

Extensive sampling for complete models of individual brains

Thomas Naselaris^{*,1}, Emily Allen², Kendrick Kay³

¹Department of Neuroscience, Medical University of South Carolina

²Department of Psychology, University of Minnesota

³Center for Magnetic Resonance Research, Department of Radiology, University of Minnesota

*Corresponding author (tnaselar@musc.edu)

1 **Highlights**

- 2 • Trade-off between sampling individual variation vs. experimental variation
- 3 • Different studies have allocated resources differently
- 4 • We argue that wide sampling of individuals provides few benefits
- 5 • We discuss design decisions made for the Natural Scenes Dataset

Abstract

In designing cognitive neuroscience experiments, resource limitations induce a fundamental trade-off between sampling variation across individual brains and sampling variation across experimental conditions. Here, we argue that extensive sampling of experimental conditions is essential for understanding how human brains process complex stimuli, that a model of how any one brain does this is likely to generalize to most other brains, and that introducing large numbers of subjects into an analysis pool is likely to introduce unnecessary and undesirable variance. Thus, contrary to conventional wisdom, we believe that sampling many individuals provides relatively few benefits and that extensive sampling of a limited number of subjects is more productive for revealing general principles. Furthermore, an emphasis on depth in individual brains is well-suited for capitalizing on the improvements in resolution and signal-to-noise ratio that are being achieved in modern neuroscientific measurement techniques.

1 **Introduction**

2
3 When can cognitive neuroscientists stop doing experiments? Each experiment costs society time,
4 money, and effort, and can pose risks to individual subjects, so it is worth considering what our
5 end-goal is and how we would know when we have achieved it.

6
7 We assume that the end-goal is a complete model of the human brain [1–3]. A complete model
8 would be both oracle and interpreter: it would predict with great accuracy the brain activity and
9 behavior that we would observe in response to arbitrary stimuli and task instructions, and would
10 connect the observed responses to meaningful evolutionary, developmental, computational,
11 psychological, and philosophical narratives.

12
13 The cognitive neuroscience of vision (where our interests lie) works toward this end-goal one
14 stimulus at a time. Experimentalists present visual stimuli while measuring brain activity using
15 one of many techniques for interrogating neural activity (e.g., fMRI, EEG, MEG, ECoG,
16 electrophysiology). The cost of each experiment is the accumulated cost of measuring these
17 responses across many stimuli. In this paper we consider whether it is better to accumulate cost
18 by sampling variation across individuals, or by sampling variation across stimuli. Specifically,
19 we ask which approach yields greater progress toward the ambitious end-goal of a complete
20 model?

21 22 23 **Fundamental trade-off between sampling individual variation and sampling** 24 **stimulus variation**

25
26 Resource constraints on experimental design impose a fundamental trade-off between sampling
27 individual variation and sampling stimulus variation. We illustrate how a number of recent fMRI
28 studies have managed this trade-off by plotting them in a two-dimensional space with number of
29 individual subjects as the y -axis and number of hours of data per subject as the x -axis (**Figure 1**).
30 We acknowledge that the number of hours of scanning per subject is only a proxy for how
31 extensively a given experiment samples variation across stimuli (or, more generally,
32 experimental conditions). We include resting-state fMRI data collection in this plot under the
33 working premise that the “conditions” that vary in this case are unmeasured endogenous (e.g.,
34 cognitive, affective) states.

In an ideal world, all experiments would reside in the upper right (many subjects and many hours per subject). However, resource limitations force experimentalists to distribute their efforts along iso-hour contours that extend from the upper left (many subjects but few hours per subject) to the lower right (many hours per subject but few subjects). Traditional fMRI studies tend to collect one session of data on about 10–20 participants (see gray box). The studies illustrated in **Figure 1** have deliberately pushed sampling strategies towards one extreme (more subjects) or the other (more data per subject).

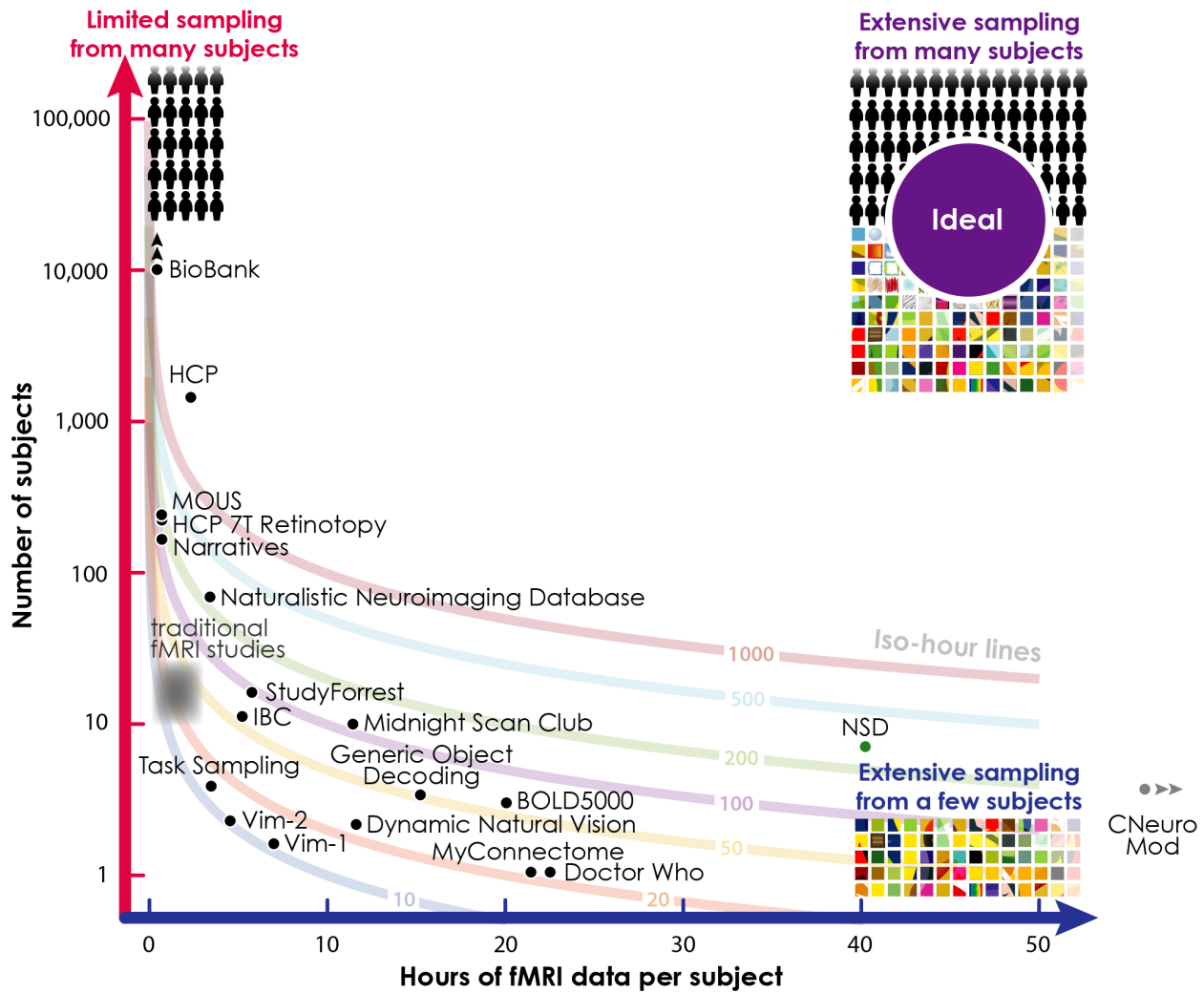


Figure 1. Trade-off between number of subjects and amount of data per subject. In the context of basic-science fMRI studies with publicly available data, we select a few representative publications, biasing towards recent and/or influential studies. Selected publications include Biobank [4], BOLD5000 [5], CNeuroMod [6], Doctor Who [7], Dynamic Natural Vision [8], Generic Object Decoding [9], HCP [10], HCP 7T Retinotopy [11], IBC [12], Midnight Scan Club [13], MOUS [14], MyConnectome [15], Narratives [16], Naturalistic Neuroimaging Database

[17], NSD [18], StudyForrest [19,20], Task Sampling [21], Vim-1 [22], and Vim-2 [23]. For each publication, we plot the number of subjects scanned and the number of hours of task-based and resting-state fMRI conducted for each subject (estimated based on information provided in the publication). The green dot denotes a soon-to-be-released study that is discussed in further detail in this article, the gray dot denotes an upcoming study, and the thick “iso-hour” lines indicate points that share the same number of total hours of fMRI data.

Efforts that involve extensive data collection across large numbers of individual brains but small numbers of experimental conditions [4,10,11,24] are in line with cognitive neuroscience’s historical penchant for studies with large (more than 20) numbers of subjects [25–27]. The intuition behind this approach seems to be that inferring a complete model of *the* human brain requires generalizing over response patterns in the brains of many individual humans. However, we argue that a complete model of *the* human brain must be able to predict and interpret activity in any *one* human brain, and that a complete model of the brain of any one individual will bring us within striking distance of a complete model that adequately describes most individuals. This line of thinking has motivated us, and presumably several others, to explore the other end of the spectrum, sampling many stimuli (or experimental conditions) in just a few individuals [5–7,15,18].

Many reasons for extensive sampling of stimulus variation

The images of the natural world that the eye transmits to the brain are ambiguous and complex [28]. While we can learn much from probing the visual system with simple stimuli [29], we ultimately want to understand how the brain solves the hard computational problem of interpreting natural scenes. We take for granted the many arguments that have been put forth in favor of using natural stimuli to probe the visual system [30–33].

Understanding how the human brain interprets natural images is likely to require extensive sampling of stimulus variation. Natural scenes are high-dimensional and, despite recent progress in modeling their statistical structure [34], it is not possible to adequately sample them by varying a small number of controllable parameters. Furthermore, visual (and auditory) features are often highly correlated in stimuli, both natural [35] and synthetic [36], and it is difficult to isolate individual features [37]. Correlation of stimulus features creates inferential challenges [38]: for example, an undersampled neuron that is strongly activated by the color yellow can easily be mistaken for a neuron that is strongly activated by bananas.

A complete model of the human visual system will have to specify how the high-dimensional space of natural stimuli is transformed and represented by the human brain. Such a model will necessarily be complex, with many potential stimulus features and free parameters [39]; fitting and adjudicating different models will therefore require substantial empirical data. Thus, we believe that to have any hope of achieving a complete model of the visual brain, extensive sampling of stimulus variation will be necessary.

Few reasons for sampling individual variation

For basic cognitive neuroscience, motivations for sampling a large number of subjects (while collecting modest data per subject) might include the following:

1. We might be concerned that what we observe to be true of one brain may not be true of others.
2. We might want to increase the statistical power of our analyses by adding more subjects.
3. We are actually interested in explaining individual variation (or variation across groups).

In this section, we consider each of these motivations in turn and suggest reasons why we nonetheless might want to prioritize the sampling of stimulus over individual variation. We acknowledge that there are other considerations that may constrain a given study: for example, extensive data collection might be impractical or too exhausting for a given subject population, or a given experiment might require novelty and cannot be repeated (e.g. one-shot experiments, drug or brain-stimulation treatments). We offer some thoughts on the issue of subject burden later in this article.

Complete models of individuals are likely to generalize

Suppose that by dedicating all available experimental resources, we are able to sample stimulus variation extensively enough to construct a complete model of an individual brain. Would this complete model of a single brain do more to mislead us about general principles than it would to reveal them?

Inferring general principles from relatively small numbers of observations is a practice taken for granted in other domains. For example, in physics, consider Kepler's deduction of the laws of planetary motion from a limited set of observations. As another example, in neuroanatomy,

consider the insights into structure-function relationships provided by Cajal’s exquisitely detailed drawings of individual neurons. Whether these types of generalizations succeed in cognitive neuroscience remains an open question. However, we point out that most of what we have learned about the organization of mammalian visual systems suggests that a complete model of an individual brain would yield useful insights. Many principles of visual organization in human brains—such as retinotopic organization [40] and face-selectivity [41]—not only generalize across individuals, but for the most part generalize across species [42,43]. Beyond sensory processing, there is even some evidence that cortical organization of high-level task representations [21] and semantic processing [44] exhibits consistent structure across individuals. We furthermore point out that generalization is an existential assumption of animal models in systems neuroscience: it is assumed that the brains of other species have much to teach us about our own.

If we were to achieve a complete model of an individual, it would certainly be important to test generalization of the model. This could be done relatively quickly and cheaply by testing a second or third subject. We would not need to gather enough samples from the additional subjects to derive complete models of their brains from scratch—we would need just enough to test the generalization of the model from the first subject, tweaking it to account for local differences.

To be clear, we are arguing here that the most useful insights for cognitive neuroscience lie in identifying general principles that govern how functional activity arises in the brain and how functional properties are anatomically organized. To achieve this, we believe that extensive sampling of individuals is essential. We acknowledge that such an approach is not suitable for assessing differences in the exact size or layout of maps or functional properties across the population at large: we feel it is more pressing to identify the axes that characterize brain function as opposed to characterizing the distribution of individuals along such axes.

More subjects = more “noise”

There are two radically different stances towards characterizing function in brains [45]. In the group-oriented stance, which perhaps is more typical in cognitive studies, statistical variability (error) is computed across subjects and effects are established at the group level. This approach lends itself well to brain measurements conducted at low spatial resolution and analyses that involve deliberate spatial smoothing of neuroimaging data and/or averaging of functional activity across subjects in a common anatomical space [46]. Accumulating samples of brain activity in

1 this fashion will certainly improve reliability of group-average measures, but is likely to preserve
2 only the simplest and most coarse patterns of activity. This is because the structural and
3 functional organization of the brain varies considerably across individuals at the fine scale [47].
4 While surface-based approaches for registering cerebral cortex across individuals are more
5 accurate than volume-based approaches [48,49], substantial variability still remains [50]. It is
6 possible, of course, to adopt strategies that avoid direct averaging of activity in a common space,
7 such as manually defining well-established regions-of-interest in individual subjects [51] or
8 characterizing activity at a more abstract level before averaging [16,52,53], but these strategies
9 come with their own limitations.

11 In the individual-oriented stance, which perhaps is more typical in sensory studies, statistical
12 variability (error) is computed across stimuli or trials (or scan sessions) for a given subject. By
13 doing so, variability is quantified at the most precise unit of measurement (e.g. voxels, neurons,
14 electrodes). Additional brain measurements contribute directly to additional information about
15 the nature of the computation occurring at each unit, and does so without incurring any cost of
16 intersubject variability. The aim is to establish effects in individual subjects and reproduce the
17 effects on additional subjects if necessary. Notably, this general approach of measurement and
18 characterization of a small number of subjects is rooted in classic studies in human
19 psychophysics [54,55], neuroimaging [56,57], and animal electrophysiology [58].

21 Our contention is that adding subjects to increase statistical power makes sense only if one
22 adopts the group-oriented stance where group-averaging of activity is performed. If we switch
23 stances and instead characterize effects in individual subjects, the problem of intersubject
24 variability is largely avoided. In a sense, the choice to study and analyze group-average activity
25 actually adds unnecessary “noise” to the analysis. We therefore suggest that increased sampling
26 of stimulus variation (within an individual) is a preferable method for improving statistical
27 power compared to increased sampling of individual variation.

29 *Complete models of individual brains may reveal the most interesting forms of individual*
30 *variation*

32 We think that the most compelling reason for sampling individual variation is if we are actually
33 interested in understanding person-to-person variation in some cognitive function [59], or
34 differences between populations grouped by a shared pathology or mental illness [60]. For these
35 purposes, studies involving large numbers of subjects are appropriate and may be highly useful
36 for developing predictive biomarkers for various behavioral phenotypes [61]. However, we

1 contend that in neurotypical subject pools, many of the most interesting forms of individual
2 variation are subjective mental events like dreams, mental images, and recalled memories
3 [62,63]. It is likely that a deep understanding of how these events are mediated by brain activity
4 will require sophisticated models of individual brains spanning perception, attention, memory,
5 and cognition—and therefore require extensive sampling of stimulus and task variation.
6 Functional brain measurements are notoriously noisy (reflecting either instrumental or true
7 neural variation), and relating brain responses to precise measures of behavior is a challenging
8 endeavor that requires substantial amounts of data, assuming we are careful not to overestimate
9 the robustness of our measures [64].

11 **A case study**

13 We have argued that extensive sampling of stimulus variation is essential for modeling brain
14 activity and that allocating samples to one or a few individual brains is more effective than
15 spreading them across many brains. But these are mere generalities—how many subjects should
16 one actually use?

18 To make our discussion more concrete, we briefly address specific choices we made for the
19 Natural Scenes Dataset (NSD, <http://naturalscenesdataset.org>). NSD is a massive sampling of
20 high-resolution whole-brain 7T fMRI responses that will soon be publicly released [18]. The
21 unique scale of the dataset makes possible analyses that are difficult to carry out in smaller
22 datasets, such as end-to-end training of a neural network-based encoding model of visual
23 responses. Indeed, one of the overarching goals of NSD was to generate a dataset that can be
24 used to develop and validate complete models of individual brains. In designing NSD, we had to
25 move from general considerations to concrete choices about the number of stimuli to sample, the
26 number of repetitions per stimulus, and the number of individual brains across which to
27 distribute these samples.

29 NSD samples brain activity in response to roughly 73,000 distinct natural scenes aggregated
30 across subjects. There is no general principle or rule one can consult to select the total number of
31 images sampled; it may be the case that neuroscientists are destined to be starved for data
32 relative to the dimensionality of visual stimuli. However, we were guided by the numbers of
33 samples needed for data-hungry deep-learning techniques [65,66]. State-of-the art techniques
34 typically require tens of thousands of samples, and so we took this rough figure as our goal.

NSD samples each image three times within a given subject. Stimulus repeats are not a strict requirement for all analyses, since there is considerable overlap in the kinds of visual features that appear across images. However, stimulus repeats make it possible to entertain analyses of variation across trials, allowing one to study issues such as repetition and memory effects [67]. In addition, stimulus repeats enable the use of cross-validation methods for data denoising [68]. Finally, in the NSD experiment, subjects were asked to report on each trial if they remembered seeing the displayed stimulus on any previous trial. This added a cognitive dimension that enriches the analysis potential of the dataset.

NSD includes eight subjects. Why this design? One motivation comes from statistical considerations. If we assume that the researcher is equipped to demonstrate effects in individual subjects, the availability of eight subjects provides statistical power to guard against incidental findings. For example, using a two-sided sign test, demonstration of a consistent sign of an effect in six out of eight subjects corresponds to a p -value of 0.03125.

Our primary consideration for recruiting eight subjects, however, was the demand on the subjects and the time required to acquire the full dataset. Although allocating the full set of 73,000 images to a single subject might not be an unreasonable proposition from the perspective of scientific value, subjects already had to commit to nearly a year of regularly scheduled scan sessions. If a single subject had been used, the dataset would have taken eight years to collect as opposed to one. Moreover, we speculate that a year of scanning may already be near the upper limit of what most subjects can be asked to tolerate, at least in the context of strenuous fMRI experiments where high performance is expected from scan subjects. Indeed, in light of these considerations, we took special care when acquiring the NSD dataset to minimize the burden of participation and to make the experience as pleasant as possible for the subjects. This included communicating clear expectations, sharing the scientific motivation and excitement of the dataset, instituting a bonus incentive structure for good performance, streamlining day-to-day scanning procedures, offering refreshments, maintaining an amicable relationship, continually assessing subject well-being, and soliciting subject feedback. While subjects had the option of halting participation at any time, all eight continued until the end.

Future outlook

Currently, there are few human fMRI datasets involving extremely deep and rich sampling of cognitive phenomena. But we imagine that the utility of and demand for such datasets will only increase as the field targets more sophisticated questions about the brain. Indeed, such datasets

1 seem a necessity if we are attempting to derive deep principles of intelligence from brain data
2 and incorporate these into artificial systems [69]. Given the expense of neuroimaging data,
3 collecting these datasets and making them publicly available are critical for scientific progress.
4 In particular, we are excited by the prospect of different research groups comparing different
5 models and analyses on common benchmark datasets [70,71]. Finally, we note that a shift to
6 extensive sampling of individuals goes hand-in-hand with improvements in our ability to collect
7 brain measurements at high spatial and temporal resolution [72], our ability to handle, store, and
8 compute on large datasets [73,74], and our ability to apply data-driven machine-learning
9 techniques that can exploit large amounts of data [75,76].

10

Competing Interests

The authors declare no conflict of interest.

Acknowledgments

We thank L. Dowdle, B. Hutchinson, and J. Winawer for comments on the manuscript. This work was supported by NSF IIS-1822683 and NSF IIS-1822929.

References

1. Kay KN: **Principles for models of neural information processing.** *Neuroimage* 2018, **180**:101–109.
2. Naselaris T, Bassett DS, Fletcher AK, Kording K, Kriegeskorte N, Nienborg H, Poldrack RA, Shohamy D, Kay K: **Cognitive Computational Neuroscience: A New Conference for an Emerging Discipline.** *Trends Cogn Sci* 2018, **22**:365–367.
3. Naselaris T, Kay KN, Nishimoto S, Gallant JL: **Encoding and decoding in fMRI.** *Neuroimage* 2011, **56**:400–410.
4. Alfaro-Almagro F, Jenkinson M, Bangerter NK, Andersson JLR, Griffanti L, Douaud G, Sotiropoulos SN, Jbabdi S, Hernandez-Fernandez M, Vallee E, et al.: **Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank.** *Neuroimage* 2018, **166**:400–424.
- This large-scale medical imaging study through the UK BioBank provides raw and pre-processed imaging results for more than 10,000 human participants. This dataset is ideal for human population-level modeling.
5. Chang N, Pyles JA, Gupta A, Tarr MJ, Aminoff EM: **BOLD5000: A public fMRI dataset of 5000 images.** *Sci Data* 2018, **6**:49.
- The aim of this slow event-related fMRI study was to present a large and diverse number of naturalistic images to a small number of human participants. With nearly 5,000 unique images and around twenty hours of neuroimaging data for each of four participants, the large size of this publicly available BOLD5000 dataset makes it amenable to data-driven analysis approaches.
6. Bellec P, Boyle JA: **Bridging the gap between perception and action: the case for neuroimaging, AI and video games.** *Psy ArXiv* 2019, doi:10.31234/osf.io/3epws.
- This opinion paper proposes training artificial neural networks on human brain activity and behavior as participants perform a variety of complex tasks. To that end, they introduce the Courtois project of neural modelling, CNeuroMod, in which 600 hours of neuroimaging data (fMRI and MEG) will be collected on each of six human participants over the course of five years as they undergo a variety of experiments, including playing video games.
7. Seeliger K, Sommers RP, Güçlü U, Bosch SE, Gerven MAJ Van: **A large single-participant fMRI dataset for probing brain responses to naturalistic stimuli in space and time.** *BioRxiv* 2019.
- Over 23 hours of fMRI data were collected on a single human participant over six months as they watched 30 episodes of BBC's *Doctor Who*. Involving rich naturalistic visual and auditory stimuli, this dataset can be used to explore a variety of questions about how the brain represents multisensory information.
8. Wen H, Shi J, Zhang Y, Lu K-H, Cao J, Liu Z: **Neural Encoding and Decoding with Deep Learning for Dynamic Natural Vision.** *Cereb Cortex* 2018, **28**:4136–4160.
9. Horikawa T, Kamitani Y: **Generic decoding of seen and imagined objects using hierarchical visual features.** *Nat Commun* 2017, **8**:15037.
10. Van Essen DC, Smith SM, Barch DM, Behrens TEJ, Yacoub E, Ugurbil K: **The WU-Minn Human Connectome Project: An overview.** *Neuroimage* 2013, **80**:62–79.
11. Benson NC, Jamison KW, Arcaro MJ, Vu AT, Glasser MF, Coalson TS, Van Essen DC, Yacoub E, Ugurbil K, Winawer J, et al.: **The Human Connectome Project 7 Tesla retinotopy dataset: Description and population receptive field analysis.** *J Vis* 2018, **18**:23.

- 1 • As part of the Human Connectome Project, this is the largest publicly available retinotopy
- 2 dataset. High-quality 7T fMRI data were collected on 181 healthy young adults and model-based
- 3 analyses were conducted at both the individual and group level in cortical and subcortical
- 4 regions.
- 5 12. Pinho AL, Amadon A, Ruest T, Fabre M, Dohmatob E, Denghien I, Ginisty C, Becuwe-
- 6 Desmidt S, Roger S, Laurier L, et al.: **Individual Brain Charting, a high-resolution**
- 7 **fMRI dataset for cognitive mapping.** *Sci data* 2018, **5**:180105.
- 8 13. Gordon EM, Laumann TO, Gilmore AW, Newbold DJ, Greene DJ, Berg JJ, Ortega M,
- 9 Hoyt-Drazen C, Gratton C, Sun H, et al.: **Precision Functional Mapping of Individual**
- 10 **Human Brains.** *Neuron* 2017, **95**:791-807.e7.
- 11 •• The Midnight Scan Club collected six hours of resting-state fMRI and five hours of task-based
- 12 fMRI from each of ten human participants in order to obtain a deeper understanding of
- 13 functional brain connectivity at the level of the individual. The study contends that precision
- 14 individual connectomics is superior to the more standard group-averaging approach.
- 15 14. Schoffelen J-M, Oostenveld R, Lam NHL, Uddén J, Hultén A, Hagoort P: **A 204-subject**
- 16 **multimodal neuroimaging dataset to study language processing.** *Sci Data* 2019, **6**:17.
- 17 • This multimodal neuroimaging dataset consists of MRI, resting-state and task-based fMRI, and
- 18 MEG data on 204 human participants. A key feature of this study was to expose half of the
- 19 participants to scrambled and unscrambled sentences in the auditory domain and the other half to
- 20 these sentences in the visual domain in order to explore the role of sensory input modality in
- 21 language processing.
- 22 15. Poldrack RA, Laumann TO, Koyejo O, Gregory B, Hover A, Chen M-Y, Gorgolewski KJ,
- 23 Luci J, Joo SJ, Boyd RL, et al.: **Long-term neural and physiological phenotyping of a**
- 24 **single human.** *Nat Commun* 2015, **6**:8885.
- 25 16. Nastase SA, Liu Y-F, Hillman H, Norman KA, Hasson U: **Leveraging shared**
- 26 **connectivity to aggregate heterogeneous datasets into a common response space.**
- 27 *Neuroimage* 2020, **217**:116865.
- 28 • Heterogeneous fMRI datasets were collected on 149 human participants over the course of
- 29 seven years as they listened to spoken stories. Datasets were aggregated and subjected to
- 30 connectivity-based hyperalignment.
- 31 17. Aliko S, Huang J, Gheorghiu F, Meliss S, Skipper JJ: **A naturalistic neuroimaging**
- 32 **database for understanding the brain using ecological stimuli.** *Sci data* 2020, **7**:347.
- 33 18. Allen EJ, Wu Y, Hutchinson JB, Naselaris T, Kay KN: **The Natural Scenes Dataset**
- 34 **(NSD): A yearlong ultra-high field whole-brain human fMRI visual perception and**
- 35 **memory study.** *J Vis* 2020, **20**:589.
- 36 •• This 7T fMRI dataset involved extensive sampling of 8 human participants as they each
- 37 viewed 9,000–10,000 distinct color natural scenes (22,500–30,000 trials) over the course of 30–
- 38 40 scan sessions. This rich dataset can be used to explore diverse neuroscientific questions with
- 39 high power at the level of individual subjects.
- 40 19. Hanke M, Adelhöfer N, Kottke D, Iacovella V, Sengupta A, Kaule FR, Nigbur R, Waite
- 41 AQ, Baumgartner F, Stadler J: **A studyforrest extension, simultaneous fMRI and eye**
- 42 **gaze recordings during prolonged natural stimulation.** *Sci data* 2016, **3**:160092.
- 43 20. Hanke M, Baumgartner FJ, Ibe P, Kaule FR, Pollmann S, Speck O, Zinke W, Stadler J: **A**
- 44 **high-resolution 7-Tesla fMRI dataset from complex natural stimulation with an**
- 45 **audio movie.** *Sci Data* 2014, **1**:140003.
- 46 21. Nakai T, Nishimoto S: **Quantitative models reveal the organization of diverse**

- cognitive functions in the brain.** *Nat Commun* 2020, **11**:1142.
- This study conducted brain activity measurements in six human participants as they each performed a large variety of naturalistic cognitive tasks. Modeling analyses achieved high prediction accuracy across a large portion of cortex, suggesting the data are probing a large portion of human cognitive space.
22. Kay KN, Naselaris T, Prenger RJ, Gallant JL: **Identifying natural images from human brain activity.** *Nature* 2008, **452**:352–355.
 23. Nishimoto S, Vu AT, Naselaris T, Benjamini Y, Yu B, Gallant JL: **Reconstructing Visual Experiences from Brain Activity Evoked by Natural Movies.** *Curr Biol* 2011, **21**:1641–1646.
 24. Bookheimer SY, Salat DH, Terpstra M, Ances BM, Barch DM, Buckner RL, Burgess GC, Curtiss SW, Diaz-Santos M, Elam JS, et al.: **The Lifespan Human Connectome Project in Aging: An overview.** *Neuroimage* 2019, **185**:335–348.
 25. Friston K: **Ten ironic rules for non-statistical reviewers.** *Neuroimage* 2012, **61**:1300–10.
 26. Desmond JE, Glover GH: **Estimating sample size in functional MRI (fMRI) neuroimaging studies: statistical power analyses.** *J Neurosci Methods* 2002, **118**:115–28.
 27. Poldrack RA, Baker CI, Durnez J, Gorgolewski KJ, Matthews PM, Munafò MR, Nichols TE, Poline J-B, Vul E, Yarkoni T: **Scanning the horizon: towards transparent and reproducible neuroimaging research.** *Nat Rev Neurosci* 2017, **18**:115–126.
 28. Yuille A, Kersten D: **Vision as Bayesian inference: analysis by synthesis?** *Trends Cogn Sci* 2006, **10**:301–308.
 29. Rust NC, Movshon JA: **In praise of artifice.** *Nat Neurosci* 2005, **8**:1647–50.
 30. Felsen G, Dan Y: **A natural approach to studying vision.** *Nat Neurosci* 2005, **8**:1643–6.
 31. Wu MCK, David S V., Gallant JL: **Complete functional characterization of sensory neurons by system identification.** *Annu Rev Neurosci* 2006, **29**:477–505.
 32. Geisler WS: **Visual perception and the statistical properties of natural scenes.** *Annu Rev Psychol* 2008, **59**:167–92.
 33. Nastase SA, Goldstein A, Hasson U: **Keep it real: rethinking the primacy of experimental control in cognitive neuroscience.** *Neuroimage* 2020, **222**:117254.
 34. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y: **Generative adversarial nets.** *Adv Neural Inf Process Syst* 2014, **3**:2672–2680.
 35. Simoncelli EP, Olshausen BA: **Natural Image Statistics and Neural Representation.** *Annu Rev Neurosci* 2001, **24**:1193–1216.
 36. Wardle SG, Ritchie JB, Seymour K, Carlson TA: **Edge-Related Activity Is Not Necessary to Explain Orientation Decoding in Human Visual Cortex.** *J Neurosci* 2017, **37**:1187–1196.
 37. Kay KN: **Understanding visual representation by developing receptive-field models.** In *Visual Population Codes: Towards a Common Multivariate Framework for Cell Recording and Functional Imaging.* . The MIT Press; 2011:133–162.
 38. Naselaris T, Kay KN: **Resolving Ambiguities of MVPA Using Explicit Models of Representation.** *Trends Cogn Sci* 2015, **19**:551–554.
 39. St-Yves G, Naselaris T: **The feature-weighted receptive field: An interpretable encoding model for complex feature spaces.** *Neuroimage* 2017, doi:10.1016/j.neuroimage.2017.06.035.

- 1 40. Wandell BA, Dumoulin SO, Brewer AA: **Visual field maps in human cortex.** *Neuron*
2 2007, **56**:366–83.
- 3 41. Kanwisher N, McDermott JH, Chun MM: **The fusiform face area: a module in human**
4 **extrastriate cortex specialized for face perception.** *J Neurosci* 1997, **17**:4302–11.
- 5 42. Van Essen DC, Lewis JW, Drury HA, Hadjikhani N, Tootell RBH, Bakircioglu M, Miller
6 MI: **Mapping visual cortex in monkeys and humans using surface-based atlases.**
7 *Vision Res* 2001, **41**:1359–78.
- 8 43. Tsao DY, Freiwald WA, Tootell RBH, Livingstone MS: **A cortical region consisting**
9 **entirely of face-selective cells.** *Science (80-)* 2006, **311**:670–4.
- 10 44. Huth AG, de Heer WA, Griffiths TL, Theunissen FE, Gallant JL: **Natural speech reveals**
11 **the semantic maps that tile human cerebral cortex.** *Nature* 2016, **532**:453–458.
- 12 45. Wandell BA, Winawer J, Kay KN: **Computational Modeling of Responses in Human**
13 **Visual Cortex.** In *Brain Mapping.* . Elsevier; 2015:651–659.
- 14 46. Scott A. Huettel, Allen W. Song GM: *Functional Magnetic Resonance Imaging.* Sinauer
15 Associates; 2009.
- 16 47. Weiner KS, Grill-Spector K: **The improbable simplicity of the fusiform face area.**
17 *Trends Cogn Sci* 2012, **16**:251–4.
- 18 48. Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, Ugurbil K,
19 Andersson J, Beckmann CF, Jenkinson M, et al.: **A multi-modal parcellation of human**
20 **cerebral cortex.** *Nature* 2016, **536**:171–8.
- 21 49. Van Essen DC, Glasser MF, Dierker DL, Harwell J, Coalson T: **Parcellations and**
22 **hemispheric asymmetries of human cerebral cortex analyzed on surface-based**
23 **atlases.** *Cereb Cortex* 2012, **22**:2241–62.
- 24 50. Frost MA, Goebel R: **Measuring structural-functional correspondence: spatial**
25 **variability of specialised brain regions after macro-anatomical alignment.**
26 *Neuroimage* 2012, **59**:1369–81.
- 27 51. Saxe R, Brett M, Kanwisher N: **Divide and conquer: a defense of functional localizers.**
28 *Neuroimage* 2006, **30**:1088–96; discussion 1097-9.
- 29 52. Haxby J V., Guntupalli JS, Nastase SA, Feilong M: **Hyperalignment: Modeling shared**
30 **information encoded in idiosyncratic cortical topographies.** *Elife* 2020, **9**:1–26.
- 31 53. Kriegeskorte N, Mur M, Bandettini P: **Representational similarity analysis - connecting**
32 **the branches of systems neuroscience.** *Front Syst Neurosci* 2008, **2**:4.
- 33 54. Campbell FW, Robson JG: **Application of Fourier analysis to the visibility of gratings.**
34 *J Physiol* 1968, **197**:551–66.
- 35 55. Edward H, Belper, Lord: **On the theory of compound colours, and the relations of the**
36 **colours of the spectrum.** *Proc R Soc London* 1860, **10**:404–409.
- 37 56. Boynton GM, Engel SA, Glover GH, Heeger DJ: **Linear systems analysis of functional**
38 **magnetic resonance imaging in human V1.** *J Neurosci* 1996, **16**:4207–21.
- 39 57. Engel SA, Rumelhart DE, Wandell BA, Lee AT, Glover GH, Chichilnisky E-J, Shadlen
40 MN: **fMRI of human visual cortex.** *Nature* 1994, **369**:525–525.
- 41 58. Newsome WT, Britten KH, Movshon JA: **Neuronal correlates of a perceptual decision.**
42 *Nature* 1989, **341**:52–4.
- 43 59. Charest I, Kievit RA, Schmitz TW, Deca D, Kriegeskorte N: **Unique semantic space in**
44 **the brain of each beholder predicts perceived similarity.** *Proc Natl Acad Sci U S A*
45 2014, **111**:14565–70.
- 46 60. Cole MW, Anticevic A, Repovs G, Barch D: **Variable Global Dysconnectivity and**

- Individual Differences in Schizophrenia.** *Biol Psychiatry* 2011, **70**:43–50.
61. Smith SM, Nichols TE, Vidaurre D, Winkler AM, Behrens TEJ, Glasser MF, Ugurbil K, Barch DM, Van Essen DC, Miller KL: **A positive-negative mode of population covariation links brain connectivity, demographics and behavior.** *Nat Neurosci* 2015, **18**:1565–7.
 62. Horikawa T, Tamaki M, Miyawaki Y, Kamitani Y: **Neural decoding of visual imagery during sleep.** *Science* 2013, **340**:639–42.
 63. Breedlove JL, St-Yves G, Olman CA, Naselaris T: **Generative Feedback Explains Distinct Brain Activity Codes for Seen and Mental Images.** *Curr Biol* 2020, **30**:2211–2224.e6.
 64. Vul E, Harris C, Winkielman P, Pashler H: **Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition.** *Perspect Psychol Sci* 2009, **4**:319–324.
 65. Jia Deng, Wei Dong, Socher R, Li-Jia Li, Kai Li, Li Fei-Fei: **ImageNet: A large-scale hierarchical image database.** In *2009 IEEE Conference on Computer Vision and Pattern Recognition*. . IEEE; 2009:248–255.
 66. Krizhevsky A, Sutskever I, Hinton GE: **ImageNet classification with deep convolutional neural networks.** *Adv Neural Inf Process Syst* 2012, doi:10.1145/3065386.
 67. Wagner AD, Schacter DL, Rotte M, Koutstaal W, Maril A, Dale AM, Rosen BR, Buckner RL: **Building memories: remembering and forgetting of verbal experiences as predicted by brain activity.** *Science (80-)* 1998, **281**:1188–91.
 68. Kay KN, Rokem A, Winawer J, Dougherty RF, Wandell BA: **GLMdenoise: A fast, automated technique for denoising task-based fMRI data.** *Front Neurosci* 2013, **7**:1–15.
 69. Ullman S: **Using neuroscience to develop artificial intelligence.** *Science (80-)* 2019, **363**:692–693.
- This perspective addresses the importance of combining neuroscience with artificial intelligence in order to achieve human-like cognitive abilities and artificial general intelligence.
70. Cichy RM, Roig G, Oliva A: **The Algonauts Project.** *Nat Mach Intell* 2019, **1**:613–613.
- The aim of the Algonauts Project is to have teams of scientists in the fields of natural and artificial intelligence come together to develop a more complete understanding of the brain. The first year’s challenge was to predict brain activity during visual object recognition, and success was quantified as the percentage of activity variance that a given model explains.
71. Schrimpf M, Kubilius J, Lee MJ, Ratan Murty NA, Ajemian R, DiCarlo JJ: **Integrative Benchmarking to Advance Neurally Mechanistic Models of Human Intelligence.** *Neuron* 2020, doi:10.1016/j.neuron.2020.07.040.
 72. Harel N: **Ultra high resolution fMRI at ultra-high field.** *Neuroimage* 2012, **62**:1024–1028.
 73. Wu R, Zhang B, Hsu M: **Clustering billions of data points using GPUs.** In *Proceedings of the combined workshops on UnConventional high performance computing workshop plus memory access workshop - UCHPC-MAW '09*. . ACM Press; 2009:1.
 74. Eklund A, Dufort P, Villani M, LaConte S: **BROCCOLI: Software for fast fMRI analysis on many-core CPUs and GPUs.** *Front Neuroinform* 2014, **8**:1–19.
 75. Sejnowski TJ, Churchland PS, Movshon JA: **Putting big data to good use in neuroscience.** *Nat Neurosci* 2014, **17**:1440–1.
 76. Seeliger K, Ambrogioni L, Güçlütürk Y, Güçlü U, Van Gerven MAJ: **End-to-end neural system identification with neural information flow.** *bioRxiv* 2019, doi:10.1101/553255.

