

# High-performance brain-to-text communication via handwriting

<https://doi.org/10.1038/s41586-021-03506-2>

Received: 2 July 2020

Accepted: 26 March 2021

Published online: 12 May 2021

 Check for updates

Francis R. Willett<sup>1,2,3</sup>, Donald T. Avansino<sup>1</sup>, Leigh R. Hochberg<sup>4,5,6,7</sup>, Jaimie M. Henderson<sup>2,8,9,12</sup> & Krishna V. Shenoy<sup>1,3,8,9,10,11,12</sup>

Brain–computer interfaces (BCIs) can restore communication to people who have lost the ability to move or speak. So far, a major focus of BCI research has been on restoring **gross motor skills**, such as reaching and grasping<sup>1–5</sup> or point-and-click typing with a computer cursor<sup>6,7</sup>. However, rapid sequences of highly **dexterous** behaviours, such as handwriting or touch typing, might enable faster rates of communication. Here we developed an **intracortical** BCI that decodes attempted handwriting movements from neural activity in the motor cortex and translates it to text in real time, using a **recurrent neural network decoding approach**. With this BCI, our study participant, whose hand was paralysed from spinal cord injury, achieved typing speeds of 90 characters per minute with 94.1% raw accuracy online, and greater than 99% accuracy offline with a general-purpose autocorrect. To our knowledge, these typing speeds exceed those reported for any other BCI, and are comparable to typical smartphone typing speeds of individuals in the age group of our participant (115 characters per minute)<sup>8</sup>. Finally, **theoretical considerations explain why temporally complex movements, such as handwriting, may be fundamentally easier to decode than point-to-point movements**. Our results open a new approach for BCIs and demonstrate the feasibility of accurately decoding rapid, dexterous movements years after paralysis.

Previous BCI studies have shown that the motor intention for gross motor skills, such as reaching, grasping or moving a computer cursor, remains neurally encoded in the **motor cortex** after paralysis<sup>1–7</sup>. However, it is still unknown whether the neural representation for a rapid and highly dexterous motor skill, such as handwriting, also remains **intact**. We tested this by recording neural activity from two **microelectrode arrays** in the hand ‘knob’ area of the precentral gyrus (a premotor area)<sup>9,10</sup> while our BrainGate study participant, T5, attempted to handwrite individual letters and symbols (Fig. 1a). T5 has a high-level spinal cord injury and was paralysed from the neck down; his hand movements were entirely non-functional and limited to twitching and micromotion. We instructed T5 to ‘attempt’ to write as if his hand were not paralysed, while imagining that he was holding a pen on a piece of ruled paper.

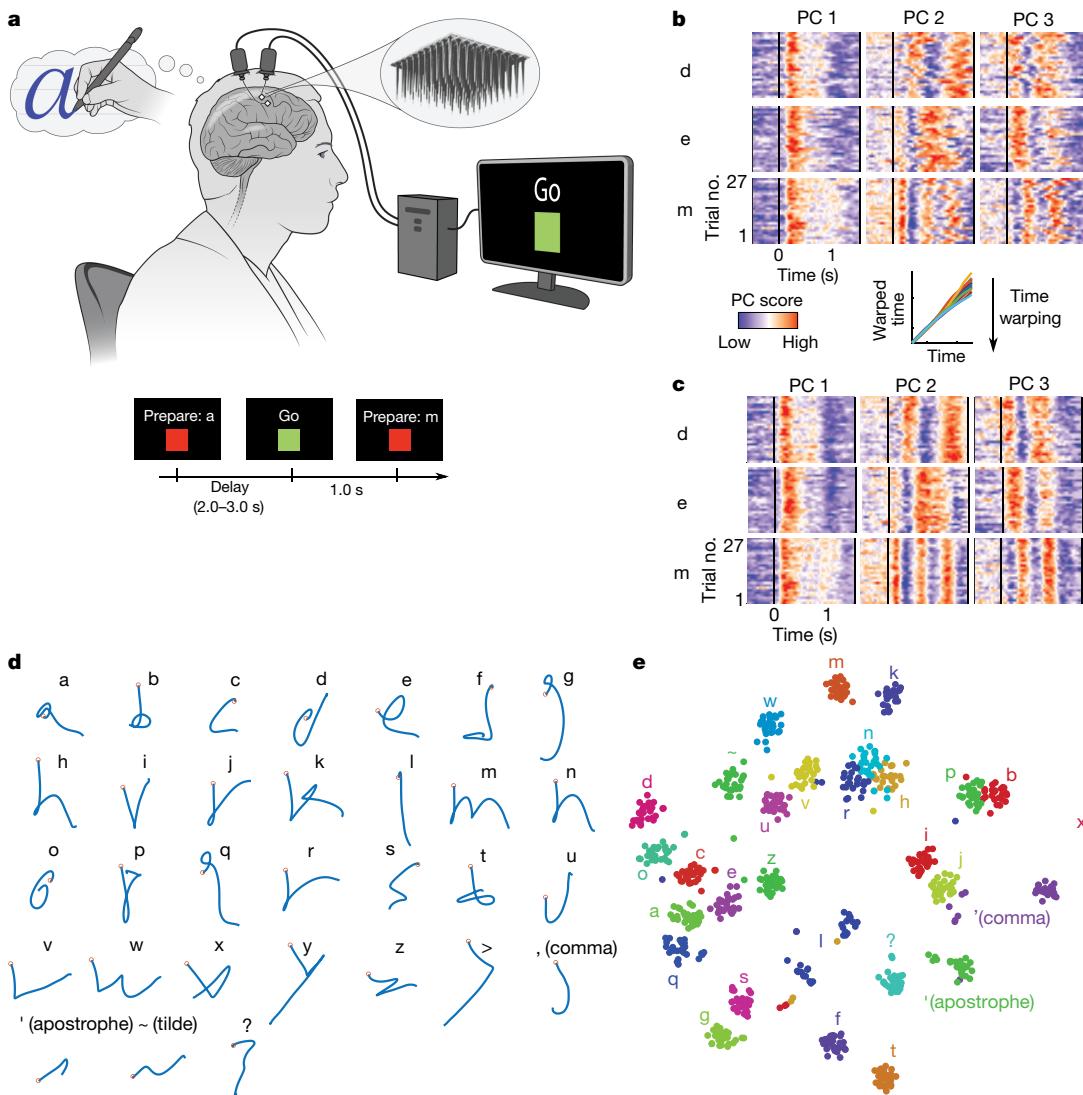
## Neural representation of handwriting

To visualize the neural activity (multiunit threshold crossing rates) recorded during attempted handwriting, we used principal components analysis to display the top three neural dimensions that contain the most variance (Fig. 1b). **The neural activity appeared to be strong and repeatable**, although the timing of its peaks and valleys varied

across trials, potentially owing to fluctuations in writing speed. We used a **time-alignment technique** to remove temporal variability<sup>11</sup>, which revealed notably consistent underlying patterns of neural activity that are unique to each character (Fig. 1c). To ascertain whether the neural activity encoded the pen movements that are needed to draw each character’s shape, we attempted to reconstruct each character by **linearly decoding** the pen-tip velocity from the trial-averaged neural activity (Fig. 1d). Readily recognizable letter shapes confirmed that **pen-tip velocity is robustly encoded**. The neural dimensions that represented pen-tip velocity accounted for 30% of the total neural variance.

Next, we used a nonlinear dimensionality reduction method (**t-distributed stochastic neighbour embedding**; t-SNE) to produce a two-dimensional (2D) visualization of each single trial’s neural activity recorded **after the ‘go’ cue** was given (Fig. 1e). The t-SNE visualization revealed tight clusters of neural activity for each character and a predominantly motoric encoding in which characters that are written similarly are closer together. Using a **k-nearest-neighbour classifier applied offline** to the neural activity, we could classify the characters with 94.1% accuracy (95% confidence interval (CI) = [92.6, 95.8]). Together, these results suggest that, even years after paralysis, the neural representation of handwriting in the motor cortex is probably strong enough to be useful for a BCI.

<sup>1</sup>Howard Hughes Medical Institute at Stanford University, Stanford, CA, USA. <sup>2</sup>Department of Neurosurgery, Stanford University School of Medicine, Stanford, CA, USA. <sup>3</sup>Department of Electrical Engineering, Stanford University, Stanford, CA, USA. <sup>4</sup>VA RR&D Center for Neurorestoration and Neurotechnology, Rehabilitation R&D Service, Providence VA Medical Center, Providence, RI, USA. <sup>5</sup>School of Engineering, Brown University, Providence, RI, USA. <sup>6</sup>Carney Institute for Brain Science, Brown University, Providence, RI, USA. <sup>7</sup>Center for Neurotechnology and Neurorecovery, Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA. <sup>8</sup>Wu Tsai Neurosciences Institute, Stanford University, Stanford, CA, USA. <sup>9</sup>Bio-X Institute, Stanford University, Stanford, CA, USA. <sup>10</sup>Department of Bioengineering, Stanford University, Stanford, CA, USA. <sup>11</sup>Department of Neurobiology, Stanford University, Stanford, CA, USA. <sup>12</sup>These authors jointly supervised this work: Jaimie M. Henderson, Krishna V. Shenoy. <sup>✉</sup>e-mail: fwillett@stanford.edu



**Fig. 1 | Neural representation of attempted handwriting.** **a**, To assess the neural representation of attempted handwriting, participant T5 attempted to handwrite each character one at a time, following the instructions given on a computer screen (bottom panels depict what is shown on the screen, following the timeline). Credit: drawing of the human silhouette created by E. Woodrum. **b**, Neural activity in the top 3 principal components (PCs) is shown for three example letters (d, e and m) and 27 repetitions of each letter (trials). The colour scale was normalized within each panel separately for visualization. **c**, Time-warping the neural activity to remove trial-to-trial changes in writing

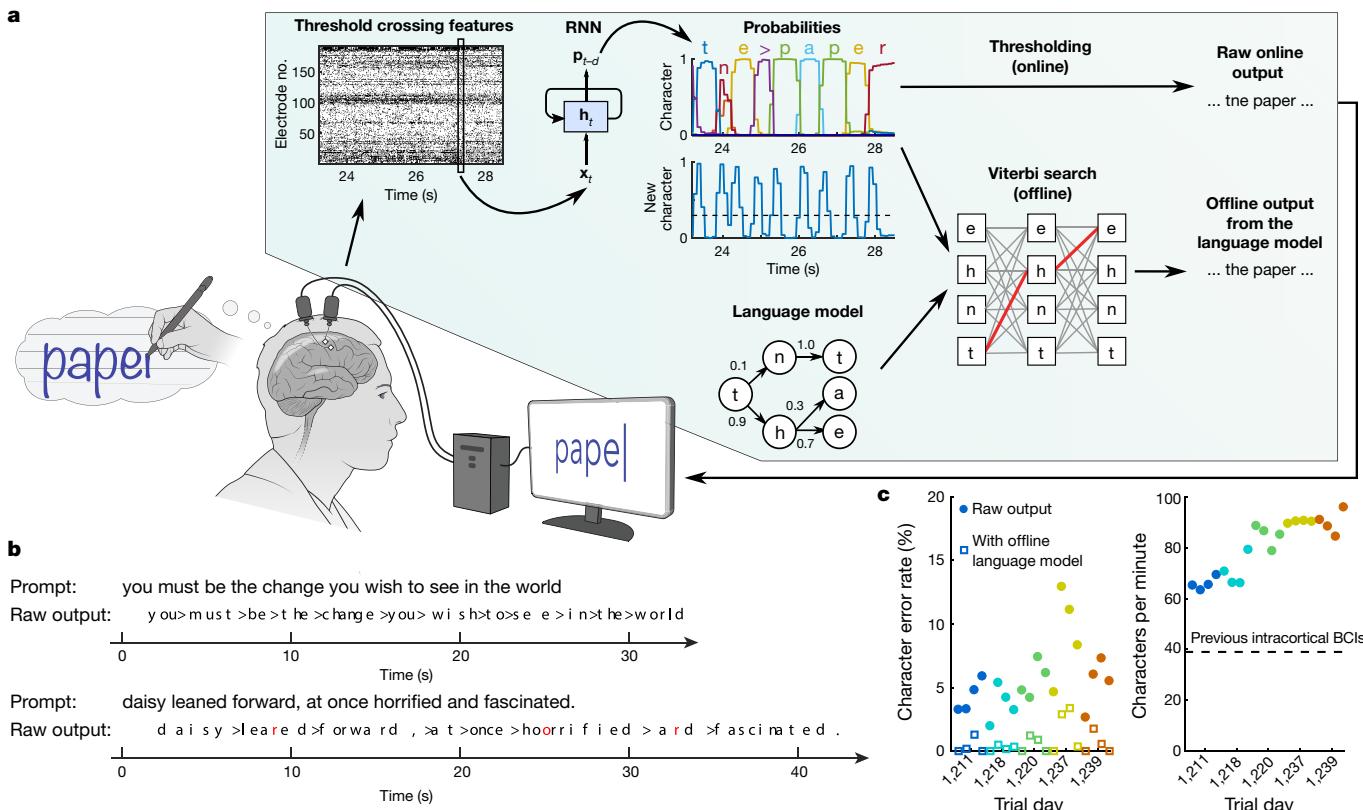
speed reveals consistent patterns of activity unique to each letter. In the inset above **c**, example time-warping functions are shown for the letter ‘m’ and lie relatively close to the identity line (the warping function of each trial is plotted with a different coloured line). **d**, Decoded pen trajectories are shown for all 31 tested characters. Intended 2D pen-tip velocity was linearly decoded from the neural activity using cross-validation (each character was held out), and decoder output was denoised by averaging across trials. Orange circles denote the start of the trajectory. **e**, A 2D visualization of the neural activity made using t-SNE. Each circle is a single trial (27 trials are shown for each of 31 characters).

## Decoding handwritten sentences

Next, we tested whether we could decode complete handwritten sentences in real time, thus enabling an individual with tetraplegia to communicate by attempting to handwrite their intended message. To do so, we trained a recurrent neural network (RNN) to convert the neural activity into probabilities describing the likelihood of each character being written at each moment in time (Fig. 2a, Extended Data Fig. 1). These probabilities could either be thresholded in a simple way to emit discrete characters, which we did for real-time decoding (‘raw online output’, Fig. 2a), or processed more extensively by a large-vocabulary language model to simulate an autocorrect feature, which we applied offline (‘offline output from a language model’, Fig. 2a). We used the limited set of 31 characters shown in Fig. 1d, consisting of the 26 lower-case letters of the alphabet, together with commas, apostrophes, question marks, full stops (written by T5 as a tilde symbol; ‘~’), and spaces (written

by T5 as a greater-than symbol; ‘>’). The ‘~’ and ‘>’ symbols were chosen to make full stops and spaces easier to detect. T5 attempted to write each character in print (not cursive), with each character printed on top of the previous one.

To collect training data for the RNN, we recorded neural activity while T5 attempted to handwrite complete sentences at his own pace, following instructions on a computer monitor. Before the first day of real-time evaluation, we collected a total of 242 sentences across 3 pilot days that were combined to train the RNN. On each subsequent day of real-time testing, additional training data were collected to recalibrate the RNN before evaluation, yielding a combined total of 572 training sentences by the last day (comprising 7.6 hours and 31,472 characters). To train the RNN, we adapted neural network methods in automatic speech recognition<sup>12–14</sup> to overcome two key challenges: (1) the time that each letter was written in the training data was unknown (as T5’s hand was paralysed), making it challenging to apply supervised learning



**Fig. 2 | Neural decoding of attempted handwriting in real time. a,** Diagram of the decoding algorithm. First, the neural activity (multiunit threshold crossings) was temporally binned and smoothed on each electrode (20-ms bins). Then, an RNN converted this neural population time series ( $x_t$ ) into a probability time series ( $p_{t-d}$ ) describing the likelihood of each character and the probability of any new character beginning. The RNN had a one second output delay ( $d$ ), giving it time to observe each character fully before deciding its identity. Finally, the character probabilities were thresholded to produce the ‘raw online output’ for real-time use (when the ‘new character’ probability crossed a threshold at time  $t$ , the most likely character at time  $t+0.3\text{s}$  was emitted and shown on the screen). In an offline retrospective analysis, the

character probabilities were combined with a large-vocabulary language model to decode the most likely text that the participant wrote (using a custom 50,000-word bigram model). Credit: drawing of the human silhouette created by E. Woodrum. **b**, Two real-time example trials are shown, demonstrating the ability of the RNN to decode readily understandable text on sentences on which it was never trained. Errors are highlighted in red and spaces are denoted with ‘>. **c**, Error rates (edit distances) and typing speeds are shown for 5 days, with 4 blocks of 7–10 sentences each (each block is indicated with a single circle and coloured according to the trial day). The speed is more than double that of the next-fastest intracortical BCI<sup>7</sup>, which is indicated with the dashed line.

techniques; and (2) the dataset was limited in size compared to typical RNN datasets, making it difficult to prevent overfitting to the training data (see Supplementary Methods, Extended Data Figs. 2, 3).

We evaluated the performance of the RNN over a series of 5 days, each day containing 4 evaluation blocks of 7–10 sentences that the RNN was never trained on (thus ensuring that the RNN could not overfit to those sentences). T5 copied each sentence from an on-screen prompt, attempting to handwrite it letter by letter, while the decoded characters appeared on the screen in real time as they were detected by the RNN (Supplementary Videos 1, 2, Extended Data Table 1). Characters appeared after they were completed by T5 with a short delay (estimated to be 0.4–0.7 s). The decoded sentences were quite legible (‘raw output’, Fig. 2b). Notably, typing speeds were high, plateauing at 90 characters per minute with a mean character error rate of 5.4% (averaged across all four blocks on the final day) (Fig. 2c). As there was no ‘backspace’ function implemented, T5 was instructed to continue writing if any decoding errors occurred.

When a language model was used to autocorrect errors offline, error rates decreased considerably (Fig. 2c, Table 1). The character error rate decreased to 0.89% and the word error rate decreased to 3.4% averaged across all days, which is comparable to state-of-the-art speech recognition systems with word error rates of 4–5%<sup>14,15</sup>, putting it well

within the range of usability. Finally, to probe the limits of possible decoding performance, we trained a new RNN offline using all available sentences to process the entire sentence in a non-causal way (comparable to other BCI studies<sup>16,17</sup>). Accuracy was extremely high in this regime (0.17% character error rate), indicating a high potential ceiling of performance, although this decoder would not be able to provide letter-by-letter feedback to the user.

Next, to evaluate performance in a less restrained setting, we collected two days of data in which T5 used the BCI to freely type answers to open-ended questions (Supplementary Video 3, Extended Data Table 2). The results confirm that high performance can also be achieved when the user writes self-generated sentences as opposed to copying on-screen prompts (73.8 characters per minute, with a character error rate of 8.54% in real time and 2.25% with a language model). To our knowledge, the previous record for free typing in intracortical BCIs is 24.4 correct characters per minute<sup>7</sup>.

## Daily decoder retraining

Following standard practice<sup>1,2,4,5,18</sup>, we retrained our handwriting decoder each day before evaluating it, with the help of ‘calibration’ data collected at the beginning of each day. Retraining helps to account

# Article

**Table 1 | Mean character and word error rates (with 95% CIs) for the handwriting BCI across all 5 days**

	Character error rate [95% CI]	Word error rate [95% CI]
Raw online output	5.9% [5.3, 6.5]	25.1% [22.5, 27.4]
Online output + offline language model	0.89% [0.61, 1.2]	3.4% [2.5, 4.4]
Offline bidirectional RNN + language model	0.17% [0, 0.36]	1.5% [0, 3.2]

'Raw online output' is what was decoded online (in real time). 'Online output + offline language model' was obtained by applying a language model retrospectively to what was decoded online (to simulate an autocorrect feature). 'Offline bidirectional RNN + language model' was obtained by retraining a bidirectional (acausal) decoder offline using all available data, in addition to applying a language model. Word error rates can be much higher than character error rates because a word is considered incorrect if any character in that word is wrong.

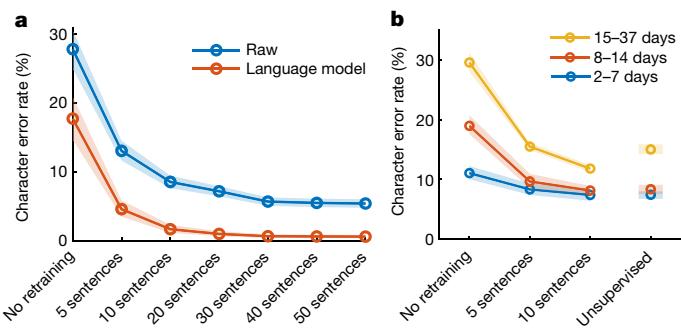
for changes in neural recordings that accrue over time, which might be caused by neural plasticity or electrode array micromotion. Ideally, to reduce the burden on the user, minimal or no calibration data would be required. In a retrospective analysis of the copy-typing data reported above in Fig. 2, we assessed whether high performance could still have been achieved using fewer than the original 50 calibration sentences per day (Fig. 3a). We found that 10 sentences (8.7 min) were enough to achieve a raw error rate of 8.5% (1.7% with a language model), although 30 sentences were needed to match the raw online error rate of 5.9%.

However, our copy-typing data were collected over a 28-day time span, possibly allowing larger changes in neural activity to accumulate. Using further offline analyses, we assessed whether sessions that are more closely spaced reduce the need for calibration data (Fig. 3b). We found that when only 2–7 days passed between sessions, performance was reasonable with no decoder retraining (11.1% raw error rate, 1.5% with a language model), as might be expected from previous work showing the short-term stability of neural recordings<sup>19–21</sup>. Finally, we tested whether decoders could be retrained in an unsupervised manner by using a language model to error-correct and retrain the decoder, thus bypassing the need to interrupt the user for calibration (by enabling automatic recalibration during normal use). Encouragingly, unsupervised retraining achieved a raw error rate of 7.3% (0.84% with a language model) when sessions were separated by a time span of 7 days or less.

Ultimately, whether decoders can be successfully retrained with minimal recalibration data depends on how quickly the neural activity changes over time. We assessed the stability of the neural patterns associated with each character and found high short-term stability (mean correlation of 0.85 when 7 days apart or less), and neural changes that seemed to accumulate at a steady and predictable rate (Extended Data Fig. 4). These results are promising for clinical viability, as they suggest that unsupervised decoder retraining, combined with more-limited supervised retraining after longer periods of inactivity, may be sufficient to achieve high performance. Nevertheless, future work must confirm this online, as offline simulations are not always predictive of online performance.

## Temporal variety improves decoding

To our knowledge, 90 characters per minute is the highest typing rate that has yet been reported for any type of BCI (see 'Discussion'). For intracortical BCIs, the best-performing method has been point-and-click typing with a 2D computer cursor, which peaks at 40 correct characters per minute<sup>7</sup> (see Supplementary Video 4 for a direct comparison). The speed of point-and-click BCIs is limited primarily by decoding accuracy. During parameter optimization, the cursor gain is increased so as to increase typing rate, until the cursor moves so quickly that it becomes uncontrollable owing to decoding errors<sup>22</sup>.



**Fig. 3 | Performance remains high when daily decoder retraining is shortened (or unsupervised).** **a**, To account for changes in neural activity that accrue over time, we retrained our handwriting decoder each day before evaluating it. Here, we simulated offline how decoding performance would have changed if fewer than the original 50 calibration sentences were used. Lines show the mean error rate across all data and shaded regions indicate 95% CIs. **b**, Copy-typing data from eight sessions were used to assess whether fewer calibration data are required if sessions occur closer in time. All session pairs ( $X, Y$ ) were considered. Decoders were first initialized using training data from session X and earlier, and then evaluated on session Y under different retraining methods (no retraining, retraining with limited calibration data, or unsupervised retraining). Lines show the average raw error rate and shaded regions indicate 95% CIs.

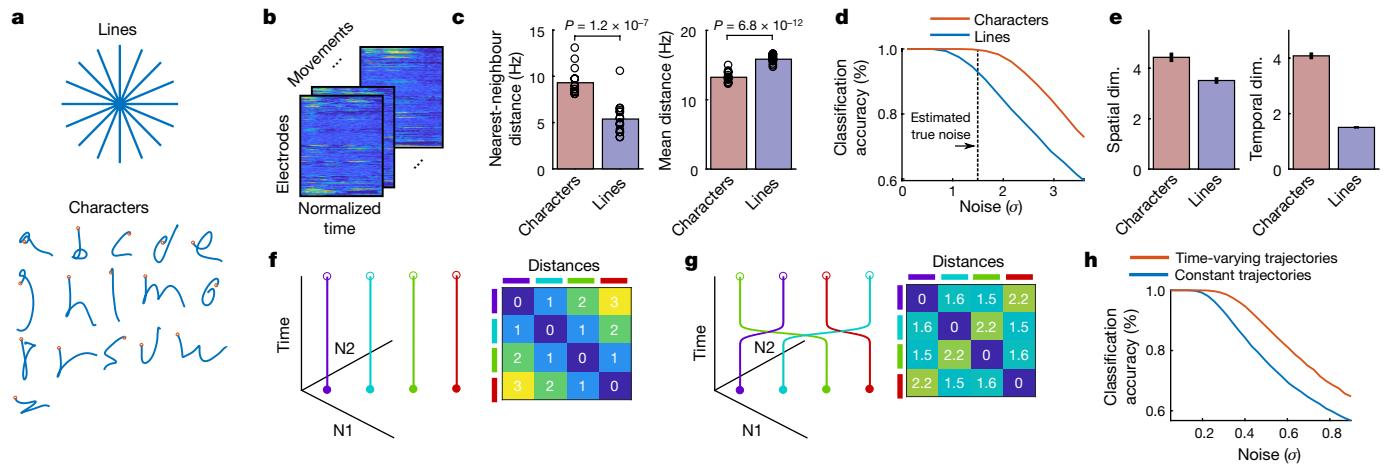
We therefore asked how handwriting movements could be decoded more than twice as fast, with similar levels of accuracy.

We theorize that handwritten letters may be easier to distinguish from each other than point-to-point movements, as letters have more variety in their spatiotemporal patterns of neural activity than do straight-line movements. To test this theory, we analysed the patterns of neural activity associated with 16 straight-line movements and 16 letters that required no lifting of the pen off the page, both performed by T5 with attempted handwriting (Fig. 4a, b).

First, we analysed the pairwise Euclidean distances between each neural activity pattern. We found that the nearest-neighbour distances for each movement were 72% larger for characters compared to straight lines (95% CI = [60%, 86%]), making it less likely for a decoder to confuse two nearby characters (Fig. 4c). To confirm this, we simulated the classification accuracy for each set of movements as a function of neural noise (Fig. 4d), which showed that characters are easier to classify than straight lines.

To gain insight into what might be responsible for the relative increase in nearest-neighbour distances for characters, we examined the spatial and temporal dimensionality of the neural patterns. Spatial and temporal dimensionality were estimated using the 'participation ratio' of the principal component analysis (PCA) eigenvalue spectrum, which quantifies approximately how many spatial or temporal dimensions are required to explain 80% of the variance in the patterns of neural activity<sup>23</sup>. We found that the spatial dimensionality was only modestly larger for characters (1.24 times larger; 95% CI = [1.19, 1.30]), but that the temporal dimensionality was much greater (2.65 times larger; 95% CI = [2.58, 2.72]), suggesting that the increased variety of temporal patterns in letter writing drives the increased separability of each movement (Fig. 4e).

To illustrate how increased temporal dimensionality can make movements more distinguishable, we constructed a toy model with four movements and two neurons, with the neural activity constrained to lie along a single dimension (Fig. 4f, g). Simply by allowing the trajectories to change in time (Fig. 4g), the nearest-neighbour distance between the neural trajectories can be increased, resulting in an increase in classification accuracy when noise levels are large enough (Fig. 4h). Although neural noise in the toy model was assumed to be independent white noise, we found that these results also hold for noise that



**Fig. 4 | Increased temporal variety can make movements easier to decode.** **a**, We analysed the spatiotemporal patterns of neural activity corresponding to 16 handwritten characters (1 s in duration) versus 16 handwritten straight-line movements (0.6 s in duration). **b**, Spatiotemporal neural patterns were found by averaging over all trials for a given movement (after time-warping to align the trials in time)<sup>11</sup>. Neural activity was resampled to equalize the duration of each set of movements, resulting in a  $192 \times 100$  matrix for each movement (192 electrodes and 100 time steps). **c**, Pairwise Euclidean distances between neural patterns were computed for each set, revealing larger nearest-neighbour distances (but not mean distances) for characters. Each circle represents a single movement and bar heights show the mean. **d**, Larger nearest-neighbour distances made the characters easier to classify than straight lines. The noise is in units of standard deviations and matches the scale

of the distances in **c**. **e**, The spatial dimensionality (dim.) was similar for characters and straight lines, but the temporal dimensionality was more than twice as high for characters, suggesting that more temporal variety underlies the increased nearest-neighbour distances and better classification performance. Error bars show 95% CIs. Dimensionality was quantified using the participation ratio. **f–h**, A toy example to give intuition for how increased temporal dimensionality can make neural trajectories more separable. Four neural trajectories are depicted (**N1** and **N2** are two hypothetical neurons, the activity of which is constrained to a single spatial dimension, the unity diagonal). Allowing the trajectories to vary in time by adding one bend, which increases the temporal dimensionality from 1 (**f**) to 2 (**g**), enables larger nearest-neighbour distances and better classification (**h**).

is correlated across time and neurons (Extended Data Fig. 5, Supplementary Note 1).

These results suggest that time-varying patterns of movement, such as handwritten letters, are fundamentally easier to decode than point-to-point movements. We think this is one—but not necessarily the only—important factor that enabled a handwriting BCI to go faster than continuous-motion point-and-click BCIs. Other discrete (classification-based) BCIs have also typically used directional movements with little temporal variety, which may have limited their accuracy and/or the size of the movement set<sup>24,25</sup>.

More generally, using the principle of maximizing the nearest-neighbour distance between movements, it should be possible to optimize a set of movements for ease of classification<sup>26</sup>. We investigated this possibility and designed an alphabet that is theoretically easier to classify than the Latin alphabet (Extended Data Fig. 6). The optimized alphabet avoids large clusters of redundant letters that are written similarly (most Latin letters begin with either a downstroke or a counter-clockwise curl).

## Discussion

Locked-in syndrome (paralysis of nearly all voluntary muscles) severely impairs or prevents communication, and is most frequently caused by brainstem stroke or late-stage amyotrophic lateral sclerosis (estimated prevalence of locked-in syndrome: 1 in 100,000<sup>27</sup>). Commonly used BCIs for restoring communication are either flashing electroencephalogram (EEG) spellers<sup>18,28–32</sup> or intracortical point-and-click BCIs<sup>6,7,33</sup>. EEG spellers based on oddball potentials or motor imagery typically achieve 1–5 characters per minute<sup>28–32</sup>. EEG spellers that use visually evoked potentials have achieved speeds of 60 characters per minute<sup>18</sup>, but have notable usability limitations, as they tie up the eyes, are not typically self-paced and require panels of flashing lights on a screen. Intracortical BCIs based on 2D cursor movements give the user more freedom to look around and set their own pace of communication, but have

yet to exceed 40 correct characters per minute in humans<sup>7</sup>. Recently, speech-decoding BCIs have shown exciting promise for restoring rapid communication<sup>16,17,34</sup>, but their accuracies and vocabulary sizes require further improvement to support general-purpose use.

Here, we introduced a new approach for communication BCIs—decoding a rapid, dexterous motor behaviour in an individual with tetraplegia—that sets a benchmark for communication rate at 90 characters per minute. The real-time system is general (the user can express any sentence), easy to use (entirely self-paced and the eyes are free to move) and accurate enough to be useful in the real-world (94.1% raw accuracy, and greater than 99% accuracy offline with a large-vocabulary language model). To achieve high performance, we developed decoding methods to work effectively with unlabelled neural sequences in data-limited regimes. These methods could be applied more generally to any sequential behaviour that cannot be observed directly (for example, decoding speech from someone who can no longer speak).

It is important to recognize that the current system is a proof of concept that a high-performance handwriting BCI is possible (in a single participant); it is not yet a complete, clinically viable system. More work is needed to demonstrate high performance in additional people, expand the character set (for example, capital letters), enable text editing and deletion, and maintain robustness to changes in neural activity without interrupting the user for decoder retraining. More broadly, intracortical microelectrode array technology is still maturing, and requires further demonstrations of longevity, safety and efficacy before widespread clinical adoption<sup>35,36</sup>. Variability in performance across participants is also a potential concern (in a previous study, T5 achieved the highest performance of three participants<sup>7</sup>).

Nevertheless, we believe that the future of intracortical BCIs is bright. Current microelectrode array technology has been shown to retain functionality for more than 1,000 days after implant<sup>37,38</sup> (including here; see Extended Data Fig. 7), and has enabled the highest BCI performance so far compared to other recording technologies (for example, EEG or electrocorticography) for restoring communication<sup>7</sup>, arm control<sup>25</sup>

and general-purpose computer use<sup>39</sup>. New developments are under way for implant designs that increase the electrode count by at least an order of magnitude, which will further improve performance and longevity<sup>35,36,40,41</sup>. Finally, we envision that a combination of algorithmic innovations<sup>42–44</sup> and improvements to device stability will continue to reduce the need for daily decoder retraining. Here, offline analyses showed the potential promise of more limited, or even unsupervised, decoder retraining (Fig. 3).

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-021-03506-2>.

1. Hochberg, L. R. et al. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature* **485**, 372–375 (2012).
2. Collinger, J. L. et al. High-performance neuroprosthetic control by an individual with tetraplegia. *Lancet* **381**, 557–564 (2013).
3. Aflalo, T. et al. Neurophysiology. Decoding motor imagery from the posterior parietal cortex of a tetraplegic human. *Science* **348**, 906–910 (2015).
4. Bouton, C. E. et al. Restoring cortical control of functional movement in a human with quadriplegia. *Nature* **533**, 247–250 (2016).
5. Ajiboye, A. B. et al. Restoration of reaching and grasping movements through brain-controlled muscle stimulation in a person with tetraplegia: a proof-of-concept demonstration. *Lancet* **389**, 1821–1830 (2017).
6. Jarosiewicz, B. et al. Virtual typing by people with tetraplegia using a self-calibrating intracortical brain–computer interface. *Sci. Transl. Med.* **7**, 313ra179 (2015).
7. Pandarinath, C. et al. High performance communication by people with paralysis using an intracortical brain–computer interface. *eLife* **6**, e18554 (2017).
8. Palin, K., Feit, A. M., Kim, S., Kristensson, P. O. & Oulasvirta, A. How do people type on mobile devices? Observations from a study with 37,000 volunteers. In *Proc. 21st International Conference on Human–Computer Interaction with Mobile Devices and Services* 1–12 (Association for Computing Machinery, 2019).
9. Yousy, T. A. et al. Localization of the motor hand area to a knob on the precentral gyrus. A new landmark. *Brain* **120**, 141–157 (1997).
10. Willett, F. R. et al. Hand knob area of premotor cortex represents the whole body in a compositional way. *Cell* **181**, 396–409 (2020).
11. Williams, A. H. et al. Discovering precise temporal patterns in large-scale neural recordings through robust and interpretable time warping. *Neuron* **105**, 246–259 (2020).
12. Hinton, G. et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. *IEEE Signal Process. Mag.* **29**, 82–97 (2012).
13. Graves, A., Mohamed, A. & Hinton, G. Speech recognition with deep recurrent neural networks. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing* 6645–6649 (2013).
14. Xiong, W. et al. The Microsoft 2017 Conversational Speech Recognition System. Preprint at <https://arxiv.org/abs/1708.06073> (2017).
15. He, Y. et al. Streaming end-to-end speech recognition for mobile devices. In *2019 IEEE International Conference on Acoustics, Speech and Signal Processing* 6381–6385 (2019).
16. Anumanchipalli, G. K., Chartier, J. & Chang, E. F. Speech synthesis from neural decoding of spoken sentences. *Nature* **568**, 493–498 (2019).
17. Makin, J. G., Moses, D. A. & Chang, E. F. Machine translation of cortical activity to text with an encoder-decoder framework. *Nat. Neurosci.* **23**, 575–582 (2020).
18. Chen, X. et al. High-speed spelling with a noninvasive brain–computer interface. *Proc. Natl Acad. Sci. USA* **112**, E6058–E6067 (2015).
19. Dickey, A. S., Suminski, A., Amit, Y. & Hatsopoulos, N. G. Single-unit stability using chronically implanted multielectrode arrays. *J. Neurophysiol.* **102**, 1331–1339 (2009).
20. Eleryan, A. et al. Tracking single units in chronic, large scale, neural recordings for brain machine interface applications. *Front. Neuroeng.* **7**, 23 (2014).
21. Downey, J. E., Schwed, N., Chase, S. M., Schwartz, A. B. & Collinger, J. L. Intracortical recording stability in human brain–computer interface users. *J. Neural Eng.* **15**, 046016 (2018).
22. Willett, F. R. et al. Signal-independent noise in intracortical brain–computer interfaces causes movement time properties inconsistent with Fitts' law. *J. Neural Eng.* **14**, 026010 (2017).
23. Gao, P. et al. A theory of multineuronal dimensionality, dynamics and measurement. Preprint at <https://doi.org/10.1101/214262> (2017).
24. Musallam, S., Corneil, B. D., Greger, B., Scherberger, H. & Andersen, R. A. Cognitive control signals for neural prosthetics. *Science* **305**, 258–262 (2004).
25. Santhanam, G., Ryu, S. I., Yu, B. M., Afshar, A. & Shenoy, K. V. A high-performance brain–computer interface. *Nature* **442**, 195–198 (2006).
26. Cunningham, J. P., Yu, B. M., Gilja, V., Ryu, S. I. & Shenoy, K. V. Toward optimal target placement for neural prosthetic devices. *J. Neurophysiol.* **100**, 3445–3457 (2008).
27. Pels, E. G. M., Aarnoutse, E. J., Ramsey, N. F. & Vansteensel, M. J. Estimated prevalence of the target population for brain–computer interface neurotechnology in the Netherlands. *Neurorehabil. Neural Repair* **31**, 677–685 (2017).
28. Vansteensel, M. J. et al. Fully implanted brain–computer interface in a locked-in patient with ALS. *N. Engl. J. Med.* **375**, 2060–2066 (2016).
29. Nijboer, F. et al. A P300-based brain–computer interface for people with amyotrophic lateral sclerosis. *Clin. Neurophysiol.* **119**, 1909–1916 (2008).
30. Townsend, G. et al. A novel P300-based brain–computer interface stimulus presentation paradigm: moving beyond rows and columns. *Clin. Neurophysiol.* **121**, 1109–1120 (2010).
31. McCane, L. M. et al. P300-based brain–computer interface (BCI) event-related potentials (ERPs): people with amyotrophic lateral sclerosis (ALS) vs. age-matched controls. *Clin. Neurophysiol.* **126**, 2124–2131 (2015).
32. Wolpaw, J. R. et al. Independent home use of a brain–computer interface by people with amyotrophic lateral sclerosis. *Neurology* **91**, e258–e267 (2018).
33. Bacher, D. et al. Neural point-and-click communication by a person with incomplete locked-in syndrome. *Neurorehabil. Neural Repair* **29**, 462–471 (2015).
34. Mugler, E. M. et al. Direct classification of all American English phonemes using signals from functional speech motor cortex. *J. Neural Eng.* **11**, 035015 (2014).
35. Nurmiikkko, A. Challenges for large-scale cortical interfaces. *Neuron* **108**, 259–269 (2020).
36. Vázquez-Guardado, A., Yang, Y., Bandodkar, A. J. & Rogers, J. A. Recent advances in neurotechnologies with broad potential for neuroscience research. *Nat. Neurosci.* **23**, 1522–1536 (2020).
37. Simner, J. D., Kim, S.-P., Black, M. J., Donoghue, J. P. & Hochberg, L. R. Neural control of cursor trajectory and click by a human with tetraplegia 1000 days after implant of an intracortical microelectrode array. *J. Neural Eng.* **8**, 025027 (2011).
38. Bullard, A. J., Hutchison, B. C., Lee, J., Chestek, C. A. & Patil, P. G. Estimating risk for future intracranial, fully implanted, modular neuroprosthetic systems: a systematic review of hardware complications in clinical deep brain stimulation and experimental human intracortical arrays. *Neuromodulation* **23**, 411–426 (2020).
39. Nuyujikian, P. et al. Cortical control of a tablet computer by people with paralysis. *PLoS One* **13**, e0204566 (2018).
40. Musk, E. An integrated brain–machine interface platform with thousands of channels. *J. Med. Internet Res.* **21**, e16194 (2019).
41. Sahasrabuddhe, K. et al. The Argo: a high channel count recording system for neural recording in vivo. *J. Neural Eng.* **18**, 015002 (2021).
42. Sussillo, D., Stavisky, S. D., Kao, J. C., Ryu, S. I. & Shenoy, K. V. Making brain–machine interfaces robust to future neural variability. *Nat. Commun.* **7**, 13749 (2016).
43. Dyer, E. L. et al. A cryptography-based approach for movement decoding. *Nat. Biomed. Eng.* **1**, 967–976 (2017).
44. Degenhart, A. D. et al. Stabilization of a brain–computer interface via the alignment of low-dimensional spaces of neural activity. *Nat. Biomed. Eng.* **4**, 672–685 (2020).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2021

## Reporting summary

Further information on research design is available in the *Nature* Research Reporting Summary linked to this paper.

## Data availability

All neural data needed to reproduce the findings in this study are publicly available at the Dryad repository (<https://doi.org/10.5061/dryad.wh70rxwmv>). The dataset contains neural activity recorded during the attempted handwriting of 1,000 sentences (43,501 characters) over 10.7 hours.

## Code availability

Code that implements an offline reproduction of the central findings in this study (high-performance neural decoding with an RNN) is publicly available on GitHub at <https://github.com/fwillett/handwritingBCI>.

**Acknowledgements** We thank participant T5 and his caregivers for their dedicated contributions to this research, N. Lam, E. Siauciunas and B. Davis for administrative support and E. Woodrum for the drawings in Figs. 1a, 2a. F.R.W. and D.T.A. acknowledge the support of the Howard Hughes Medical Institute. L.R.H. acknowledges the support of the Office of Research and Development, Rehabilitation R&D Service, US Department of Veterans Affairs (A2295R, N2864C); the National Institute of Neurological Disorders and Stroke and BRAIN Initiative (UH2NS095548); and the National Institute on Deafness and Other

Communication Disorders (R01DC009899, U01DC017844). K.V.S. and J.M.H. acknowledge the support of the National Institute on Deafness and Other Communication Disorders (R01-DC014034, U01-DC017844); the National Institute of Neurological Disorders and Stroke (UH2-NS095548, U01-NS098968); L. and P. Garlick; S. and B. Reeves; and the Wu Tsai Neurosciences Institute at Stanford. K.V.S. acknowledges the support of the Simons Foundation Collaboration on the Global Brain 543045 and the Howard Hughes Medical Institute (K.V.S. is a Howard Hughes Medical Institute Investigator). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

**Author contributions** F.R.W. conceived the study, built the real-time decoder, analysed the data and wrote the manuscript. F.R.W. and D.T.A. collected the data. L.R.H. is the sponsor-investigator of the multi-site clinical trial. J.M.H. planned and performed T5's array placement surgery and was responsible for all clinical-trial-related activities at Stanford. J.M.H. and K.V.S. supervised and guided the study. All authors reviewed and edited the manuscript.

**Competing interests** The MGH Translational Research Center has a clinical research support agreement with Neuralink, Paradromics and Synchron, for which L.R.H. provides consultative input. J.M.H. is a consultant for Neuralink, and serves on the Medical Advisory Board of Enspire DBS. K.V.S. consults for Neuralink and CTRL-Labs (part of Facebook Reality Labs) and is on the scientific advisory boards of MIND-X, Inscopix and Heal. F.R.W., J.M.H. and K.V.S. are inventors on patent application US 2021/0064135 A1 (the applicant is Stanford University), which covers the neural decoding approach taken in this work. All other authors have no competing interests.

## Additional information

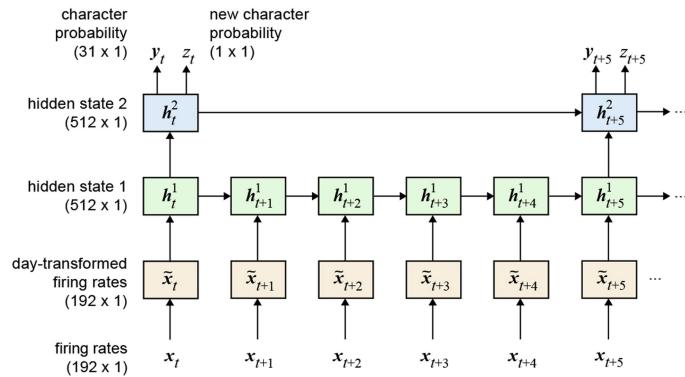
**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41586-021-03506-2>.

**Correspondence and requests for materials** should be addressed to F.R.W.

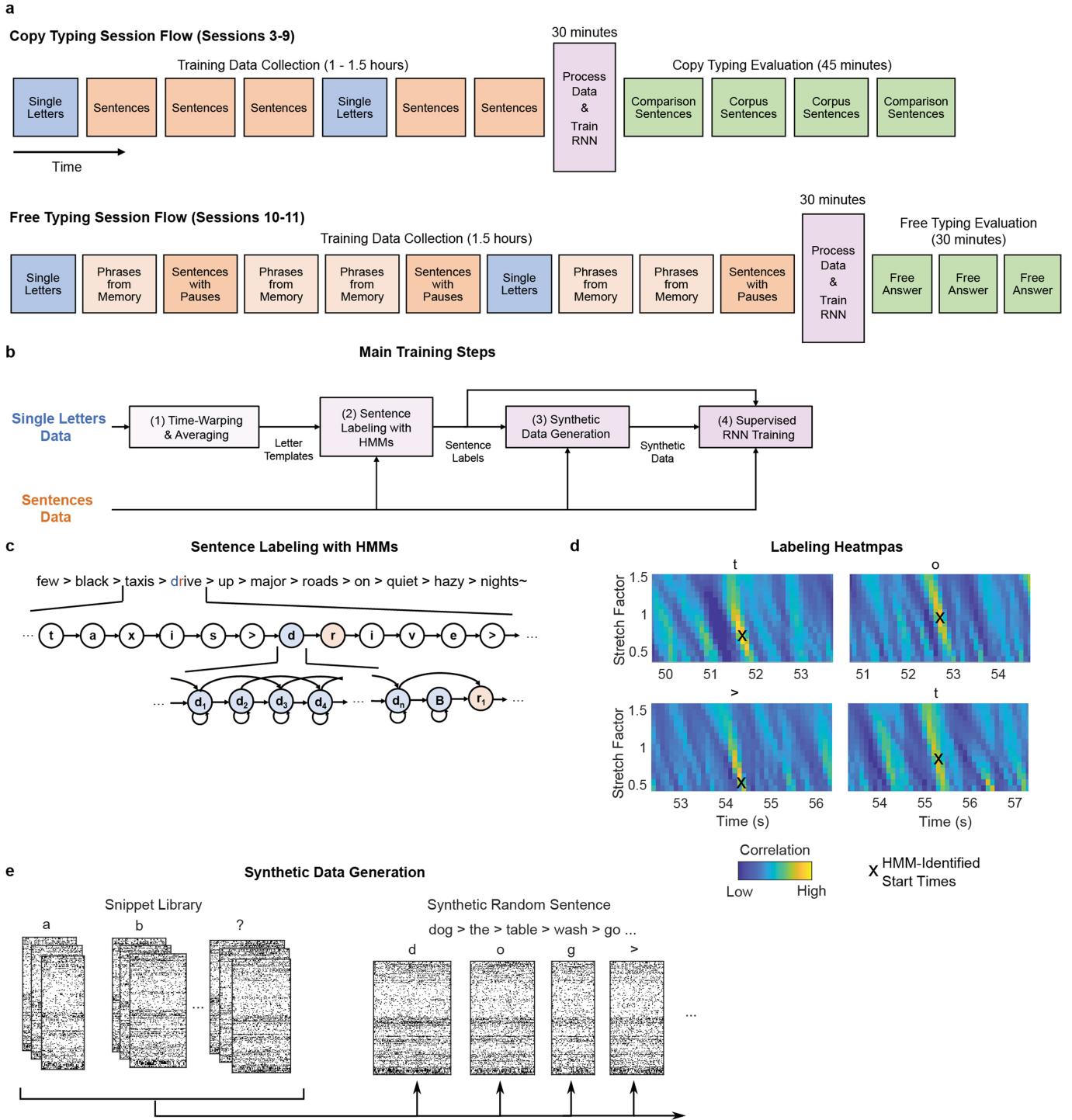
**Peer review information** *Nature* thanks Karim Oweiss and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

# Article

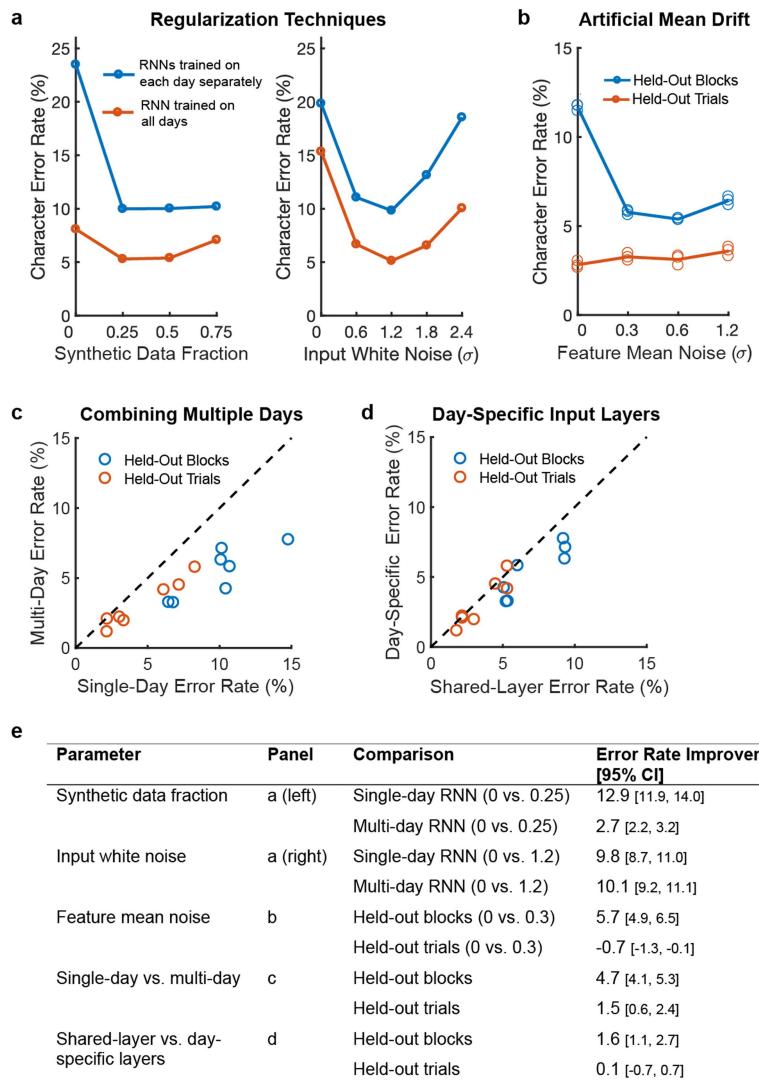


**Extended Data Fig. 1 | Diagram of the RNN architecture.** We used a two-layer gated recurrent unit (GRU) recurrent neural network architecture to convert sequences of neural **firing rate** vectors  $x_t$  (which were temporally smoothed and binned at 20 ms) into sequences of character probability vectors  $y_t$  and ‘new character’ probability scalars  $z_t$ . The  $y_t$  vectors describe the probability of each character being written at that moment in time, and the  $z_t$  scalars go high whenever the RNN detects that T5 is beginning to write any new character. Note that the **top RNN layer runs at a slower frequency than the bottom layer**, which we found improved the speed of training by making it easier to hold information in memory for long time periods. Thus, the RNN outputs are updated only once every 100 ms. Also, note that we used a **day-specific affine transform** to account for day-to-day changes in the neural activity (bottom row)—this helps the RNN to account for changes in neural tuning caused by electrode array micromotion or brain plasticity when training data are combined across multiple days.



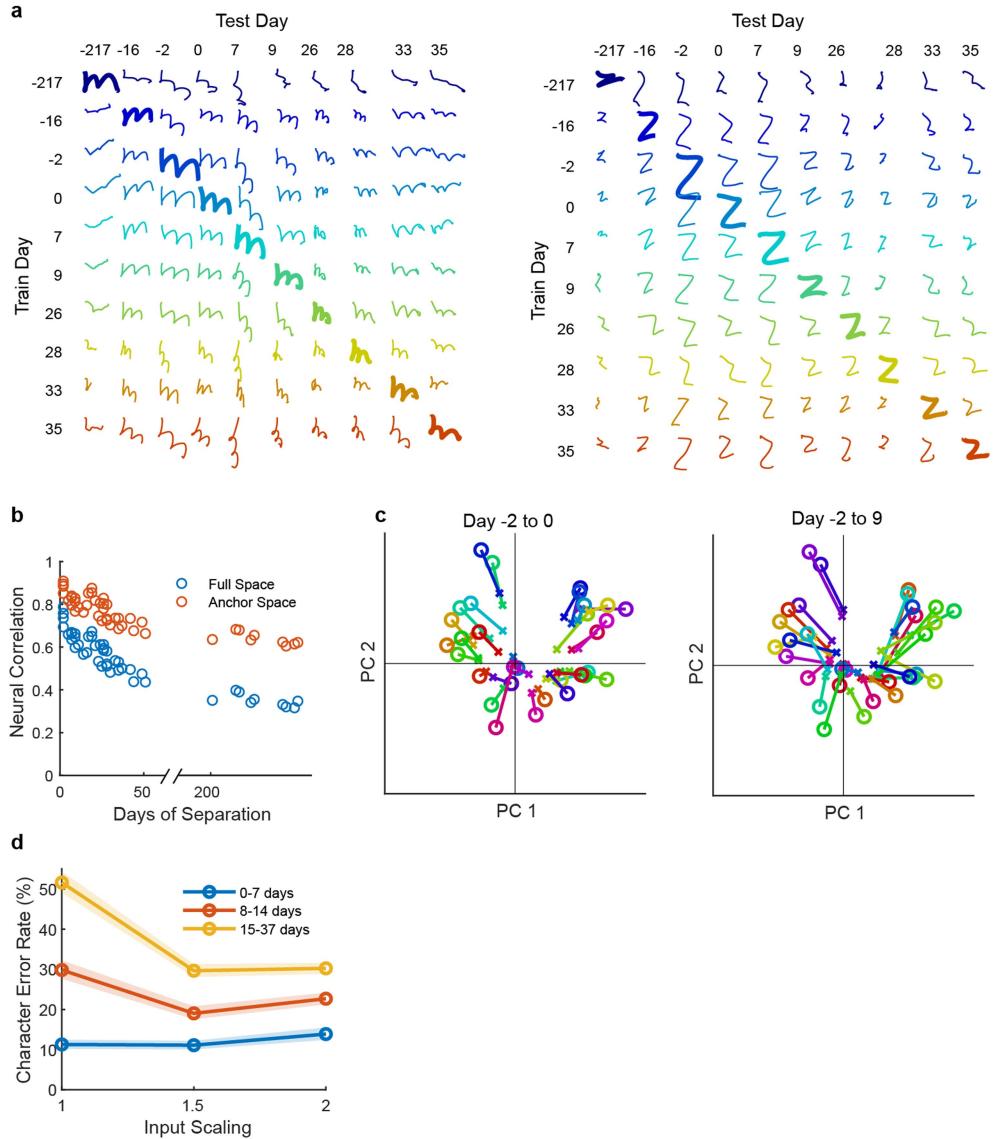
**Extended Data Fig. 2 | Overview of RNN training methods.** **a**, Diagram of the session flow for copy-typing and free-typing sessions (each rectangle corresponds to one block of data). First, single-letter and sentences training data are collected (blue and red blocks). Next, the RNN is trained using the newly collected data plus all of the previous days' data (purple block). Finally, the RNN is held fixed and evaluated (green blocks). **b**, Diagram of the data processing and RNN training process (purple block in **a**). First, the single-letter data are time-warped and averaged to create spatiotemporal templates of neural activity for each character. These templates are used to initialize the hidden Markov models (HMMs) for sentence labelling. After labelling, the observed data are cut apart and rearranged into new sequences of characters to make synthetic sentences. Finally, the synthetic sentences are combined with the real sentences to train the RNN. **c**, Diagram of a forced-alignment HMM

used to label the sentence ‘few black taxis drive up major roads on quiet hazy nights’. The HMM states correspond to the sequence of characters in the sentence. **d**, The label quality can be verified with cross-correlation heat maps made by correlating the single character neural templates with the real data. The HMM-identified character start times form clear hotspots on the heat maps. Note that these heat maps are depicted only to qualitatively show label quality and aren't used for training (only the character start times are needed to generate the targets for RNN training). **e**, To generate new synthetic sentences, the neural data corresponding to each labelled character in the real data are cut out of the data stream and put into a snippet library. These snippets are then pulled from the library at random, stretched or compressed in time by up to 30% (to add more artificial timing variability) and combined into new sentences.

**Extended Data Fig. 3 | The effect of key RNN parameters on performance.**

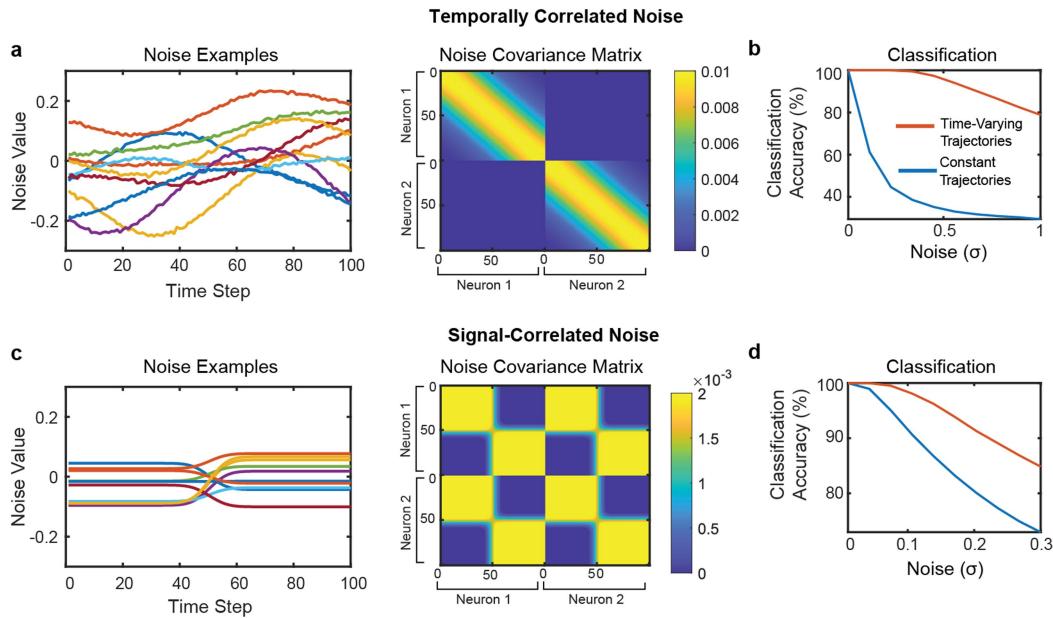
**a**, Training with synthetic data (left) and artificial white noise added to the inputs (right) were both essential for high performance. Data are shown from a grid search over both parameters, and lines show performance at the best value for the other parameter. Results indicate that both parameters are needed for high performance, even when the other is at the best value. Using synthetic data is more important when the size of the dataset is highly limited, as is the case when training on only a single day of data (blue line). Note that the inputs given to the RNN were  $z$ -scored, so the input white noise is in units of standard deviations of the input features. **b**, Artificial noise added to the feature means (random offsets and slow changes in the baseline firing rate) greatly improves the ability of the RNN to generalize to new blocks of data that occur later in the

session, but does not help the RNN to generalize to new trials within blocks of data on which it was already trained. This is because feature means change slowly over time. For each parameter setting, three separate RNNs were trained (circles); results show low variability across RNN training runs. **c**, Training an RNN with all of the datasets combined improves performance relative to training on each day separately. Each circle shows the performance on one of seven days. **d**, Using separate input layers for each day is better than using a single layer across all days. **e**, Improvements in character error rates are summarized for each parameter. 95% CIs were computed with bootstrap resampling of single trials ( $n=10,000$ ). As shown in the table, all parameters show a statistically significant improvement for at least one condition (CIs do not intersect zero).



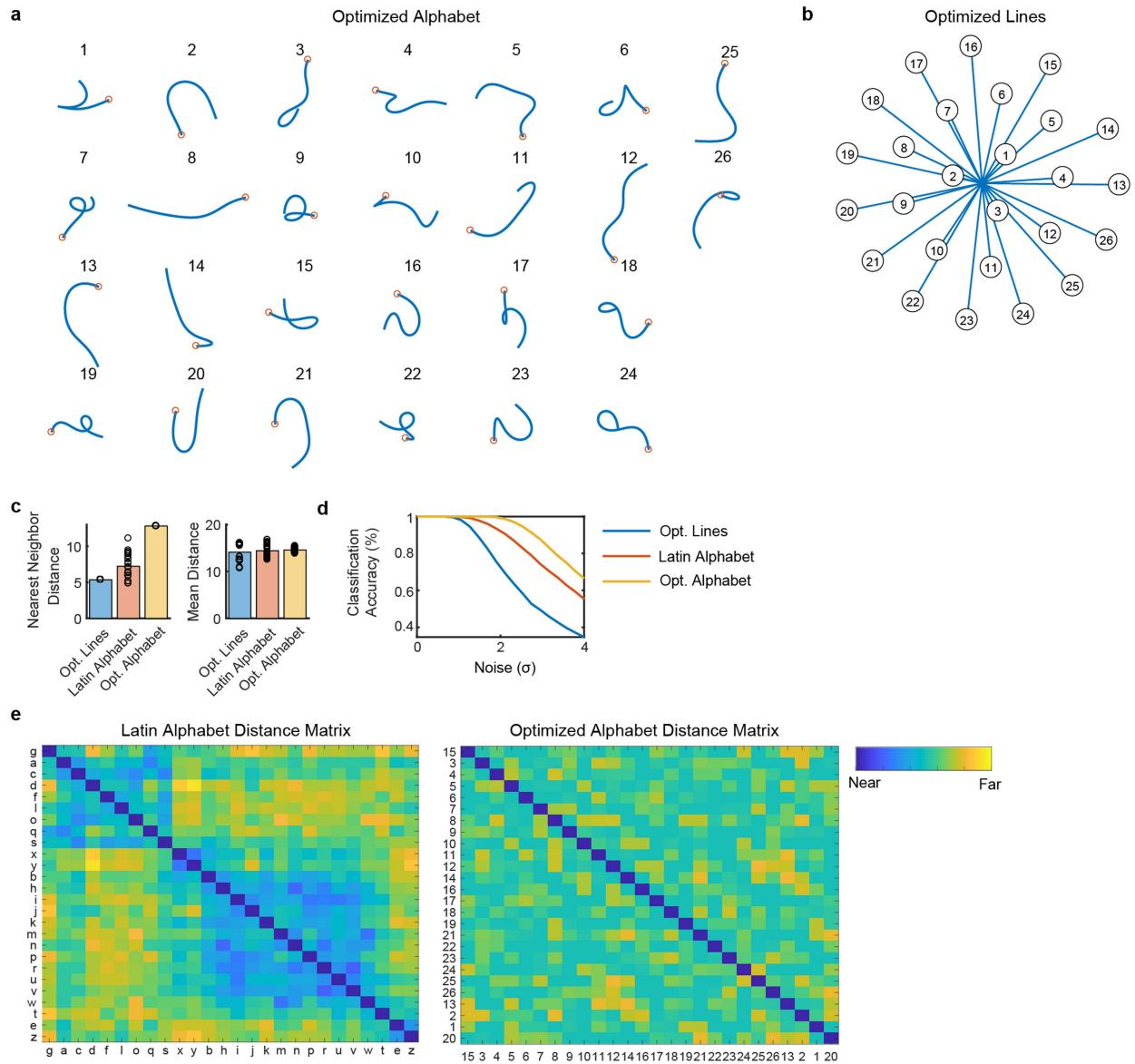
**Extended Data Fig. 4 | Changes in neural recordings across days.** **a**, To visualize how much the neural recordings changed across time, decoded pen-tip trajectories were plotted for two example letters (m and z) for all 10 days of data (columns), using decoders trained on all other days (rows). Each session is labelled according to the number of days passed relative to 9 December 2019 (day 4). Results show that although patterns of neural activity clearly change over time, their essential structure is largely conserved (as decoders trained on past days transfer readily to future days). **b**, The correlation (Pearson's  $r$ ) between the neural activity patterns of each session was computed for each pair of sessions and plotted as a function of the number of days separating each pair. Blue circles show the correlation computed in the full neural space (all 192 electrodes), whereas red circles show the correlation in the ‘anchor’ space (top 10 principal components of the earlier session). High values indicate a high similarity in how characters are neurally encoded across days. The fact that correlations are higher in the anchor space suggests that the structure of the neural patterns stays largely the same as it slowly rotates into a new space, causing shrinkage in the original space but little change in structure. **c**, A visualization of how each character’s neural representation changes over time, as viewed through the top two PCs of the original ‘anchor’ space. Each circle represents the neural activity pattern for a single character,

and each x symbol shows that same character on a later day (lines connect matching characters). Left, a pair of sessions with only two days between them (day -2 to 0); right, a pair of sessions with 11 days between them (day -2 to 9). The relative positioning of the neural patterns remains similar across days, but most conditions shrink noticeably towards the origin. This is consistent with the neural representations slowly rotating out of the original space into a new space, and suggests that scaling-up the input features may help a decoder to transfer more accurately to a future session (by counteracting this shrinkage effect). **d**, Similar to Fig. 3b, copy-typing data from eight sessions were used to assess offline whether scaling-up the decoder inputs improves performance when evaluating the decoder on a future session (when no decoder retraining is used). All session pairs ( $X, Y$ ) were considered. Decoders were first initialized using all data from session X and earlier, then evaluated on session Y under different input-scaling factors (for example, an input scale of 1.5 means that input features were scaled up by 50%). Lines indicate the mean raw character error rate and shaded regions show 95% CIs. Results show that when long periods of time pass between sessions, input scaling improves performance. We therefore used an input-scaling factor of 1.5 when assessing decoder performance in the ‘no retraining’ conditions of Fig. 3.



**Extended Data Fig. 5 | Effect of correlated noise on the toy model of temporal dimensionality.** **a**, Example noise vectors and covariance matrix for temporally correlated noise. On the left, example noise vectors are plotted (each line depicts a single example). Noise vectors are shown for all 100 time steps of neuron 1. On the right, the covariance matrix used to generate temporally correlated noise is plotted (dimensions =  $200 \times 200$ ). The first 100 time steps describe the noise of neuron 1 and the last 100 time steps describe the noise of neuron 2. The diagonal band creates noise that is temporally correlated within each simulated neuron (but the two neurons are uncorrelated with each other). **b**, Classification accuracy when using a maximum likelihood classifier to classify between all four possible trajectories

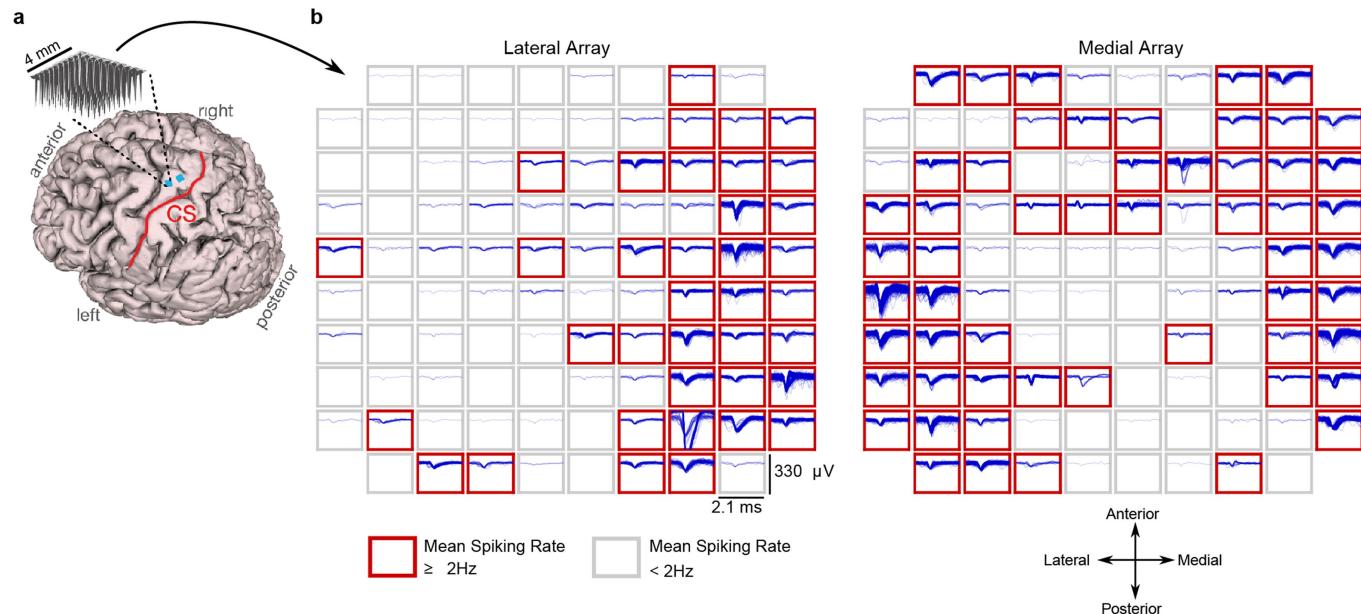
in the presence of temporally correlated noise. Even in the presence of temporally correlated noise, the time-varying trajectories are still much easier to classify. **c**, Example noise vectors and noise covariance matrix for noise that is correlated with the signal (that is, noise that is concentrated only in spatiotemporal dimensions that span the class means). Unlike the temporally correlated noise, this covariance matrix generates spatiotemporal noise that has correlations between time steps and neurons. **d**, Classification accuracy in the presence of signal-correlated noise. Again, time-varying trajectories are easier to classify than constant trajectories. See Supplementary Note 1 for a detailed interpretation of this figure.



**Extended Data Fig. 6 | An artificial alphabet optimized to maximize neural decodability.** **a**, Using the principle of maximizing the nearest-neighbour distance, we optimized a set of pen trajectories that are theoretically easier to classify than the Latin alphabet (using standard assumptions of linear neural tuning to pen-tip velocity). **b**, For comparison, we also optimized a set of 26 straight lines that maximize the nearest-neighbour distance. **c**, Pairwise Euclidean distances between pen-tip trajectories were computed for each set, revealing a larger nearest-neighbour distance (but not mean distance) for the optimized alphabet compared to the Latin alphabet. Each circle represents a single movement and bar heights show the mean. **d**, Simulated classification

accuracy as a function of the amount of artificial noise added. Results confirm that the optimized alphabet is indeed easier to classify than the Latin alphabet, and that the Latin alphabet is much easier to classify than straight lines, even when the lines have been optimized. **e**, Distance matrices for the Latin alphabet and optimized alphabets show the pairwise Euclidean distances between the pen trajectories. The distance matrices were sorted into seven clusters using single-linkage hierarchical clustering. The distance matrix for the optimized alphabet has no apparent structure; by contrast, the Latin alphabet has two large clusters of similar letters (letters that begin with a counter-clockwise curl, and letters that begin with a downstroke).

# Article



**Extended Data Fig. 7 | Example spiking activity recorded from each microelectrode array.** **a**, Magnetic resonance imaging (MRI)-derived brain anatomy of participant T5. Microelectrode array locations (blue squares) were determined by co-registration of postoperative CT images with preoperative MRI images. **b**, Example spike waveforms detected during a 10-s time window are plotted for each electrode (data were recorded on post-implant day 1,218). Each rectangular panel corresponds to a single electrode and each blue trace is a single spike waveform (2.1-ms duration). Spiking events were detected using a  $-4.5$  root mean square (RMS) threshold, thereby excluding almost all

background activity. Electrodes with a mean threshold crossing rate of at least 2 Hz were considered to have 'spiking activity' and are outlined in red (note that this is a conservative estimate that is meant to include only spiking activity that could be from single neurons, as opposed to multiunit 'hash'). The results show that many electrodes still record large spiking waveforms that are well above the noise floor (the y-axis of each panel spans 330  $\mu$ V, whereas the background activity has an average RMS value of only 6.4  $\mu$ V). On this day, 92 electrodes out of 192 had a threshold crossing rate of at least 2 Hz.

**Extended Data Table 1 | Example decoded sentences for one block of copy typing**

Prompt	Real-Time Output	Real-Time Output (Annotated)	Offline Language Model (Annotated)
infected adults develop a cough and their skin and ears turn blue.	infected adults dercep a cough and thir skin and ears tunn blue	infected adults dercep a cough and thir skin and ears tunn blue.	infected adults develop a cough and their skin and ears turn blue.
i interrupted, unable to keep silent.	i interrupted, unabee to keep silent.	i interrupted, unabee to keep silent.	i interrupted, unable to keep silent.
i dumped the tools in the hut.	i dumped the tools in the hut.	i dumped the tools in the hut.	i dumped the tools in the hut.
within thirty seconds the army had landed.	within thirty seconds the armmy had landed	within thirty seconds the armmy had landed.	within thirty seconds the army had landed.
that's when i threw up on the carpet.	that's when i threw up on the carpet.	that's when i threw up on the carpet.	that's when i threw up on the carpet.
he didn't want to rub salt into her wounds.	he didn't want to rub salt into her wounds.	he didn't want to rub salt into her wounds.	he didn't want to rub salt into her wounds.
shouting and swearing, i yelled for an epidural.	shouting and swearin, I yeled for an epidural.	shouting and swearing, i yelled for ■ an epidural.	shouting and swearing, i yelled for a an epidermal.
you wish to purchase something?	you wish to purchase something?	you wish to purchase somathing?	you wish to purchase something?
lowell felt like a soldier on a battlefield, stripped of ammunition.	lowel felt like a soldier on a battlefield, stripped ef ammunition.	lowell felt like a soldier on a battlefield, stripped ef ammunition.	lowell felt like a soldier on a battlefield, stripped of ammunition.
there are only one or two minor casualties.	thee ane only on or two minor cafaulties.	there ane only one or two minor cafaulties.	there are only one or two minor casualties.

In the rightmost columns, errors are highlighted in red (extra spaces are denoted with a red square, and omitted letters are indicated with a strikethrough). Note that our language model substitutes 'epidermal' for 'epidural', because 'epidural' is out of vocabulary. The mean typing rate for this block was 86.47 characters per minute and the character error rates were 4.18% (real-time output) and 1.22% (language model). Sentence prompts were taken from the British National Corpus according to a random selection process (see Supplementary Methods for details).

# Article

**Extended Data Table 2 | Example decoded sentences for one block of free typing**

Prompt	Real-Time Output	Real-Time Output (Annotated)	Offline Language Model (Annotated)
what made you first get interested in machining?	iit was the famil business	iiit was the family business	it was the family business
what is the hardest part of machining?	forming the tooling	forming the tooling	forming the tooling
how much spice do you like in your food?	lots and lots	lots and lots	lots and lots
what type of music do you enjoy most?	ii like loud rock musica	ii like loud rock musica	i like loud rock music
what are some of your favorite games to play?	i like to play cards	i like to play cards	i like to play cards
what has taken you the longest to get good or decent at?	i worked for years to perfecet my photography	i worked for years to perfecet my photography	i worked for years to perfect my photography
what food do you love that a lot of people might find a little odd?	very few poope can eat smeked sailfish	very few poople can eat smeked sailfish	very few people can eat smoked salt fish
if you could start a charity, what would it be for?	i woud proovide furdng to smal families	i woud proovide furdng to small families	i would provide funding to small families
what advice would you give to your younger self?	be patientt it wil got better	be patientt it will got better	be patient, it will get better

In the rightmost columns, errors are highlighted in red (omitted letters are indicated with a strikethrough). Note that our language model substitutes ‘salt fish’ for ‘sailfish’, because ‘sailfish’ is out of vocabulary. The mean typing rate for this block was 73.8 characters per minute and the character error rates were 6.82% (real-time output) and 1.14% (language model).

## Reporting Summary

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

### Statistics

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

n/a Confirmed

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

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection	The software for running the experimental tasks, recording data, and implementing the real-time decoding system was custom developed using MATLAB and Simulink Real-Time (MathWorks, Natick, MA).
Data analysis	Data was analyzed using custom MATLAB and python code. Custom code will be made publicly available on GitHub after acceptance.

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

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

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

Data will be made publicly available on Dryad upon acceptance. Relevant links and identifiers will be included in the final manuscript.

# Field-specific reporting

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

- Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

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

Sample size	No sample-size calculation was performed. Data were collected in a single participant to characterize the performance of a brain-computer interface. Uncertainty in performance estimates were quantified with confidence intervals, and show a robust result.
Data exclusions	This study is based on brain-computer interface performance evaluation data collected over a series of days. All days are reported in the study and all relevant data is included.
Replication	This study assessed brain-computer interface performance in a single participant. Results were replicated across multiple days of performance evaluation.
Randomization	Randomization into groups is not relevant for this study - only one participant was included in this study.
Blinding	Blinding is not relevant to this study, which assessed the performance of a brain-computer interface in a single individual.

## Reporting for specific materials, systems and methods

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

### Materials & experimental systems

- |                                     |   |
|-------------------------------------|---|
| n/a                                 | Involved in the study   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies                             |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines                  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology          |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms            |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Human research participants |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                          |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern           |

### Methods

- |                                     |   |
|-------------------------------------|---|
| n/a                                 | Involved in the study                           |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

## Human research participants

Policy information about [studies involving human research participants](#)

### Population characteristics

This study includes data from one participant (identified as T5) who gave informed consent and was enrolled in the BrainGate2 Neural Interface System clinical trial (ClinicalTrials.gov Identifier: NCT00912041, registered June 3, 2009) - but note that this study does not report clinical trial results. T5 is a right-handed man, 65 years old at the time of data collection, with a C4 AIS C spinal cord injury that occurred approximately 9 years prior to study enrollment.

### Recruitment

Participant T5 was enrolled in the BrainGate2 pilot clinical trial prior to the design and execution of this study, after meeting inclusion criteria based in part on disease characteristics. Inclusion and exclusion criteria are available online (ClinicalTrials.gov).

### Ethics oversight

The BrainGate2 Neural Interface System clinical trial was approved under an Investigational Device Exemption (IDE) by the US Food and Drug Administration (Investigational Device Exemption #G090003). Permission was also granted by the Institutional Review Boards of Stanford University (protocol #20804). All research was performed in accordance with relevant guidelines/regulations.

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