antigen-presenting cells or cytokines. In either case, the current observations exclude models in which each naïve T cell yields progeny with the same distribution of cells of either short- or long-term potential. Thus, although our data do not evaluate the role of asymmetric division as a mechanism to generate daughter cells with different fates, they do show that asymmetric division by itself cannot explain the disparity between individual T cell families that is experimentally observed. Rather, a strong variation between families in the expansion of proximal and distal daughter cells needs to be invoked to arrive at the heterogeneity observed here. Lastly, the observation of strong heterogeneity at the single-cell level indicates that T cell responses are made up of "averages," a behavior reminiscent of recent analyses of stem-cell renewal (19). Thus, although the differentiation and expansion of the combined T cell population follow a uniform course, the fate of individual naïve T cells is unpredictable.

References and Notes

1. J. J. Obar, K. M. Khanna, L. Lefrançois, *Immunity* **28**, 859 (2008).
2. J. N. Blattman et al., *J. Exp. Med.* **195**, 657 (2002).
3. K. Murali-Krishna et al., *Immunity* **8**, 177 (1998).
4. E. A. Butz, M. J. Bevan, *Immunity* **8**, 167 (1998).
5. M. A. Williams, M. J. Bevan, *Annu. Rev. Immunol.* **25**, 171 (2007).
6. S. Sarkar et al., *J. Exp. Med.* **205**, 625 (2008).
7. J. J. Obar, L. Lefrançois, *Int. Immunol.* **22**, 619 (2010).
8. C. Stemberger et al., *Immunity* **27**, 985 (2007).
9. C. Gerlach et al., *J. Exp. Med.* **207**, 1235 (2010).
10. J. T. Chang et al., *Science* **315**, 1687 (2007).
11. J. W. J. van Heijst et al., *Science* **325**, 1265 (2009).
12. T. N. M. Schumacher, C. Gerlach, J. W. J. van Heijst, *Nat. Rev. Immunol.* **10**, 621 (2010).
13. K. Schepers et al., *J. Exp. Med.* **205**, 2309 (2008).
14. Materials and methods are available as supplementary materials on *Science* Online.
15. W. N. D'Souza, S. M. Hedrick, *J. Immunol.* **177**, 777 (2006).
16. T. E. Schlub et al., *Eur. J. Immunol.* **39**, 67 (2009).
17. V. R. Buchholz et al., *Science* **340**, 630 (2013).
18. K. R. Duffy et al., *Science* **335**, 338 (2012).
19. B. D. Simons, H. Clevers, *Cell* **145**, 851 (2011).

Neural Decoding of Visual Imagery During Sleep

T. Horikawa,[1,2] M. Tamaki,[1]* Y. Miyawaki,[3,1]† Y. Kamitani[1,2]‡

Visual imagery during sleep has long been a topic of persistent speculation, but its private nature has hampered objective analysis. Here we present a neural decoding approach in which machine-learning models predict the contents of visual imagery during the sleep-onset period, given measured brain activity, by discovering links between human functional magnetic resonance imaging patterns and verbal reports with the assistance of lexical and image databases. Decoding models trained on stimulus-induced brain activity in visual cortical areas showed accurate classification, detection, and identification of contents. Our findings demonstrate that specific visual experience during sleep is represented by brain activity patterns shared by stimulus perception, providing a means to uncover subjective contents of dreaming using objective neural measurement.

Dreaming is a subjective experience during sleep often accompanied by vivid visual contents. Previous research has attempted to link physiological states with dreaming (1–3), but none has demonstrated how specific visual contents are represented in brain activity. The advent of machine-learning–based analysis allows for the decoding of stimulus- and task-induced brain activity patterns to reveal visual contents (4–9). We extended this approach to the decoding of spontaneous brain activity during sleep (Fig. 1A). Although dreaming has often been associated with the rapid-eye movement (REM) sleep stage, recent studies have demonstrated that dreaming is dissociable from REM sleep and can be experienced during non-REM periods (10). We focused on visual imagery (hallucination) experienced during the sleep-onset (hypnagogic) period (sleep stage 1 or 2) (11, 12) because it allowed us to collect many observations by repeating awakenings and recording participants' verbal reports of visual experience. Reports at awakenings in sleep-onset and REM periods share general features such as frequency, length, and contents, while differing in several aspects, including the affective component (13–15). We analyzed verbal reports using a lexical database to create systematic labels for visual contents. We hypothesized that contents of visual imagery during sleep are represented at least partly by visual cortical activity patterns shared by stimulus representation. Thus, we trained decoders on brain activity induced by natural images from Web image databases.

Three people participated in the functional magnetic resonance imaging (fMRI) sleep experiments (Fig. 1A), in which they were woken when an electroencephalogram signature was detected (16) (fig. S1), and they were asked to give a verbal report freely describing their visual experience before awakening [table S1; duration, 34 ± 19 s (mean ± SD)]. We repeated this procedure to attain at least 200 awakenings with a visual report for each participant. On average, we awakened participants every 342.0 s, and visual contents were reported in over 75% of the awakenings (Fig. 1B). Offline sleep stage scoring (fig. S2) further selected awakenings to exclude contamination from the wake stage in the period immediately before awakening (235, 198, and 186 awakenings for participants 1 to 3, respectively, used for decoding analyses) (16).

From the collected reports, words describing visual objects or scenes were manually extracted and mapped to WordNet, a lexical database in which semantically similar words are grouped as "synsets" in a hierarchical structure (17, 18) (Fig. 2A). Using a semantic hierarchy, we grouped extracted visual words into base synsets that appeared in at least 10 reports from each participant (26, 18, and 16 synsets for participants 1 to 3, respectively; tables S2 to S4) (16). The fMRI data obtained before each awakening were labeled with a visual content vector, each element of which indicated the presence or absence of a base synset in the subsequent report (Fig. 2B and fig. S3). We also collected images depicting each base synset from ImageNet (19), an image database in which Web images are grouped according to WordNet, or from Google Images, for decoder training.

We constructed decoders by training linear support vector machines (SVMs) (20) on fMRI data measured while each person viewed Web images for each base synset. Multivoxel patterns in the higher visual cortex [HVC; the ventral region covering the lateral occipital complex (LOC), fusiform face area (FFA), and parahippocampal place area (PPA); 1000 voxels], the lower visual

[1]ATR Computational Neuroscience Laboratories, Kyoto 619-0288, Japan. [2]Nara Institute of Science and Technology, Nara 630-0192, Japan. [3]National Institute of Information and Communications Technology, Kyoto 619-0288, Japan.

*Present address: Brown University, 190 Thayer Street, Providence, RI 02912, USA.
†Present address: The University of Electro-Communications, Tokyo 182-8585, Japan.
‡Corresponding author. E-mail: kmtn@atr.jp

cortex (LVC; V1 to V3 combined; 1000 voxels), or the subareas (400 voxels for each area) were used as the input for the decoders (*16*).

First, a binary classifier was trained on the fMRI responses to stimulus images of two base synsets (three-volume averaged data corresponding to the 9-s stimulus block) and tested on the sleep samples [three-volume (9-s) averaged data immediately before awakening] that contained exclusively one of the two synsets while ignoring other concurrent synsets (*16*) (Fig. 3A). We only used synset pairs in which one of the synsets appeared in at least 10 reports without co-occurrence with the other (201, 118, and 86 pairs for participants 1 to 3, respectively). The distribution of the pairwise decoding accuracies for the HVC is shown together with that from the decoders trained on the same stimulus-induced fMRI data with randomly shuffled synset labels (Fig. 3B; fig. S4, individual participants). The mean decoding accuracy was 60.0%, 95% confidence interval (CI) [(59.0, 61.0%); three participants pooled], which

was significantly higher than that of the label-shuffled decoders with both Wilcoxon rank-sum and permutation tests ($P < 0.001$).

To look into the commonality of brain activity between perception and sleep-onset imagery, we focused on the synset pairs that produced content-specific patterns in each of the stimulus and sleep experiments (pairs with high cross-validation classification accuracy within each of the stimulus and sleep data sets; figs. S5 and S6) (*16*). With the selected pairs, even higher accuracies were obtained [mean = 70.3%, CI (68.5, 72.1); Fig. 3B, dark blue; fig. S4, individual participants; tables S5 to S7, lists of the selected pairs], indicating that content-specific patterns are highly consistent between perception and sleep-onset imagery. The selection of synset pairs, which used knowledge of the test (sleep) data, does not bias the null distribution by the label-shuffled decoders (Fig. 3B, black), because the content specificity in the sleep data set alone does not imply commonality between the two data sets.

Additional analyses revealed that the multi-voxel pattern, rather than the average activity level, was critical for decoding (figs. S7 and S8). We also found that the variability of decoding performance among synset pairs can be accounted for at least partly by the semantic differences between paired synsets. The decoding accuracy for synsets paired across meta-categories (human, object, scene, and others; tables S2 to S4) was significantly higher than that for synsets within meta-categories (Wilcoxon rank-sum test, $P < 0.001$; Fig. 3C and fig. S9). However, even within a meta-category, the mean decoding accuracy significantly exceeded chance level, indicating specificity to fine object categories.

The mean decoding accuracies for different visual areas are shown in Fig. 3D (fig. S10, individual participants). The LVC scored 54.3%, CI (53.4, 55.2) for all pairs, and 57.2%, CI (54.2, 60.2) for selected pairs (three participants pooled). The performance was significantly above chance level but worse than that for the HVC. Individual
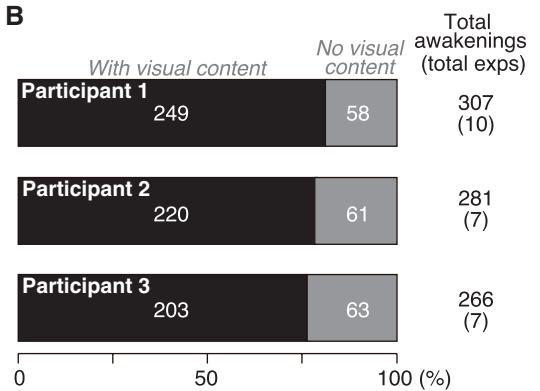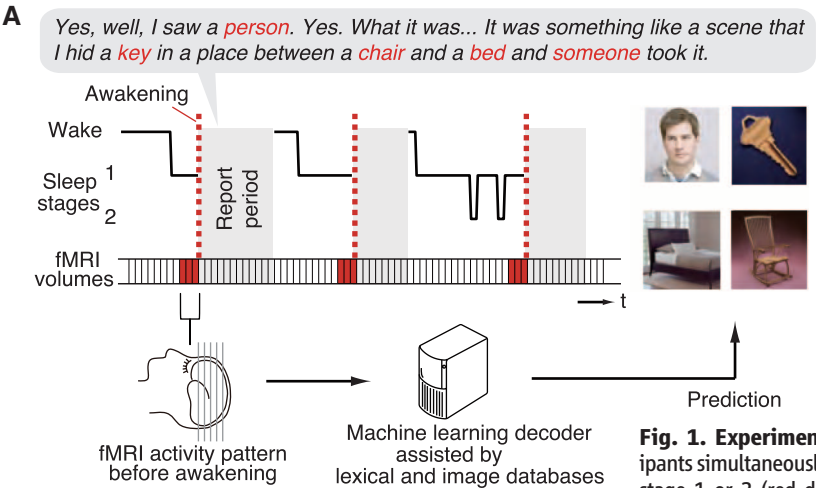
Fig. 1. Experimental overview. (**A**) fMRI data were acquired from sleeping participants simultaneously with polysomnography. Participants were awakened during sleep stage 1 or 2 (red dashed line) and verbally reported their visual experience during sleep. fMRI data immediately before awakening [an average of three volumes (= 9 s)] were used as the input for main decoding analyses (sliding time windows were used for time course analyses). Words describing visual objects or scenes (red letters) were extracted. The visual contents were predicted using machine-learning decoders trained on fMRI responses to natural images. (**B**) The numbers of awakenings with and without visual contents are shown for each participant (with numbers of experiments in parentheses).
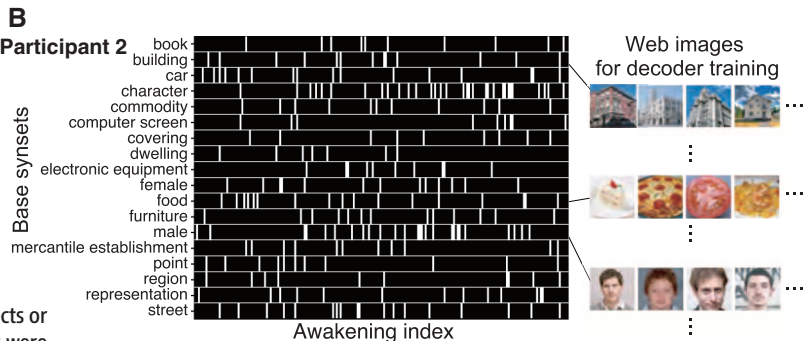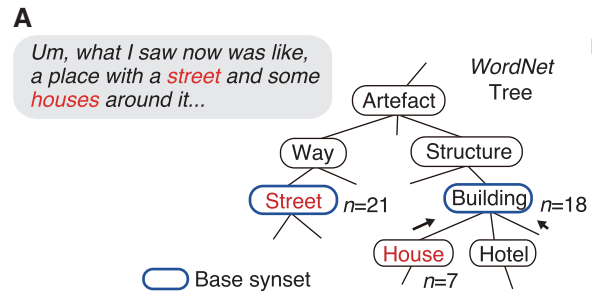

Fig. 2. Visual content labeling. (**A**) Words describing visual objects or scenes (red) were mapped onto synsets of the WordNet tree. Synsets were grouped into base synsets (blue frames) located higher in the tree. (**B**) Visual reports (participant 2) are represented by visual content vectors, in which the presence or absence of the base synsets in the report at each awakening is indicated by white or black, respectively. Examples of images used for decoder training are shown for some of the base synsets.

areas (V1 to V3, LOC, FFA, and PPA) showed a gradual increase in accuracy along the visual processing pathway, mirroring the progressively complex response properties from low-level image features to object-level features (*21*). When the time window was shifted, the decoding accuracy peaked around 0 to 10 s before awakening (Fig. 3E and fig. S11; no correction for hemodynamic delay). The high accuracies after awakening may be due to hemodynamic delay and the large time window. Thus, verbal reports are likely to reflect brain activity immediately before awakening.

To read out richer contents given arbitrary sleep data, we next performed a multilabel decoding analysis in which the presence or absence
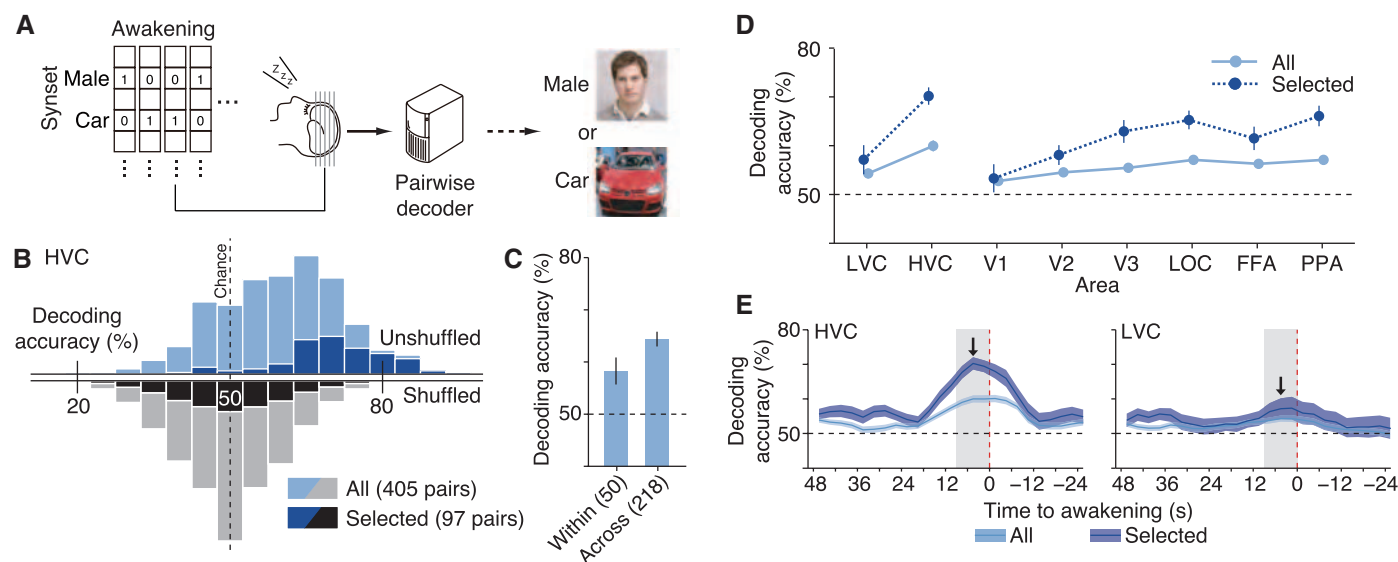
Fig. 3. Pairwise decoding. (**A**) Schematic overview. (**B**) Distributions of decoding accuracies with original and label-shuffled data for all pairs (light blue and gray) and selected pairs (dark blue and black) (three participants pooled). (**C**) Mean accuracies for the pairs within and across meta-categories (synsets in others were excluded; numbers of pairs are in parentheses). (**D**) Accuracies across visual areas (numbers of selected pairs for V1, V2, V3, LOC, FFA, PPA, LVC, and HVC: 45, 50, 55, 70, 48, 78, 55, and 97). (**E**) Time course (HVC and LVC; averaged across pairs and participants). The plot shows the performance with the 9-s (three-volume) time window centered at each point (gray window and arrow for main analyses). For all results, error bars or shadings indicate 95% CI, and dashed lines denote chance level.
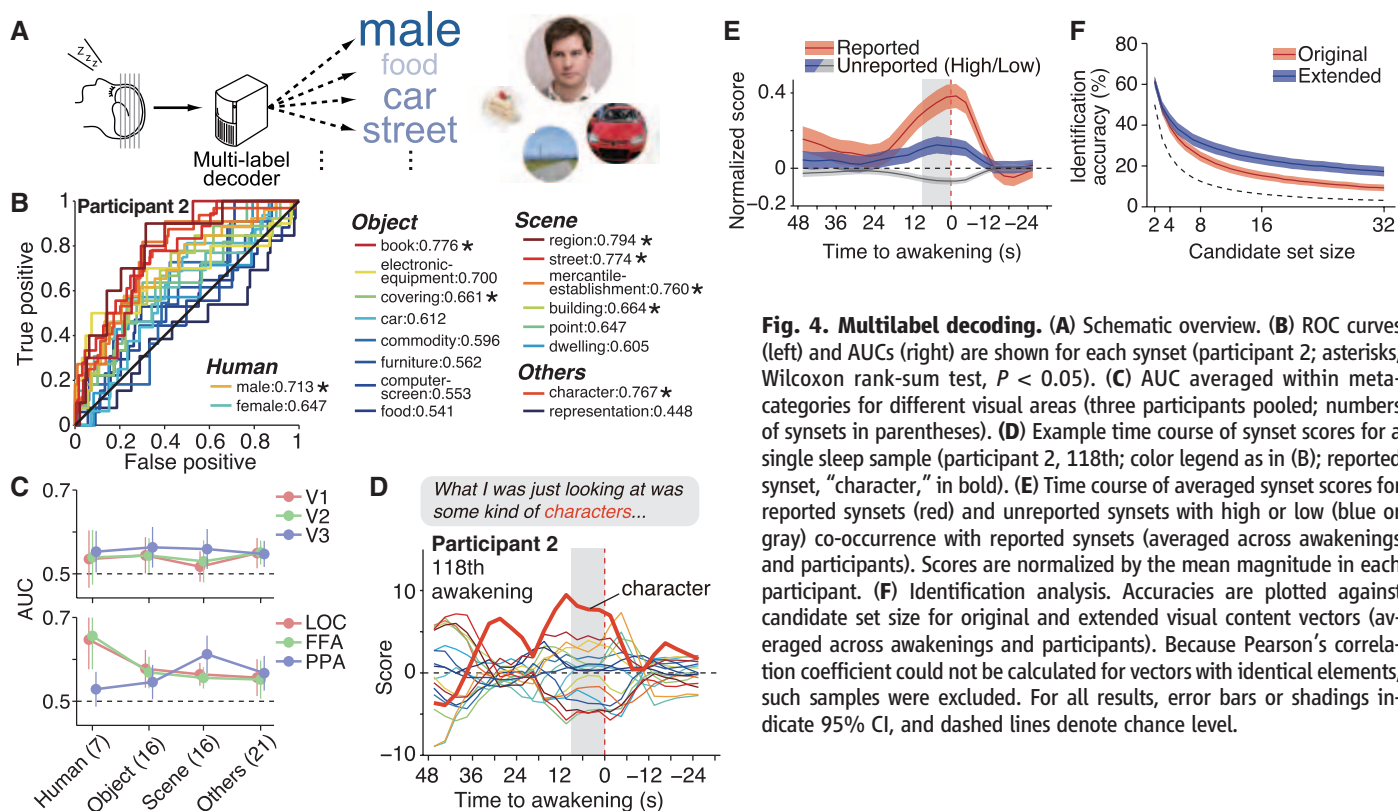

Fig. 4. Multilabel decoding. (**A**) Schematic overview. (**B**) ROC curves (left) and AUCs (right) are shown for each synset (participant 2; asterisks, Wilcoxon rank-sum test, *P* < 0.05). (**C**) AUC averaged within meta-categories for different visual areas (three participants pooled; numbers of synsets in parentheses). (**D**) Example time course of synset scores for a single sleep sample (participant 2, 118th; color legend as in (B); reported synset, "character," in bold). (**E**) Time course of averaged synset scores for reported synsets (red) and unreported synsets with high or low (blue or gray) co-occurrence with reported synsets (averaged across awakenings and participants). Scores are normalized by the mean magnitude in each participant. (**F**) Identification analysis. Accuracies are plotted against candidate set size for original and extended visual content vectors (averaged across awakenings and participants). Because Pearson's correlation coefficient could not be calculated for vectors with identical elements, such samples were excluded. For all results, error bars or shadings indicate 95% CI, and dashed lines denote chance level.

of each base synset was predicted by a synset detector constructed from a combination of pairwise decoders (Fig. 4A) (*16*). The synset detector provided a continuous score indicating how likely the synset is to be present in each report. We calculated receiver operating characteristic (ROC) curves for each base synset by shifting the detection threshold for the output score (Fig. 4B, the HVC in participant 2, time window immediately before awakening; fig. S12, all participants), and the detection performance was quantified by the area under the curve (AUC). Although the performance varied across synsets, 18 out of the total 60 synsets were detected with above-chance levels (Wilcoxon rank-sum test, uncorrected $P < 0.05$), greatly exceeding the number of synsets expected by chance ($0.05 \times 60 = 3$).

Using the AUC, we compared the decoding performance for individual synsets grouped into meta-categories in different visual areas. Overall, the performance was better in the HVC than in the LVC, consistent with the pairwise decoding performance [fig. S13; three participants pooled; analysis of variance (ANOVA), $P = 0.003$]. Although V1 to V3 did not show different performances across meta-categories, the higher visual areas showed a marked dependence on meta-categories (Fig. 4C and fig. S13). In particular, the FFA showed better performance with human synsets, whereas the PPA showed better performance with scene synsets [ANOVA (interaction), $P = 0.001$], consistent with the known response characteristics of these areas (*22, 23*). The LOC and FFA showed similar results, presumably because our functional localizers selected partially overlapping voxels.

The output scores for individual synsets showed diverse and dynamic profiles in each sleep sample (Fig. 4D, fig. S14, and movies S1 and S2) (*16*). These profiles may reflect a dynamic variation of visual contents, including those experienced even before the period near awakening. On average, there was a general tendency for the scores for reported synsets to increase toward the time of awakening (Fig. 4E and fig. S15). Synsets that did not appear in reports showed greater scores if they had a high co-occurrence relationship with reported synsets (Fig. 4E; synsets with the top 15% conditional probabilities given a reported synset, calculated from the whole-content vectors in each participant). The effect of co-occurrence is rather independent of that of semantic similarity (Fig. 3C), because both factors (high/low co-occurrence and within/across meta-categories) had highly significant effects on the scores of unreported synsets (time window immediately before awakening; two-way ANOVA, $P < 0.001$, three participants pooled) with moderate interaction ($P = 0.016$). The scores for reported synsets were significantly higher than those for unreported synsets even within the same meta-category (Wilcoxon rank-sum test, $P < 0.001$). Verbal reports are unlikely to describe full details of visual experience during sleep, and it is possible that contents with high general co-occurrence

(such as street and car) tend to be experienced together even when all are not reported. Therefore, high scores for the unreported synsets may indicate unreported but actual visual contents during sleep.

Finally, to explore the potential of multilabel decoding to distinguish numerous contents, we performed identification analysis (*7, 8*). The output scores (score vector) were used to identify the true visual content vector among a variable number of candidates (true vector + random vectors with matched probabilities for each synset) by selecting the candidate most correlated with the score vector (repeated 100 times for each sleep sample to obtain the correct identification rate) (*16*). The performance exceeded chance level across all set sizes (Fig. 4F, HVC, three participants pooled; fig. S16, individual participants), although the accuracies were not as high as those achieved using stimulus-induced brain activity in previous studies (*7, 8*). The same analysis was performed with extended visual content vectors in which unreported synsets having a high co-occurrence with reported synsets (top 15% conditional probability) were assumed to be present. The results showed that extended visual content vectors were better identified (Fig. 4F and fig. S16), suggesting that multilabel decoding outputs may represent both reported and unreported contents.

Together, our findings provide evidence that specific contents of visual experience during sleep are represented by, and can be read out from, visual cortical activity patterns shared with stimulus representation. Our approach extends previous research on the (re)activation of the brain during sleep (*24–27*) and the relationship between dreaming and brain activity (*2, 3, 28*) by discovering links between complex brain activity patterns and unstructured verbal reports using database-assisted machine-learning decoders. The results suggest that the principle of perceptual equivalence (*29*), which postulates a common neural substrate for perception and imagery, generalizes to spontaneously generated visual experience during sleep. Although we have demonstrated semantic decoding with the HVC, this does not rule out the possibility of decoding low-level features with the LVC. The decoding presented here is retrospective in nature: Decoders were constructed after sleep experiments based on the collected reports. However, because reported synsets largely overlap between the first and the last halves of the experiments (59 out of 60 base synsets appeared in both), the same decoders may apply to future sleep data. The similarity between REM and sleep-onset reports (*13–15*) and the visual cortical activation during the REM sleep (*24, 25, 28*) suggest that the same decoders could also be used to decode REM imagery. Our method may further work beyond the bounds of sleep stages and reportable experience to uncover the dynamics of spontaneous brain activity in association with stimulus representation. We expect that it will lead to a better understanding of

the functions of dreaming and spontaneous neural events (*10, 30*).

References and Notes

1. W. Dement, N. Kleitman, *J. Exp. Psychol.* **53**, 339 (1957).
2. M. Dresler et al., *Curr. Biol.* **21**, 1833 (2011).
3. C. Marzano et al., *J. Neurosci.* **31**, 6674 (2011).
4. J. V. Haxby et al., *Science* **293**, 2425 (2001).
5. Y. Kamitani, F. Tong, *Nat. Neurosci.* **8**, 679 (2005).
6. S. A. Harrison, F. Tong, *Nature* **458**, 632 (2009).
7. K. N. Kay, T. Naselaris, R. J. Prenger, J. L. Gallant, *Nature* **452**, 352 (2008).
8. Y. Miyawaki et al., *Neuron* **60**, 915 (2008).
9. M. Stokes, R. Thompson, R. Cusack, J. Duncan, *J. Neurosci.* **29**, 1565 (2009).
10. Y. Nir, G. Tononi, *Trends Cogn. Sci.* **14**, 88 (2010).
11. T. Hori, M. Hayashi, T. Morikawa, in *Sleep Onset: Normal and Abnormal Processes*, R. D. Ogilvie, J. R. Harsh, Eds. (American Psychological Association, Washington, DC, 1994), pp. 237–253.
12. R. Stickgold, A. Malia, D. Maguire, D. Roddenberry, M. O'Connor, *Science* **290**, 350 (2000).
13. D. Foulkes, G. Vogel, *J. Abnorm. Psychol.* **70**, 231 (1965).
14. G. W. Vogel, B. Barrowclough, D. D. Giesler, *Arch. Gen. Psychiatry* **26**, 449 (1972).
15. D. Oudiette et al., *Conscious. Cogn.* **21**, 1129 (2012).
16. Materials and methods are available as supplementary materials on *Science* Online.
17. C. Fellbaum, Ed., *WordNet: An Electronic Lexical Database* (MIT Press, Cambridge, MA, 1998).
18. A. G. Huth, S. Nishimoto, A. T. Vu, J. L. Gallant, *Neuron* **76**, 1210 (2012).
19. J. Deng et al., *IEEE Conference on Computer Vision and Pattern Recognition,* 248 (2009).
20. V. N. Vapnik, *Statistical Learning Theory* (Wiley, New York, 1998).
21. E. Kobatake, K. Tanaka, *J. Neurophysiol.* **71**, 856 (1994).
22. R. Epstein, N. Kanwisher, *Nature* **392**, 598 (1998).
23. N. Kanwisher, J. McDermott, M. M. Chun, *J. Neurosci.* **17**, 4302 (1997).
24. A. R. Braun et al., *Science* **279**, 91 (1998).
25. P. Maquet, *J. Sleep Res.* **9**, 207 (2000).
26. Y. Yotsumoto et al., *Curr. Biol.* **19**, 1278 (2009).
27. M. A. Wilson, B. L. McNaughton, *Science* **265**, 676 (1994).
28. S. Miyauchi, M. Misaki, S. Kan, T. Fukunaga, T. Koike, *Exp. Brain Res.* **192**, 657 (2009).
29. R. A. Finke, *Principles of Mental Imagery* (MIT Press, Cambridge, MA, 1989).
30. J. A. Hobson, *Nat. Rev. Neurosci.* **10**, 803 (2009).

Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1234330/DC1
Materials and Methods
Figs. S1 to S16
Tables S1 to S7
References (*31–43*)
Movies S1 and S2

Science

Neural Decoding of Visual Imagery During Sleep

T. Horikawa, M. Tamaki, Y. Miyawaki and Y. Kamitani

Reading Dreams

How specific visual dream contents are represented by brain activity is unclear. Machine-learning–based analyses can decode the stimulus- and task-induced brain activity patterns that represent specific visual contents. **Horikawa *et al.*** (p. 639, published online 4 April) examined patterns of brain activity during dreaming and compared these to waking responses to visual stimuli. The findings suggest that the visual content of dreams is represented by the same neural substrate as observed during awake perception.

ARTICLE TOOLS	http://science.sciencemag.org/content/340/6132/639
SUPPLEMENTARY MATERIALS	http://science.sciencemag.org/content/suppl/2013/04/04/science.1234330.DC2 http://science.sciencemag.org/content/suppl/2013/04/03/science.1234330.DC1
RELATED CONTENT	http://science.sciencemag.org/content/sci/340/6128/21.full http://science.sciencemag.org/content/sci/340/6128/99.2.full
REFERENCES	This article cites 34 articles, 9 of which you can access for free http://science.sciencemag.org/content/340/6132/639#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service