STUDY PROTOCOL

Stage 1 Registered Report: Refinement of tickling protocols to improve positive animal welfare in laboratory rats [version 2; peer review: 1 approved, 1 approved with reservations]

Vincent Bombail1, Sarah M. Brown2, Jessica E. Martin2, Simone L. Meddle2, Michael Mendl3, Emma S.J. Robinson4, Tayla J. Hammond1,2, Birte L. Nielsen5, Megan R. LaFollette6, Ignacio Vinuela-Fernandez7, Emma K.L. Tivey2, Alistair B. Lawrence1,2

1Animal Behaviour and Welfare group, Scotland's Rural College (SRUC), Edinburgh, EH9 3RG, UK
2The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, EH25 9RG, UK
3Bristol Veterinary School, University of Bristol, Bristol, BS40 5DU, UK
4School of Physiology, Pharmacology and Neuroscience, University of Bristol, Bristol, BS8 1TD, UK
5Universities Federation for Animal Welfare, Wheathampstead, AL4 8AN, UK
6The North American 3Rs Collaborative, Denver, CO, 80202, USA
7Bioresearch Veterinary Services, University of Edinburgh, Edinburgh, EH16 4SB, UK

First published: 15 Sep 2022, 11:1053
https://doi.org/10.12688/f1000research.125649.1
Latest published: 18 Nov 2022, 11:1053
https://doi.org/10.12688/f1000research.125649.2

Open Peer Review

Approval Status ?

version 2
(revision)
18 Nov 2022

version 1
15 Sep 2022

1. Timothy M. Errington, Centre for Open Science, Charlottesville, USA
2. Theresa M. Kisko, KU Leuven, Leuven, Belgium

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

Abstract

Rat tickling is a heterospecific interaction for experimenters to mimic the interactions of rat play, where they produce 50 kHz ultrasonic vocalisations (USV), symptoms of positive affect; tickling can improve laboratory rat welfare. The standard rat tickling protocol involves gently pinning the rat in a supine position. However, individual response to this protocol varies. This suggests there is a risk that some rats may perceive tickling as only a neutral experience, while others as a positive one, depending on how tickling is performed. Based on our research experiences of the standard tickling protocol we have developed a playful handling (PH) protocol, with reduced emphasis on pinning, intended to mimic more closely the dynamic nature of play.

We will test whether our PH protocol gives rise to more uniform increases in positive affect across individuals relative to protocols involving pinning. We will compare the response of juvenile male and female Wistar rats as: Control (hand remains still against the side of the test arena), P0 (PH with no pinning), P1 (PH with one pin), P4 (PH with four pins). P1 and P4 consist of a background of PH, with treatments involving administration of an increasing dosage of
pining per PH session.

We hypothesise that rats exposed to handling protocols that maximise playful interactions (where pinning number per session decreases) will show an overall increase in total 50 kHz USV as an indicator of positive affect, with less variability. We will explore whether behavioural and physiological changes associated with alterations in PH experience are less variable.

We propose that maximising the numbers of rats experiencing tickling as a positive experience will reduce the variation in response variables affected by tickling and increase the repeatability of research where tickling is applied either as a social enrichment or as a treatment.

**Keywords**
rat tickling, playful handling, protocol, positive animal welfare, affective neuroscience, laboratory animal welfare

This article is included in the NC3Rs gateway.
Research highlights

Scientific benefits
Heterospecific play with rats can be used to study the physiology of positive affect and as a positive laboratory animal welfare measure.

3Rs benefits
Refinement: This work constitutes an attempt to (a) refine the tickling methods and provide further scientific credibility to tickling as a positive social enrichment for laboratory rats by potentially increasing the number of rats that enjoy tickling; (b) encourage wider uptake of tickling by studying the implications of tickling for experimental repeatability.

Practical benefits
From our experience there is general interest in tickling as a form of social enrichment but potential users might have reservations. We anticipate that a better understanding of biological consequences of welfare improvement measures could enhance uptake in the wider research community.

Current applications
None.

Potential applications
We aim to promote a validated playful handling protocol as a refinement to the tickling methodologies currently used.

Introduction

Research question: background, importance and relevance
During 2020 in the UK, 201,600 experimental procedures were carried out on rats, the second most investigated laboratory species (UK Home Office, 2021). Worldwide there has recently been an increasing emphasis on positive welfare, and the introduction of positive experiences to improve animals’ quality of life (Lawrence et al., 2019; Makowska & Weary, 2021). The use of manipulations to enhance positive welfare in laboratory animals has been shown not to be a source of variability (Kentner et al., 2021), and a possible contribution towards improved model validity and data reproducibility (Loss et al., 2021; van de Weerd et al., 2004; Würbel & Garner, 2007). It is therefore likely that ‘happy animals make good science’ (Grimm, 2018; Poole, 1997). For instance, positive welfare might reduce stress caused by aversive procedures (Hinchcliffe et al., 2022). For laboratory rats, a positive handling approach referred to as rat tickling is proposed as an effective and practical approach to positively improve their welfare (https://www.nc3rs.org.uk/rat-tickling; https://na3rsc.org/rodent-handling/). Rat tickling with the human hand was developed to study positive emotions and ultrasonic vocalisations (USVs) in rats by mimicking playful interactions between rats (Panksepp, 1998, Figure 1i). The standard tickling protocol involves initial finger contact with the nape of the neck before flipping the rat and gently pinning it in a supine position whilst making rapid finger movements focusing on the ventral surface as used in human tickling (Cloutier et al., 2018; Figure 1ii).

Research (see Bombail et al., 2021; LaFollette et al., 2017 for reviews) has shown that when rats are tickled by the human hand, they often produce 50 kHz USVs, which are understood to indicate positive affect in rats (e.g., Barker, 2018, Figure 2i). Other evidence suggests that rats find tickling rewarding as in addition to production of USVs, tickled rats show a reduced latency for approaching the human hand and can be trained to perform operant responses for tickling (Burgdorf & Panksepp, 2001; LaFollette et al., 2017).

However, individual responses to the standard tickling protocol indicate that there is a risk that some rats may not perceive tickling to be a positive experience. We summarise that evidence below (see also Bombail et al., 2021).
It is well accepted that rats vary in their response to tickling: The NC3Rs web-site on tickling is clear on this point: “But I hate being tickled; shouldn’t rats hate it too? This is a common concern, and it is true that some rats enjoy it more than others” (https://www.nc3rs.org.uk/tickling-rats-improved-welfare). We initially developed doubts over aspects of the standard tickling protocol (Bombail et al., 2019) with its regular use of pinning due to the individual variation in USVs.
and behaviour we observed in response to this protocol (e.g., Hinchcliffe et al., 2020; see Figure 2ii). Other work has reported significant levels of individual variation in response to tickling (LaFollette et al., 2017).

Individual responses to tickling have welfare implications: Rat tickling response is associated with responses indicating affective states ranging from neutral to positive (e.g., Hinchcliffe et al., 2020), and hence tickling may not possibly always be welfare enhancing. We have shown that USVs in response to tickling are correlated with an independent test of affective state (an affective bias test) indicating the degree of emotional valence. Each data point represents, for each of the 16 rats, the 50 kHz USV response to tickling and the substrate preference in the Affective Bias Test, indicative of the degree of emotional valence. (ii) USVs in playfully handled (PH) rats compared to a control group on the 1st and 5th session of treatment (redrawn with data available online and reproduced with permission from Bombail et al., 2019). By the 5th session, all PH rats showed an increase in USV production in contrast to panel (ii), which used the standard protocol where some rats show little or no increase in USVs.

We therefore propose that individual variation in response to the standard tickling protocol is a significant issue when promoting the widespread use of tickling as an approach to improving rat welfare. Due to our experiences using the standard tickling protocol we have developed a playful handling (PH) protocol intended to better mimic the dynamic nature of play, giving the rat more choice over how it is tickled (Bombail et al., 2019, Figure 1iii). The PH protocol is described by Bombail et al. (2019) as the handler using one hand to touch, tickle, and play with the rat for 20 second periods interspersed with 20 second pauses. During the active periods, the handler mimicked the rough-and-tumble play seen in adolescent rats, with the hand tickling, chasing and pinning the rat, depending on its response. Our data suggest that the PH protocol resulted in a more uniform increase in USVs across individuals relative to the standard protocol. Figure 2iii adapted from Bombail and colleagues (Bombail et al., 2019) shows USVs in playfully handled rats vs. a control group (who did not play in the day of recording) on the 1st and 5th session of treatment. By the 5th session, all playfully handled rats showed an increase in USV production in contrast to Figure 2ii, which used the standard protocol and where some rats show little or no increase in USVs.
We conclude that this evidence calls for a scientifically based refinement of tickling to ensure that the positive effects of tickling are maximised and that individual differences in response to tickling are minimised with benefits to rat welfare and the reproducibility of research where tickling is applied as a treatment or more generally as an enrichment. In this report, we define tickling as the standard protocol (e.g., Cloutier et al., 2018) and, although it was initially coined as an equivalent term to tickling (Cloutier et al., 2018), PH as the approach outlined in Bombail et al. (2019).

Hypotheses

PH consists in mirroring playful behaviour by allowing for more flexibility between the human handler and the rat and consequently we hypothesise PH will induce a more homogeneous positive affective response. We also hypothesise PH will lead to reduced variability of biological outcomes.

Confirmatory Hypotheses: (Confirmatory Hypothesis 1) We hypothesise that rats exposed to handling protocols that maximise playful interactions (i.e., where pinning number per session decreases) will show an overall increase in total 50 kHz USVs as an indicator of positive affect. (Confirmatory Hypothesis 2) We also predict that 50kHz USV measures will be less variable in treatments with less pinning.

Exploratory Hypothesis 1: We hypothesise that the behavioural and physiological changes associated with alterations in PH experience will be less variable when rats are exposed to handling protocols that maximise playful interactions (where pinning number per session decreases). There is no previous data on this question, we cannot generate power calculations and pre-register this hypothesis, this is therefore exploratory work.

Study timeline

The diagram presented in Figure 3 synthesises the timeline for each three-week long cohort studied. The experimental part of this study will be conducted using eight such cohorts, each made up of 16 animals. The timing of these cohorts will be staggered, so experimental cycles overlap, with 16 animals brought into the facility every week, following successful peer review of this Stage 1 Registered Report submission. A step-by-step protocol is available in Box 1. The experimental work for this study should be completed by 21 December 2022, the data analysis should be completed by 1 July 2023.

Methods

Basic methodological framework/identification strategy

Our experimental design is a pseudorandomised control trial where every week, each of the eight cages of two male or two female rats will be assigned to one of four treatments (control, P0, P1, P4). Treatment allocation is described below.

Prior to the experiment and from day 1 to day 6 of the first cohort, various optimisation and preparation test recordings will be carried out, this will result in a general level of noise as throughout the experiment (when several cohorts are housed and tested simultaneously)

Day 1. Reception of rats and group formation (morning).
- Temporary mark on tail,
- Weigh each rat,
- Form eight cage pairs based on: (1) sex, (2) provenance (always housed with an individual from a different litter), (3) similar cage weight (monitor total treatment group average as the values must balance by the 8th cohort).

Day 1 to day 6. Rats habituate to facility environment.

Day 7. Habituation to PH testing arena.
- Collect faecal pellets for faecal corticosterone metabolite (FCM) analysis. Weigh rats. Habituation 1: in the morning (start of dark phase), for each of the eight cages, transfer pair into the play arena and leave together for five minutes. Remove any faeces from arena if necessary.
- Habituation 2: in the afternoon, for each of the eight cages, transfer pair into the play arena and leave together for five minutes. Remove any faeces from arena if necessary.

Day 8. Habituation to PH testing arena.
- Preparation of the next cohort (Day 1)
- Habituation 3: in the morning, for each of the 16 individuals, place into the play arena and leave for five minutes. Remove any faeces from arena if necessary.
- Habituation 4: in the afternoon, for each of the 16 individuals, place into the play arena and leave for five minutes. Remove any faeces from arena if necessary.
- Cage bedding is changed.

Day 9 to day 11. Experimental treatment.
- The cage will be moved close to the testing arena five minutes prior to the first of the pair of rats being introduced in the test arena. Both rats will be filmed during this anticipatory phase, to investigate their behaviour.
- From each cage, a rat taken individually into the arena (order of passage pseudo-randomised), recorded for a one min anticipation phase and treated according to the schedule (control, P0, P1 or P4) for 30 seconds. Ultrasonic vocalisation (USV) are recorded. Arena wood shavings stirred between each animals (for odour homogenisation).
- Following this treatment, rats will be gently picked up and returned to the home cage. Their cage mate will then be immediately placed in the arena for testing and the same tickling treatment repeated.
- The home cage is covered with a cardboard screen, behind the plastic curtains protecting other cages from emotional contagion via USV.
- Two hours later, each cages are placed back on the housing rack.

Days 12 and 13 are the weekend, animals are undisturbed except for welfare checks.

- Morning of Day 14 only: Weigh rats
- Day 15 only: preparation of the next cohort (Day 1)
- The cage will be moved close to the testing arena five minutes prior to the first of the pair of rats being introduced in the test arena. Both rats will be filmed during this anticipatory phase, to investigate their behaviour.
- From each cage, a rat taken individually into the arena (order of passage pseudo-randomised), recorded for a one-minute anticipation phase and treated according to the schedule (control, P0, P1 or P4) for 30 seconds. USV are recorded. Arena wood shavings stirred between each animals (for odour homogenisation).
- Following this treatment, rats will be gently picked up and returned to the home cage. Their cage mate will then be immediately placed in the arena for testing and the same tickling treatment repeated.
- The home cage is covered with a cardboard screen, behind the plastic curtains protecting other cages from emotional contagion via USV.
- Two hours later, each cages are placed back on the housing rack.

Day 17. Collect faecal pellets for FCM analysis. Elevated Plus Maze (EPM) testing late morning (middle of dark phase). Standard EPM protocol, EPM will be cleaned between each animal.

Day 18. Open Field (OF) testing.
- Rats are tested in the OF in the morning, Standard OF protocol, OF will be cleaned between each animal.
- Rat are weighed
- Rats are injected with an overdose of pentobarbital and tissues sampled for molecular studies to investigate physiological impact of experimental treatments. Collect faecal pellets for FCM analysis.
Within a cage, individuals cannot be considered independent, as they might influence each other in response to the treatment through emotional contagion, we will therefore fit the cage as a random effect for all treatments and both sexes.

**Ethical approval**

All animal work will be carried out at the Roslin Institute, Edinburgh, U.K., in accordance with the U.K. Animals (Scientific Procedures) Act 1986 and reported in compliance with the ARRIVE guidelines (Bombail, 2022). This work does not involve aversive procedures (and is deemed less harmful and distressful than skilled insertion of a hypodermic needle according to good veterinary practice) and therefore not legally required to have UK Home Office approval. As recommended (Olsson et al., 2022), we collectively discussed welfare and ethical considerations with the Named Animal Care Welfare Officers, veterinarians and scientists involved in the study. Following consideration by the University of Edinburgh’s Animal Welfare and Ethical Review Body (reference NC/W001209/1, 26 July 2022, Establishment number X212DDDBD), it was confirmed the work was subthreshold to regulation under the U.K. Animals (Scientific Procedures) Act 1986, and the committee approved it as detailed in this paper. This study was also approved by the Veterinary Ethical Review Committee of the Royal (Dick) School of Veterinary Studies, University of Edinburgh (reference 92.22, 2 September 2022).

**Experimental animals**

Practical and financial constraints require that we limit this study to a single strain. We will use Wistar Han rats (Research Resource Identification, RRID:RGD_2308816) purchased from Charles River Laboratories (CRL, Margate, UK). This outbred strain has been shown to be reliably responsive to tickling and PH (Bombail et al., 2019; Hammond et al., 2019) and is one of the commonly used strains for such experiments (23% of rat tickling studies used Wistar, Lafollette et al., 2017). We will purchase the rats from CRL (16 at a time), aged 3-4 weeks and will complete the research prior to their puberty around 7 weeks of age. Lafollette et al. (2017) found that over 60% of papers on tickling used rats within this time window. We have also recently found evidence of sex differences with female rats producing more USVs than males in response to tickling (Tivey, 2022) and each cohort will consist of equal numbers of males and females.

We aim to carry out the experiment in eight cohorts of n=8 cages (with two rats per cage) and equal numbers of males and females in each cohort. A total of 128 rats (64 males and 64 females) will be used in this study, in accordance with our power analysis detailed below.

To the greatest extent possible we will aim for a matching number of males and females coming from the same litter. Depending on availability, the number of litters sampled will be two to four per cohort (we are aiming for two litters per cohort, four males and four females from the same litter). Rat treatment allocation will be spread evenly across all litters to prevent bias due to genetic background and early life experience (i.e., each of the four individuals from each sex will be assigned to each of the four treatments). We have arranged for CRL to only provide animals from experienced parents (sire/dam will have been sired/nursed at least one previous litter). This will reduce variability in pup early life environment caused by differences in parental experience.

**Acclimation, treatment allocation, housing and husbandry**

On arrival at the Roslin Institute facility, the rats will be housed in a bespoke behaviour research room with an average room temperature of 22°C and relative humidity of 43%. The rats will be on a 12 h:12 h reverse light cycle (lights off at 06:00), to allow testing of the rats during their dark photoperiod when they are most active and to accommodate the preference of albino rats for lower light intensities (see one of our previous reports (Hammond et al., 2019)). All treatments and behavioural tests will be carried out between 8:00 and 16:00 (2 hours from lights on or off). The allocation of rats to cages will be carried out in the following order: sex, litter of origin and body weight. Rats will be weighed and allocated to home cage as pairs of the same sex (but never from the same litter). As much as feasibly possible, we will aim to house together rats of similar weight. Pairs will be assigned to a treatment using a random list generator. We aim to homogenise weights between groups. For each sex, rat weight will be averaged for each treatment, and to avoid potential bias caused by uneven distribution of body weight between treatment groups, adjustments in treatment distribution will be made to balance evenly weights across treatments (e.g., we will assign heavier pairs of animals to treatments that previously contained lighter animals). Rats will be left to acclimatise for up to six days prior to handling.

The home cages are made of clear plastic with a metal mesh open-top lid (l x d x h: 612 mm x 435 mm x 216 mm), filled with 2 cm-deep woodchip bedding, shredded paper, a 20 cm-long red plastic tunnel, a 10 cm wooden chewstick and a 2 cm diameter wooden ball as enrichment (all as per normal practice in our animal facility). This size cage allows animals to rear and stand on their hind limbs. Rats will have *ad libitum* access to standard rat chow (current supplier: Teklad Global Rodent Maintenance Diet (14% protein); Envigo, UK) and tap water. The cages will be distributed across four tiers of a standard cage rack. Due to the varying lux levels and other potentially confounding variables, position of the home cages in the rack will be randomised across treatments (as in Hammond et al., 2019). The rats will be checked daily,
and nitrile gloves worn when handling the animals. To minimise handling stress, rats will be picked up gently by holding them behind their forelegs and then cupping them with both hands. Cage bedding will be changed once, halfway through the experiment (on the afternoon of day 8), according to the routine protocol (a small volume of soiled litter without faeces will be kept in the cage to maintain a familiar olfactory environment). Body weight will be measured routinely across the study (on days 1, 7, 14, 18).

**Experimental treatments**

A unique male experimenter (with prior experience in all rat playful interactions tested) will perform all treatments and behavioural tests. All experiments will be performed in the same room as where the animals are housed, as in previous studies (e.g., Hammond et al., 2019). Using curtains and physical distance, we will prevent emotional contagion phenomena involving USV (i.e., we have confirmed USV cannot be detected by other animals, using our recording equipment). Rats are used to a degree of audible activity in the room throughout the experiment. The play arena, placed behind fabric curtains 6 m away from the home cages, will be a Plexiglas square (60 cm) of depth 25 cm, filled with 3 cm-deep woodchip litter in which Wistar rats have played previously (over 200 sessions by 16 males rats took place by the start of the experiment). We keep the same litter for all experiments and remove the rare faeces produced by the animals (in our experience, exclusively produced during the habituation sessions). Litter is regularly stirred between animals. We use this method to guarantee the litter contains odorants from several rats, thereby standardising, diluting and minimising the impact of individual olfactory cues. This can be replicated in other laboratories by using a mixture of soiled litter and clean litter. Following experimental treatment (described below), cages will be left, loosely covered with cardboard screens, for 2 h in an area behind plastic curtains (a switched off Lateral Air Flow cabinet), before being returned to their usual home cage rack position. In the absence of this ‘cool down’ period, in our experience a marked rise in USV production over time can be otherwise be measured in control rats (Lam, 2017).

The rats will be habituated to handling and the experimental recording arena, to lessen any anxiety or fear due to novelty, as negative affective states inhibit play behaviour (reviewed in Ahloy-Dallaire et al., 2018). For habituation to the treatment arena, rats will be moved in their cage to be beside the treatment arena before being lifted gently into the arena, and left there to explore, first as a pair of cage mates (2 × 5 minutes), on the first day, and then alone (2 × 5 minutes) on the second day.

The aim of PH is to give all rats a similar amount of hand contact (we approximate 70–80% of human-animal interaction time, e.g., about 24 out of 30 seconds), while being responsive to differences between individuals. The similar amount of physical contact for all individuals is important, since activation of the cortical area in response to touch has been shown to influence USV production (Ishiyama & Brecht, 2016). When not touching the rat, the hand can also approach the animal while the hand is in contact with the wood shavings, to provide sensory cues consistent with a play partner approaching (e.g., about 6 out of 30 seconds).

Another key component of PH is to be responsive to individual differences, which can range from more playful individuals who will quickly move (accelerate or turn) to face the hand, and display playful behaviour (e.g., joy jumps (Reinhold et al., 2019)) to less playful individuals who move less and more slowly and show little or no play behaviour. An example of PH with a responsive individual is available to watch online here. PH starts with presentation of the back of the experimenter’s hand for a brief sniff and exploration by the rat. The experimenter then starts playing with the rat using a rapid tickling finger motion on the sides, back and front of the rat (alternating region). The experimenter should focus on being playful (unexpected moves, fast change of hand movements, change of pace) rather than simply the physical act of tickling (e.g., Figure 1iii).

Pinning consists in approaching the rat during PH as described above. Thumb and middle fingers are then placed behind the front legs, and the animal is gently but swiftly moved onto its back, and upon landing, while the rat is pinned down, it is tickled on the front of its thorax and abdomen involving rapid tickling movements of fingers (e.g., see Cloutier et al., 2018).

In both PH and standard tickling, movements should reflect that rat play is boisterous (Cloutier et al., 2018) with finger pressure being firm and movements fast.

**Objective 1: Testing protocols varying in PH and pinning events against USV as a measure of affective state**

**Confirmatory Hypotheses**: (Confirmation Hypothesis 1) We hypothesise that rats exposed to handling protocols that maximise playful interactions (i.e., where pinning number per session decreases) will show an overall increase in total
50 kHz USVs as an indicator of positive affect. (Confirmation Hypothesis 2) We also predict that 50 kHz USV measures will be less variable in treatments with less pinning.

Objective 1 will be addressed in Phase 1 of the experiment (see Figure 3).

**Experimental design:** The aim of this objective is to refine the standard tickling protocol based on the rats’ affective response to treatments that vary in PH and number of pins. All rats will be exposed to six sessions of experimental human-animal interaction treatment, over eight days, interrupted by a two-day weekend break. We have chosen a standard session length of 30 seconds that reflects our previous work (Hinchcliffe et al., 2020) and also the practical reality that if tickling is to be used widely in practice it needs to be applied in a time efficient manner (LaFollette et al., 2019). The order of testing of cages, and rats within cage will be pseudo-randomised (a sequence for cage and individuals will be generated from a random list generator). The cage will be moved close to the testing arena five minutes prior to the first of the pair of rats being introduced in the test arena. Both rats will be filmed during this anticipatory phase, to investigate their behaviour.

Rats will be removed individually from the home cage, placed into the handling arena and, following an additional one-minute recording of anticipatory play behaviour, the experimental treatments applied:

P0: ‘PH’, where zero pins will be applied; the experimenter using one hand will touch, tickle and chase the rat in a manner that mimics rough and tumble play (Bombail et al., 2019; Figure 1iii).

P1: One pin will be applied at approximately 15 seconds from the start of the tickling session. The rat will be held and flipped into a supine position and the experimenter will gently pin the rat in this position whilst moving their fingers quickly and vigorously but gently on the belly, as is commonly used in human tickling. The pin will last for between 2–4 seconds and the rat will be allowed to right itself at the end of pinning (Cloutier et al., 2018; Figure 1ii). On either side of the pin the experimenter will engage in PH as in P0.

P4: Four pins will be applied from five seconds into the tickling session (four pins being the number delivered during a 15 second pinning session in the standard protocol (Cloutier et al., 2018)); pins will be applied as in P1 and the rat will be allowed to right itself after each pin with a short gap before the next pin. On either side of the pinnings the experimenter will engage in PH as in P0.

Control: We will apply the same control treatment as in our previous work (Bombail et al., 2019; Hammond et al., 2019) where experimenter places their hand against the inside of the arena for the entire session.

Following this treatment, rats will be gently picked up and returned to the home cage. Their cage mate will then be immediately placed in the arena for testing and the same tickling treatment repeated.
**Data collection**

We will record the behaviour during the tickling session using a digital HD camcorder. We will record USV using a high-quality microphone designed for recording ultrasonic vocalisations produced by bats (Pettersson M500-384 USB Ultrasound microphone) and Audacity recording software. The microphone will be placed over the centre of the arena, pointing downwards at the arena floor. Total USV per unit of time will be quantified by spectrogram analysis.

**Measures:**  (i) Vocalisations will be recorded during the handling sessions as in our previous work (Bombail et al., 2019; Hammond et al., 2019) and categorised into 50 kHz USVs and 22 kHz USVs; (ii) We will also record other quantitative behavioural responses to tickling as per our previous work (Hammond et al., 2019).

Our primary response variables for Confirmatory Hypotheses 1 and 2 will be:

(a) Total 50 kHz USV: We will analyse 50 kHz USV as indicators of positive affective state (e.g., Barker, 2018; Brudzynski, 2013). We define 50 kHz USV as calls with a peak frequency between 30 and 80 kHz and a short duration between 10–400 ms (based on Brudzynski, 2013 and our own empirical observations). We predict 50 kHz will be significantly increased in P0 and P1 treatments, where pinning is minimised, relative to P4, where play is restricted relative to pinning, and the Control treatment, where there is no play.

(b) Variability of 50 kHz USV production: we will compare individual variability across the different treatments. We predict a greater individual variability in positive affective states in treatments that maximise pinning.

Our secondary response variables will be:

(a) Categorisation of 50 kHz calls: we will analyse sub-types of USV following segregation based on their acoustic properties, according to published criteria (Wright et al., 2010). We will investigate whether certain USV types might be preferably associated with certain treatments.

(b) Anticipatory solitary play behaviours: We will also analyse anticipatory play behaviour, as this has been shown to be an indicator of positive affect. We will use the ethogram in Table 1, based on observations from our earlier work (Bombail et al., 2019; Champeil-Potokar et al., 2021; Hammond et al., 2019).

(c) Behavioural transitions: we will analyse the number of behavioural transitions (including play) that occur prior to treatment, as indicators of reward anticipation (Anderson et al., 2015).

(d) 22 kHz USV: We will also analyse 22 kHz USVs as indicators of negative affective state (Brudzynski, 2013). We define 22 kHz USV as calls with a frequency between 20–30 kHz, and a long duration between 500–3000 ms (Brudzynski, 2013). We predict overall very low number of events, clustered in treatments that maximise pinning.

**Analyses**

Data will be analysed in R through General Linear Mixed Models using the lme4 package (Bates et al., 2015), to test our Confirmatory Hypothesis that indicators of positive affect are increased in treatments P0, P1 where PH is increased and amount of pinning reduced, relative to P4 and Control (i.e., no play) treatments. Our fixed effects are tickling treatment (Control, P0, P1, P4), sex (male, female) and the tickling*Sex interaction. In addition to testing differences in average USV we will also test whether treatments P0, P1 have reduced the variability of response relative to P4 as we anticipate (using the Levene’s test for equality of variance or equivalent). As individuals from a pair cannot be considered independent, as they might influence each other in response to the treatment, we will fit the cage as a random effect for all treatments and both sexes.

**Objective 2: Testing the behavioural and physiological effects of protocols varying in PH and pinning events**

**Exploratory Hypothesis:** We hypothesise that the behavioural and physiological consequences of tickling will be less variable and more repeatable when rats are exposed to handling protocols that maximise playful interactions and where pinning episode number per session decreases. Due to the paucity of data on the impact PH might have on the...
variables we will measure, we cannot generate power calculations and pre-register this hypothesis, this is therefore exploratory work.

Objective 2 will be addressed in Phase 2 of the experiment (see Figure 3).

Experimental design: In Phase 2, the next day after the final playful or control interaction, rats will undergo behaviour testing with Elevated Plus Maze (EPM) and Open Field (OF), their order of passage will be pseudo-randomised. The EPM and OF test consist in quantifying exploration behaviour in 2 different structures, that both consist in an area with open spaces (the centre of the OF and the open arms of the EPM) and an area with more enclosed spaces (the periphery of the OF and the closed arms of the EPM). For rats, this ambiguity creates a conflict between the thigmotaxis (preference for proximity to walls) and the desire to explore open spaces, which is used to assess the emotional response to those novel experiences, and to test for anxiety levels. Following these tests, the rats will be humanely killed by anaesthetic overdose and tissues collected for analysis. Post mortem tissue measures will be critical to assess the impact of our treatments on variability and repeatability of commonly used measures of physiological function. Faecal pellets will be collected on days 17 (morning) and 18 (post cull) for non-invasive corticosterone metabolite analysis, in comparison to pellets collected before treatment starts on day seven.

Measures: (i) Behavioural measures of response to OF and EPM (Ethovision) will be collected to quantify anxiety behaviour prior to culling; (ii) Physiological measures of stress and inflammation: Faecal samples will be collected on day 18, at the end of Phase 2 (and compared to levels at the start and end of Phase 1) and assayed for corticosterone (Lepschy et al., 2007) using a commercially available corticosterone ELISA (Enzo LifeSciences). Point of cull plasma will be assayed for a set of inflammatory markers including the pro-inflammatory cytokines IL1a, IL1b, IL2, IL6, IFNγ and TNFα and the anti-inflammatory cytokines IL4, IL10, IL13, using a magnetic bead based multiplex ELISA (FIRELEX); our selection of inflammatory markers includes cytokines known to be sensitive in humans to negative and positive affect (Anisman & Merali, 2003; Stellar et al., 2015); (iii) Other tissues such as the brain, pituitary gland, heart, liver, spleen and gut will be collected at culling and flash frozen on dry ice and stored at -80°C for future analyses.

Our primary response variables will be:

a) EPM: increased markers of open arm exploration (% time and distance exploring the open arm) reflect reduced anxiety, based on rat aversion of open spaces (Kraeuter et al., 2019).

b) OF: decreased markers of thigmotaxis (% time spent and distance covered close to the peripheral walled area) reflect decreased anxiety, based on rat aversion of open spaces. Markers of activity (rearings, total distance covered, bouts of locomotion) will also be recorded since, although interpretation can be challenging (Denenberg, 1969), they do reflect arousal upon OF arena exploration.

Our secondary response variables will be:

a) Physiological and inflammation related variables: we will analyse markers regulated in affective states, and their variability around the treatment mean/median. We expect P0 and P1 to be more effective at inducing anti-inflammatory cytokines and reducing pro-inflammatory cytokines, and faecal corticosterone (Anisman & Merali, 2003; Lepschy et al., 2007; Stellar et al., 2015), in comparison to P4 and control treatments.

b) Body weight: we will keep a record of weight as an indicator of general body condition, we do not expect any treatment-related changes, as per our previous studies (Bombail et al., 2019; Hammond et al., 2019).

Analyses

Using the same approach as described for Objective 1, data will be analysed in R using Generalised Linear Mixed Models to ascertain the effects of tickling treatments and sex on behavioural and physiological measures. To test our Exploratory Hypothesis that the biological consequences of tickling will be less variable among rats and more repeatable within rats when rats are tickled using more playful protocols (P1, P0), we will analyse the residual variance from the undertaken models and explore intraclasse correlation coefficients of combined random and fixed effects.

Diagram of the experimental plan

A diagram of the experimental procedure is presented in Figure 3i. We also present a graphical representation of our predicted results, which will be tested in our experiments (Figure 3ii).
Further exploratory analyses
Brain tissue will be collected for future analysis of neurobiological consequences of the handling treatments. When funding and collaborations can be raised, it is our intention to further explore the biology of positive affective states through analysis of physiological functions that are associated with emotional experiences (e.g., immunology or metabolism). We will also analyse USV production in relation to specific USV types (Wright et al., 2010) using spectrograms from our recordings, with a view to correlating specific USV types to treatments and behaviours.

Statistical analysis
Specific aspects of data analysis are described in the descriptions of objectives 1 and 2 above. Prior to analysis, we will visualise and test the data for the assumptions of the cognate statistical approaches (data distribution normality or heteroskedasticity), and when necessary, we will use non-parametric alternatives or data transformation. To account for potential type I errors that may arise from multiple comparisons, we will use Bonferroni correction or equivalent (depending on data distribution). We will consider the p-value threshold of 0.05 as significant in our statistical comparisons.

Power analysis: Based on our previous data, the most variable response measure is total 50 kHz USVs with data taken from our own and other work (LaFollette et al., 2018; Tivey, 2022) where treatments come closest to matching our proposed treatments. We have used these data to provide us with means per primary fixed effect combinations (i.e., treatment x sex) and pooled standard deviations, with resulting simulations based on the smallest effect size (0.09). Power analysis and sample size estimations were conducted using GLIMMPSE software. The desired power was set at 0.8, the type I error rate at 0.05, the means scale factor was set at 1.0 and the variability scale factor at 1.3 using the Hotelling Lawley Trace test (Hotelling, 1931). Simulations incorporated the non-independence of rats within a cage (cage fitted as a random effect), via a calculated intraclass correlation of 0.17 based on previous data (LaFollette et al., 2018; Tivey, 2022). Additionally, repeated measures over six tickling treatment events/exposures were included in the power analysis. Our fixed effects are treatment (Control, P0, P1, P4), sex (male, female) and the tickling*Sex interaction. Based on simulations detailed above, a sample size of n=16 rats (i.e., n=8 cages) per treatment and sex combinations, or n=64 total experimental units (128 total animals), provides a power of 0.829. In the spirit of transparency, we will report effect sizes and uncertainty in our analyses.

Blinding: We will not be able to blind handlers during the preparatory and experimental phases, as the handlers need to know which handling method they will be using with each rat. We will also not be able to blind collection of quantitative behaviour during tickling sessions, as it will be obvious which handling method is being used in each session. We will blind all other aspects including collection of USVs, analysis of physiological samples, and collation and analysis of behavioural data.

Outlier extraction: We are aware of the high variability in USV response among rats, and are experienced in recording it, and therefore will not exclude rats based on their USV production. For physiological markers we will use data points within the dynamic range of the assay and will be checking replicate quality. In the unlikely event of outlier removal (based on the Grubb’s test (Grubs, 1969)), this will be explicitly described in the result section. Additionally, should any animal be excluded from the study for health reasons (e.g. injury or infection), we will report it in order to achieve full transparency about the fate of our experimental rats.

Study status
We are currently awaiting publication of this peer-reviewed protocol to start experimental data collection.

Study limitations
We will assign animals to treatments according to preset criteria to homogenise factors we assume might affect treatment response (sex, litter of origin and social cage environment, body weight). Our pseudo-randomisation procedure will therefore be more deterministic than the computer-run randomisation algorithm; cryptic biases might emerge upon data interpretation. Additionally, the experimenter will be aware of the treatments administered as the rats are handled. He will therefore not be blinded to the treatments. This is limited by the lack of discernible patterns (order of passage pseudo-randomisation) and impossibility to memorise treatments for each 64 cages at the time of analysis. Finally, we accept use of a single male experiment constitutes a limitation, especially in the light of recent reports about experimenter sex affecting rodents. We have to start somewhere, and although this will not the ultimate testing of the hypothesis, this will be a first attempt that might next be transposed to other laboratories and rat strains.

Data availability
Underlying data
No data are associated with this article.
**Reporting guidelines**


Data are available under the terms of the Creative Commons Zero “No rights reserved” data waiver (CC0 1.0 Public domain dedication).

**Acknowledgements**

We are extremely grateful for conversations with colleagues at the National Centre for Replacement, Refinement and Reduction of Animals in Research, which brought us substantive and substantial enlightenment.

**References**


Grimm D: Are happy lab animals better for science? Science. 2018. Published Full Text


Hinchcliffe JK, Jackson MG, Robinson ES: The use of ball pits and playpens in laboratory Lister Hooded male rats induces ultrasonic vocalisations indicating a more positive affective state and can reduce the welfare impacts of aversive procedures. Lab. Anim. 2022; 236772211065920. Published Full Text


Page 14 of 27


Würbel H, Garner JP: NC3Rs #9 Environmental enrichment and systematic randomization Refinement of rodent research through environmental enrichment and systematic randomization Available at. 2007. Authors. Reference Source
Open Peer Review

Current Peer Review Status: 

Version 1

Reviewer Report 24 October 2022

https://doi.org/10.5256/f1000research.137980.r153996

© 2022 Kisko T. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Theresa M. Kisko
Social and Affective Neuroscience – Brain and Cognition Research Unit, KU Leuven, Leuven, Belgium

Summary: The goal of this protocol is to evaluate the standard “Panksepp” tickling protocol against the authors' own playful handling (PH) protocol and assess which is a more ideal way to ensure the maximum positive experiences for the animals. This is done by varying the amount of pinning and providing a more interactive/reciprocal type of heterospecific play. The authors hypothesise that less pinning and more “playful interactions” will create an overall more positive experience – assessed through primary response variables Total 50-kHz USV and variability of 50-kHz USV across treatments. The second objective is to assess the physiological and behavioural effects of the various pinning and playful handling protocols. The hypothesis is that less pinning and more PH will create less variable behavioural and physiological responses. The authors will assess this through an open field, elevated plus maze and post-mortem tissue measures (pro-inflammatory cytokines) and faecal corticosterone.

Evaluation: The rationale for refining the tickling protocol is justly described and the authors provide compelling evidence from their previous studies, with descriptions and example figures and graphs, that their playful handling protocol is efficient and that it reduces the variability across individuals. Their methods suggest that more rats are finding the playful handling experience and interaction positive.

I quite like how the authors have considered the reciprocal nature of playful interactions and incorporated this into their playful handling protocol as the standard tickling does not take this into account enough. For play interactions to remain playful, there needs to be a certain amount of reciprocity and I think the way that the authors can assess the individual rats' responses to the experimenter hand and interact with the rat in a more dynamic way is an excellent way to ensure that the result is a positive interaction rather than a forced USV release that may not actually be enjoyable. Naturally creating a more fluid emission of 50-kHz USV in contrast to high rates of pinning and a more “coerced” 50-kHz USV emission can provide less confounds in experimental testing and ensure a more positive welfare condition. When Panksepp created and published the standard tickling protocol the field of affective USV research I would argue was still in its infancy.
However, over the last two decades, we have learned much more and, therefore, I quite agree that the “older” standardized protocols should be re-evaluated and updated to reflect the knowledge we now know, especially for the benefit of the animals in our care.

The experimental methods outlined for objectives 1 and 2 are straightforward and well thought out. I have only a couple of minor queries for them:

1. I would question the use of Audacity software in USV analysis though, as in my experience with the USV analysis it is not a common software for assessment. I have never heard of this software being used for USVs. There are some that are more specifically targeted for ultrasonic vocalizations and might possibly provide a more accurate assessment, especially when it comes to the categorization of 50-kHz USV subtypes. There are even several freeware versions that can be used that are user-friendly and intuitive (i.e. Deepsqueek). Have the authors considered comparing one or two files assessed with Audacity to one or two assessed with a specific vocalization software typically used in USV research, i.e. Avisoft or Deepsqueek?

2. In terms of the secondary response variable (a) in objective 1, what is the question you are exploring with subtypes? Essentially, I would ask why are you wanting to look at subtypes? Are you looking to see if certain subtypes reflect a more positive interaction? For example, would one expect that with less pinning the rats emit less trills or frequency-modulated subtypes, or just the variability of subtypes emitted during tickling? Are you planning to coordinate the subtype emissions with the behaviour occurring, i.e. while chasing the experimenter’s hand, what calls do they primarily emit? This can be quite a complicated assessment and requires accurate synchronization of behaviour and USV times.

3. How do you plan to do the categorization? By an independent observer manually categorizing or do you plan to use software that is trained for automated categorizations? There are obviously pros and cons to both, but I think this is an important point to consider when planning categorical USV analysis.

In terms of general experimental overview, controls are provided for each instance and the experimental treatments are balanced and pseudorandomized between groups of animals. The collection of tissue and faecal samples to assess physiological responses as well as using standard behavioural assays to further evaluate the effects or consequences of the playful handling protocols provides a nice balance to purely behavioural and vocalization data. This assessment could strengthen results and provide compelling evidence for an effective protocol.

Diagrams and figures in the protocol are well described and include ample details to allow for sufficient replication studies. All the statistical analyses are well outlined and have been justly described based on the variables and research questions.

Study limitations are outlined and clearly have been thought out.

In general, I quite like the protocol refinement and I agree that it is justified, and the authors provide compelling reasons and evidence to support their hypothesis. I look forward to the outcome of the experiments.

Have the authors pre-specified sufficient outcome-neutral tests for ensuring that the results obtained can test the stated hypotheses, including positive controls and quality checks?
Yes

Are the 3Rs implications of the work described accurately?
Yes

Are a suitable application and appropriate end-users identified?
Yes

Is the rationale for, and objectives of, the study clearly described?
Yes

Is the study design appropriate for the research question?
Yes

Are sufficient details of the methods provided to allow replication by others?
Yes

Are the datasets clearly presented in a useable and accessible format?
Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Social play behaviour in rats, Ultrasonic vocalizations, Social and affective behaviours and communication.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

---

**Author Response 15 Nov 2022**

**Vincent Bombail**, Scotland's Rural College (SRUC), Edinburgh, UK

*We wish to thank this reviewer for their encouragements and positive feedback. Some good points are raised, we have addressed the questions below.*

I would question the use of Audacity software in USV analysis though, as in my experience with the USV analysis it is not a common software for assessment. I have never heard of this software being used for USVs. There are some that are more specifically targeted for ultrasonic vocalizations and might possibly provide a more accurate assessment, especially when it comes to the categorization of 50-kHz USV subtypes. There are even several freeware versions that can be used that are user-friendly and intuitive (i.e. Deepsqueek). Have the authors considered comparing one or two files assessed with Audacity to one or two assessed with a specific vocalization software typically used in USV research, i.e. Avisoft or Deepsqueek?

*This is a fair point, we have previously used Audacity for visual detection of USV, but we will hopefully gain enough expertise in the use of Deepsqueak, to compare both detection*
methods.

In terms of the secondary response variable (a) in objective 1, what is the question you are exploring with subtypes? Essentially, I would ask why are you wanting to look at subtypes? Are you looking to see if certain subtypes reflect a more positive interaction? For example, would one expect that with less pinning the rats emit less trills or frequency-modulated subtypes, or just the variability of subtypes emitted during tickling? Are you planning to coordinate the subtype emissions with the behaviour occurring, i.e. while chasing the experimenter’s hand, what calls do they primarily emit? This can be quite a complicated assessment and requires accurate synchronization of behaviour and USV times.[SB1]

We have added a sentence in Data collection/Confirmatory hypothesis, to clarify this use of USV subtypes (We will investigate whether certain USV types might be preferably associated with certain treatments).

How do you plan to do the categorization? By an independent observer manually categorizing or do you plan to use software that is trained for automated categorizations[SB2]? There are obviously pros and cons to both, but I think this is an important point to consider when planning categorical USV analysis.

We aim to categorise our USVs according to the previously published description (Wright et al., 2010), as we feel this common referential will contributes toward an improved understanding of rat USV production.

Competing Interests: No competing interests were disclosed.
for the authors to increase transparency of the Stage 1 Registered Report and experimental factors to consider in their design or to include as limitations, or discussion.

1. In Hypotheses section, suggest listing Hypothesis 1 as Confirmatory Hypothesis 1a and Confirmatory Hypothesis 1b (or just making them 1 and 2?). And Hypothesis 2 as Exploratory Hypothesis 1. That way it is even clearer to readers what is confirmatory vs exploratory.

2. The authors should include a brief statement of why EMP and OF are the tests they are using.

3. How many pins are normally done in the standard tickling procedure? Is P4 equivalent to it? It sounds like there is quite a bit of variation in how the standard tickling procedure is done including number per treatment and number of treatments. On number of treatments, the authors will record and analysis after just 1 treatment (Phase 1) and not also at Phase 2. Since the data are recorded it seems useful to analysis at both points (possibly even incorporating it in the statistical design?)

4. Related to point 3, what is the need for a gap between Phase 1 and Phase 2? That is, why not just move from the treatment on days 9-11 into the EPM and OF testing? Is it because the rats should have two treatment sessions spaced apart by 5 days before the behavioural and physiological tests? Either way can the authors elaborate this design decision.

5. Maybe I missed it, but will the cages assigned to one treatment (e.g., H0) in Phase 1 also be in the same treatment in Phase 2 or will they be randomized?

6. It is nice to see the pseudo-randomization. I was going to recommend also introducing the randomization for the order of cages treated (in addition to the order of rats tested in each cage), but it looks like this is part of the design. However, it is only stated once in Objective 1 experimental design. I suggest moving this up earlier in the protocol and possibly repeating it where appropriate, such as when discussing the pseudo-randomization design and rat ordering as it was unclear until this single sentence.

7. Related to the need to randomize where possible, especially since the rat handler will not be blind to treatments, I think an unfortunate weakness of this protocol, although I understand why from a practical sense, is the lack of variability of researcher. That is, a single researcher at a single institution with a single rat strain is being used. Three related thoughts on this:

   a) The authors might consider discussing the impact of features that are not being accounted for that might impact the outcome (e.g., experience of the handler performing the standard tickling protocol, experience of the handler performing the PH protocol, sex of the researcher (e.g., Sorge et al., 2014)).

   b) I am not an expert in this technique, but having a strong proponent of the standard rat tickling protocol could likely provide some quite useful critique – essentially an adversarial collaboration that help with pre-commitment (Brian Nosek and I wrote a piece about it here: https://www.nature.com/articles/d41586-020-02142-6 - and note, I'm not suggesting you cite this, only providing for background with the concept if helpful).
c) A consideration to keep in mind for the discussion of the outcomes of this (yes, getting ahead for a moment), is one way to address this would be through a ‘many-labs’ style replication of these findings, essentially introducing the variation of researcher, strain, laboratory, etc. Again, this study would need to be completed with results to warrant a ‘many-labs’ approach, but sharing now for future consideration.

8. Another potential factor to consider is time of day of the treatment. It is unclear if that is held constant. That is, will the treatment always occur at the same time of day relative to the 12h:12h light cycle of the rats or if it will also be variable across cohorts? I am unsure if this could have an impact on behaviour, but raising as something potentially to track and share at the completion of the study.

9. For forming cage pairs for each cohort, while you will always achieve 4 cages of female and 4 cages of male (by design), it seems it might not always be possible to form pairs based on provenance since that will depend on litter size. It is less of a concern if there are rats from many litters, but is there a chance of receiving more than 4 rats of one sex from the same litter during any of the cohorts? It sounds like the authors have already discussed this with the supplier of the rats, but it’s unclear how this is taken into account in the randomization strategy if this issue arises.

10. Is there any consideration of whether a rat is unable to continue? For example, if a rat forms an infection during the habitation caused by something else, will the authors exclude the rat from the experiment? If any of this is known a priori it would be good to mention, if not, then transparency during report will be sufficient.

11. I’d recommend adding RRIDs where appropriate (e.g., rat strain).

12. It sounds like for the secondary variables of interest (e.g., when coding the anticipatory play behaviour), these will be coded by a researcher from the video recordings and that the researchers will be blinded to the behavioural and physiological tests. This is again great to see, however if the authors could share a little more detail of how they will do this, that would be helpful for transparency.

13. For analyses, the authors should include more transparency and more explicitness for confirmatory tests. For example:

   a) The p-value threshold (assuming 0.05 based on power analysis) they will consider as significant.

   b) The comparisons they will perform, which is especially important for the primary response variables. So for example, is the test for total 50 kHz USV a contrast of P0 and P1 compared to P4 and Control? Or are the authors suggesting multiple contrasts (e.g., P0 vs P4, P1 vs P4, etc). For example, the legend for Figure 3ii suggests it would be P0 and P1 vs P4 (excluding control).

   c) If relevant, how are multiple tests being corrected for?
d) What approaches will be used to test data fit the assumptions of the tests (e.g., what specific tests for normality), and same with data transformations (e.g., what are viable options to consider and what is the workflow to consider it sufficient) – this can help make the decision tree the authors will use on whether to use one test vs an alternative more transparent a priori.

e) And while not necessarily needed now, it would be beneficial to know the authors plan to report effect sizes and uncertainty in their analyses.

References

Have the authors pre-specified sufficient outcome-neutral tests for ensuring that the results obtained can test the stated hypotheses, including positive controls and quality checks?
Yes

Are the 3Rs implications of the work described accurately?
Yes

Are a suitable application and appropriate end-users identified?
Yes

Is the rationale for, and objectives of, the study clearly described?
Yes

Is the study design appropriate for the research question?
Yes

Are sufficient details of the methods provided to allow replication by others?
Yes

Are the datasets clearly presented in a useable and accessible format?
Yes

**Competing Interests:** I am an employee of the non-profit Center for Open Science (COS) that has a mission to increase openness, integrity, and reproducibility of research. COS offers support to journals, editors, and researchers in adopting and conducting Registered Reports.

**Reviewer Expertise:** I have expertise in replicability, reproducibility, metascience, cellular and biochemical assays (e.g., PCR, imaging), and some animal protocols (e.g., tail-vein injections, xerograph models). I do not have expertise in rat tickling (or rat handling specifically) nor expertise in behavioural test, such as EPM and OF described in this Stage 1 Registered Report.
I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 15 Nov 2022

Vincent Bombail, Scotland's Rural College (SRUC), Edinburgh, UK

We wish to thank this reviewer for their thorough and constructive criticism, and we hope our attempt at addressing their questions has brought more clarity to our manuscript.

The Stage 1 Registered Report outlines the rationale, study design, hypotheses, and analysis strategy for testing whether updated rat tickling protocol (playful handling: PH) leads to a higher and less variable induction of ultrasonic vocalisation (USV) compared to the standard tickling protocol, indicative of an increased positive playful interaction. The authors also describe exploratory experimentation to examine whether the difference in tickling protocols leads to less variation in behavioural and physiological tests

Overall, this is a well thought out and described protocol. Suggestions below are considerations for the authors to increase transparency of the Stage 1 Registered Report and experimental factors to consider in their design or to include as limitations, or discussion.

In Hypotheses section, suggest listing Hypothesis 1 as Confirmatory Hypothesis 1a and Confirmatory Hypothesis 1b (or just making them 1 and 2?). And Hypothesis 2 as Exploratory Hypothesis 1. That way it is even clearer to readers what is confirmatory vs exploratory.

This is a good idea and we have updated the text to follow this advice.

The authors should include a brief statement of why EMP and OF are the tests they are using.

We have added explanatory sentences in the methods section: ‘The EPM and OF tests consist in quantifying exploration behaviour in 2 different structures, that both consist in an area with open spaces (the centre of the OF and the open arms of the EPM) and an area with relatively more enclosed spaces (the periphery of the OF and the closed arms of the EPM). For rats, this ambiguity creates a conflict between the thigmotaxis (preference for proximity to walls) and the motivation to explore open spaces, which is used to assess the emotional response to those novel experiences, and to test for anxiety levels.’

How many pins are normally done in the standard tickling procedure? Is P4 equivalent to it?
It sounds like there is quite a bit of variation in how the standard tickling procedure is done including number per treatment and number of treatments. On number of treatments, the authors will record and analysis after just 1 treatment (Phase 1) and not also at Phase 2. Since the data are recorded it seems useful to analysis at both points (possibly even incorporating it in the statistical design?)

The protocol from Cloutier et al. (2018) describes there should be 4 to 5 pins per 15 s (2-4s
per pin). We also postulated a 30 s treatment would be a practical and effective duration to achieve induction of most individual rat positive affect (based on LaFollette et al., 2018, and our own findings). Our protocol and treatment design synthesise these requirements.

Related to point 3, what is the need for a gap between Phase 1 and Phase 2? That is, why not just move from the treatment on days 9-11 into the EPM and OF testing? Is it because the rats should have two treatment sessions spaced apart by 5 days before the behavioural and physiological tests? Either way can the authors elaborate this design decision. 

*There is no gap between Phases 1 and 2, phase 2 starts the next day. To clarify we have added the word ‘next’, in experimental design for objective 2.*

Maybe I missed it, but will the cages assigned to one treatment (e.g., H0) in Phase 1 also be in the same treatment in Phase 2 or will they be randomized?

**For phase 2, to investigate the behavioural and physiological correlates of positive experiences, the animals and their treatments will be the same as in phase 1.**

It is nice to see the pseudo-randomization. I was going to recommend also introducing the randomization for the order of cages treated (in addition to the order of rats tested in each cage), but it looks like this is part of the design. However, it is only stated once in Objective 1 experimental design. I suggest moving this up earlier in the protocol and possibly repeating it where appropriate, such as when discussing the pseudo-randomization design and rat ordering as it was unclear until this single sentence.

*We have added this to the testing in objective 2: ‘their order of passage will be pseudo-randomised’. *

Related to the need to randomize where possible, especially since the rat handler will not be blind to treatments, I think an unfortunate weakness of this protocol, although I understand why from a practical sense, is the lack of variability of researcher. That is, a single researcher at a single institution with a single rat strain is being used. Three related thoughts on this:

a) The authors might consider discussing the impact of features that are not being accounted for that might impact the outcome (e.g., experience of the handler performing the standard tickling protocol, experience of the handler performing the PH protocol, sex of the researcher (e.g., Sorge et al., 20141).

The inclusion of further experimenters in the design could constitute another source of variation, but we agree the involvement of a single experienced male experimenter might constitute a limitation. Over time there will be other sources of variation (e.g. dietary molecules excreted through sweat, fluctuation in mood and playfulness), but the study of human-animal interactions requires such human intervention. We have to start somewhere, and although this will not the ultimate testing of the hypothesis, this will be a first attempt that might next be transposed to other laboratories and rat strains. We have added these points in the study limitations section.
b) I am not an expert in this technique, but having a strong proponent of the standard rat tickling protocol could likely provide some quite useful critique – essentially an adversarial collaboration that help with pre-commitment (Brian Nosek and I wrote a piece about it here: https://www.nature.com/articles/d41586-020-02142-6 - and note, I'm not suggesting you cite this, only providing for background with the concept if helpful).

Thank you for this interesting read. This protocol, and the study in general, were designed inclusive of different perspectives on tickling, and in collaboration with our co-author Megan LaFollette, who is a proponent of the standard protocol (i.e. the Panksepp protocol with pinning described in Cloutier et al., 2018).

c) A consideration to keep in mind for the discussion of the outcomes of this (yes, getting ahead for a moment), is one way to address this would be through a ‘many-labs’ style replication of these findings, essentially introducing the variation of researcher, strain, laboratory, etc. Again, this study would need to be completed with results to warrant a ‘many-labs’ approach, but sharing now for future consideration.

This is very interesting. This has been incorporated in the response to point a) above, in the study limitations sections.

Another potential factor to consider is time of day of the treatment. It is unclear if that is held constant. That is, will the treatment always occur at the same time of day relative to the 12h:12h light cycle of the rats or if it will also be variable across cohorts? I am unsure if this could have an impact on behaviour, but raising as something potentially to track and share at the completion of the study.

We have specified time of treatments in the ‘Acclimation, treatment allocation, housing and husbandry section’.

For forming cage pairs for each cohort, while you will always achieve 4 cages of female and 4 cages of male (by design), it seems it might not always be possible to form pairs based on provenance since that will depend on litter size. It is less of a concern if there are rats from many litters, but is there a chance of receiving more than 4 rats of one sex from the same litter during any of the cohorts? It sounds like the authors have already discussed this with the supplier of the rats, but it’s unclear how this is taken into account in the randomization strategy if this issue arises.

We have discussed this with the supplier (CRL) and we will not receive more than 4 rats of one sex from the same litter.

Is there any consideration of whether a rat is unable to continue? For example, if a rat forms an infection during the habitation caused by something else, will the authors exclude the rat from the experiment? If any of this is known a priori it would be good to mention, if not, then transparency during report will be sufficient.

We have updated the text to include this information. We will follow local guidelines should rat injury occur. If this were to happen, this information would be described in the methods.
section, for full transparency about the fate of our experimental rats.

I’d recommend adding RRIDs where appropriate (e.g., rat strain).

We have added the following information to the materials section: Research Resource Identification, RRID:RGD_2308816

It sounds like for the secondary variables of interest (e.g., when coding the anticipatory play behaviour), these will be coded by a researcher from the video recordings and that the researchers will be blinded to the behavioural and physiological tests. This is again great to see, however if the authors could share a little more detail of how they will do this, that would be helpful for transparency.

It will be impossible to deduce the nature of the treatment, since all treatment orders will have been pseudo randomised, and none of the experimenters involved has a good enough memory to remember the treatment group of 64 different cages.

For analyses, the authors should include more transparency and more explicitness for confirmatory tests. For example:

a) The p-value threshold (assuming 0.05 based on power analysis) they will consider as significant.

This has now been specified in the manuscript.

b) The comparisons they will perform, which is especially important for the primary response variables. So for example, is the test for total 50 kHz USV a contrast of P0 and P1 compared to P4 and Control? Or are the authors suggesting multiple contrasts (e.g., P0 vs P4, P1 vs P4, etc). For example, the legend for Figure 3ii suggests it would be P0 and P1 vs P4 (excluding control).

All treatments will be compared

c) If relevant, how are multiple tests being corrected for?

We have amended the text accordingly: ‘To account for potential type I errors that may arise from multiple comparisons, we will use Bonferroni correction or equivalent (depending on data distribution).’

d) What approaches will be used to test data fit the assumptions of the tests (e.g., what specific tests for normality), and same with data transformations (e.g., what are viable options to consider and what is the workflow to consider it sufficient) – this can help make the decision tree the authors will use on whether to use one test vs an alternative more transparent a priori.

We will use classical tests such as the Bartlett’s test for homogeneity of variances and Anderson-Darling test for normality. If necessary, data might be transformed (e.g. square root transformation)
e) And while not necessarily needed now, it would be beneficial to know the authors plan to report effect sizes and uncertainty in their analyses.

*We have added: 'In the spirit of transparency, we will report effect sizes and uncertainty in our analyses.'*

*Competing Interests:* No competing interests were disclosed.

---

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com