A mapping review of refinements to laboratory rat housing and husbandry

Vikki Neville*, James Lind, Elsa Mendl, Nathalie E. Cozma, Elizabeth S. Paul and Michael Mendl

Animal Welfare and Behaviour Research Group, Bristol Veterinary School, University of Bristol, UK

*e-mail: vikki.neville@bristol.ac.uk

Abstract

Refining the housing and husbandry of laboratory rats is an important goal, both for ethical reasons and to allow better quality research. We conducted a mapping review of 1,017 studies investigating potential refinements of housing and husbandry of the laboratory rat to assess what refinements have, and haven't, been studied, and to briefly assess whether there is evidence to support any impact on rat welfare. Among the many refinements studied, the majority involve changes to the cage, but some also involve alterations to the wider environment. The effects of these refinements were assessed using a range of readouts, many of which are difficult to interpret from a welfare perspective. Preference studies, which are easier to interpret, provide evidence that rats prefer complex environments, including shelters and multiple objects, which offer different areas/resources allowing the rat to engage in diverse behaviours. The reporting of methodology in papers was often poor, indicating that studies were potentially subject to biases. Given that many refinements co-occurred, it was often difficult to tease apart which ones were most beneficial for rat welfare. Effects of refinements were also moderated by a number of factors including age, sex, strain, and photoperiod. Altogether our findings show that a one-size-fits-all approach to refinements is not appropriate, because different refinements will impact different rats in different ways. Our review has also produced a database of >1,000 articles that can be used for further and more detailed analyses. Our findings have also highlighted areas where future research is likely to be valuable, including refinements to rat transport, handling, and the use of training.

Introduction

The welfare of rats used in scientific research is a major concern. In 2017, a total of 1,146,299 rats were used in scientific procedures in the European Union alone[1]; refining the housing and husbandry of these animals to enhance their welfare is thus an important goal. Aside from an ethical obligation to ensure good welfare, it has been suggested that improving the welfare of research animals might help to improve the quality of science[2–4]. In addition, the general public have indicated that their support for animal research is dependent on the good welfare of research animals[5].

A potentially large number of refinements can be applied with the aim of improving rat welfare and it is important to know which are effective, which have little impact, and which may even be detrimental. However, poor reproducibility is a widely accepted problem in animal research[6,7], and single studies of refinements can vary greatly in their reliability and generalisability. Evidence synthesis is the cornerstone of evidence-based assessment, and involves comprehensively compiling and reviewing results from multiple studies to allow the identification of potential biases or areas of poor knowledge, and ultimately the drawing of solid conclusions with external validity[8,9]. Despite their importance, such syntheses and reviews remain relatively rare in animal welfare science. A complete review of potential refinements to the housing and husbandry of laboratory rats has not previously been undertaken**;** yet such a review is likely to be extremely valuable in identifying robust findings, detecting knowledge gaps and weaknesses in study designs, and summarizing current knowledge without the need of using additional experimental animals.

To this end, we aimed to collect, compile and review research on refinements to the housing and husbandry of laboratory rats to assess the available evidence on the types of refinements that might improve laboratory rat welfare and identify potential knowledge gaps warranting further research. For this purpose, we searched the literature to identify studies with rats that involved manipulations to housing or husbandry, which could potentially act as 3Rs refinements. Interpreting study findings in terms of changes in welfare depends on how welfare is conceptualised. While welfare has been defined in a number of ways, the experience of positive and negative mental states is arguably a key determinant of welfare[10–14]. Accordingly, we considered that animals have good welfare when they experience positive affective states, while animals with poor welfare experience negative states. However, measuring affective states and hence welfare is challenging because we cannot simply ask animals how they feel. To date, two main approaches have been used for animal welfare assessment[15].

One approach involves interpreting changes in behavioural, physiological, or neurophysiological 'welfare indicators' in affective terms[10,16–18]. In the existing literature, such inferences typically rely on assumptions or intuitions about what is likely or unlikely to induce positive or negative affective states, or about human-to-animal translatability. For example, an indicator may be considered to measure a particular emotional state if consistent changes in that indicator occur when animals are exposed to manipulations that we assume will induce that state (for example, assuming that chronic mild stress will induce a depression-like state)[19–21]. Additionally, indicators that change consistently in animals receiving pharmacological substances that have antidepressant/anxiolytic properties in humans are generally considered to reflect a relatively positive state[22–24]. Finally, behavioural, neurophysiological or cognitive changes that are reliably observed in humans reporting negative moods or emotions may also be translated to animals as markers of negative affective states[18,20,23]. There are a vast number of outcome measures based on these approaches, with many new measures in development. However, because of the assumptions these measure rely on, outlined above, and the varying levels of validation they have received, drawing conclusions about affective states based on these measures is not always straightforward.

The other approach – assessing 'motivation and preference'– 'asks' animals what they want, for example by testing preference or evaluating motivation to access or avoid specific resources[25–29]. Following operational definitions of animal affect[17,30,31], it can be inferred that resources (in our case refinements) that animals prefer or work to access are rewarding and generate positive affective states (and hence good welfare), whilst those they work to avoid are aversive and generate negative affective states (and hence poor welfare). Preferences can be assessed by measuring the number of times animals visit, or how much time they spend in, different environments when presented with two or more choices. The preferred environment is assumed to be the one(s) that they choose most or spend the most time in. Consumer demand tasks can be used to measure how hard animals work to access a particular resource, with the assumption that animals will work harder to obtain more preferred resources. Conditioned place preferences can be used to assess how much time an animal will spend in an area that is associated with, but doesn't contain, a particular resource, with the assumption that animals spend more time in areas associated with preferred resources.

Accordingly, we aimed to collect and review the literature on potential refinements to rat housing and husbandry to answer the following questions: (1) What refinements for rats have, and haven't, been studied? (2) How have these refinements been studied? (3) What conclusions can be drawn about these refinements from studies (particularly from those assessing motivation or preference)? Given the breadth of these research questions and the size of the housing and husbandry literature, we took a mapping review approach to achieve these aims. Mapping reviews are high-level reviews that tackle broad questions; they contain more global and brief analyses of the resulting data than other types of review, such as systematic reviews[32].

Results

The initial search revealed 84,288 articles, which reduced to 70,372 articles after de-duplication. There were 1,105 articles that met the criteria for inclusion following abstract-based selection, and a total of 1,017 articles that met the criteria for inclusion and from which data were extracted (Fig 1) – these 1,017 articles are detailed in Supplementary Table 1. Reasons for exclusion at the full-text stage were that: the research did not investigate a refinement (n=46); the research was not empirical (n=17); the research was not conducted on rats (n=8); the refinement did not relate to housing or husbandry (n=7); there was insufficient information about the refinement (n=4); the research did not investigate preference or compare a control and treatment group (n=4); the full article was not in English (n=2). These articles spanned nine decades, with the earliest article published in 1944 and most recent in 2020. The number of published articles was the greatest in 2015 – with 53 articles published (Fig. 2).

The majority of studies used only male rats (60.2%); just 10.5% of studies used female animals only, and the remainder either used both (27.4%) or did not report this information (1.87%). Albino rats (for example Sprague Dawley, Wistar or Fischer 344) were the most commonly used (77.6% of the studies); hooded rats (such as Long Evans, Lister Hooded or PVG/c) were used in 16.3% of studies.

The PREPARE and ARRIVE guidelines for planning and reporting of research involving animals highlight the importance of methods to reduce bias (including randomisation and blinding) and to assess power (such as running power calculations) in experimental design[33,34]. However, reporting of these important components of reliable research was largely the exception rather than rule in the articles: 50.6% reported randomisation (with evidence that the proportion of articles doing so decreased over time; rho=-0.254, P=0.040), 14.1% reported blinding (with evidence that the proportion of articles doing so increased over time; rho=0.619, P<0.001), and only 3.8% of articles provided a justification for their sample size (with evidence that the proportion of articles doing so increased over time; rho=749, P<0.001). Thus, while there is some evidence for an improvement in the quality and reliability of research over time, there is a risk that many of the studies were affected by some degree of biased data collection and/or were under-powered.

Refinements that have been studied?

A wide range of putative refinements have been tested (these are summarised in Table 1 and Fig. 3). They most commonly involve manipulations of the rat's home cage, such as the provision of objects, shelter, or increased space. But they also include changes to many different aspects of housing and husbandry such as the colony room environment, handling, and social conditions. Many manipulations were studied concurrently with others, particularly cage-mate number, cage type, and cage contents.

Outcome measures used to study the refinements. The 30 most common 'welfare indicator' outcome measures identified here (Fig. 4), excluding preference which is discussed in more detail below, spanned behavioural tests (such as open field, elevated plus maze), physiology (such as body weight, CORT), and neurophysiology (such as neurotrophin levels, neurogenesis). In more than 50% of cases, several outcomes were measured. Some outcome measures were commonly assessed in the same article. For example, we found that 70.7% of articles measuring food intake also measured body weight; 51.3% of articles that used a novel object test also used an open field test; 64.7% of articles that measured heart rate also measured blood pressure; 59.0% of articles that measured adrenal weight also measured body weight; 65.4% of articles that measured water intake also measured body weight; 69.2% of studies that measured water intake also measured food intake; 72.7% of articles that measured organ weight also measured body weight; 75.9% of articles that measured blood pressure also measured heart rate, and 50.0% of articles that assessed maternal behaviour also used an open field test.

A summary of the number of studies using each outcome measure for each refinement is provided in Supplementary Table 3 to allow identification of areas where further study and/or meta-analysis might be valuable. Many of the refinements have been studied using a broad range of outcome measures. Cage contents, cage type, cagemate number, foraging opportunities, and neonatal handling have been studied using all of the 30 most common outcome measures, whereas cage controllability, group composition, reiki, temperature, and transportation have been studied using fewer than five of these outcome measures.

Studies assessing preference or motivation for refinements. In total, 60 articles studied preference or motivation, the large majority of which involved assessing rats' choices (85.0%) – such as the resource/food they opted to use/eat, time spent engaging with a particular resource compared with other resources, or time spent in particular areas with different resources. To assess preference or motivation, 10.0% of these studies used consumer demand tests and 5.0% used conditioned place preference tests (Fig. 5).

Preference for cage contents. Preference studies investigating the provision of cage contents largely found positive results, revealing that rats generally show a preference for objects within their home-cage. Specifically, one study found that rats showed a significant preference for a range of items (plastic pipe, wood platform, wood chips, and paper) compared to no items[35]. When presented with different items (wooden block, plumbing fixture, white light, large soup can, metal walls, golf ball, two acrylic calls, two acrylic blocks, sandwiched mango pit, small soup can, bone-shaped rawhide, wood block with holes, caged peach pit, large wood ball, wood dowel or small wood balls), rats showed a preference for wooden blocks[36]. Rats also showed a preference for a 'high complexity' cage that contained multiple different types of chains[37]. However, some studies showed that the preference depended on different aspects of these objects: when given a choice between different objects made from Lego, rats preferred a shelter to a box, which was preferred to a post and car, and when given a choice between large and small Lego objects, rats preferred the larger objects[38,39]. Rats were also found to spend more time in contact with enrichment items when provided with multiple items simultaneously compared to individual provision of each item (nylabone, retreat, wooden block, ladder, crawl ball)[40]. Similarly, rats were found to spend more time interacting with natural objects (rocks, sticks, dirt) as opposed to artificial objects (plastic toys)[41], and spent more time interacting with objects in a cage containing multiple different objects of different types as opposed to multiple objects of the same type[42]. Yet, a preference for objects within the cage was not always observed, with one study finding that rats showed a preference for an empty compartment compared with a compartment containing six different objects (such as chains, ladders and blocks of wood)[43].

The studies also identified that rats showed a preference for cages with a shelter. In particular, they preferred a cage with a shelter compared to an empty cage[44,45]; a cage enriched with biting stick, nesting material, and shelter compared to a cage without these items – although there was no difference in preference between this type of enrichment and an 'extra-enriched' cage that also contained a shelf and raised lid [46]; another study found that rats preferred an empty solid plastic box to a semi opaque Perspex one for use as a shelter (which, in turn, was preferred to a clear Perspex box, which was preferred to a vertical partition)[47]. In terms of substrate, rats showed a preference for wide soft paper strips (40 cm x 10 mm) over narrow soft paper strips (40 cm x 5 mm), which in turn were preferred to coarse paper strips[47]. The two consumer demand studies showed that rats will work to obtain access to a nest and running wheel[48], and will work harder to access a cage with a nest-box (with or without nesting material) than to access an empty cage[49]. The one conditioned place preference study found that rats develop a preference for the after-effects of wheel running[50].

Several articles also identified preferences that were moderated by factors relating to the rat or their environment, although these moderators were only found by studies assessing preference for different substrates within the cage. Specifically, when given a range of bedding materials (wire mesh, wood chips and filter paper, sawdust (small), sawdust (large), wood shavings (small), wood shavings (large)), both male and female rats showed an avoidance of smaller particle sawdust; but male rats showed a preference for large wood shavings, while females showed no preference [51]. Similarly, preference for bedding material (100% aspen, 25% corncob: 75% aspen, 50% corncob:50% aspen, 100% corncob) depended on time of day, with rats showing avoidance of 100% corncob during the light phase and a preference for 25% corncob during the dark phase[52]. When given a choice between aspen vs. corncob, and paper strips vs. corn husk, rats raised on corncob bedding showed no preference in terms of overall dwelling time but did have a preference for defecation on corncob bedding, whereas rats raised on aspen showed a preference for aspen bedding compared to corncob bedding[53]. Rats also showed a preference for paper strips in terms of dwelling time, but a preference for defecating on corn-husk[53]. In one study, Brown Norway males and females, Wistar females, but not Wistar males showed a significant preference for cages with either paper particles or wood shavings over those with wire mesh or sawdust[54].

Two of the twenty-four identified studies considering cage contents provision found a null result. One study, using a consumer demand test, found that motivation to access a larger cage, a cage with pillars, and a cage with novel objects was not significantly higher than motivation to access a standard cage[55]. The other study, using a choice test, found that rats did not have a preference for a metabolic cage containing an enrichment device, which provided shelter and increased floor area, compared to a metabolic cage without this device; instead, the rats showed a preference for the tunnel connecting these cages[56].

Preference for temperature. Some studies found a clear and significant preference for a particular temperature in rats, such as a preference for 27°C (versus 30°C and 33°C)[57,58], a preference for 24°C (versus 18°C and 30°C)[59], and a preference for 19.2+/-5.2°C when placed in a temperature gradient from 10°C to 40°C[60]. However, the majority of the identified studies investigating thermopreference found that preference depended on a moderating factor. Photoperiod moderated temperature preference in three choice tests, with rats preferring cooler temperatures during the dark phase[61–63]. Three choice studies also found strain differences in temperature preferences: Sprague Dawleys preferred lower temperature than Fischer rats, Long Evans preferred lower temperatures than Sprague Dawley and Fischer rats[63–65]. Body weight[66], previous experience[67], and the presence of nesting material also altered temperature preferences in choice tests[68]. One study also found that rats showed no preference for an area of a cage that contained a warm plate[69].

Preference for cage type. Rats were found to prefer cages with solid floors rather than grid floors, particularly during periods of rest[70], and another study found that rats would work as hard to access a solid floor as they would to explore a novel environment – with rats lifting up to 83% of their own body weight to access these areas[71]. However, a number of moderators for cage type preference were identified, which included strain, methods, age, and other housing or husbandry factors. In choice tests, rats largely preferred solid to mesh flooring, but Brown Norway rats preferred cages with wire mesh floor for eating and drinking, and juveniles more often spent time on wire floors[44]. Similarly, female rats showed a preference for an area with bedding material over an area with a grid floor both during the day and at night, while male rats only showed a preference for bedding material during the day and preferred the grid floor at night [72]. When given an option between a small, medium, and large arena in a choice test, the time spent in each arena depended on the size of the rat's home-cage and whether or not they were housed in isolation; paired-housed rats in large cages showed a preference for larger areas, whereas isolated rats housed in large cages showed a preference for the smaller area[73]. One study found that in the light phase, stainless steel cages – which offered increased darkness – were preferred, but in the dark phase, animals raised in stainless steel cages showed a preference for steel, while animals raised in polycarbonate cages showed no preference[74]. Likewise, another study found that preference for polycarbonate versus steel cages was complex, with moderators including the presence and location of a cage hopper, transparency of the cage walls, and photoperiod; the main conclusion of this study was that preference for low illumination in the cage drives preference[75]. One of the identified studies found a null result: when rats were provided with areas of different heights (80 mm, 160 mm, 240 mm, and 320 mm), they showed no clear preference and instead split their time fairly uniformly between the different areas[76].

Preference for foraging opportunities. With regard to foraging opportunities, the studies showed that rats exhibited a preference for: different flavours of food presented simultaneously[77], pelleted over powdered food[78], many small pellets compared to one large pellet[79], smaller compared to larger food items more generally[80], flattened food (covering a larger surface area) compared to rounded food (covering a smaller surface area)[79], red coloured food compared to yellow, green, black, or white food[80], and novel food[81]. Rats were also found to prefer feeding from a foraging device (a raised metal dish which required rats to move gravel to access the food) in terms of both time spent feeding and amount of food, compared to feeding from a standard food hopper, or from a food hopper that limited the feeding area to 4.5% of that available in a standard food hopper, and from an area with a standard food hopper in addition to gnawing sticks; however, no preference for total dwelling time was observed between these four areas[82]. Preference for rough food compared to pureed food was found to change over time, with rats initially preferring a rough food mixture and then preferring pureed food[83]. Likewise, one study found that rats accustomed to standard, hard food pellets showed a preference for soft food compared to hard food, but this preference was weaker in rats accustomed to powdered food[84].

Preference for other refinements. One study showed that rats would press a lever that resulted in tickling significantly more than they would press on a bar that had no consequence, suggesting rats were motivated to be tickled[85]. However, while rats showed a conditioned place preference for one odour associated with tickling, no conditioned place preference was observed when a different odour was used[86]. Similarly, another study found no conditioned place preference for tickling[87].

Regarding ventilation, rats showed avoidance of cages with 120 air changes per hour compared to 50 and 80 air changes per hour[88] (although it is possible that this manipulation altered other aspects of the rats' environment such as temperature and humidity), and preferred a CO_2 concentration of 3% compared to 1% and 5% in choice tests[89]. Rats showed no preference for an air speed of either 0.2m/s or 0.5m/s in a choice test[88].

Rats were found to work harder to access a group of three familiar rats than to access an empty standard home cage[55]. For cage-mate number, a quadratic relationship was identified by one study with preference, according to a consumer demand test, peaking at 5 rats compared to the other options tested: 0,1,3 and 11 cage-mates[90]. Sex moderated preference for social contact in a choice test: female rats showed a greater preference for contact with another rat while males did not show this preference[91].

Preference for different types of lighting was moderated by photoperiod, strain, and other aspects of housing and husbandry. One study found that a preference for opaque areas was only observed when rats were inactive in the light period, and when laboratory personnel were present[92]. Another study found that overall rats had a preference for the darkest cage (when none, one quarter, half, or three quarters of cage covered with black tape) but that this preference was most pronounced in albino rats in the light phase[76]. The one preference study for noise identified that silence was preferred to speech, which was preferred to radio, which was preferred to white noise[93]. Regarding olfactory stimulation, rats showed an overall preference for other rats that smelled similar to themselves in a choice test: cologne-odoured group preferred cologne-odoured rat, normal-odoured rats preferred normal-odoured rats, and methyl salicylate-odoured rats showed no preference[94].

Discussion

Refining the housing and husbandry of laboratory rats is an important goal. The implementation of any refinement should have a strong evidence base because it makes no sense to spend time and money on refinements that have no benefits to rats, and we need to be certain that any changes we make are not detrimental to welfare. However, given that one swallow does not make a summer, data synthesis is critical to draw conclusions. Here, we conducted a mapping review of studies investigating potential refinements of housing and husbandry to assess what refinements have, and haven't, been studied, how they have been studied, and to examine what conclusions can be drawn from these articles. To this end, we performed an extensive search of the literature that identified over 85,000 initial records and resulted in 1,017 articles for review.

A summary of the refinements studied in the identified articles. We found that a broad range of refinements have been studied. The majority of these refinements focussed on changes to the cage itself, such as adding enrichment objects, increasing the cage size, or changing the number of cage-mates, but there were also many other types of refinements, including those that altered aspects of the rats' environment such as noise or lighting, human–rat interactions such as tickling, gentling, and neonatal handling, and provision of playpens. However, many of the refinements (particularly those relating to housing such as cage contents and cage type) co-occurred, making it difficult to tease apart the specific aspects that are most beneficial to rat welfare.

The rationale for the manipulations studied was often unclear. For example, altering cage contents was the most common refinement, with arbitrary objects most commonly used as enrichment; however, it is unclear why such objects might be beneficial to rats beyond the novelty or perceptual/tactile stimulation they provide. Although its potential benefit to rats seems clearer, addition of shelters to the cages, which serve a particular function by providing rats with a refuge, was less commonly studied. Likewise, it is unclear why two studies investigated the performance of reiki on rats given that reiki is considered to be a pseudoscience. In many cases, the poor or unclear rationale of the study may reflect the fact that the primary aim of the studies was not necessarily to study potential refinements to rat welfare (but to instead study pup development, mood disorders, or learning for example, with a view to translating those findings to humans). However, to apply the 3Rs when studying rat welfare and to reduce the number of rats used for experiments, we suggest developing a good scientific rationale to select potential refinements for a study.

A summary of the results of studies of preference / motivation. Overall, the results of the preference studies reviewed here suggest that rats prefer complex environments that provide them with different areas to engage in different behaviours (in order to fulfil different functions). It is clear that rats require a shelter (an enclosed area that is entirely dark), given that several studies showed that rats have a preference for and will work to access shelters, and a study that manipulated lighting in the cage showed that rats prefer to spend time in a dark area when laboratory personnel are present and during the inactive period. In addition to a shelter, studies that looked at preferences across multiple objects showed that rats preferred larger objects, which provide climbing opportunities; and hence climbing opportunities of some kind should be provided. Multiple types of smaller objects might also help to increase complexity in the cage, as suggested by two studies showing that rats prefer these refinements. Areas without any objects may also be valuable given that one study found that rats prefer a cage with no objects. One preference study suggested that rats may benefit from provision of a running wheel. Providing rats with different areas of different sizes would also be ideal to account for differences in preference for open spaces across rats. Likewise, our result show that the substrate in the cage clearly fulfilled different purposes; therefore it may be valuable to provide multiple types of substrate in different parts of the cage, so that these substrates can be used in different ways – for example for nesting (where longer soft paper strips may be preferred), for defecating, and for general locomotion. Providing rats with multiple options in terms of substrate and cage contents would also account for preferences that shift with photoperiod, sex, age, or just individual differences, and allow for different levels of behavioural thermoregulation to be achieved in accordance with different temperature preferences. Due to the low numbers of studies and the lack of clear results, it is difficult to draw any solid conclusions about the number of cage-mates or degree of social contact that might be preferable, about the use of odours as potential enrichment, or the use of tickling.

A summary of the non-preference outcome measures used to study the refinements, and their relevance to welfare. A broad range of non-preference outcome measures were used in the reviewed studies to assess the effects of housing and husbandry interventions on the affective states and welfare of rats. The most common of these outcome measures included behavioural measures of anxiety- or depression- like states: elevated plus maze (where an increased proportion of time in the open arms is considered to reflect reduced anxiety[95]), open field test (where an increased proportion of time in central area is considered to reflect reduced anxiety[96]), light/dark box (where an increased proportion of time in the light area is considered to reflect reduced anxiety[97]), forced swim test (where increased immobility is thought to reflect decreased depression[98]), and sucrose preference test (where increased consumption of sucrose solution relative to water is thought to reflect decreased depression). There is some evidence supporting the use of these measures. For example, some anxiolytic drugs alter behaviour in the elevated plus maze[99] open field[100], and light/dark box[97], and some antidepressant drugs alter sucrose preference and behaviour in the forced swim test. However, the use of these measures is also subject to a number of criticisms. For example, antidepressants that are used to treat generalised anxiety disorder in humans do not consistently produce changes in elevated plus maze, open field, or light/dark box behaviour[100,101], suggesting that these tasks may not capture all features of anxiety-like states. Similarly, immobility in the forced swim test is now considered to reflect an adaptive and learnt response that allows the animal to conserve energy and hence prolong survival, as opposed to reflecting depression-induced helplessness[102–104]. Moreover, the ethics of the forced swim test has been under increasing scrutiny[105] recently and studies have challenged the validity of the sucrose preference test by showing that humans with depression show no reduction in their preference for sucrose over water[106,107].

Physiological indicators of stress-like states, such as cortisol levels, heart rate, and blood pressure, were also among the most commonly used outcome measures. However, while these measures may rise in stressful situations and hence may reflect stress-like states, they can also increase in situations that are neutral or even rewarding[16,18,108–110], which complicates their interpretation in terms of measure. Some of the most common outcome measures used were neurobiological measures that have been implicated in human depression, such as changes in the dopaminergic or serotonergic system, neurotrophin levels, or neurogenesis. Yet, the relationship between dopaminergic and serotonergic system activity, neurotrophin levels and affect are complex[111,112], meaning that there is no straightforward way to interpret changes in these measures in terms of welfare. Similarly, many factors can influence neurogenesis including those that may be less related to welfare such as physical activity and diet[113].

As a result of these limitations, individually, each of these outcome measures may not provide reliable information about the impact of the housing and husbandry refinements. However, considering the effect of a refinement on multiple outcome measures, and assessing whether results across these outcome measures align (c.f. [114]), might provide a more complete and comprehensive insight into the impact of the refinement on the welfare of the animals. We found a number of potential refinements that had been studied using a wide range of outcome measures, such as cage contents, cage type, cagemate number, and foraging opportunities. Further and more detailed analyses of these refinements across multiple outcome measures (for example by using a meta-analytic approach) might therefore be warranted. Similarly, single studies that use multiple outcome measures to assess welfare are likely to be the most informative in the future.

Given the limitations outlined above, we did not conduct in-depth analyses of the putative effects of refinements on non-preference outcome measures. However, we provide a database of the reviewed articles that includes information about the refinement studies and outcome measures used (Supplementary Table 1 and 3). These resources offer a springboard for future meta-analyses of specific refinements assessed with a smaller number of outcome measures, should this be considered valuable by future researchers.

Limitations of this mapping review. It is possible that our methodology may have resulted in the exclusion of some relevant studies. Specifically, by restricting articles to those written in English, relevant but locally published articles may have been excluded. Additionally, as part of the search strategy, some search terms were excluded (such as ‘oxidative stress’, ‘stress-induced’ and ‘stress induced’, although the term ‘stress’ was included) in line with our inclusion/exclusion criteria: “the research described in the article involves a manipulation relating to housing (including within the colony room) or husbandry that has potential to act as a 3Rs refinement (therefore,. manipulations solely aimed at inducing stress or negative affect - ‘isolation stress’/‘chronic mild stress’/‘crowding stress’ - were excluded)”. While this approach streamlined the literature search to make the review both focused and feasible, it may have resulted in the exclusion of some articles, in which the authors explicitly stated in their title or abstract that they were "not examining stress induced…". Thus, while we have identified a large number of relevant studies, we cannot guarantee that this review provides a complete and fully comprehensive overview of how rat housing and husbandry could be improved. Nonetheless, the review does provide the broadest and most complete review of housing and husbandry refinements to rat welfare to date.

A further limitation of this review is that we were unable to provide information about sample sizes or ethical approval in the articles due to poor agreement between our reviewers regarding these features in quality checks. The poor agreement could partly be attributed to improper, confusing, or misleading reporting of these data in the original articles. For example, reporting sample size information in figure legends rather than the methodology section (sometimes with different sample sizes for different outcome measures), only providing the total sample size without stating that the number of rats assigned to each treatment was equal, or stating that ethical guidelines were followed without stating that ethical approval was obtained. It may additionally reflect the fact that, when extracting information, reviewers did not implement identical rules when encountering data in an unexpected format.

Future directions. The review process highlighted a number of issues that should be considered when interpreting results of the studies analysed, and when designing future research in this area. Reporting of methodology was often poor and it is possible that many studies had weak designs or were subject to biases. The replication crisis has demonstrated the need for more rigorous methodology[6] and the ARRIVE and PREPARE guidelines have been developed to promote and support higher quality research. While we observed an increase in the proportion of studies reporting blinding and justification of sample sizes over time, which is promising, the decrease in the proportion of studies reporting randomisation and the overall low levels of reporting of these important components of reliable research are concerning. More work is clearly needed to promote more reliable and better quality research, a sentiment that has been echoed by other reviews[115,116].

Potential refinements for certain aspects of husbandry were also clearly understudied, especially when considering their potential impact on animal welfare, and the positive effects of such refinements in other species. For example, most laboratory rats will be transported at some point in their life, such as from a supplier to the institution, or from a holding room to a procedure room within a facility. Transportation is likely to be stressful[117], but we only identified one study that investigated a potential refinement to transportation. Likewise, most laboratory animals will need to be handled by laboratory personnel, and a number of studies in mice have studied potential refinements to handling (such as tunnel handling or cupping, as opposed to tail handling[118–121]). Yet, we again only identified one study that had investigated specific types of handling and their impact on rats. Training rats to cooperate with laboratory personnel seems to be understudied: we only identified one article on this topic (in which rats were trained to move from a dirty to a clean cage[122]). However, we believe that rats can be trained to cooperate with laboratory personnel in several regards, for example, to climb onto arms for handling or to go to specific locations in a home-cage to aid different husbandry procedures. Such refinements have been very successful in non-human primate species in a laboratory setting[123,124]. Finally, while there are suggestions from research in humans and other species that controllable environments with predictable punishers and unpredictable rewards should lead to more positive and less negative emotional states[125–127], this has received little attention in the context of rat welfare, with only one study assessing the impact of more controllable environments (in the form of providing operant control over delivery of food, water, and lighting)[128].

Therefore, future interesting areas of research include examining the welfare impact of scooping, tunnel handling, or training rats to move into a transport cage or onto a human hand. An assessment of the many components of transport (such as noise, vibration, crowding) and their likely impact on welfare would also be a valuable, similar to those conducted for farm species[129,130]. Potential refinements to these different components could then be devised and investigated (e.g., development of transport cages with material that attenuates noise and vibration). Future research could also include studies examining the effect of conducting potentially aversive husbandry procedures (such as cleaning) at a predictable time and conducting potentially positive procedures (such as provision of treats) at unpredictable times.

Finally, several promising measures of animal welfare, which have been less commonly used, may be valuable in future studies of rat housing and husbandry. For example, judgement bias, which measures decision-making under ambiguity , has recently been validated as a measure of animal affect (although these validation attempts also indicated that this measure has some limitations, and may not apply to all species)[21,24]. In the past few years, the use of rat ultrasonic vocalisations to infer welfare has also received growing attention, and deep-learning methods have now been developed to parse and analyse these vocalisations[131–133]. Of note, the most-common measures identified by this review are largely focussed on poor welfare states (those associated with depression, anxiety, or stress), and might therefore provide little or no information about good welfare states. Ultimately, good welfare for laboratory rats should be the goal of any refinement and the importance of assessing the presence of positive affective states, as opposed to just the absence of negative affective states, cannot be understated. Judgement bias and ultrasonic vocalisations may provide information about both positive and negative affect[21,24,131].

Conclusions

Despite the large amount of research that has been conducted over the past several decades, clear gaps remain in our knowledge of the housing and husbandry interventions that influence rat welfare. This review highlights areas where further research would be valuable, such as refinements to rat transportation, handling, the use of training to increase cooperation between rats and laboratory personnel, and the provision of more controllable environments with more predictable punishers and less predictable rewards. The review also highlights specific interventions that have been found to be preferred by rats, and that are likely to generate welfare improvements. Importantly, the review also identified a number of moderating factors, such as age, sex, strain, and photoperiod, affecting the effectiveness of modifications and interventions. Our review clearly shows that a one-size-fits-all approach to refinements is not likely to be appropriate, given that different refinements can impact different rats in different ways. For this reason, we recommend giving rats a heterogeneous habitat; and provide them with areas that vary in size and complexity, a range of substrates, and different ways to obtain food. A key aim to improve laboratory rat welfare should be to develop heterogeneous in-cage habitats (with areas varying in size, complexity, substrates, and food provision method) to cater for the varied needs of individual rats.

Methods

Literature search. A literature search was conducted on 30th April 2020 using Scopus and Web of Science to identify articles reporting studies that had investigated potential refinements to rat welfare. The search terms were developed around three concepts: rats, welfare, and housing/husbandry (Supplementary Table 4**)** with the specific search terms designed to capture as many articles relating to these concepts as possible. Additionally, we searched for papers cited by, or citing, review articles relating to rat welfare. Further details on the literature search, including search terms used, can be found in Supplementary Information.

Supplementary Table 4| Concepts used to develop the search terms for the literature search: rats, welfare, and housing/husbandry

Inclusion and exclusion criteria. The criteria for inclusion were that: 1) The article describes an empirical study (which means that review papers were excluded); 2) the article describes a study using rats, specifically *Rattus norvegicus* (which means that studies using mice or other rat species were excluded); 3) The research described in the article involves a manipulation relating to housing (including within the colony room) or husbandry that has potential to act as a 3Rs refinement (which means that manipulations solely aimed at inducing stress or negative affect – 'isolation stress'/'chronic mild stress'/'crowding stress' — were excluded); 4) The research either: (a) investigates the preference / motivations of rats for that manipulation compared with a control; or (b) compares the effects of manipulated and control group(s) on other outcome measures; and 5) The article was written in English.

Abstract screening. Following deduplication, articles underwent abstract-based selection which was conducted using Rayyan[134]. Abstract-screening was conducted by one researcher (V.N.) who was highly familiar with the literature on laboratory rat behaviour, biology, and welfare. To reduce the possibility that articles were wrongly excluded during abstract-screening, only articles that unambiguously did not meet the inclusion criteria were excluded (for example the study was conducted with mice, not with rats; the study was a review paper, not an experimental paper). If there was any doubt or ambiguity about whether an article should be excluded, it was not excluded at this stage.

Full-text screening and data extraction. Full-text screening and data extraction were conducted by four researchers (V.N., J.L., E.M., N.C.C.) who worked independently. Each article was screened by one of these researchers; quality checks were then made to ensure consistency between these researchers once all articles had been screened and data extracted (see below). These researchers first assessed whether each article met the inclusion criteria and then extracted the following data from each article (actual prompts used shown in brackets): title (what is the title of the article?); authors (who authored the article?); publication year (when was the article published?); journal (in which journal was the article published?); strain (what is the strain of rat used?); age or weight (what are the ages of the rats used or their weight); treatment sample (how many animals are there per treatment?); sample justification (was there any justification for the number of animals used?); sex (what is the sex of the rats?);ethical approval (does the article state that ethical approval was received for the study?); manipulation details (what are the details of the more positive manipulation, according to the hypotheses, or set of preferences being tested?); control details (what are the details of the more negative/control manipulation, according to the hypotheses?); randomisation (does the article state that animals were randomly assigned to treatments?); blinding (does the article state that the experimenter was blind to the treatments?); outcomes (what are the specific outcome measures?); results obtained (what were the findings of the experiment?). At a later stage, the dataset was split in two according to the outcome measures: preference studies (where the outcome measure was 'choice', 'consumer demand', or 'conditioned place preference') and potential welfare indicators (for all other outcome measures).

Following full-text screening and data extraction, the data for approximately one third (n=329) of articles were selected at random and checked by one of three researchers; these checks were made on researchers' own and others' extracted data. Data were found to be correct with no omission of important details in 100% of the publication dates, 99.7% of the author names, 99.7% of the strain details, 98.2% of the age/weight details, 74.5% of the treatment sample details, 99.4% of the sample size justification details, 99.4% of the sex, 94.3% of ethical approval details, 98.2% of manipulation details, 98.5% of randomisation details, 99.0% of blinding details, 92.7% of outcome measure details, 97.6% of results. Consequently, we did not use data about the treatment sample size or ethical approval (because they were < 95% correct). We also conducted further checks on the outcome measures and results data, as detailed below. To ensure that no articles were erroneously excluded at full-text screening, all articles excluded at this stage were checked by one person (V.N.); if there was any uncertainty about whether the article should be excluded it was discussed between two people (V.N. and M.M.) who jointly made a decision about its exclusion.

The data then underwent further processing and cleaning, which included correcting typographic errors, grouping together highly similar or identical outcome measures (for example 'elevated plus maze' and 'elevated zero maze' were included together as 'EPM'; 'serotonin turnover', '5HT levels' and 'tryptophan levels' were included together as 'serotonergic activity'), and grouping together highly similar or identical refinements (for example 'playing Mozart', 'decreasing noise in colony room' were grouped as 'noise'). Refinements were also further categorised into 'sub-refinements' where this was relevant (for example changes in 'cage contents' were divided into-sub-categories including providing or altering 'objects', 'shelters', 'wheels', and/or 'substrates' (see Table 1**).**

Given the large size of the dataset, it was not possible to analyse all the results extracted. We therefore opted to only analyse the results of preference studies and the results of studies using the 30 most common 'welfare indicator' outcome measures (excluding preference). Data were summarised by tallying the number of articles for each refinement using each of these outcome measures (Supplementary Table 3). Given that there were a non-negligible number of instances where extracted information about the outcome measure was inconsistent between researchers extracting the data, we also checked the outcome measure data against the extracted results data. This involved assessing whether all extracted outcome measure data had a corresponding result in the extracted results data and vice versa (e.g., if the outcome measure was 'open field test', checking that the extracted results included information about the outcome of the open field test). We referred to the original article where there were inconsistencies**.**

Data analysis. To assess whether the number of articles reporting specific methodologies changed over time, we tested the correlation between publication year and proportion of articles in each year reporting randomisation, blinding, or sample size justification using Spearman's rank.

Acknowledgements

This work was supported by the UK Biotechnology and Biological Sciences Research Council (BBSRC) grant BB/T002654/1 (Principal Investigator: MM) and a Universities Federation for Animal Welfare (UFAW) small project award (Principal Investigator: VN).

Author contributions

Conceptualization and methodology: V.N., E.S.P, M.M.; Data collection: V.N., J.L, E.M, N.C.C; Data analysis: V.N., Writing – original draft: V.N., Writing – reviewing and editing: all authors.

Competing interests

The author(s) declare no competing interests.

Data availability

Data are provided in Supplementary Table 1.

References

1. European Commision. Report from the Commision to the European Parliament nad the Council on the implementation of Directive 2010/63/EU on the protection of animals used for scientific purposes in the Member States of the European Union le. (2020).
2. Poole, T. Happy animals make good science. *Lab. Anim.* **31**, 116–124 (1997).
3. Prescott, M. J. & Lidster, K. Improving quality of science through better animal welfare: The NC3Rs strategy. *Lab Anim. (NY).* **46**, 152–156 (2017).
4. Cait, J., Cait, A., Scott, R. W., Winder, C. B. & Mason, G. J. Conventional laboratory housing increases morbidity and mortality in research rodents: results of a meta-analysis. *BMC Biol.* **20**, 1–22 (2022).
5. Ormandy, E. H. & Schuppli, C. A. Public attitudes toward animal research: A review. *Animals* **4**, 391–408 (2014).
6. Voelkl, B. *et al.* Reproducibility of animal research in light of biological variation. *Nat. Rev. Neurosci.* **21**, 384–393 (2020).
7. Jilka, R. L. The Road to Reproducibility in Animal Research. *J. Bone Miner. Res.* **31**, 1317–1319 (2016).
8. Meerpohl, J. J., Herrle, F., Antes, G. & von Elm, E. Scientific value of systematic reviews: Survey of editors of core clinical journals. *PLoS One* **7**, 3–7 (2012).
9. O'Hagan, E. C., Matalon, S. & Riesenberg, L. A. Systematic reviews of the literature: A better way of addressing basic science controversies. *Am. J. Physiol. - Lung Cell. Mol. Physiol.* **314**, L439–L442 (2018).
10. Mason, G. & Mendl, M. Why is there no simple way of Measuring Animal Welfare? *Anim. Welf.* **2**, 301–319 (1993).
11. Broom, D. M. Considering animals' feelings. *Anim. Sentience* **5**, (2016).
12. Dawkins, M. S. The science of animal suffering. *Ethology* **114**, 937–945 (2008).
13. Mendl, M. & Paul, E. S. Consciousness, emotion and animal welfare: Insights from cognitive science. *Anim. Welf.* **13**, 17–25 (2004).
14. Browning, H. The Natural Behavior Debate: Two Conceptions of Animal Welfare. *J. Appl. Anim. Welf. Sci.* **23**,

325–337 (2020).

15. Mendl, M. Assessing the Welfare State. *Nature* **410**, 31–32 (2001).

16. Broom, D. M. The scientific assessment of animal welfare. *Appl. Anim. Behav. Sci.* **20**, 5–19 (1988).

17. Mendl, M. & Paul, E. S. Animal affect and decision-making. *Neurosci. \& Biobehav. Rev.* **112**, 144–163 (2020).

18. Paul, E. S., Harding, E. J. & Mendl, M. Measuring emotional processes in animals: The utility of a cognitive approach. *Neurosci. Biobehav. Rev.* **29**, 469–491 (2005).

19. Overstreet, D. H. Modeling depression in animal models. *Methods Mol. Biol.* **829**, 125–144 (2012).

20. Willner, P., Muscat, R. & Papp, M. Chronic mild stress-induced anhedonia: A realistic animal model of depression. *Neurosci. Biobehav. Rev.* **16**, 525–534 (1992).

21. Lagisz, M. *et al.* Optimism, pessimism and judgement bias in animals: A systematic review and meta-analysis. *Neurosci. Biobehav. Rev.* **118**, 3–17 (2020).

22. McArthur, R. & Borsini, F. Animal models of depression in drug discovery: A historical perspective. *Pharmacol. Biochem. Behav.* **84**, 436–452 (2006).

23. Willner, P. The validity of animal models of depression. *Psychopharmacology (Berl).* **83**, 1–16 (1984).

24. Neville, V. *et al.* Pharmacological manipulations of judgement bias: A systematic review and meta-analysis. *Neurosci. Biobehav. Rev.* **108**, 269–286 (2020).

25. Fraser, D. & Matthews, L. R. Preference and motivation testing. *Anim. Welf.* 159–173 (1997).

26. Hughes, B. O. & Duncan, I. J. H. The notion of ethological 'need', models of motivation and animal welfare. *Anim. Behav.* **36**, 1696–1707 (1988).

27. Sherwin, C. M. & Nicol, C. J. Behavioural demand functions of caged laboratory mice for additional space. *Anim. Behav.* **53**, 67–74 (1997).

28. Sherwin, C. & Nicol, C. A demanding task: using economic techniques to assess animal priorities. A reply to Mason et al. *Anim. Behav.* **55**, 1079–1081 (1998).

29. Dawkins, M. S. From an animal's point of view: Motivation, fitness, and animal welfare. *Behav. Brain Sci.* 1–61 (1990).

30. Rolls, E. T. *Emotion and Decision-Making Explained*. (Oxford University Press, 2014).

31. Paul, E. S. & Mendl, M. T. Animal emotion: Descriptive and prescriptive definitions and their implications for a comparative perspective. *Appl. Anim. Behav. Sci.* **205**, 202–209 (2018).

32. Leenaars, C. *et al.* Reviewing the animal literature: how to describe and choose between different types of literature reviews. *Lab. Anim.* **55**, 129–141 (2021).

33. Smith, A. J., Clutton, R. E., Lilley, E., Hansen, K. E. A. & Brattelid, T. PREPARE: guidelines for planning animal research and testing. *Lab. Anim.* **52**, 135–141 (2018).

34. du Sert, N. P. *et al. Reporting animal research: Explanation and elaboration for the arrive guidelines 2.0*. *PLoS Biology* vol. 18 (2020).

35. Bradshaw, A. L. & Poling, A. Choice by rats for enriched versus standard home cages: plastic pipes, wood platforms, wood chips, and paper towels as enrichment items. *J. Exp. Anal. Behav.* **55**, 245–250 (1991).

36. Chmiel Jr., D. J. *et al.* Preference of laboratory rats for potentially enriching stimulus objects. *Lab. Anim.* **30**, 97–101 (1996).

37. Denny, M. S. The rat's long-term preference for complexity in its environment. *Anim. Learn. Behav.* **3**, 245–249 (1975).

38. Williams, C. M., Riddell, P. M. & Scott, L. A. Comparison of preferences for object properties in the rat using

paired- and free-choice paradigms. *Appl. Anim. Behav. Sci.* **112**, 146–157 (2008).

39. Williams, C. M., Hanmer, L. A. & Riddell, P. M. The effect of the functional attributes of objects within the caged environment on interaction time in laboratory rats. *Appl. Anim. Behav. Sci.* **120**, 208–215 (2009).

40. Abou-Ismail, U. A. Are the effects of enrichment due to the presence of multiple items or a particular item in the cages of laboratory rat? *Appl. Anim. Behav. Sci.* **134**, 72–82 (2011).

41. Lambert, K. *et al.* Natural-enriched environments lead to enhanced environmental engagement and altered neurobiological resilience. *Neuroscience* **330**, 386–394 (2016).

42. Abou-Ismail, U. A. & Mendl, M. T. The effects of enrichment novelty versus complexity in cages of group-housed rats (Rattus norvegicus). *Appl. Anim. Behav. Sci.* **180**, 130–139 (2016).

43. Will, B., Pallaud, B., Ungerer, A. & Ropartz, P. Effects of rearing in different environments on subsequent environmental preference in rats. *Dev. Psychobiol.* **12**, 151–160 (1979).

44. De Villiers, C. & Seal, J. V. Solid floor versus wire floor housing for rats revisited: Influence of age and the presence of nest boxes. *Anim. Technol. Welf.* **9**, 31–36 (2010).

45. Townsend, P. Use of in-cage shelters by laboratory rats. *Anim. Welf.* **6**, 95–103 (1997).

46. Krohn, T. C. *et al.* The ability of SD-rats to distinguish between three different housing environments. *Scand. J. Lab. Anim. Sci.* **38**, 21 (2011).

47. Manser, C. E., Broom, D. M., Overend, P. & Morris, T. H. Investigations into the preferences of laboratory rats for nest-boxes and nesting materials. *Lab. Anim.* **32**, 23–35 (1998).

48. Collier, G. H., Johnson, D. F., Cybulski, K. A. & McHale, C. A. Activity patterns in rats (Rattus norvegicus) as a function of the cost of access to four resources. *J. Comp. Psychol.* **104**, 53–65 (1990).

49. Manser, C. E., Broom, D. M., Overend, P. & Morris, T. H. Operant studies to determine the strength of preference in laboratory rats for nest-boxes and nesting materials. *Lab. Anim.* **32**, 36–41 (1998).

50. Greenwood, B. N. *et al.* Long-term voluntary wheel running is rewarding and produces plasticity in the mesolimbic reward pathway. *Behav. Brain Res.* **217**, 354–362 (2011).

51. Blom, H. J. M., Van Tintelen, G., Van Vorstenbosch, C. J. A. H. V., Baumans, V. & Beynen, A. C. Preferences of mice and rats for types of bedding material. *Lab. Anim.* **30**, 234–244 (1996).

52. Krohn, T. C. & Hansen, A. K. Evaluation of corncob as bedding for rodents. *Scand. J. Lab. Anim. Sci.* **35**, 231–236 (2008).

53. Ras, T. *et al.* Rats' preferences for corn versus wood-based bedding and nesting materials. *Lab. Anim.* **36**, 420–425 (2002).

54. Van De Weerd, H. A., Van Den Broek, F. A. R. & Baumans, V. Preference for different types of flooring in two rat strains. *Appl. Anim. Behav. Sci.* **46**, 251–261 (1996).

55. Patterson-Kane, E. G., Hunt, M. & Harper, D. Rats demand social contact. *Anim. Welf.* **11**, 327–332 (2002).

56. Sørensen, D. B. *et al.* Enriching the metabolic cage: effects on rat physiology and behaviour. *Anim. Welf.* **17**, 395–403 (2008).

57. Kumar, D., Kumar, V. M. & Mallick, H. N. Ambient temperature-dependent thermoregulatory role of REM sleep. *J. Therm. Biol.* **37**, 392–396 (2012).

58. Kumar, D., Mallick, H. N. & Kumar, V. M. Ambient temperature that induces maximum sleep in rats. *Physiol. Behav.* **98**, 186–191 (2009).

59. Ray, B., Mallick, H. & Kumar, V. M. Role of the medial preoptic area in thermal preference of rats. *Indian J. Physiol. Pharmacol.* **45**, 445–450 (2001).

60. Refinetti, R. & Horvath, S. M. Thermopreferendum of the rat: Inter- and intra-subject variabilities. *Behav.*

Neural Biol. **52**, 87–94 (1989).

61. Briese, E. Rats prefer ambient temperatures out of phase with their body temperature circadian rhythm. *Brain Res.* **345**, 389–393 (1985).

62. Gordon, C. J. 24-Hour control of body temperature in rats. I. Integration of behavioral and autonomic effectors. *Am. J. Physiol. - Regul. Integr. Comp. Physiol.* **267**, (1994).

63. Brown, J. W. & Pham-Le, N. M. The effect of thermopreference on circadian thermoregulation in sprague-dawley and fisher 344 rats. *J. Therm. Biol.* **37**, 309–315 (2012).

64. Gordon, C. J. Relationship between preferred ambient temperature and autonomic thermoregulatory function in rat. *Am. J. Physiol. - Regul. Integr. Comp. Physiol.* **252**, R1130–R1137 (1987).

65. Gordon, C. J., Lee, K. L. A., Chen, T. L. A., Killough, P. & Ali, J. S. Dynamics of behavioral thermoregulation in the rat. *Am. J. Physiol. - Regul. Integr. Comp. Physiol.* **261**, 705–711 (1991).

66. Gordon, C. J. Simultaneous measurement of preferred ambient temperature and metabolism in rats. *Am. J. Physiol. - Regul. Integr. Comp. Physiol.* **254**, (1988).

67. Villarreal, J. A. *et al.* Thermal environment affects morphological and behavioral development of Rattus norvegicus. *Physiol. Behav.* **91**, 26–35 (2007).

68. Jans, J. E. & Leon, M. Determinants of mother-young contact in Norway rats. *Physiol. Behav.* **30**, 919–935 (1983).

69. Hillebrand, J. J. G. *et al.* Voluntary access to a warm plate reduces hyperactivity in activity-based anorexia. *Physiol. Behav.* **85**, 151–157 (2005).

70. Manser, C. E., Morris, T. H. & Broom, D. M. An investigation into the effects of solid or grid cage flooring on the welfare of laboratory rats. *Lab. Anim.* **29**, 353–363 (1995).

71. Manser, C. E., Elliott, H., Morris, H. & Broom, D. M. The use of a novel operant test to determine the strength of preference for flooring in laboratory rats. *Lab. Anim.* **30**, 1–6 (1996).

72. Krohn, T. C. & Hansen, A. K. Weighing used for the automatic registration of preferences when testing rats. *Scand. J. Lab. Anim. Sci.* **28**, 223–229 (2001).

73. Syme, L. A. & Syme, G. J. Spatial interaction between cage and test environments: Position preferences of young isolated and pair-housed rats. *Anim. Learn. Behav.* **4**, 396–400 (1976).

74. Heikkilä, M., Sarkanen, R., Voipio, H. M., Mering, S. & Nevalainen, T. Cage position preferences of rats. *Scand. J. Lab. Anim. Sci.* **28**, 65–74 (2001).

75. Voipio, H. M. *et al.* Cage Material and Food Hopper as Determinants. *Scand. J. Lab Anim. Sci.* **35**, 69–77 (2008).

76. Blom, H. J. M., Tintelen, G. Van, Baumans, V., Broek, J. Van Den & Beynen, A. C. Development and application of a preference test system to evaluate housing conditions for laboratory rats. *Appl. Anim. Behav. Sci.* **43**, 279–290 (1995).

77. Le Magnen, J. Effect of a multiplicity of food stimuli on the amount eaten by the rat. *Appetite* **33**, 36–39 (1999).

78. Naim, M., Brand, J. G., Christensen, C. M., Kare, M. R. & Van Buren, S. Preference of rats for food flavors and texture in nutritionally controlled semi-purified diets. *Physiol. Behav.* **37**, 15–21 (1986).

79. Wadhera, D. *et al.* The rewarding effects of number and surface area of food in rats. *Learn. Behav.* **46**, 242–255 (2018).

80. Hegab, I. M. *et al.* Effects of physical attributes and chemical composition of novel foods on food selection by Norway rats (Rattus norvegicus). *J. Pest Sci. (2004).* **87**, 99–106 (2014).

81. Welker, W. I. & King, W. A. Effects of stimulus novelty on gnawing and eating by rats. *J. Comp. Physiol. Psychol.* **55**, 838–842 (1962).

82. Johnson, S. R., Patterson-Kane, E. G. & Niel, L. Foraging enrichment for laboratory rats. *Anim. Welf.* **13**, 305–312 (2004).

83. Labouré, H., Saux, S. & Nicolaidis, S. Effects of food texture change on metabolic parameters: Short- and long-term feeding patterns and body weight. *Am. J. Physiol. - Regul. Integr. Comp. Physiol.* **280**, 780–789 (2001).

84. Sako, N., Okamoto, K., Mori, T. & Yamamoto, T. The hardness of food plays an important role in food selection behavior in rats. *Behav. Brain Res.* **133**, 377–382 (2002).

85. Burgdorf, J. & Panksepp, J. Tickling induces reward in adolescent rats. *Physiol. Behav.* **72**, 167–173 (2001).

86. Bombail, V. *et al.* Odour conditioning of positive affective states: Rats can learn to associate an odour with being tickled. *PLoS One* **14**, 1–17 (2019).

87. Paredes-Ramos, P. *et al.* Tickling in juvenile but not adult female rats conditions sexual partner preference. *Physiol. Behav.* **107**, 17–25 (2012).

88. Krohn, T. C., Hansen, A. K. & Dragsted, N. The impact of cage ventilation on rats housed in IVC systems. *Lab. Anim.* **37**, 85–93 (2003).

89. Krohn, T. C., Hansen, A. K. & Dragsted, N. The impact of low levels of carbon dioxide on rats. *Lab. Anim.* **37**, 94–99 (2003).

90. Patterson-Kane, E. G., Hunt, M. & Harper, D. Short Communication: Rat's Demand for Group Size. *J. Appl. Anim. Welf. Sci.* **7**, 267–272 (2004).

91. Sorensen, D. *et al.* Preferences for limited versus no contact in SD rats. *Lab. Anim.* **44**, 274–277 (2010).

92. Cloutier, S. & Newberry, R. C. Physiological and behavioural responses of laboratory rats housed at different tier levels and levels of visual contact with conspecifics and humans. *Appl. Anim. Behav. Sci.* **125**, 69–79 (2010).

93. Krohn, T. C., Salling, B. & Hansen, A. K. How do rats respond to playing radio in the animal facility? *Lab. Anim.* **45**, 141–144 (2011).

94. Marr, J. N. & Gardner, L. E. Early olfactory experience and the later social behavior in the rat: preference, sexual responsiveness, and care of young. *J. Genet. Psychol.* **107**, 167–174 (1965).

95. Montgomery, K. C. The relation between fear induced by novel stimulation and exploratory drive. *J. Comp. Physiol. Psychol.* **48**, 254 (1955).

96. Hall, C. S. Emotional behavior in the rat. I. Defecation and urination as measures of individual differences in emotionality. *J. Comp. Psychol.* **18**, 385 (1934).

97. Pich, E. M. & Samanin, R. A two-compartment exploratory model to study anxiolytic/anxiogenic effects of drugs in the rat. *Pharmacol. Res.* **21**, 595–602 (1989).

98. Porsolt, R. D., Anton, G., Blavet, N. & Jalfre, M. Behavioural despair in rats: A new model sensitive to antidepressant treatments. *Eur. J. Pharmacol.* **47**, 379–391 (1978).

99. Pellow, S., Chopin, P., File, S. E. & Briley, M. 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 (1985).

100. Prut, L. & Belzung, C. The open field as a paradigm to measure the effects of drugs on anxiety-like behaviors: A review. *Eur. J. Pharmacol.* **463**, 3–33 (2003).

101. Borsini, F., Podhorna, J. & Marazziti, D. Do animal models of anxiety predict anxiolytic-like effects of antidepressants? *Psychopharmacology (Berl).* **163**, 121–141 (2002).

102. Molendijk, M. L. & de Kloet, E. R. Immobility in the forced swim test is adaptive and does not reflect

depression. *Psychoneuroendocrinology* **62**, 389–391 (2015).

103. Borsini, F. & Meli, A. Is the forced swimming test a suitable model for revealing antidepressant activity? *Psychopharmacology (Berl).* **94**, 147–160 (1988).

104. West, A. P. Neurobehavioral studies of forced swimming: The role of learning and memory in the forced swim test. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **14**, 863–877 (1990).

105. Carvalho, C., Herrmann, K., Marques, T. A. & Knight, A. Time to Abolish the Forced Swim Test in Rats for Depression Research? *J. Appl. Anim. Ethics Res.* 1–9 (2021) doi:10.1163/25889567-bja10026.

106. Dichter, G. S., Smoski, M. J., Kampov-Polevoy, A. B., Gallop, R. & Garbutt, J. C. Unipolar depression does not moderate responses to the sweet taste test. *Depress. Anxiety* **27**, 859–863 (2010).

107. Berlin, I., Givry-Steiner, L., Lecrubier, Y. & Puech, A. J. Measures of anhedonia and hedonic responses to sucrose in depressive and schizophrenic patients in comparison with healthy subjects. *Eur. Psychiatry* **13**, 303–309 (1998).

108. Campos, A. C., Fogaça, M. V., Aguiar, D. C. & Guimarães, F. S. Animal models of anxiety disorders and stress. *Rev. Bras. Psiquiatr.* **35**, 101–111 (2013).

109. Krohn, T. C., Hejgaard, K. & Hansen, A. K. Methods for General Assessment of the Welfare of Laboratory Rats. *Acta Agric. Scand. A Anim. Sci.* **51**, 118–123 (2001).

110. Ralph, C. R. & Tilbrook, A. J. Invited Review: The usefulness of measuring glucocorticoids for assessing animal welfare. *J. Anim. Sci.* **94**, 457–470 (2016).

111. Ruhé, H. G., Mason, N. S. & Schene, A. H. Mood is indirectly related to serotonin, norepinephrine and dopamine levels in humans: A meta-analysis of monoamine depletion studies. *Mol. Psychiatry* **12**, 331–359 (2007).

112. Perona, M. T. G. *et al.* Animal models of depression in dopamine, serotonin, and norepinephrine transporter knockout mice: Prominent effects of dopamine transporter deletions. *Behav. Pharmacol.* **19**, 566–574 (2008).

113. Shohayeb, B., Diab, M., Ahmed, M. & Ng, D. C. H. Factors that influence adult neurogenesis as potential therapy. *Transl. Neurodegener.* **7**, (2018).

114. Paul, E. S. *et al.* Assessing animal welfare: a triangulation of preference, judgement bias and other candidate welfare indicators. *Anim. Behav.* **186**, 151–177 (2022).

115. Cait, J., Cait, A., Scott, R. W., Winder, C. B. & Mason, G. J. Conventional laboratory housing increases morbidity and mortality in research rodents: results of a meta-analysis. *BMC Biol.* **20**, (2022).

116. Leung, V., Rousseau-Blass, F., Beauchamp, G. & Pang, D. S. J. Arrive has not arrived: Support for the arrive (animal research: Reporting of in vivo experiments) guidelines does not improve the reporting quality of papers in animal welfare, analgesia or anesthesia. *PLoS One* **13**, 1–13 (2018).

117. United Kingdom Home Office. Code of Practice for the Housing and Care of Animals Bred, Supplied or Used for Scientific Purposes. (2014).

118. Clarkson, J. M., Dwyer, D. M., Flecknell, P. A., Leach, M. C. & Rowe, C. Handling method alters the hedonic value of reward in laboratory mice. *Sci. Rep.* **8**, 1–8 (2018).

119. Gouveia, K. & Hurst, J. L. Reducing Mouse Anxiety during Handling: Effect of Experience with Handling Tunnels. *PLoS One* **8**, (2013).

120. Gouveia, K. & Hurst, J. L. Improving the practicality of using non-aversive handling methods to reduce background stress and anxiety in laboratory mice. *Sci. Rep.* **9**, 1–19 (2019).

121. Hurst, J. L. & West, R. S. Taming anxiety in laboratory mice. *Nat. Methods* **7**, 825–826 (2010).

122. Leidinger, C. S., Kaiser, N., Baumgart, N. & Baumgart, J. Using Clicker Training and Social Observation to Teach Rats to Voluntarily Change Cages. *J. Vis. Exp.* (2018) doi:10.3791/58511.

123. Schapiro, S. J., Perlman, J. E., Thiele, E. & Lambeth, S. Training nonhuman primates to perform behaviors useful in biomedical research. *Lab Anim. (NY).* **34**, 37–42 (2005).

124. Laule, G. & Whittaker, M. Enhancing nonhuman primate care and welfare through the use of positive reinforcement training. *J. Appl. Anim. Welf. Sci.* **10**, 31–38 (2007).

125. Rutledge, R. B., Skandali, N., Dayan, P. & Dolan, R. J. A computational and neural model of momentary subjective well-being. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 12252–12257 (2014).

126. Eldar, E., Rutledge, R. B., Dolan, R. J. & Niv, Y. Mood as Representation of Momentum. *Trends Cogn. Sci.* **xx**, 1–10 (2015).

127. Bassett, L. & Buchanan-Smith, H. M. Effects of predictability on the welfare of captive animals. *Appl. Anim. Behav. Sci.* **102**, 223–245 (2007).

128. Joffe, J. M., Rawson, R. A. & Mulick, J. A. Control of their environment reduces emotionality in rats. *Science (80-.).* **180**, 1383–1384 (1973).

129. Swanson, J. C. & Morrow-Tesch, J. Cattle transport: Historical, research, and future perspectives. *J. Anim. Sci.* **79**, E102 (2001).

130. de Witte, K. Development of the Australian Animal Welfare Standards and Guidelines for the Land Transport of Livestock: Process and philosophical considerations. *J. Vet. Behav. Clin. Appl. Res.* **4**, 148–156 (2009).

131. Knutson, B., Burgdorf, J. & Panksepp, J. Ultrasonic vocalizations as indices of affective states in rats. *Psychol. Bull.* **128**, 961–977 (2002).

132. Brudzynski, S. M. Biological Functions of Rat Ultrasonic Vocalizations, Arousal Mechanisms, and Call Initiation. *BRAIN Sci.* **11**, (2021).

133. Coffey, K. R., Marx, R. G. & Neumaier, J. F. DeepSqueak: a deep learning-based system for detection and analysis of ultrasonic vocalizations. *Neuropsychopharmacology* **44**, 859–868 (2019).

134. Ouzzani, M., Hammady, H., Fedorowicz, Z. & Elmagarmid, A. Rayyan-a web and mobile app for systematic reviews. *Syst. Rev.* **5**, 1–10 (2016).

Fig. 1| Flow Diagram illustrating the number (n) of articles included at each stage of the literature review

Fig. 2| A histogram of the number of articles that met the inclusion criteria and their year of publication

Fig. 3| The number of articles using different types of refinements

Fig. 4| The number of articles using each of the top 30 most common outcome measures (excluding preference). The following abbreviations are used: OF for open field test, EPM for elevated plus maze, CORT for corticosteroids, MWM for Morris water maze, FST for forced swim test, NOT for novel object test, HR for heart rate, BP for blood pressure.

Fig. 5| The number of articles using different measures of preference across the different refinements. Conditioned place preference is abbreviated as CPP.

Table 1| The refinements identified by the review: each (putative) refinement type; the number of articles identified for each refinement type; variants of each refinement (where relevant), including the proportion of articles within the refinement using each variant; other refinements that were manipulated concurrently with the target refinement (where this was >50%) and the percentage of articles using those additional refinements. Supplementary Table 2 provides ID numbers for these articles for cross-referencing with the full dataset provided in Supplementary Table 1.

Refinement type	Number of studies	Variants of the refinement type (% of the studies)	Refinements co-occurring (>50% of studies) with the target refinement type
Cage contents	533	Provision of/change to cage objects (68.3%) Provision of/change to shelters (53.8%) Provision of running wheels (46.9%) Provision of/change to substrates (27.8%)	Cage type (63.6%) Cagemate number (51.4%)
Cage type	403	Increased cage dimensions (95.5%) Provision of compartments/levels (45.2%) Change to cage material (21.8%).	Cage contents (84.1%) Cagemate number (69.5%).
Cagemate number	363	2 v 1 (9.9%) 4 v 1 (8.0%) 3 v 1 (6.3%) 10 v 2 (2.8%) 4 v 2 (2.8%) 6 v 1 (2.8%) 8 v 1 (2.8%) Other (64.7%)	Cage contents (75.8%) Cage type (77.4%)
Neonatal handling	216	Dam and pup separation (69.9%) Dam and pup separation & pup tactile stimulation, e.g. with paintbrush (30.1%)	-
Foraging opportunities	72	Provision of treats (48.6%) Changing food location (16.7%) Provision of foraging device (15.3%) Multiple food locations (11.1%) Change to food flavour (2.8%) Change to food shape (2.8%)	Cage contents (72.2%) Cage type (68.1%) Cagemate number (54.2%)

		Change to food texture (2.8%) Change to food colour (1.4%) Change to food size (1.4%) Change to food type (1.4%) Change to food provision timing (1.4%)	
Handling type	58	Habituation (94.8%) Handler familiarity (1.7%) Handler identity (1.7%) Handling method (1.7%)	Cage contents (62.1%) Cage type (58.6%) Cagemate number (55.2%)
Playpen	57	Objects (89.5%) Additional space (87.7%) Companions (40.4%) Foraging opportunities (5.3%) Water maze training (1.8%)	-
Noise	32	Acoustic stimulation (90.6%) Noise reduction (6.3%) Noise type (3.1%)	-
Lighting	25	Photoperiod duration (40.0%) Lighting intensity (36.0%) Light colour (16.0%) Lighting as enrichment (8.0%) Timing of procedures (8.0%)	-
Temperature	19	Temperature gradient/areas of different temperatures (94.7%) Provision of warm plate (5.3%)	-
Gentling	16	-	-
Tickling	15	-	-
Cleaning regime	12	Cleaning frequency (50.0%) Soiled materials in clean cage (41.7%) Training rats to move to a clean cage (8.3%)	-
Olfactory stimulation	12	-	Cage contents (66.7%) Cage type (58.3%) Noise (50.0%)
Cage position	9	Position of the cage on a rack (33.3%) Locations varying in footfall (33.3%) Housing away from the main colony room (33.3%)	Cage type (55.6%) Cagemate number (55.6%)
Social contact	9	Cage barrier allowing degree of contact between rats (77.8%)	

		Proximity of cage to other rats (22.2%)	Cagemate number (66.7%)
Group composition	5	Male:Female ratio (40.0%) Exposure rats to pups (40.0%) Same/different strains (20.0%)	-
Ventilation	4	IVC vs. open-topped cages (50.0%) CO_2 levels (25.0%) Air changes per hour (25.0%) Air speed (25.0%)	-
Extended weaning age	3	21-22 vs. 30 (33.3%) 21-22 vs. 35 (33.3%) 21-22 vs. 51 (33.3%)	-
Reiki (a form of alternative medicine in which practitioners transfer 'energy' from themselves to the individual to be healed)	2	-	-
Cage controllability	1	Food, water, and lighting controlled by lever vs. yoked control (100%)	-
Interspecific colony room	1	Rat-only colony room vs. mouse and rat colony room (100%)	-
Transportation	1	Transporting by hand vs. cart (100%)	Cagemate number (100%)