (1999a). Strain-specific effects of cage enrichment in male laboratory mice (*Mus Musculus*). *Anim. Welf.* 8, 361–379.
- Nevison, C. M., Hurst, J. L., and Barnard, C. J. (1999b). Why do male ICR(CD-1) mice perform bar-related (stereotypic) behaviour? *Behav. Processes* 47, 95–111. doi: 10.1016/s0376-6357(99)00053-4
- Newberry, R. C. (1995). Environmental enrichment: increasing the biological relevance of captive environments. *Appl. Anim. Behav. Sci.* 44, 229–243. doi: 10.1016/0168-1591(95)00616-z
- Nicol, C. J., Brocklebank, S., Mendl, M., and Sherwin, C. M. (2008). A targeted approach to developing environmental enrichment for two strains of laboratory mice. *Appl. Anim. Behav. Sci.* 110, 341–353. doi: 10.1016/j.applanim.2007.05.006
- Nithianantharajah, J., and Hannan, A. J. (2006). Enriched environments, experience-dependent plasticity and disorders of the nervous system. *Nat. Rev. Neurosci.* 7, 697–709. doi: 10.1038/nrn1970
- Novak, J., Bailoo, J. D., Melotti, L., and Würbel, H. (2016). Effect of cage-induced stereotypies on measures of affective state and recurrent perseveration in CD-1 and C57BL/6 mice. *PLoS One* 11:e0153203. doi: 10.1371/journal.pone.0153203
- Olsson, I. A. S., and Dahlborn, K. (2002). Improving housing conditions for laboratory mice: a review of "environmental enrichment". *Lab. Anim.* 36, 243–270. doi: 10.1258/002367702320162379

- Pellow, S., Chopin, P., File, S., and Briley, M. (1985). Validation of open-closed arm entries in an elevated plus-maze as a measure of anxiety in the rat. *J. Neurosci. Methods* 14, 149–167. doi: 10.1016/0165-0270(85)90031-7
- Richter, S. H., Schick, A., Hoyer, C., Lankisch, K., Gass, P., and Vollmayr, B. (2012). A glass full of optimism: enrichment effects on cognitive bias in a rat model of depression. *Cogn. Affect. Behav. Neurosci.* 12, 527–542. doi: 10.3758/s13415-012-0101-2
- Roy, V., Belzung, C., Delarue, C., and Chapillon, P. (2001). Environmental enrichment in BALB/c mice: effects in classical tests of anxiety and exposure to a predatory odor. *Physiol. Behav.* 74, 313–320. doi: 10.1016/s0031-9384(01)00561-3
- Sherwin, C. M. (1997). Observations on the prevalence of nest-building in non-breeding TO strain mice and their use of two nesting materials. *Lab. Anim.* 31, 125–132. doi: 10.1258/002367797780600134
- Smith, A. L., and Corrow, D. J. (2005). Modifications to husbandry and housing conditions of laboratory rodents for improved well-being. *ILAR J.* 46, 140–147. doi: 10.1093/ilar.46.2.140
- Stauffacher, M. (1992). Tiergerechte haltung von hauskaninchen: neue konzepte für die zucht und haltung von labor-und fleischmastkaninchen. *Deutsch. Tierarztl. Wochenschr.* 99, 9–15.
- Stauffacher, M. (1994). Ethologische konzepte zur entwicklung tiergerechter haltungssysteme und haltungsnormen für versuchstiere. *Tierarztl. Umsch.* 49, 560–569.
- Stolba, A., and Wood-Gush, D. G. (1984). The identification of behavioural key features and their incorporation into a housing design for pigs. *Ann. Rech. Vet.* 15, 287–299.
- Sztainberg, Y., and Chen, A. (2010). An environmental enrichment model for mice. *Nat. Protoc.* 5, 1535–1539. doi: 10.1038/nprot.2010.114
- The Swiss Federal Council. (2008). Animal Welfare Ordinance, 455.1. (English version) 1–152. Available online at: [https://www.blv.admin.ch/dam/blv/en/dokumente/tiere/rechts-und-vollzugsgrundlagen/animal-welfare-ordinance-tschr.pdf.download.pdf/Animal\\_Welfare\\_Ordinance\\_\(TSchV\)\\_position\\_as\\_at\\_1.4.2011.pdf](https://www.blv.admin.ch/dam/blv/en/dokumente/tiere/rechts-und-vollzugsgrundlagen/animal-welfare-ordinance-tschr.pdf.download.pdf/Animal_Welfare_Ordinance_(TSchV)_position_as_at_1.4.2011.pdf) [Accessed on 11 October 2018].
- Tilly, S., Dallaire, J., and Mason, G. J. (2010). Middle-aged mice with enrichment-resistant stereotypic behaviour show reduced motivation for enrichment. *Anim. Behav.* 80, 363–373. doi: 10.1016/j.anbehav.2010.06.008
- Torres, S. J., and Nowson, C. A. (2007). Relationship between stress, eating behavior, and obesity. *Nutrition* 23, 887–894. doi: 10.1016/j.nut.2007.08.008
- Toth, L. A., Kregel, K., Leon, L., and Musch, T. I. (2011). Environmental enrichment of laboratory rodents: the answer depends on the question. *Comp. Med.* 61, 314–321.
- Touma, C., and Palme, R. (2005). Measuring fecal glucocorticoid metabolites in mammals and birds: the importance of validation. *Ann. N Y Acad. Sci.* 1046, 54–74. doi: 10.1196/annals.1343.006
- Touma, C., Palme, R., and Sachser, N. (2004). Analyzing corticosterone metabolites in fecal samples of mice: a noninvasive technique to monitor stress hormones. *Horm. Behav.* 45, 10–22. doi: 10.1016/j.yhbeh.2003.07.002
- Touma, C., Sachser, N., Möstl, E., and Palme, R. (2003). Effects of sex and time of day on metabolism and excretion of corticosterone in urine and feces of mice. *Gen. Comp. Endocrinol.* 130, 267–278. doi: 10.1016/s0016-6480(02)00620-2
- Tsai, P.-P., Pachowsky, U., Stelzer, H. D., and Hackbarth, H. (2002). Impact of environmental enrichment in mice. 1: effect of housing conditions on body weight, organ weights and haematology in different strains. *Lab. Anim.* 36, 411–419. doi: 10.1258/002367702320389071
- Turner, M. (1997). “Towards an executive dysfunction account of repetitive behaviour in autism,” in *Autism as an Executive Disorder*, ed. J. Russell (New York, NY: Oxford University Press), 57–100.
- van de Weerd, H. A., Aarsen, E. L., Mulder, A., Kruitwagen, C. L., Hendriksen, C. F., and Baumans, V. (2010). Effects of environmental enrichment for mice: variation in experimental results. *J. Appl. Anim. Welf. Sci.* 5, 87–109. doi: 10.1207/s15327604jaws0502\_01
- van de Weerd, H. A., Baumans, V., Koolhaas, J. M., and van Zutphen, L. F. (1994). Strain specific behavioural response to environmental enrichment in the mouse. *J. Exp. Anim. Sci.* 36, 117–127.
- van de Weerd, H. A., Van Loo, P. L. P., Van Zutphen, L. F. M., Koolhaas, J. M., and Baumans, V. (1997). Preferences for nesting material as environmental enrichment for laboratory mice. *Lab. Anim.* 31, 133–143. doi: 10.1258/002367797780600152
- van de Weerd, H. A., Van Loo, P. L. P., Van Zutphen, L. F. M., Koolhaas, J. M., and Baumans, V. (1998). Strength of preference for nesting material as environmental enrichment for laboratory mice. *Appl. Anim. Behav. Sci.* 55, 369–382. doi: 10.1016/s0168-1591(97)00043-9
- Van Loo, P. L. P., Van Zutphen, L. F. M., and Baumans, V. (2003). Male management: coping with aggression problems in male laboratory mice. *Lab. Anim.* 37, 300–313. doi: 10.1258/002367703322389870
- van Praag, H., Kempermann, G., and Gage, F. (2000). Neural consequences of environmental enrichment. *Nat. Rev. Neurosci.* 1, 191–198. doi: 10.1038/35044558
- Walf, A. A., and Frye, C. A. (2007). The use of the elevated plus maze as an assay of anxiety-related behavior in rodents. *Nat. Protoc.* 2, 322–328. doi: 10.1038/nprot.2007.44
- Whimberly, A. E., and Denenberg, V. (1967). Two independent behavioral dimensions in open-field performance. *J. Comp. Physiol. Psychol.* 63, 500–504. doi: 10.1037/h0024620
- Whittaker, A., Howarth, G., and Hickman, D. (2012). Effects of space allocation and housing density on measures of wellbeing in laboratory mice: a review. *Lab. Anim.* 46, 3–13. doi: 10.1258/la.2011.011049
- Wolfer, D. P., Litvin, O., Morf, S., Nitsch, R. M., Lipp, H.-P., and Würbel, H. (2004). Laboratory animal welfare: cage enrichment and mouse behaviour. *Nature* 432, 821–822. doi: 10.1038/432821a
- Würbel, H. (2001). Ideal homes? Housing effects on rodent brain and behaviour. *Trends Neurosci.* 24, 207–211. doi: 10.1016/s0166-2236(00)01718-5
- Würbel, H. (2006). “The motivational basis of caged rodents’ stereotypies,” in *Stereotypic Animal Behaviour-Fundamentals and Applications to Welfare*, 2nd Edn. eds G. J. Mason and J. Rushen (Wallingford: CAB International), 86–120.
- Würbel, H., Chapman, R., and Rutland, C. (1998). Effect of feed and environmental enrichment on development of stereotypic wire-gnawing in laboratory mice. *Appl. Anim. Behav. Sci.* 60, 69–81. doi: 10.1016/s0168-1591(98)00150-6
- Zhu, S. W., Yee, B. K., Nyffeler, M., Winblad, B., Feldon, J., and Mohammed, A. H. (2006). Influence of differential housing on emotional behaviour and neurotrophin levels in mice. *Behav. Brain Res.* 169, 10–20. doi: 10.1016/j.bbr.2005.11.024

**Conflict of Interest Statement:** 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.

Copyright © 2018 Bailoo, Murphy, Boada-Saña, Varholick, Hintze, Baussièrre, Hahn, Göpfert, Palme, Voelkl and Würbel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.