

1 **Title**

2 THINGS-EEG: Human electroencephalography recordings for 1,854 concepts presented in
3 rapid serial visual presentation streams

4 **Authors**

5 Tijl Grootswagers^{1,2}, Ivy Zhou², Amanda K. Robinson², Martin N. Hebart³, Thomas A. Carlson²
6

7 **Affiliations**

8 1. The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Sydney,
9 Australia

10 2. School of Psychology, The University of Sydney, Sydney, Australia

11 3. Vision and Computational Cognition Group, Max Planck Institute for Human Cognitive and Brain
12 Sciences, Leipzig, Germany

13 corresponding author: Tijl Grootswagers (t.grootswagers@westernsydney.edu.au)

14 **Abstract**

15 The neural basis of object recognition and semantic knowledge have been the focus of a large
16 body of research but given the high dimensionality of object space, it is challenging to develop
17 an overarching theory on how brain organises object knowledge. To help understand how the
18 brain allows us to recognise, categorise, and represent objects and object categories, there is
19 a growing interest in using large-scale image databases for neuroimaging experiments.
20 Traditional image databases are based on manually selected object concepts and often single
21 images per concept. In contrast, ‘big data’ stimulus sets typically consist of images that can
22 vary significantly in quality and may be biased in content. To address this issue, recent work
23 developed THINGS: a large stimulus set of 1,854 object concepts and 26,107 associated
24 images. In the current paper, we present THINGS-EEG, a dataset containing human
25 electroencephalography responses from 50 subjects to all concepts and 22,248 images in the
26 THINGS stimulus set. The THINGS-EEG dataset provides neuroimaging recordings to a
27 systematic collection of objects and concepts and can therefore support a wide array of
28 research to understand visual object processing in the human brain.

29 **Background & Summary**

30 Humans are able to visually recognise and meaningfully interact with a large number of
31 different objects, despite drastic changes in retinal projection, lighting or viewing angle, and
32 the objects being positioned in cluttered visual environments. Object recognition and
33 semantic knowledge, our ability to make sense of the objects around us, have been the subject
34 of a large amount of cognitive neuroscience research^{1–3}. However, previous neuroimaging
35 research in this field has often relied on a manual selection of a small set of images^{1,4,5}. In
36 contrast, recent developments in computer vision have produced very large image sets for
37 training artificial intelligence, but the individual images in these sets are minimally curated and
38 therefore make them often unsuitable for research in psychology and neuroscience. To
39 overcome these issues, recent work has created large, curated image sets that are designed
40 for studying the cognitive and neural basis of human vision^{5,6}. One of these is THINGS⁵, which
41 is an image set containing 1,854 object concepts representing the most frequent concepts
42 used in the English language, accompanied by 26,107 associated manually-curated high-
43 quality image exemplars and human behavioural annotations. This rich collection of stimuli
44 and behavioural data has already been used to study the core representational dimensions
45 underlying human similarity judgements⁷. The next phase is to collate corresponding neural
46 responses to stimuli in THINGS. This would contribute to an emerging landscape of large
47 datasets of neural responses to curated image sets that accelerate research in visual,
48 computational, and cognitive neuroscience⁸.

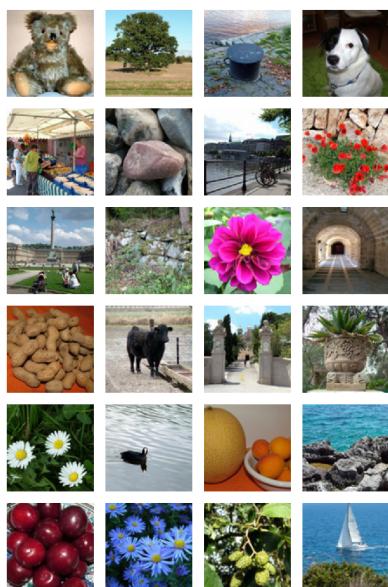
49
50 Collecting neurophysiological data for the THINGS dataset, with over 26,000 images, is
51 unachievable in a traditional neuroimaging experiment: Typically, classic object vision
52 experiments present around one image per second ^{e.g., 9–11}. Collecting one trial for each image
53 in THINGS would take more than seven hours, which is infeasible to achieve with a single
54 session and would require a complex design involving multiple scanning sessions. However,
55 we have recently shown that it is possible to uncover detailed information about visual stimuli
56 presented in rapid serial visual presentation (RSVP) streams using electroencephalography
57 (EEG)^{12–15}. In one of these studies¹², participants viewed over 16,000 visual object
58 presentations at 5 and 20 images per second, in a single 40-minute EEG session. Results from
59 multivariate pattern classification and representational similarity analysis revealed detailed
60 temporal dynamics of object processing that were similar to work that used slower
61 presentation speeds (around one image per second). Therefore, fast presentation paradigms
62 are highly suitable for collecting neural responses to the large number of visual object stimuli
63 in the THINGS database.
64
65 Here, we recorded human brain responses of 50 participants to all 1,854 object concepts in
66 THINGS using 22,248 of the THINGS stimuli (12 images per concept). Participants were also
67 presented with a separate validation set of 200 images that were repeated 20 times. Here, we
68 describe the THINGS-EEG dataset and recording procedure, and present initial technical
69 validation analyses. The THINGS-EEG dataset presents a rich resource of time-varying neural
70 recordings that we hope will be of great value for studying the temporal dynamics of human
71 object processing.

72 **Methods**

73 50 individuals volunteered to take part in the experiment in return for course credit. This
74 comprised 36 females and 14 males, mean age 20.44 (sd 2.72), age range 17 – 30. Participants
75 had different language profiles, with 26 native English speakers, 24 non-native speakers, 24
76 monolinguals, and 25 bilinguals. All participants reported normal or corrected-to-normal
77 vision. There are 4 participants marked for potential exclusion due to notably poor signal
78 quality or equipment failure (marked in the *participants.tsv* file). These participants are
79 included in the release for completeness. The study was approved by the University of Sydney
80 ethics committee. Informed consent was obtained from all participants at the start of the
81 experiment.
82

83 Stimuli were obtained from the THINGS database⁵ (Figure 1A). THINGS contains 1,854 objects
84 concepts, with 12 or more images per concept. The first 12 images for each concept were used
85 for this experiment, resulting in 22,248 different visual images, which we divided into 72
86 sequences of 309 stimuli. Individual concepts were never repeated within one sequence. To
87 be able to assess within- and between-subject variance on the same images, we presented
88 another 200 validation images from the THINGS database 12 times to every subject after the
89 main part of the experiment. For every subject we used the same 200 validation images (listed
90 in the *test_images.csv* file). We repeated these images in 12 sequences of 200, where we
91 presented them in random order. The experiment thus contained 84 sequences in total. The
92 subjects were not explicitly made aware of the two different parts.
93

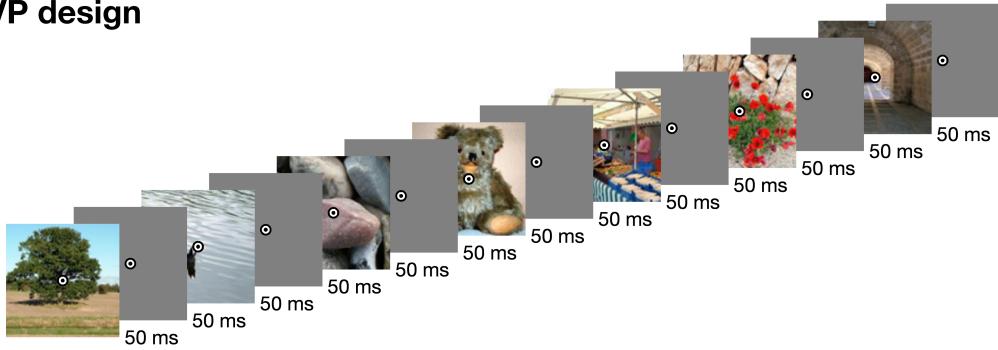
A Example images



B EEG setup



C RSVP design



94

95 **Figure 1:** Example stimuli, design, and EEG setup image. A) Example stimuli similar to the
96 stimuli used in the experiment. B) EEG experimental setup (photo credit AKR). C) Rapid serial
97 visual presentation design. For illustration purposes, only part of the sequence is shown. For
98 this figure, all images were replaced by public domain images with similar appearance
99 (obtained from PublicDomainPictures.net: Brunhilde Reinig).

100

101 The experiment was programmed in Python, using the Psychopy¹⁶ library. The sequences were
102 presented at 10Hz, with a 50% duty cycle (Figure 1C). That is, each image was presented for
103 50ms, followed by a 50ms blank screen. Participants were seated about 57cm from the screen,
104 and the stimuli subtended approximately 10 degrees visual angle. Overlaid at the centre of
105 each image was a bullseye (0.5 degrees visual angle) to help participants maintain fixation. To
106 increase attention and engagement, each sequence contained 2 to 5 random target events,
107 where the bullseye turned red for 100ms, and the participants were instructed to press a
108 button on a button box. At the end of each sequence, the display showed the progress through
109 the experiment, and participants were able to start the next sequence with a button press.
110 Participants were asked to sit still and minimise eye movements during the sequences and to
111 use the time between sequences as breaks and relax, and start the next sequence using a
112 button press when they were ready. The experiment lasted around one hour.

113

114 We used a BrainVision ActiChamp system to record continuous data while participants viewed
115 the sequences (Figure 1B). Conductive gel was used to reduce impedance at each electrode
116 site below 10 kOhm where possible. The median electrode impedance was under 18 kOhm in

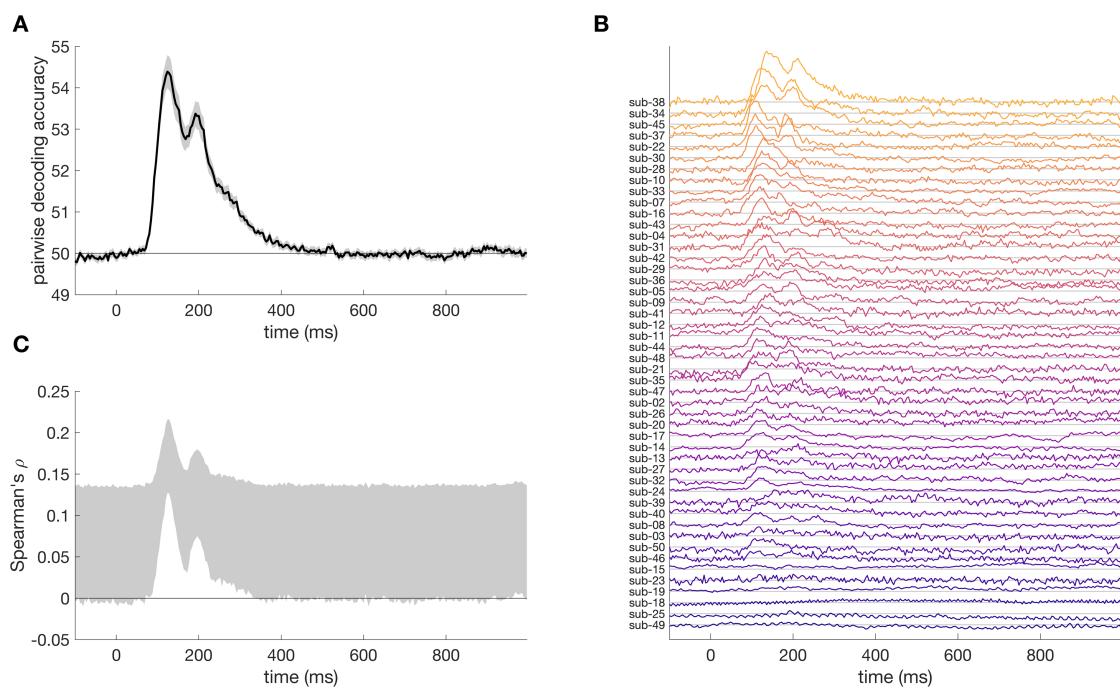
117 40/50 participants and under 60 kOhm in all participants. We used 64 electrodes, arranged
118 according to the international standard 10–10 system for electrode placement^{17,18}. The signal
119 was digitised at a 1000-Hz sample rate with a resolution of 0.0488281μV. Electrodes were
120 referenced online to Cz. An event trigger was sent over the parallel port at the start of each
121 sequence, and at every stimulus onset and offset event.
122

123 To perform basic quality checks and technical validation, for each subject, we ran a standard
124 decoding analysis. We decoded pairwise images for the 200 validation images, and we created
125 the full time-varying 1,854×1,854 Representational Dissimilarity Matrix^{19,20} reflecting the
126 pairwise decoding accuracies between all 1,854 object concepts. We used a minimal
127 preprocessing pipeline derived from our previous RSVP-MVPA studies^{12–15}. Using Matlab and
128 the EEGLab toolbox²¹, data were filtered using a Hamming windowed FIR filter with 0.1Hz
129 highpass and 100Hz lowpass filters, re-referenced to the average reference, and downsampled
130 to 250Hz. Epochs were created for each individual stimulus presentation ranging from [-100
131 to 1000ms] relative to stimulus onset. No further preprocessing steps were applied (e.g.,
132 baseline correction or epoch rejection), as in our previous work using similar presentation
133 paradigms^{12–15}. The channel voltages at each time point served as input to the decoding
134 analysis.
135

136 Decoding analyses were performed in Matlab using the CoSMoMVPA toolbox²². We first
137 decoded between the 200 validation images. For a given pairs of images, we used a leave-one-
138 sequence out (total: 12 sequences) cross-validation procedure and trained a regularised
139 ($\lambda=0.01$) linear discriminant classifier to distinguish between the images. The mean
140 classification accuracy on the image pair in the left-out sequences were stored in a
141 200×200×275×50 (image×image×time point×subject) RDM, which is symmetrical across its
142 first diagonal. A similar procedure was performed for the main experiment, using the 1,854
143 different image concepts. This resulted in an 1,854×1,854×275×50 (concept×concept×time
144 point×subject) RDM. For the 200 validation images, we also computed noise ceilings by
145 comparing between subject RDMs, as described in previous work²³. The noise ceilings estimate
146 the lower and upper bound of the highest achievable performance of a model that attempts
147 to explain variance in the data.

148 Data & Code Availability

149 All data and code are publicly available from the Open Science Framework
150 (<https://osf.io/hd6zk/>). The raw EEG recordings are hosted on Figshare in BIDS^{24,25} format
151 (<https://doi.org/10.6084/m9.figshare.14721282>). The preprocessed (Matlab/EEGLAB²¹
152 format), and epoched (Matlab/CoSMoMVPA²² format) data are included for convenience. The
153 code to reproduce the technical validation analyses and figures presented in this manuscript,
154 as well as the RDMs for the full set and the RDMs for the validation images (Matlab format)
155 are also available from the Open Science Framework (<https://osf.io/hd6zk/>), which also
156 contains links to the above repositories.



157

158 **Figure 2:** Results for the 200 validation images that were repeated 12 times at the end of the
159 session. A) mean pairwise classification accuracy over time. B) Mean pairwise decoding over
160 time, per subject, sorted by peak classification accuracy. Subjects 1 and 6 are not shown as
161 they did not have data on the validation images. C) Noise ceiling over time shows the expected
162 correlation of the ‘true’ model with the RDMs of the validation images and reflects the
163 between-subject variance in the RDMs.

164

Technical Validation Analyses

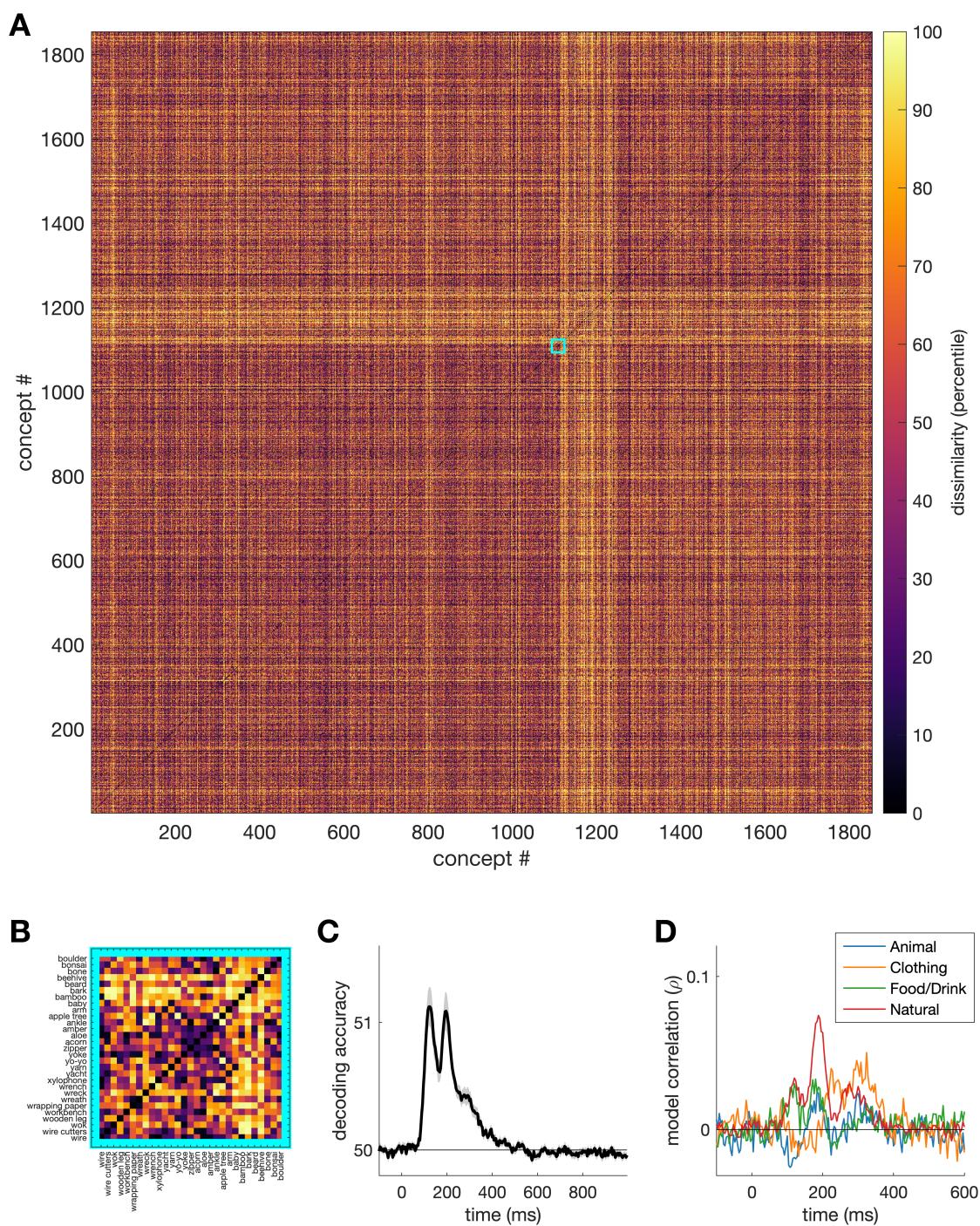
165

We computed representational dissimilarity matrices for the 200 validation images, by calculating time-varying decoding accuracy between all pairs of images. Mean, subject-wise decoding accuracy (Figure 2A) showed an initial peak around 100ms, and a second, lower peak around 200ms after stimulus onset. The shape of the time-varying decoding was similar to previous object decoding studies^{e.g., 9,10}, and was also similar to previous results on images presented in fast succession^{12–14}, indicating data quality was similar to these studies. For most subjects, this shape was apparent from their individual data (Figure 2B). The noise ceiling (Figure 2C) indicates an average similarity (correlation) of up to 0.2 between the subject-specific dissimilarity matrices.

174

175

Next, we computed the full RDM for all 1,854 images/concepts (1,717,731 pairs). The average accuracy within this RDM (Figure 3A) was lower than for the validation images, which is likely due to the fact that accuracy reflects generalised concept-similarity across images. Figure 3B shows the full RDM at one time point (200ms). To test if the values in the RDM contain meaningful information, we computed the correlation between the full RDM and four example categorical models (Figure 3C). The models coded for the presence of a certain category (e.g., animal). Figure 3C shows each model reaches an above-zero correlation, with the ‘natural’ model reaching the highest correlation, around 200ms.



183

184 **Figure 3:** Results for the 1,854 image concepts that were repeated 12 times (using a different
185 image each time). A) Full 1,854×1,854 RDM at 200ms, arranged by high-level category. B)
186 Zoomed in section of the full RDM. C) Mean pairwise classification accuracy between concepts
187 over time. D) Correlation over time between the neural RDM and four high-level categorical
188 models.
189

190 **Discussion**

191 Here, we presented THINGS-EEG, a dataset of human electro-encephalography responses to
192 22248 images from all 1,854 concepts in the THINGS object database. In the main session,
193 each image was repeated once, and in the second part, 200 validation images were repeated
194 12 times each to be able to assess data quality and compare the data to future datasets
195 acquired with other modalities. In total, 26,248 visual images were presented in a 1-hour EEG
196 session. This was achieved using a rapid serial visual processing paradigm. Technical validation
197 results indicated the dataset contains detailed neural responses to images, which shows that
198 the dataset is a high-quality resource for future investigations into the neurobiology of visual
199 object recognition.

200

201 In this study, we presented over 25,000 trials in a 1-hour EEG experiment. While this is an
202 exceptionally large number in a visual object perception study, the paradigm has several
203 limitations. Firstly, by presenting the images in rapid succession at 10Hz, new information is
204 being presented while previous trials are still being processed. While our previous work has
205 shown that a great amount of detail about objects can be extracted from brain responses to
206 RSVP streams^{12–15}, the images are being forward and backward masked, and therefore the
207 data does not capture the full brain response to each image¹³. For example, cognitive functions
208 such as memories or emotions may not have enough time to be instantiated at such rapid
209 presentation rates. Our design also involved one presentation per image, which makes image-
210 specific analyses challenging, placing the focus of this work at the level of the 1,854 object
211 concepts. The benefit of this is that the data has a built-in control for image-level confounds.
212 For example, visual regularities that are specific to an image will not affect the analysis at the
213 concept level.

214

215 The THINGS-EEG dataset has a lot of potential for investigating the neurobiology of visual
216 object recognition and semantic knowledge. The dataset presents a very large set of non-
217 invasive neural responses to visual stimuli in human participants. We foresee many possible
218 uses of this dataset. For example, the dataset could be used to test models of visual object
219 representation, such as different semantic models, or deep neural networks. The rapid
220 presentation paradigm allows to examine sequential effects in the data, such as how a specific
221 object concept influences the encoding of the subsequent presentations. Finally, the data
222 could be used to test the generalisability of previous studies that were limited by small
223 stimulus sets⁴. In sum, as THINGS is a high-quality stimulus set of record size, THINGS-EEG
224 accompanies this resource with a comprehensive set of human neuroimaging recordings.

225 **Acknowledgements**

226 This research was supported by ARC DP160101300 (TAC), ARC DP200101787 (TAC), and ARC
227 DE200101159 (AKR). The authors acknowledge the University of Sydney HPC service for
228 providing High Performance Computing resources.

229 **Author contributions**

230 TG: Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Data
231 Curation, Writing – Original Draft, Writing – Review & Editing, Supervision, Project
232 administration.

233 IZ: Conceptualization, Investigation, Data Curation, Writing – Review & Editing, Project
234 administration.

235 AKR: Conceptualization, Methodology, Writing – Review & Editing.

236 MNH: Conceptualization, Methodology, Writing – Review & Editing.

237 TAC: Conceptualization, Methodology, Writing – Review & Editing, Supervision, Funding
238 acquisition, Project administration.

239 **Competing interests**

240 The authors declare no competing interests.

241 **References**

- 242 1. Wardle, S. G. & Baker, C. Recent advances in understanding object recognition in the
243 human brain: deep neural networks, temporal dynamics, and context. *F1000Research* **9**, 590
244 (2020).
- 245 2. DiCarlo, J. J. & Cox, D. D. Untangling invariant object recognition. *Trends Cogn. Sci.* **11**,
246 333–341 (2007).
- 247 3. Gauthier, I. & Tarr, M. J. Visual Object Recognition: Do We (Finally) Know More Now
248 Than We Did? <http://dx.doi.org/10.1146/annurev-vision-111815-114621>
249 <http://www.annualreviews.org/doi/10.1146/annurev-vision-111815-114621> (2016).
- 250 4. Grootswagers, T. & Robinson, A. K. Overfitting the literature to one set of stimuli and
251 data. *ArXiv210209729 Q-Bio* (2021).
- 252 5. Hebart, M. N. *et al.* THINGS: A database of 1,854 object concepts and more than
253 26,000 naturalistic object images. *PLOS ONE* **14**, e0223792 (2019).
- 254 6. Mehrer, J., Spoerer, C. J., Jones, E. C., Kriegeskorte, N. & Kietzmann, T. C. An
255 ecologically motivated image dataset for deep learning yields better models of human vision.
256 *Proc. Natl. Acad. Sci.* **118**, (2021).
- 257 7. Hebart, M. N., Zheng, C. Y., Pereira, F. & Baker, C. I. Revealing the multidimensional
258 mental representations of natural objects underlying human similarity judgements. *Nat. Hum.
259 Behav.* **4**, 1173–1185 (2020).
- 260 8. Naselaris, T., Allen, E. & Kay, K. Extensive sampling for complete models of individual
261 brains. *Curr. Opin. Behav. Sci.* **40**, 45–51 (2021).
- 262 9. Carlson, T. A., Tovar, D. A., Alink, A. & Kriegeskorte, N. Representational dynamics of
263 object vision: The first 1000 ms. *J. Vis.* **13**, 1 (2013).
- 264 10. Cichy, R. M., Pantazis, D. & Oliva, A. Resolving human object recognition in space and
265 time. *Nat. Neurosci.* **17**, 455–462 (2014).
- 266 11. Kaneshiro, B., Guimaraes, M. P., Kim, H.-S., Norcia, A. M. & Suppes, P. A
267 Representational Similarity Analysis of the Dynamics of Object Processing Using Single-Trial
268 EEG Classification. *PLOS ONE* **10**, e0135697 (2015).
- 269 12. Grootswagers, T., Robinson, A. K. & Carlson, T. A. The representational dynamics of
270 visual objects in rapid serial visual processing streams. *NeuroImage* **188**, 668–679 (2019).
- 271 13. Robinson, A. K., Grootswagers, T. & Carlson, T. A. The influence of image masking on
272 object representations during rapid serial visual presentation. *NeuroImage* **197**, 224–231
273 (2019).
- 274 14. Grootswagers, T., Robinson, A. K., Shatek, S. M. & Carlson, T. A. Untangling featural
275 and conceptual object representations. *NeuroImage* **202**, 116083 (2019).
- 276 15. Grootswagers, T., Robinson, A. K., Shatek, S. M. & Carlson, T. A. The neural dynamics
277 underlying prioritisation of task-relevant information. *Neurons Behav. Data Anal. Theory* **5**, 1–
278 17 (2021).
- 279 16. Peirce, J. *et al.* PsychoPy2: Experiments in behavior made easy. *Behav. Res. Methods*
280 **51**, 195–203 (2019).
- 281 17. Jasper, H. H. The ten-twenty electrode system of the International Federation.
282 *Electroencephalogr Clin Neurophysiol* **10**, 371–375 (1958).
- 283 18. Oostenveld, R. & Praamstra, P. The five percent electrode system for high-resolution
284 EEG and ERP measurements. *Clin. Neurophysiol.* **112**, 713–719 (2001).
- 285 19. Grootswagers, T., Wardle, S. G. & Carlson, T. A. Decoding Dynamic Brain Patterns from
286 Evoked Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series
287 Neuroimaging Data. *J. Cogn. Neurosci.* **29**, 677–697 (2017).

- 288 20. Kriegeskorte, N., Mur, M. & Bandettini, P. A. Representational Similarity Analysis -
289 Connecting the Branches of Systems Neuroscience. *Front. Syst. Neurosci.* **2**, 4 (2008).
- 290 21. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial
291 EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21
292 (2004).
- 293 22. Oosterhof, N. N., Connolly, A. C. & Haxby, J. V. CoSMoMVPA: Multi-Modal Multivariate
294 Pattern Analysis of Neuroimaging Data in Matlab/GNU Octave. *Front. Neuroinformatics* **10**,
295 (2016).
- 296 23. Nili, H. *et al.* A Toolbox for Representational Similarity Analysis. *PLoS Comput Biol* **10**,
297 e1003553 (2014).
- 298 24. Gorgolewski, K. J. *et al.* The brain imaging data structure, a format for organizing and
299 describing outputs of neuroimaging experiments. *Sci. Data* **3**, 160044 (2016).
- 300 25. Pernet, C. R. *et al.* EEG-BIDS, an extension to the brain imaging data structure for
301 electroencephalography. *Sci. Data* **6**, 103 (2019).
- 302