Method Article

The cage test as an easy way to screen and evaluate spontaneous activity in preclinical neuroscience studies

Mijail Rojas-Carvajal\textsuperscript{a, b,*}, Daniel Quesada-Yamasaki\textsuperscript{b}, Juan C. Brenes\textsuperscript{a, b}

\textsuperscript{a}Institute for Psychological Research, University of Costa Rica Costa Rica
\textsuperscript{b}Neuroscience Research Center, University of Costa Rica Costa Rica

A B S T R A C T

Within behavioral neuroscience, subjects used to be randomly assigned to the experimental groups based on the premise that interindividual variability will be homogeneously distributed. However, the equivalence offered by randomization diminishes in small samples, which is the case for most experiments in the field. In rodents, it is well-recognized that individual differences in psychomotor reactivity, risk-assessment behaviors, and emotional responsiveness modulate the effects of different pharmacological and non-pharmacological treatments. For that reason, knowing such differences before the experiment provides highly valuable information for balancing the groups so that the interindividual variability is equally distributed within the groups without excluding subjects as far as possible. Because unconditioned anxiety tests such as the open-field (OF) and the elevated plus-maze are commonly used within experimental procedures, we developed a strategy to explore the rat’s behavioral phenotype by assessing it in a very innocuous testing context: a housing cage.

- We offer a very straightforward protocol for assessing spontaneous, novelty-induced reactivity in rodents.
- We describe its implementation, analysis, and use, as well as some suggestions about key behavioral readouts for the group allocation procedure.
- The current protocol provides an alternative strategy to assess a reasonably wide range of behavioral outcomes, most of which are of great interest in modeling different neuropsychiatric disorders.

© 2021 The Author(s). Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

A R T I C L E   I N F O

Method name: Cage test
Keywords: Screening test, Unconditioned anxiety, Behavioral analysis, Ultrasonic vocalizations, Risk-assessment, Grooming, Balanced group assignment
Article history: Received 29 September 2020; Accepted 10 February 2021; Available online 17 February 2021

* Corresponding author.
E-mail address: mijail.rojas@ucr.ac.cr (M. Rojas-Carvajal).

https://doi.org/10.1016/j.mex.2021.101271
2215-0161/© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Specifications table

<table>
<thead>
<tr>
<th>Subject Area</th>
<th>Psychology</th>
</tr>
</thead>
<tbody>
<tr>
<td>More specific subject area</td>
<td>Behavioral Psychology</td>
</tr>
<tr>
<td>Method name</td>
<td>Cage test</td>
</tr>
<tr>
<td>Resource availability</td>
<td>In the article</td>
</tr>
</tbody>
</table>

Method details

The equivalence offered by the randomization of experimental groups diminishes considerably in small samples [39], which is the case for most behavioral neuroscience experiments. Considering the high costs associated with large samples, and the ethical commitment for using the minimum possible number of animals per experiment [2], it is unlikely that researchers will ever have a sample big enough for the random selection to thoroughly homogenize the influence of individual differences and variability on group’s average scores. It is then possible that part of the high within-group variability seen in some experiments and the inconsistent results observed within and between labs might result from uneven distribution of individuals’ traits between groups. This is particularly important for those experiments with small sample size, independent variables with middle-to-low effect sizes, and dependent parameters with high interindividual variability. In turn, addressing this issue would contribute to reducing the reproducibility crisis faced in preclinical research.

Different strategies have been developed to assess and control individual variability on the biobehavioral phenotype resulting from experimental treatments [26]. One of those strategies consists of balancing the groups based on behavioral parameters with empirical and theoretical relevance for the following experiment. This approach has been used for decades in clinical and health research in humans, where the criteria for group assignment include sociodemographic (e.g., educational level, gender, age, origin, ethnic, economic level, religion), psychological (e.g., emotional traits, intelligence, particular cognitive abilities), and clinical variables (e.g., the background of diseases, medications, and other treatments). However, it is still seldom implemented in preclinical models.

To assess individual differences in rat’s unconditioned responses to novelty, we structured and implemented a simple but effective strategy to evoke a broad spectrum of behavioral responses within a very innocuous testing context: a housing cage. The cage test (CT) consists of singly exposing the animals to a novel housing cage for a short time [27,34,40]. The testing cage is filled with clean bedding material and placed in an unfamiliar, mildly illuminated testing room. The cage’s overall features are highly recognizable for the animals (e.g., bedding material, the form, and texture of the wall), but the absence of familiar olfactory cues and the testing room’s novelty effectively evoke exploratory and risk-assessment behaviors (e.g., locomotion, rearing, sniffing, and digging). Because the CT naturally induces this behavioral repertoire, it becomes a valuable form to assess emotionality and spontaneous psychomotor activity in rats [7,20,34]. Implementing the CT brings a considerable amount of information about the rat’s behavioral phenotype for a relatively small investment of resources and time. The information could be used for allocating subjects to the groups, as covariates during statistical analyses, and for classifying subjects based on particular traits in studies of individual differences. Based on our previous experience with this test [7,28,31], we offer a comprehensive protocol for its implementation, a suggested strategy for its analysis, and a procedure
for allocating and balancing the subjects to the experimental groups. All experimental procedures
reported in this protocol were done in accordance with the guidelines of the Costa Rican Ministry
of Science and Technology for the Care and Use of Laboratory Animals and were approved by the
Institutional Committee for Animal Care and Use of the University of Costa Rica (CICUA-047-17).

**Materials and reagents**

**Materials**

1. Small whiteboard and whiteboard markers.
2. Paper towels.
3. Wash bottles filled with ethanol (70%) or other disinfectant solution.
4. Spray bottle filled with odorless disinfectant.
5. Latex or nitrile gloves.
6. Three copies of the experimental lists. Experimentation lists should fulfill the researcher’s needs
   and particularities, but must contain at least the following information:
   - Apparatus number: Refers to which of the apparatuses will be used to evaluate the animal.
   - Identification mark: Refers to the marks used for identifying the animal. It must be carefully
     checked out to ensure that all animals have been assessed and that none of them repeat the
     test.
   - Home cage number: Indicates in which cage the animal should be taken out and placed back
     after testing. Ideally, after the experiment, rats should be housed in a different cage to avoid
     disturbing the conspecifics that have not been tested yet.
   - ID code: It refers to the unique identification number of the animal in the context of a particular
     experiment. This should be written down on the whiteboard so that testing animals can be
     identified on the video recording. It is critical to check the correspondence of the IDs with the
     animal marks.
   - Corroboration space: It refers to a checkbox to verify that the animal has been successfully tested
     and returned to the animal room.
   - Observations space: Some blank lines should be left for reporting incidences.
7. Ensure to have at least 1 L of clean bedding material (e.g., sawdust) per animal per test.
8. Eventually, a heavy object might be useful to keep the cage grid in place if needed.

**Reagents**

1. Male, albino rats bred for scientific purposes. Note: This protocol has been implemented and
   validated in group-housed Sprague-Dawley and Wistar rats, ranging from 21 post-natal days (PND)
   to 330 PND of age, which are usually tested within two weeks upon arrival to our facilities. For
   other strains, species or housing conditions, the appropriate adjustments and validation should
   be made. Before testing, avoid any kind of handling other than the one required for relocating animals
   between cages or imprinting ID marks in their bodies. If required, handling procedures can start
   the next day after the CT.
2. Ethanol at 70%, to clean individual cages between sessions.
3. Odorless disinfectant.

**Equipment**

1. A wooden or acrylic open-field (OF) arena (55 cm x 55 cm, and 40 cm high). The wooden OF
   version could be covered with black Formica® or a similar material for easy cleaning. For a better
   rat/background contrast, any enclosed apparatus with a dark background would be useful. Note:
   many apparatuses may be used to run several CT simultaneously to save time and reduce the
   effects of circadian variations on behavior.
2. Individual housing cages (transparent, 42 × 26.5 × 15 cm, with 1L of fresh bedding) covered with
   flat grids (e.g., Eurostandard type III). For animals older than 150 PND, a taller testing cage should
be used (e.g., Eurostandard Type III H: 42.5 × 26.6 × 18.5 cm). Note: additional cages may be necessary for simultaneous CT sessions. Ensure to have enough clean cages availability for testing.
3. Transporting cages (opaque, 46 × 22 × 21 cm, with fresh bedding and a grid on top) will speed up the process of moving animals back and forward between animal and testing rooms. Note: additional cages may be necessary for simultaneous CT sessions.
4. Group housing cages with food, water, and clean bedding. Note: These are additional cages used for housing the animals after tested, so their number and type should be the same as those used for regular housing. Consider that cage size and number of subjects per cage will depend on local regulations about the recommendations for minimal housing area per animal.
5. Video cameras: We suggest using GoPro Hero cameras (or equivalent models) for recording from the side of the testing cage (See Fig. 1). To ensure the best image quality, the cameras should be configured at high sensitivity to light and with a resolution of ≥960p.
6. Luxmeter and decibelmeter to measure the light and sound intensity within the testing cage, respectively.
7. White light bulbs plugged into dimmable sockets.
8. Computer with the following minimum requirements: (1) 10GB free space, (2) Intel® Core i3 or similar processor; (3) 4GB RAM memory; (4) 18” monitor with >960p resolution. Note: External storage can also be used, but high-speed transmission ports will be required (e.g., USB 3.0).

To record rat’s ultrasonic vocalizations (USVs), the following equipment (or equivalent) is necessary:
9. Microphones for USVs (e.g., Condenser Microphones CM16, Avisoft Bioacoustics, Berlin, Germany).
10. Microphone cables (Avisoft Bioacoustics, Berlin, Germany). Note: Ensure to have at least one set of microphones and cables per each testing apparatus.
11. SoundGate Recorder (Avisoft Bioacoustics, Berlin, Germany).
12. USB cable to connect the SoundGate to a computer.
13. Software for recording, visualization, and analyzing USVs already installed (e.g., Avisoft-Recorder USGH and Avisoft-SASLAb pro; http://www.avisoft.com/).

Software
1. Video play software (recommended: VLC®; https://www.videolan.org/)
2. Microsoft Office Excel® or similar software for data manipulation.
3. Software for automatic behavioral detection (e.g., Any-Maze® or Nodulus®)
4. Software for manual behavioral assessment (e.g., Solomon Coder®; download free from https://solomoncoder.com/download.php).
Procedures

Preparation of the testing rooms

1. Clean the entire testing room with disinfectant.
2. Place the apparatuses under the light source.
3. Put a wash bottle full of 70% ethanol and a roll of paper towels next to the apparatus.
4. Deeply clean the apparatus, first with disinfectant and then with 70% ethanol.
5. If USVs are recorded, ensure the apparatus or the room is sound insulated.
6. If more than one apparatus is used, number them to be identified in the video recordings (Fig. 1). Note: We strongly recommend the use of in-video identification when simultaneous tests are running.
7. Place the individual housing cage inside the apparatus and locate it at one corner against the two perpendicular walls (see Fig. 1).
8. Ensure that the light bulbs illuminate all regions of the cage homogeneously.
9. Put the luxmeter inside the cage and place back the upper grid.
10. Dim the lights at 10 lumens.
11. Place the camera in front of the long side of the testing cage (as shown in Fig. 1).
12. Adjust the cameras’ inclination so that all the videos’ recordings are as homogeneous as possible. Ensure that cameras are firmly fixed in their final position to prevent them from being accidentally moved during the experimental procedures. Keep the lenses clean, the memory cards with enough space, and the batteries charged.
13. If USVs are recorded:
   - Connect the microphone to its cable and hang it at 35 to 40 cm over the testing cage’s center.
   - Adjust the cable so that it does not affect the researcher’s circulation or the testing process.
   - Place the SoundGate next to a computer with the AviSoft Recorder 2.7 software installed.
   - Connect the microphone cable to the SoundGate and use the USB cable to connect the computer to the SoundGate.
   - In a clockwise motion, set the gain of the SoundGate to 3”, which is suggested for detecting 50 kHz and 22 kHz calls in rats with this particular setting. Note: Ensure that the knobs’ numbers correspond with the microphones’ numbers by checking them at the back of the SoundGate.
   - Place and turn on the computer with the USV recording software. Ensure there is enough storage space and that the power source is connected to an electrical outlet.
   - Open the AviSoft Recorder 2.7 program. Set the name of the recordings in a separate folder for each channel and adjust the following parameters. Sampling rate: 214.285 Hz; Format: 16 bits; Frequency resolution: 0.488 kHz; Time resolution: 0.512 ms; Fourier transformation: 512 FFT-length; Frame: 100%, Window: Hamming; Time window overlap: 75%.
   - Run a recording test in the program to verify that a signal is correctly detected. Shake a bunch of keys gently under each microphone to ensure that a clear ultrasound signal is observed in the corresponding channels in the spectrogram (high-intensity black spikes across all the frequency spectrum should be observed). If necessary, adjust the gain in the SoundGate. Once adjusted, leave the equipment installed and untouched during the entire evaluation period.
14. Print 3 copies of the experimental lists, place one in the vivarium, one in the testing room, and keep an additional one if you need it to make notes or as a replacement if any of the other lists are damaged. Note: You can use your own system to organize your experiment. We recommend taking all the necessary steps to guarantee that the animals are manipulated and tested in the right order and under the correct procedures and that their information can be easily tracked down during all the experimental steps.
15. Label the transporting cages with identification numbers to correspond them with the testing rooms or apparatuses. Prepare two sets of transporting cages per apparatus. Then, fill them with fresh bedding.
16. Label the testing cages with identification numbers to correspond them with the testing rooms or apparatuses. Prepare two sets of testing cages per apparatus and fill them with 1L of fresh bedding material.

17. Label extra housing cages by duplicating the identification tag used in the original cages. Fill them with fresh bedding material (~4–5L) and place them in the designated room.

18. Keep a complete set of charged camera batteries and empty microSD cards available for replacement.

**Executing cage test: experimental session**

1. Make sure all the equipment and materials are ready to use.
2. Write the information about the experiment on the acrylic board. We recommend including at least the following information: Apparatus ID, animal ID, date.
3. Find the testing animal by verifying its ID, identification marks, and housing cage. Note: Animals could be transported directly from their vivarium to the testing room, but to reduce animals’ permanence in the transporting cage, they might be temporarily transferred to an adjacent room to the testing area at least one hour before assessments. It will make testing more efficient.
4. Grab the animal from its back, place it in the transporting cage, and immediately transfer it to the testing room.
5. Right before the animal arrives, start the video recording, and show the acrylic board to the camera.
6. Once in the testing room, place the transporting cage next to the apparatus and put the animal in the testing cage. For that purpose:
   - Hold the animal gently from its back and place it in the center of the cage. Note: If the animal is unsettled, do not chase it around the cage. Surround the animal with your hands, approach it slowly, and try to grab it gently.
   - The animal must be placed carefully until its four paws are on the ground. It should never be thrown.
   - If the animal is unsettled and constantly fleeing from the experimenter’s hand, hold it from its root tail to placing it in the apparatus. Note: Placed the animal on the cage and released it as soon as its four legs are on the ground. Do not hold nor drag it from the tail while it is on the testing cage.
   - Animals’ permanence in the transporting cage should be shortened as much as possible.
7. Cover the testing cage with the grid and secure it by placing a heavy object on top, if necessary. Note: An adjustable, flat lid can also be used to cover the cage. We discourage using lids with troughs (e.g., for food pellets and water bottles) since they change animals’ behavior.
8. Leave the room, start the chronometer, and run the test for the desired time (e.g., five minutes). If USVs are recorded, start the audio recording immediately after leaving the room.
9. Enter the room about one minute after the test finished. Note: The extra minute will prevent disturbing animals’ activity when the experimenter approaches the testing room.
10. Stop the video and sound recording.
11. Remove the grid, place the animal back into the transporting cage, and return it to the corresponding extra housing cage placed in the vivarium or designated room. Note: To avoid alterations in untested cage mates, tested animals should not be returned to their original cage until all animals of the same cage have been assessed.
12. Mark in your lists that the animal finished its test.
13. Clean the OF apparatus and grid with 70% ethanol. Then, place the new testing cage filled with fresh bedding material previously prepared for that purpose.
14. Start the procedure all over again until all animals are tested.
15. Once finished, ensure transferring all video and audio data from the recording equipment to a safe storage device.
16. Once all the animals of a given housing cage have been tested, they can remain in the new cage or being returned to the former cage. Apply the same manipulation homogeneously to all testing animals.
17. Return all housing cages to the vivarium in case they were placed on a different room during testing.
18. If testing cages are reused within the same experimental session, ensure removing urine, feces, and any other residues by cleaning the cage with abundant 70% alcohol and with a paper towel. Once clean and dry, fill the cage with 1L of fresh bedding.

Behavioral analysis

1. Select a video to carry out the behavioral analysis.
2. Open the software and load the configuration. Then, choose the destination folder for the registered results. Note: For manual scoring of behaviors, it is recommended a program frame-rate resolution of ≤500 ms, since some behaviors (e.g., rearing and grooming) might last for <1 s.
3. Register at least the following behaviors: (1) Locomotion (by automatic tracking), (2) rearing, (3) grooming, (4), and USVs (see Table 1 for details). To reach an acceptable degree of reliability in the manual behavioral scoring, the intra- and inter-subject variability should not exceed 5%. Note: The following behaviors can also be observed during testing but at lower rates: digging, burying (i.e., using the forepaws, the animal throws bedding material forward, usually, accumulating it against the corners), eating (i.e., feces or bedding material), hanging (from the lid), running, and diving (i.e., the animal dives below the bedding material).
4. For USVs analysis, use your particular protocol of counting and classification. Note: Previous reports indicate that rats naturally emit 50 kHz calls when facing novel contexts filled with bedding material [27]. For that reason, the CT is a reliable and straightforward method to elucidate interindividual differences in the emission of 50 kHz USVs. Nevertheless, stress experiences can also induce 22 kHz alarm calls while suppressing 50 kHz calls [28].

Data analysis for balanced group assignment

1. Check that all the videos were adequately analyzed and that the information in the data sheets is correct.
2. Create a data matrix using rows for subjects and columns for variables (Fig. 2, 3). Use a Microsoft Office Excel® datasheet or similar.
3. Incorporate cumulative data per subject as corresponds to the data matrix (Fig. 2.A).
4. Create a new empty matrix per each intended experimental group. The number of rows will depend on the number of subjects per group. Add a row at the bottom of the charts for average scores (Fig. 2.B-C).
5. Configure each group chart’s last row to show the average score per column corresponding to each variable. For that purpose, use the “AVERAGE” formula or an analogous.
6. Create a new matrix using columns for variables and as many rows as intended groups. Each row will be designated for the comparison of average scores between each group (Fig. 2.D).
   a. Configure each row to display the group’s difference percentage per variable using the following formula: \([(\text{Mean of Group 1} / \text{Mean of Group 2}) - 1] \times 100\). Note: The inclusion of other statistics, such as variance and ranges, will contribute to a better balancing distribution. We are only showing variable means, but we strongly recommend considering other descriptive statistics for group allocation.
7. Consider the most important parameters for the groups to be homogenized. This decision should be based on the following criteria: (1) the theoretical framework of the study, (2) the aims of the experiment, and (3) its dependent variables. Note: If only exploratory and risk-assessment behaviors are the variables of interest, locomotion and rearing may be enough to balance between groups.
<table>
<thead>
<tr>
<th>Behavior</th>
<th>Description</th>
<th>Ethological interpretation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locomotion</td>
<td>It can be registered automatically by video tracking softwares (e.g., Any-Maze®, Nodulus®). It is essential that the starting minute matches between the manual and automatic behavioral scoring. To see the details of the program configuration, refer to the corresponding tracking software's user manual. The locomotor activity can also be manually assessed by dividing the testing cage into two halves and then counting the number of lines crossed as the animal moves laterally.</td>
<td>Horizontal activity (i.e., locomotion) increases when animals face a novel situation or context serving exploratory and information-gathering purposes. Locomotion is a general indicator of psychomotor reactivity and an index of anxiety (the motivation to escape from unknown and unpredictable stimuli).</td>
<td>[8,12,19,28,30,31,36].</td>
</tr>
<tr>
<td>Rearing</td>
<td>Rearing consists of a bipedal posture (&gt;45° from the floor), in which the animal extends its head upwards, executing a series of lateral movements using its vibrissae (whiskers) to sense the surroundings (by smelling and touching out). These exploratory movements with the head and vibrissae are referred to as sniffing. The animals may or may not lean against the walls to get up on their hind legs. Rearing should not be counted if the animals rise on its hind legs but keep its forepaws and head parallel to the ground, with an inclination of its torso &lt;45°.</td>
<td>Rearing is considered a risk assessment (RA) behavior involved in evaluating possible threat sources during potentially or ambiguous situations. RA serves to predict the likelihood of threats and choose the best coping behaviors. Because uncertainty and ambiguity about possible threats naturally evoke rearing, it has been interpreted as an anxiety-like behavior.</td>
<td>[4,5,28,31].</td>
</tr>
<tr>
<td>Grooming</td>
<td>Grooming consists of a set of complex, self-directed movements aimed at the care of the body surface. It includes (1) hand rubbing, (2) face washing, (3) unilateral and bilateral strokes over the head and ears, (4) body licking, (5) anus-genital licking, (6) head and body scratching, and (7) tail licking. Micro grooming (&lt;1s) and scratching events are sometimes excluded from the analysis. However, we strongly suggest that those sequences should not be neglected. The scratching events are counted as grooming only when they are chained with other sequences.</td>
<td>Self-grooming is a widespread behavior in the animal kingdom. Besides its cleaning purpose, grooming's biological functions involve communication, parasite control, thermoregulation, pain-relief, and many others. Grooming has also been associated with anxiety and stress and constitutes a critical behavioral marker in many models of neuropsychiatric diseases. Evidence suggests that different grooming subtypes have different interpretations and, therefore, may be analyzed separately. From our perspective, short, head directed grooming strokes are mainly related to on-going negative arousal, whereas the more complex sequences of grooming seem to be involved in emotional self-regulation. In that sense, we suggest that grooming would be part of an arousal-inhibiting system serving for homeostatic behavioral control. However, grooming interpretation in the context of behavioral sciences is still a vividly open discussion.</td>
<td>[1,8,12,16–18, 21,22,25,28,30–32,37,38].</td>
</tr>
<tr>
<td>USVs</td>
<td>Most rodent species, including laboratory rats and mice, emit USVs (i.e., &gt;20 kHz) produced by a whistle mechanism through their larynx's constriction. In contrast to sounds generated by vocal fold vibration—which are readily audible to humans—these whistle-driven USVs are inaudible without specialized equipment. According to their average frequency, three groups of USVs have been identified so far: (1) 22 kHz (emitted by young and adult rats), (2) 40 kHz (emitted only by rat pups), and (3) 50 kHz USVs (emitted by young and adult rats). When the emission of 22 kHz USVs occurs, a particularly evident body posture is displayed (which does not occur for 50 kHz USVs). Stereotyped diaphragm contractions characterize this posture while the rat is relatively immobile.</td>
<td>USVs are a form of social-affective communication naturally emitted by rats under emotionally related situations. 50-kHz USVs are emitted during conspecific and hetero-specific play, reward anticipation, mating-related activities, and the exposure to different familiar or novel contexts. Animals exposed to a cage filled with fresh bedding material, such as the CT, also emit 50 kHz calls. In consequence, 50 kHz calls are related to positive affect and social coordination. Conversely, 22 kHz calls are emitted when the animals are exposed to acute stress events such as conspecific aggression, inescapable pain, or predator-related olfactory cues. Further, rats exposed to 22 kHz USVs show a substantial increase in defensive behaviors such as freezing. Therefore, 22 kHz USVs are considered as an index of a negative affective state.</td>
<td>[9–11, 14,24,27,33–35].</td>
</tr>
</tbody>
</table>
Fig. 2. Balancing process based on fictional data. (A) Cumulative scores of each variable should be incorporated into the data matrix. (B-C) Then, one empty data matrix per group should be created, and the subjects’ data should be added following the sorting order in the source matrix. (D) Once the group differences reach the lowest scores in all the variables of interest, the researcher should perform a One-way ANOVA to identify possible significant differences. The variability of the scores should be carefully observed and aimed to keep it homogenous between groups. Notice that in our fictional data, there is only one littermate in each group. However, because litter information is usually unavailable, balancing littermates between groups is not always possible. As behavioral responses to the CT greatly differs between individuals, their assignment should be carefully performed to avoid a biased or unbalanced distribution to the experimental groups.
Fig. 3. Counterbalanced group assignment based on actual data. (A) A data matrix containing scores from all subjects was sorted in descending order based on locomotion and rearing levels. (B-E) Subjects were equally distributed in four data matrices guided by their sorting position in the source data matrix. (F) Based on groups’ difference percentages per behavior, subjects were swapped between groups until acceptable levels were achieved. No group differences were found after the one-way ANOVA comparisons.
8. In your first data matrix, use your selected variables’ cumulative scores to sort subjects in descending order.
   a. Customize the sorting preferences by adding as many levels as variables you decide to include. For Microsoft Office Excel 365®, use the following path in the Home tab: Sort & filter → Custom sort → Add Level.
   b. Add as many levels as variables you include. Note: For this step, we recommend using only one or two levels.
   c. Indicate the priority order of the levels.
   d. Execute the sorting function.
9. Based on the sorting order, copy the subject’s information placed in the first row of the matrix and paste it in the first row of one of the group’s data matrix. Then, copy the second row’s information and paste it in the first row of the next group’s data matrix.
10. Repeat the process following the sorting order.
11. When all subjects were assigned to the groups, check groups’ difference percentages in the corresponding data matrix. Differences should remain below 10%, but slightly higher differences (≤15%) could be accepted. Note: We recommend making such a decision based on groups’ data range and variance.
12. Interchange subjects between groups to adjust the differences between groups. If you select locomotion and rearing as the most important sorting parameters, it is expected that they will be the most equally distributed variables. Follow the next steps to balance the differences observed in other variables:
   a. Identify the unbalanced distribution source by looking at the average scores of each groups’ data matrix. According to the previous formula, negative percentages correspond to higher scores in the second group, whereas positive percentages indicate the opposite.
   b. Search for animals with similar scores in the most homogeneous variables (e.g., locomotion and rearing).
   c. Identify which of those animals have scores in the unbalanced variables that, when interchanged between groups, equalize the differences in those specific variables.
   d. Interchange subjects’ rows between groups’ data matrixes and check the change in the different percentages.
   e. Follow these steps until you reach the appropriate homogeneity between groups.
13. Consider the following statements during subjects’ distribution:
   a. When litter information is available, reduce the number of littermates within groups and assign –as much as possible– an equal number of littermates to each group. Note: Despite it is well-reported that differences in maternal care affects brain and behavior (e.g., [29]), litter provenance from commercially-bred animals is not always available to take it into account in the experimental design. In those cases, omit this parameter and base subjects’ distribution on CT data.
   b. Body weight should also be balanced between groups, especially in psychopharmacological studies. However, as body weight does not correlate with CT parameters, it distributes almost randomly after balancing the subjects based on CT data.
   c. In behavioral studies, the homogeneity of behavioral parameters should be a priority, but the ultimate decision must be based on theoretical and empirical criteria.
14. Once difference percentages reach the lowest values across all variables, run one-way ANOVAs to assess group differences per variable. Based on the sample size, the Student’s t-test can also be used. Significant differences should not be observed in any of the variables.
15. Finally, number the groups and use a random number generator for their assignment to the experimental conditions. Note: There are many different random generators online. We recommend https://www.randomlists.com/.

Method validation

A balanced assignment is a laborious task requiring patience, practice, and objectivity, but their benefits pay-off its implementation and, to a greater extent for those paradigms using small sample
sizes and studying dependent variables with high interindividual variability (e.g., USVs or psychomotor response to drugs of abuse). In Fig. 3, we illustrate the process by which actual data from 35 male Wistar rats (~150 PND) was used for their balanced assignment into four experimental groups (unpublished data). These animals were tested as described in this protocol, and their behaviors were scored according to Table 1. In this case, we aimed to study the effect of different drug doses on OF habituation—defined as the gradual reduction in locomotion and rearing levels throughout the test. Consequently, we prioritize the homogeneity of those parameters while keeping other behaviors as equally distributed as possible. An additional challenge from this data set was that an odd number of subjects had to be split into an even number of groups. In our experience, this situation requires more subjects’ swapping iterations between groups. After assigning subjects randomly to the groups, the next decision is which group will be designated as the control group. Such a decision could also be made with a random-assignment method, but the researchers could deliberately choose the group with the smallest number of subjects as the control group. A closer look to the data set reveals that interindividual variability varies substantially among behavioral parameters (Fig. 3A), with differences in the data range going from ~550% and ~250% for locomotion and rearing, respectively; to an impressive ~2500% and ~15000% range for 50 kHz USVs and grooming, respectively. In this sample, locomotion (skewness = .03; kurtosis = -.70) and rearing (skewness = .37; kurtosis = -.40) followed a normal distribution and were highly related to each other (Pearson’s r = .77, P < .001), whereas the 50 kHz USVs showed an non-normal distribution (skewness = 2.72; kurtosis = 8.03) and were unrelated with any other parameter. Grooming was also unrelated to other behaviors but showed a relatively normal distribution (skewness = 1.23; kurtosis = 1.32). It is worth noting that an uneven subjects’ allocation could negatively affect statistical analyses, which might be especially problematic when key dependent variables show wide interindividual variability, such as 50 kHz USVs [13,34]. In this example, our balanced assignment leads into group differences ≤6% in locomotion and rearing and ≤12% in other behaviors (Fig. 3F), which is ideal for the intended experiment. Litter homogeneity was not particularly relevant in this test, but it was still balanced to the greatest extent possible without compromising the equal distribution of behavioral parameters. Since these animals were drug-treated later on in the experiment, body weight was also balanced between groups (group differences ≤5%). We recently reported two different works where CT was used for balancing subjects between groups. In the first report, differences in locomotion, rearing, and grooming were practically nil between unhandled animals tested in an OF taking place one week after the CT [28]. In the second report, levels of physical activity, social interaction, and 50-kHz USVs were indistinguishable between two groups of rats exposed by the first time to an environmental enrichment cage [31]. These groups were housed under two different enrichment protocols for one month and then tested in four one-day apart OF for 15 min. During their first OF assessment, locomotion, rearing, grooming (i.e., cephalic grooming subtype), and 50 kHz USVs were indistinguishable between both groups. Therefore, this evidence suggests that counterbalanced assignments based on CT parameters equalize group differences assessed on further tests.

Considering that other unconditioned anxiety tests such as the OF and the elevated plus-maze are typically used as readouts in a vast array of rodent models, the current approach constitutes an alternative providing similar information but minimizing possible carry-over effects (e.g., [6]). In this regard, we consider that CT could be used not just as a screening tool but also for studying individual differences. For example, animals showing high spontaneous, novelty-induced exploratory activity also display increased responsiveness to amphetamine administration [3,15,23]. These paradigms commonly use the OF for animals’ classification and later assessments, but we suggest that CT could replace the first screening evaluation, which might prevent novelty habituation caused by repeated OF [30]. When a group of unmanipulated animals was classified based on locomotor activity during their first testing minute, high phenotypical differences in their exploratory patterns were observed during the whole assessment (Fig. 4). Rats above one standard deviation (SD) displayed high locomotion levels that progressively decreased throughout the testing minutes (i.e., novelty habituation). Conversely, rats below one SD showed stable, low levels of locomotion with no significant test changes. This evidence highlights that individual differences in novelty-induced reactivity could be detected even in a straightforward assessment such as the CT, which in turn provides useful information for a balanced subjects’ assignment into the experimental groups.
Fig. 4. Classification based on the animal’s distance traveled during the first testing minute in the CT. Male Wistar rats (N=44; ~250g; ~100 PND) were grouped in three categories based on locomotion displayed during the first testing minute: animals within one standard deviation (1 SD) (blue line, n = 35), animals <1 SD (red line, n = 5), and animals >1 SD (green line, n = 4). Then, repeated-measures ANOVA was run with minute as within-group factor and group as a between-group factor. Locomotion decreased over testing minutes (Minute: $F_{(4,364)}=42.05, \ p<.001, \ \eta^2=.51$) but in a group-dependent manner (Minute Group: $F_{(8,364)}=5.42, \ p<.001, \ \eta^2=.20$): >1 SD ($F_{(4,12)}=21.00, \ p<.001, \ \eta^2=.87$) and 1 SD animals ($F_{(4,136)}=69.16, \ p<.001, \ \eta^2=.67$) showed a significant reduction thru test while <1 SD showed no significant changes along testing minutes. Considering this evidence, temporal analyses of behavioral data could be also useful for balancing groups and for the analysis of individual differences.

Concluding remarks

Even when a balanced-group assignment is seldom practiced in the field, it is even more uncommon that experiments using this approach report the followed procedure. Here, we offered a simple, straightforward behavioral assessment method to analyze novelty-induced reactivity and a detailed group formation strategy based on the studied parameters. Considering that pre-treatment conditions leading to individual differences usually escape researchers’ control, balanced group formation constitutes an excellent strategy to reduce the influence of confounding variables in the experiments. Furthermore, this approach increases the internal validity of experimental designs in which dependent variables show high interindividual variability. We believe that a detailed description of different methods and protocols strongly contributes to the sophistication of experimental design and behavioral analysis and reduces replicability and reproducibility problems in the field. Nevertheless, it should be acknowledged that the current method might not be suitable for all experiments using rats, which by no means indicate a less rigorous or valid approach. In fact, many different paradigms with solid independent variables might not be affected by such interindividual variability. However, experiments with small sample sizes, independent variables with middle-to-low effect sizes, and behavioral or physiological readouts with high interindividual variability could benefit most from this approach.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.
Acknowledgment

This protocol was custom-designed based on different studies (e.g., [7,10,11]), and was partially supported by the projects 723-B9-197, 837-B7-603, and 837-B8-123, Vice-rectory of Research, University of Costa Rica. A special acknowledgment to Daniela Sandoval, B.A who consecrated the artwork used of the manuscript. We would also like to thank Ph.D. Jaime Fornaguera Trías for his support.

References


