

1 Title

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

4 Authors

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

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 has been extensively studied
16 but the high dimensionality of object space makes it challenging to develop overarching
17 theories on how the brain organises object knowledge. To help understand how the brain
18 allows us to recognise, categorise, and represent objects and object categories, there is a
19 growing interest in using large-scale image databases for neuroimaging experiments. In the
20 current paper, we present THINGS-EEG, a dataset containing human electroencephalography
21 responses from 50 subjects to 1,854 object concepts and 22,248 images in the THINGS
22 stimulus set, a manually curated and high-quality image database that was specifically
23 designed for studying human vision. The THINGS-EEG dataset provides neuroimaging
24 recordings to a systematic collection of objects and concepts and can therefore support a
25 wide array of research to understand visual object processing in the human brain.

26 Background & Summary

27 Humans are able to visually recognise and meaningfully interact with a large number of
28 different objects, despite drastic changes in retinal projection, lighting or viewing angle, and
29 the objects being positioned in cluttered visual environments. Object recognition and
30 semantic knowledge, our ability to make sense of the objects around us, have been the
31 subject of a large amount of cognitive neuroscience research¹⁻³. However, previous
32 neuroimaging research in this field has often relied on a manual selection of a small set of
33 images^{1,4,5}. In contrast, recent developments in computer vision have produced very large
34 image sets for training artificial intelligence, but the individual images in these sets are
35 minimally curated and therefore make them often unsuitable for research in psychology and
36 neuroscience. To overcome these issues, recent work has created large, curated image sets
37 that are designed for studying the cognitive and neural basis of human vision^{4,6}. One of these
38 is THINGS⁴, which is an image set containing 1,854 object concepts representing a
39 comprehensive set of nameable concepts used in the English language, accompanied by
40 26,107 associated manually-curated high-quality image exemplars and human behavioural
41 annotations. This rich collection of stimuli and behavioural data has already been used to
42 study the core representational dimensions underlying human similarity judgements⁷. The
43 next phase is to collate corresponding neural responses to stimuli in THINGS. This would
44 contribute to an emerging landscape of large datasets of neural responses to curated image
45 sets that accelerate research in visual, computational, and cognitive neuroscience⁸.

46
47 Collecting neurophysiological data for the THINGS dataset, with over 26,000 images, is
48 unachievable in a traditional neuroimaging experiment: Typically, classic object vision

49 experiments present around one image per second^{e.g., 9-11}. Collecting one trial for each
50 image in THINGS would take more than seven hours, which is infeasible to achieve with a
51 single session and would require a complex design involving multiple scanning sessions.
52 However, we have recently shown that it is possible to uncover detailed information about
53 visual stimuli presented in rapid serial visual presentation (RSVP) streams using
54 electroencephalography (EEG)¹²⁻¹⁵. In one of these studies¹², participants viewed over 16,000
55 visual object presentations at 5 and 20 images per second, in a single 40-minute EEG session.
56 Results from multivariate pattern classification and representational similarity analysis
57 revealed detailed temporal dynamics of object processing that were similar to work that
58 used slower presentation speeds (around one image per second). Therefore, fast
59 presentation paradigms are highly suitable for collecting neural responses to the large
60 number of visual object stimuli in the THINGS database.

61
62 Here, we present THINGS-EEG, a dataset of human (n=50) electro-encephalography
63 responses to 22248 images from all 1,854 concepts in the THINGS object database. In the
64 main session, each image was repeated once, and in the second part, 200 validation images
65 were repeated 12 times each to be able to assess data quality and compare the data to
66 future datasets acquired with other modalities. In total, 26,248 visual images were presented
67 in a 1-hour EEG session. This was achieved using a rapid serial visual presentation paradigm.
68 Technical validation results indicated the dataset contains detailed neural responses to
69 images, which shows that the dataset is a high-quality resource for future investigations into
70 the neurobiology of visual object recognition.

71
72 In this study, we presented over 25,000 trials in a 1-hour EEG experiment. While this is an
73 exceptionally large number in a visual object perception study, the paradigm has several
74 limitations. Firstly, by presenting the images in rapid succession at 10Hz, new information is
75 being presented while previous trials are still being processed. While our previous work has
76 shown that a great amount of detail about objects can be extracted from brain responses to
77 RSVP streams¹²⁻¹⁵, the images are being forward and backward masked, and therefore the
78 data does not capture the full brain response to each image¹³. For example, cognitive
79 functions such as memories or emotions may not have enough time to be instantiated at
80 such rapid presentation rates. Our design also involved one presentation per image, which
81 makes image-specific analyses challenging, placing the focus of this work at the level of the
82 1,854 object concepts. The benefit of this is that the data has a built-in control for image-
83 level confounds. For example, visual regularities that are specific to an image but vary across
84 a concept will not generalise to the concept level. Visual statistics that reliably differ at the
85 concept level of course may need to be accounted for, depending on the goals of the
86 experimenter. For example, recent work has pointed out concept-level differences in mean
87 luminance between images in the THINGS image set¹⁶, which can be controlled for in future
88 analyses of the THINGS-EEG dataset. Another point to consider is that our recording setup
89 did not include EOG or EMG channels, which means the dataset does not contain external
90 recordings of eye or other muscle movements. These movement patterns are unlikely to
91 contain informative stimulus-specific information, due to the fast presentation paradigm.
92 However, they still cause noise artefacts in the EEG data. Future users could consider
93 detecting and correcting for eye movements using the frontal EEG channels.

94
95 The THINGS-EEG dataset has strong potential for investigating the neurobiology of visual
96 object recognition and semantic knowledge. The dataset presents a very large set of non-
97 invasive neural responses to visual stimuli in human participants. We foresee many possible
98 uses of this dataset. For example, the dataset could be used to test models of visual object
99 representation, such as different semantic models, or deep neural networks. It could be used
100 to test the generalisability of previous studies that were limited by small stimulus sets⁵. The

101 rapid presentation paradigm allows to examine sequential effects in the data, such as how a
102 specific object concept influences the encoding of the subsequent presentations. The
103 consistent presentation frequency also lends itself to separate the data into oscillatory
104 components. Instead of the RSA and classification analyses presented here, it is also possible
105 to analyse the data in an encoding framework, for example by creating an encoding model
106 from the semantic information in the THINGS-dataset. In sum, as THINGS is a high-quality
107 stimulus set of record size, THINGS-EEG accompanies this resource with a comprehensive set
108 of human neuroimaging recordings.
109

110 **Methods**

111 50 individuals volunteered to take part in the experiment in return for course credit.
112 Participants were recruited from the undergraduate student population at the University of
113 Sydney. They were 36 females and 14 males, mean age 20.44 (sd 2.72), age range 17 – 30.
114 Participants had different language profiles, with 26 native English speakers, 24 non-native
115 speakers, 24 monolinguals, and 25 bilinguals. Handedness was not recorded. All participants
116 reported normal or corrected-to-normal vision and reported no neurological or psychiatric
117 disorders. There are 4 participants marked for potential exclusion due to notably poor signal
118 quality or equipment failure (marked in the *participants.tsv* file). These participants are
119 included in the release for completeness. The study was approved by the University of
120 Sydney ethics committee. Informed consent was obtained from all participants at the start of
121 the experiment.
122

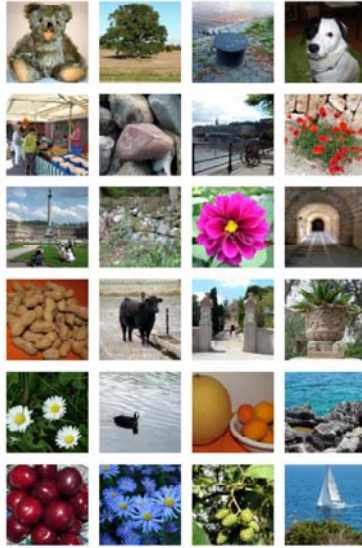
123 Stimuli were obtained from the THINGS database⁴ (Figure 1A). For detailed information on
124 the contents and organisation of this stimulus database, readers are referred to the
125 accompanying publication⁴. THINGS contains 1,854 objects concepts, with 12 or more images
126 per concept. The first 12 images for each concept were used for this experiment, resulting in
127 22,248 different visual images, which we divided into 72 sequences of 309 stimuli. Individual
128 concepts were never repeated within one sequence. To be able to assess within- and
129 between-subject variance on the same images, we presented another 200 validation images
130 from the THINGS database 12 times to every subject after the main part of the experiment.
131 For every subject we used the same 200 validation images (listed in the *test_images.csv* file).
132 We repeated these images in 12 sequences of 200, where we presented them in random
133 order. The experiment thus contained 84 sequences in total. The subjects were not explicitly
134 made aware of the two different parts.
135

136 The experiment was programmed in Python (v3.7), using the Psychopy¹⁷ library (version
137 3.0.5). The sequences were presented at 10Hz, with a 50% duty cycle (Figure 1C). That is,
138 each image was presented for 50ms, followed by a 50ms blank screen. Participants were
139 seated about 57cm from the screen, and the stimuli subtended approximately 10 degrees
140 visual angle. Overlaid at the centre of each image was a bullseye (0.5 degrees visual angle) to
141 help participants maintain fixation. To increase attention and engagement, each sequence
142 contained 2 to 5 random target events, where the bullseye turned red for 100ms, and the
143 participants were instructed to press a button on a button box using their right hand. These
144 target events are marked in the dataset, as researchers may want to consider excluding
145 these target events, depending on the aim of their analysis. At the end of each sequence, the
146 display showed the progress through the experiment, and participants were able to start the
147 next sequence with a button press. Participants were asked to sit still and minimise eye
148 movements during the sequences and to use the time between sequences as breaks and
149 relax, and start the next sequence using a button press when they were ready. The
150 experiment lasted around one hour.
151

152 We used a BrainVision ActiChamp system to record continuous data while participants
153 viewed the sequences (Figure 1B). Conductive gel was used to reduce impedance at each
154 electrode site below 10 kOhm where possible. The median electrode impedance was under
155 18 kOhm in 40/50 participants and under 60 kOhm in all participants. We used 64 electrodes,
156 arranged according to the international standard 10–10 system for electrode placement^{18,19}.
157 The signal was digitised at a 1000-Hz sample rate with a resolution of 0.0488281µV.
158 Electrodes were referenced online to Cz. An event trigger was sent over the parallel port at
159 the start of each sequence (trigger code E3), and at every stimulus onset event (trigger code
160 E1) and stimulus offset event (trigger code E2).

161
162 To perform basic quality checks and technical validation, for each subject, we ran a standard
163 decoding analysis. We decoded pairwise images for the 200 validation images, and we
164 created the full time-varying 1,854×1,854 Representational Dissimilarity Matrix^{20,21} reflecting
165 the pairwise decoding accuracies between all 1,854 object concepts. We used a minimal
166 preprocessing pipeline derived from our previous RSVP-MVPA studies^{12–15}. Using Matlab
167 (R2020b) and the EEGLab (v14.0.0b) toolbox²², data were filtered using a Hamming
168 windowed FIR filter with 0.1Hz highpass and 100Hz lowpass filters, re-referenced to the
169 average reference, and downsampled to 250Hz. Epochs were created for each individual
170 stimulus presentation ranging from [-100 to 1000ms] relative to stimulus onset. No further
171 preprocessing steps were applied for the technical validation analysis presented here (as in
172 our previous work using similar presentation paradigms^{12–15}). Researchers may want to
173 consider popular preprocessing steps such as baseline correction or eye movement
174 correction. The channel voltages at each time point served as input to the decoding analysis.
175

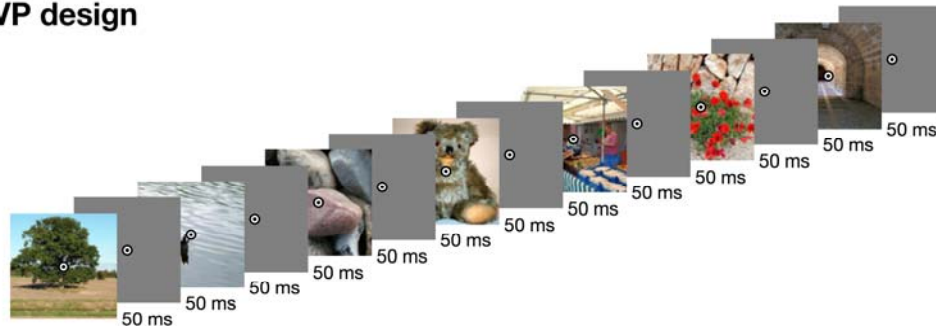
A Example images



B EEG setup



C RSVP design



176

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

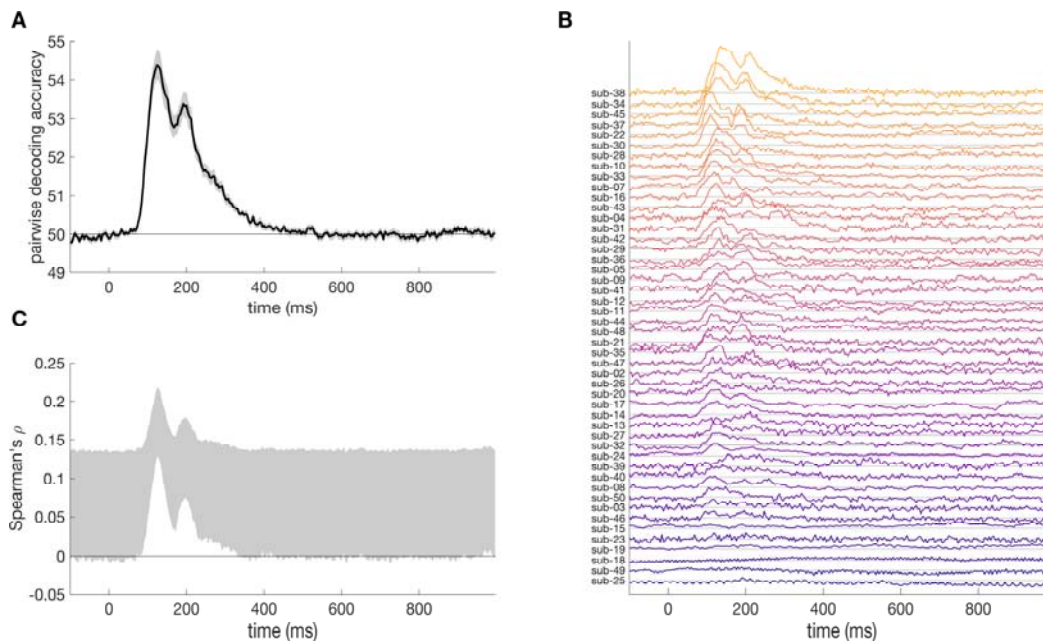
182

183 Decoding analyses were performed in Matlab using the CoSMoMvpa toolbox²³. We first
184 decoded between the 200 validation images. For a given pairs of images, we used a leave-
185 one-sequence out (total: 12 sequences) cross-validation procedure and trained a regularised
186 ($\lambda=0.01$) linear discriminant classifier to distinguish between the images. The mean
187 classification accuracy on the image pair in the left-out sequences were stored in a
188 $200 \times 200 \times 275 \times 50$ (image \times image \times time point \times subject) RDM, which is symmetrical across its
189 first diagonal. A similar procedure was performed for the main experiment, using the 1,854
190 different image concepts. This resulted in an $1,854 \times 1,854 \times 275 \times 50$ (concept \times concept \times time
191 point \times subject) RDM. For the 200 validation images, we also computed noise ceilings by
192 comparing between subject RDMs, as described in previous work²⁴. The noise ceilings
193 estimate the lower and upper bound of the highest achievable performance of a model that
194 attempts to explain variance in the data.

195 Data Records

196 All data and code are publicly available. The raw EEG recordings are hosted in BIDS^{25,26}
197 format on OpenNeuro (<https://doi.org/10.18112/openneuro.ds003825.v1.1.0>)²⁷. The

198 preprocessed (Matlab/EEGLAB²² format) data, and group-average RDMs are included for
199 convenience (in the data/derivatives directory). The RDMs for individual subjects are hosted
200 in Matlab format in a separate repository on Figshare
201 (<https://doi.org/10.6084/m9.figshare.14721282>)²⁸. All custom code is available from the
202 Open Science Framework (<https://doi.org/10.17605/OSF.IO/HD6ZK>)²⁹, which also contains
203 links to the above repositories.
204

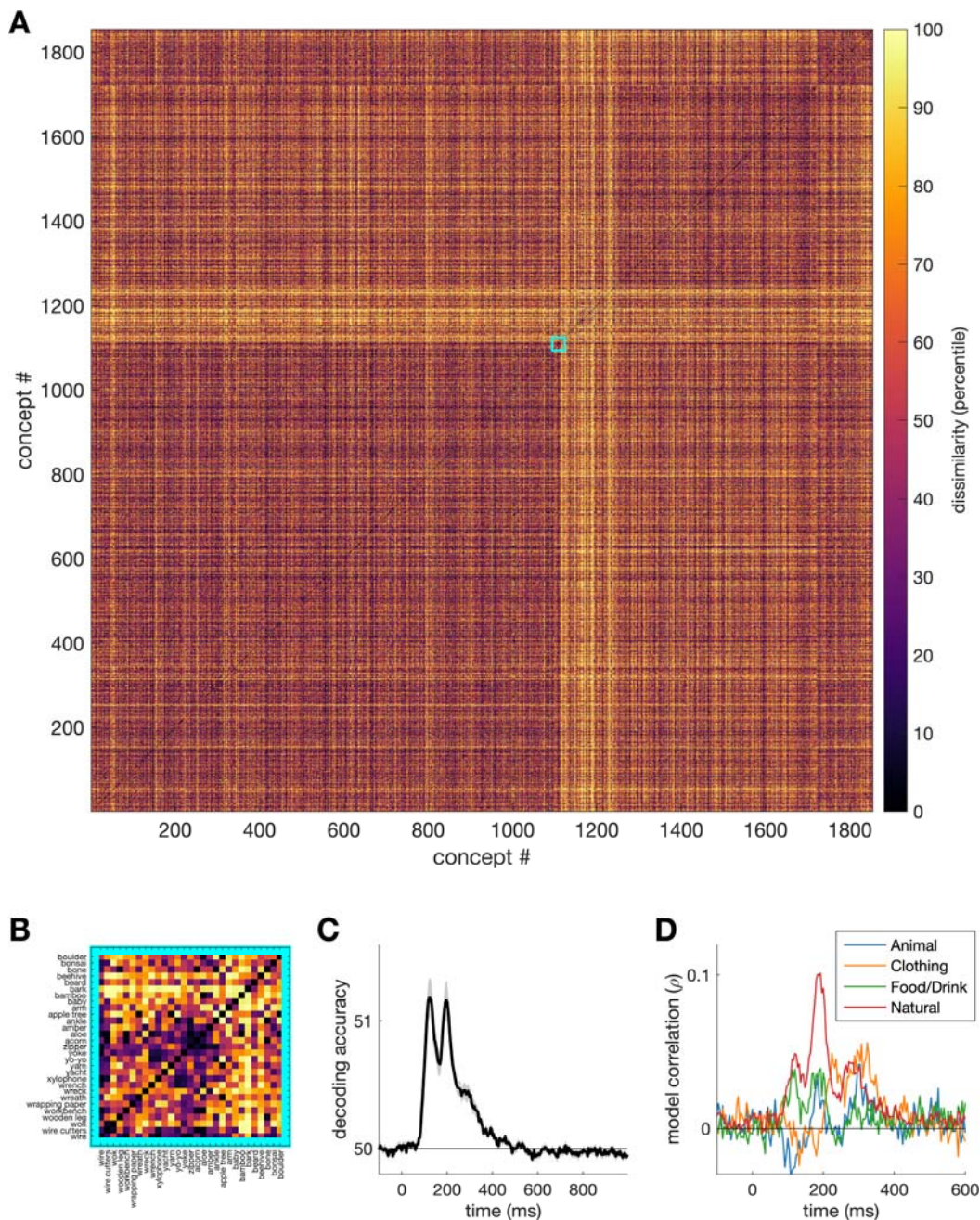


205
206 **Figure 2:** Results for the 200 validation images that were repeated 12 times at the end of the
207 session. A) mean pairwise classification accuracy over time. B) Mean pairwise decoding over
208 time, per subject, sorted by peak classification accuracy. Subjects 1 and 6 are not shown as
209 they did not have data on the validation images. C) Noise ceiling over time shows the
210 expected correlation of the 'true' model with the RDMs of the validation images and reflects
211 the between-subject variance in the RDMs.

212 Technical Validation

213 We computed representational dissimilarity matrices for the 200 validation images, by
214 calculating time-varying decoding accuracy between all pairs of images. Mean, subject-wise
215 decoding accuracy (Figure 2A) showed an initial peak around 100ms, and a second, lower
216 peak around 200ms after stimulus onset. The shape of the time-varying decoding was similar
217 to previous object decoding studies^{e.g., 9,10}, and was also similar to previous results on images
218 presented in fast succession¹²⁻¹⁴, indicating data quality was similar to these studies. For
219 most subjects, this shape was apparent from their individual data (Figure 2B). The noise
220 ceiling (Figure 2C) indicates an average similarity (correlation) of up to 0.2 between the
221 subject-specific dissimilarity matrices.
222

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



231

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

237 Code Availability

238 Code and detailed instructions to reproduce the technical validation analyses and figures
239 presented in this manuscript are available from the Open Science Framework
240 (<https://doi.org/10.17605/OSF.IO/HD6ZK>)²⁹, which also contains links to the data
241 repositories.
242

243 **Acknowledgements**

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

247 **Author contributions**

248 TG: Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Data
249 Curation, Writing – Original Draft, Writing – Review & Editing, Supervision, Project
250 administration.

251 IZ: Conceptualization, Investigation, Data Curation, Writing – Review & Editing, Project
252 administration.

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

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

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

257 **Competing interests**

258 The authors declare no competing interests.

259 **References**

260 1. Wardle, S. G. & Baker, C. Recent advances in understanding object recognition in the

261 human brain: deep neural networks, temporal dynamics, and context. *F1000Research* **9**,

262 590 (2020).

263 2. DiCarlo, J. J. & Cox, D. D. Untangling invariant object recognition. *Trends Cogn. Sci.* **11**,

264 333–341 (2007).

265 3. Gauthier, I. & Tarr, M. J. Visual Object Recognition: Do We (Finally) Know More Now

266 Than We Did? <http://dx.doi.org/10.1146/annurev-vision-111815-114621>

267 <http://www.annualreviews.org/doi/10.1146/annurev-vision-111815-114621> (2016).

268 4. Hebart, M. N. *et al.* THINGS: A database of 1,854 object concepts and more than 26,000

269 naturalistic object images. *PLOS ONE* **14**, e0223792 (2019).

270 5. Grootswagers, T. & Robinson, A. K. Overfitting the Literature to One Set of Stimuli and

271 Data. *Front. Hum. Neurosci.* **15**, (2021).

272 6. Mehrer, J., Spoerer, C. J., Jones, E. C., Kriegeskorte, N. & Kietzmann, T. C. An ecologically

273 motivated image dataset for deep learning yields better models of human vision. *Proc.*

274 *Natl. Acad. Sci.* **118**, (2021).

- 275 7. Hebart, M. N., Zheng, C. Y., Pereira, F. & Baker, C. I. Revealing the multidimensional
276 mental representations of natural objects underlying human similarity judgements. *Nat.*
277 *Hum. Behav.* **4**, 1173–1185 (2020).
- 278 8. Naselaris, T., Allen, E. & Kay, K. Extensive sampling for complete models of individual
279 brains. *Curr. Opin. Behav. Sci.* **40**, 45–51 (2021).
- 280 9. Carlson, T. A., Tovar, D. A., Alink, A. & Kriegeskorte, N. Representational dynamics of
281 object vision: The first 1000 ms. *J. Vis.* **13**, 1 (2013).
- 282 10. Cichy, R. M., Pantazis, D. & Oliva, A. Resolving human object recognition in space and
283 time. *Nat. Neurosci.* **17**, 455–462 (2014).
- 284 11. Kaneshiro, B., Guimaraes, M. P., Kim, H.-S., Norcia, A. M. & Suppes, P. A
285 Representational Similarity Analysis of the Dynamics of Object Processing Using Single-
286 Trial EEG Classification. *PLOS ONE* **10**, e0135697 (2015).
- 287 12. Grootswagers, T., Robinson, A. K. & Carlson, T. A. The representational dynamics of
288 visual objects in rapid serial visual processing streams. *NeuroImage* **188**, 668–679 (2019).
- 289 13. Robinson, A. K., Grootswagers, T. & Carlson, T. A. The influence of image masking on
290 object representations during rapid serial visual presentation. *NeuroImage* **197**, 224–231
291 (2019).
- 292 14. Grootswagers, T., Robinson, A. K., Shatek, S. M. & Carlson, T. A. Untangling featural and
293 conceptual object representations. *NeuroImage* **202**, 116083 (2019).
- 294 15. Grootswagers, T., Robinson, A. K., Shatek, S. M. & Carlson, T. A. The neural dynamics
295 underlying prioritisation of task-relevant information. *Neurons Behav. Data Anal. Theory*
296 **5**, 1–17 (2021).
- 297 16. Harrison, W. J. Luminance and contrast of images in the THINGS database. *bioRxiv*
298 2021.07.08.451706 (2021) doi:10.1101/2021.07.08.451706.
- 299 17. Peirce, J. *et al.* PsychoPy2: Experiments in behavior made easy. *Behav. Res. Methods* **51**,
300 195–203 (2019).

- 301 18. Jasper, H. H. The ten-twenty electrode system of the International Federation.
302 *Electroencephalogr Clin Neurophysiol* **10**, 371–375 (1958).
- 303 19. Oostenveld, R. & Praamstra, P. The five percent electrode system for high-resolution EEG
304 and ERP measurements. *Clin. Neurophysiol.* **112**, 713–719 (2001).
- 305 20. Grootswagers, T., Wardle, S. G. & Carlson, T. A. Decoding Dynamic Brain Patterns from
306 Evoked Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series
307 Neuroimaging Data. *J. Cogn. Neurosci.* **29**, 677–697 (2017).
- 308 21. Kriegeskorte, N., Mur, M. & Bandettini, P. A. Representational Similarity Analysis -
309 Connecting the Branches of Systems Neuroscience. *Front. Syst. Neurosci.* **2**, 4 (2008).
- 310 22. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG
311 dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21
312 (2004).
- 313 23. Oosterhof, N. N., Connolly, A. C. & Haxby, J. V. CoSMoMVPA: Multi-Modal Multivariate
314 Pattern Analysis of Neuroimaging Data in Matlab/GNU Octave. *Front. Neuroinformatics*
315 **10**, (2016).
- 316 24. Nili, H. *et al.* A Toolbox for Representational Similarity Analysis. *PLoS Comput Biol* **10**,
317 e1003553 (2014).
- 318 25. Gorgolewski, K. J. *et al.* The brain imaging data structure, a format for organizing and
319 describing outputs of neuroimaging experiments. *Sci. Data* **3**, 160044 (2016).
- 320 26. Pernet, C. R. *et al.* EEG-BIDS, an extension to the brain imaging data structure for
321 electroencephalography. *Sci. Data* **6**, 103 (2019).
- 322 27. Grootswagers, T., Zhou, I., Robinson, A., Hebart, M. N. & Carlson, T. A. [dataset] Human
323 electroencephalography recordings from 50 subjects for 22,248 images from 1,854
324 object concepts. (2021) doi:10.18112/openneuro.ds003825.v1.1.0.

- 325 28. Grootswagers, T., Zhou, I., Robinson, A., Hebart, M. N. & Carlson, T. A. [dataset] THINGS-
326 EEG: Human electroencephalography recordings for 22,248 images from 1,854 object
327 concepts. (2021) doi:<https://doi.org/10.6084/m9.figshare.14721282>.
- 328 29. Grootswagers, T., Zhou, I., Robinson, A., Hebart, M. N. & Carlson, T. A. [dataset] THINGS-
329 EEG: Human electroencephalography recordings from 50 subjects for 22,248 images
330 from 1,854 object concepts. (2021) doi:[10.17605/OSF.IO/HD6ZK](https://doi.org/10.17605/OSF.IO/HD6ZK).
- 331