nature research

Corresponding author(s): Andreas T Schaefer

Last updated by author(s): 3/14/2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Odour transmission data was collected with commercially available Spike 2 software (Version 8.10). Behavioral data was collected with custom open-source Python (3.6) code that controlled the automated behavior system: github.com/RoboDoig/autonomouse-control, with Qt5 as the framework for GUI generation. Other custom Python modules used to run this software are also availale at github.com/RoboDoig. 2-photon data was acquired with commercially available Scientifica software (SciScan 1.3). Extracellular unit recordings data was collected using the Open Ephys system (Version 0.5.3.1) and analysed using Kilosort 2. Experiment control and odour stimulation in 2-photon and electrophysiology experiments was orchestrated with custom open-source Python software: github.com/RoboDoig/PulseBoy and http://github.com/warnerwarner/PulseBoy Sound recordings were acquired via a Focusrite Scarlett 18i8 USB audio interface and saved using Audacity 2.4.2. Code and related data for the OSN model can be found at https://github.com/stootoon/crick-osn-model-release.
Data analysis	All data analysis was performed with Python (3.6 and 3.8), ImageJ (1.52), suite2p, Igor Pro 6 or MATLAB (2017b-2020a). Code related to the glomerular classifier analysis is available at https://github.com/stootoon/crick-osn-decoding-release.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data related to the OSN model is available at https://github.com/stootoon/crick-osn-model-release.
Data related to the glomerular classifier analysis is available at https://github.com/stootoon/crick-osn-decoding-release.
The remaining data that support the findings of this study are available from the corresponding author upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Maximum sample sizes for each cohort were determined by the availability of simultaneously weaned mice to allow for large groups of male mice to be housed simultaneously for long periods of time without significant aggression within the group. Minimum sample sizes for each cohort were determined by requiring 3+ mice be in each subgroup within a cohort. Three separate experimental cohorts of mice were used for AutonoMouse experiments (correlation discrimination: cohort 1: n = 14 mice, cohort 2: n = 25 mice; cohort 3: plume discrimination: n = 24 mice). For physiological experiments a total of 49 mice were used (glomerular imaging: n = 9 mice, Mitra I/Tufted cell imaging: 9 mice, extracellular unit: n = 6 mice and whole-cell patch clamp recordings n = 25 mice). We found that due to the consistency in the structure of odour representations apparent across individuals, the number of animals sufficed for each experimental approach.
Data exclusions	For behavioral experiments, in the case that animals did not complete the pre-training phase (did not learn to gain sufficient daily water through self-initiated tasks) they were removed from the experiment. This was a pre-determined condition for experimental animal removal. For correlation discrimination experiments, one animal did not successfully pass the pre-training (1/39). For plume discrimination experiments, the second half of the cohort (n = 12/24 mice) with reversed valence (one source (S+, unrewarded), from separated sources (S-, rewarded) were not included in the analysis or carried forward to probe trials as they did not pass the performance criterion within the given timeframe. For imaging experiments, exclusion criteria were pre-established to select for experiments where fields of view could be imaged continuously for at least 2 hours with minimal drift and motion artifacts and where odour-evoked activity could be detected over the course of the entire imaging session. For extracellular recordings, units were classified as 'good' if they displayed a well defined inter-spike interval, a stable waveform, and a minimally varying baseline firing rate over at least half of the recording. Units were not selected by their odour responsiveness. For whole-cell patch recordings, cells with series resistance <25 MOhm were used for further analysis. No other data was excluded from our analyses.
Replication	For the odour transmission data, results were gathered from 3 environmental scenarios over 3 separate days and experiments. The principle result was confirmed across all these experiments and was reproducible from day-to-day. In the behavioral experiments, many of the key results (correlation detection, psychophysical threshold for correlation detection) were initially reported in one cohort of mice and then confirmed in an entirely separate cohort containing more animals and more detailed controls. Evidence for olfactory bulb input and output responses encoding odour correlation structure or paired pulses was found in 49 animals (imaging: n = 18 mice, extracellular recordings: n = 6 mice, whole-cell patch recordings: n = 25 mice) across 49 independent experiments.
Randomization	In all behavioural experiments, animals were first trained on a simple odour discrimination go/no-go task. Based on the performance levels in this task, animals were randomly assigned to different test subgroups until performance levels between the subgroups were statistically indistinguishable by a 1-way ANOVA.
Blinding	Due to the group selection method, experimenters could not be completely blinded to group allocation, as different experimental parameters had to be manually assigned to different animals based on their subgroup, in particular for test vs. control (scramble) mice. However, group allocation was fixed at the start of the experiment based on performance in a simple go/no-go task, and investigators did not move mice between groups after this initial choice. Also, since trial allocation was done entirely in software and the only distinguishing feature between mice was their RFID chip code (which could only be viewed in software), there was no possibility of investigators handling or otherwise treating mice differently due to their group selection. Therefore, although investigators could read-out the identity of the mice in the groups, it is very unlikely that this had any effect on subgroup differences.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	All mice used for behavioural experiments were 6-8 week old C57/BI6 males at experiment initiation, and were used in the experiment for a maximum of 78 weeks. All mice used for in vivo calcium imaging experiments were 12-20 week old Tbet-cre (Haddad et al., 2013) or OMP-cre (Ishii et al., 2003) mice crossed with a GCaMP6f reporter line (Otazu et al., 2015) of either sex. All mice used for extracellular and whole-cell recordings were 5-8 week old C57 /BI6 males. Mice were housed up to 5 per cage in a 12/12h light/dark cycle with food and water provided ad libitum.
Wild animals	The study did not involve wild animals.
Field-collected samples	The study did not involve field-collected samples.
Ethics oversight	All animal procedures performed in this study were approved by the UK government (Home Office) and by the Institutional Animal Welfare Ethical Review Panel.

Note that full information on the approval of the study protocol must also be provided in the manuscript.