Learning along the manifold of human brain activity via real-time neurofeedback

Erica L. Busch (erica.busch@yale.edu)
Department of Psychology, Yale University
100 College Street, New Haven, CT 06511 USA

E. Chandra Fincke (chandra.fincke@yale.edu)
Department of Psychology, Yale University
100 College Street, New Haven, CT 06511 USA

Guillaume Lajoie (guillaume.lajoie@mila.quebec)
Mila – Québec AI Institute
Mathematics and Statistics Department, Université de Montréal
2920 chemin de la Tour, Montréal, QC, H3T 1J4 Canada

Smita Krishnaswamy (smita.krishnaswamy@yale.edu)
Departments of Genetics, Computer Science, & Applied Mathematics
Wu Tsai Institute, Yale University
100 College Street, New Haven, CT 06511 USA

Nicholas B. Turk-Browne (nicholas.turk-browne@yale.edu)
Department of Psychology & Wu Tsai Institute, Yale University
100 College Street, New Haven, CT 06511 USA

Abstract

Learning to perform a new behavior is constrained by the geometry, or intrinsic manifold, of the neural population activity supporting that behavior. Recent work highlights the importance of manifolds capturing low-dimensional neural dynamics for learning to control brain-computer interfaces (BCIs). In non-human primate studies, BCI learning has been expedited and stabilized by mapping neural recordings from motor cortex through a low-dimensional manifold and then to a feedback display. In macaque motor cortex, the manifold uncovers more concise and plastic neural signals. Here, we investigate the manifold constraints on human learning in brain regions associated with higher-order cognitive processes using a non-invasive BCI. Using a custom neural manifold learning framework for real-time fMRI neurofeedback and a virtual reality stimulus, we trained participants in a multi-session study to perform a navigation task using their brain activity. Task performance was significantly improved by feedback based on the brain's intrinsic relative to lower-ranked ("off") manifold activity. Neural activity was modulated along the manifold over the course of neurofeedback training, such that neural activity became better aligned with the components of the manifold determining the feedback as performance improved.

Keywords: Brain-computer interface; learning; manifold learning; closed-loop

Introduction

Learning new behaviors requires neural populations and networks to learn to generate new activity patterns. Some activity patterns are more challenging to learn to generate than others, thus constraining the behaviors an organism can learn to produce. BCI studies using non-human primates have shown that the activity a neural population can learn to generate is computationally constrained by its intrinsic connectivity and co-modulation (Sadtler et al., 2014; Golub et al., 2018; Oby et al., 2019), which creates a low-dimensional "neural manifold." These studies have shown that, within motor cortex, activity sharing low-dimensional structure with what the population already represents ("on-manifold" activity) is easier to learn to modulate (Sadtler et al., 2014), but neural activity farther "off-manifold" is harder or impossible to learn (Oby et al., 2019).

Can neural manifolds of higher-order cognitive regions be harnessed to enhance human learning? Neurofeedback is a form of BCI where neural activity is measured in real-time and presented to participants (e.g., visually on a screen) to facilitate self-regulation of a specific neural state (Sitaram et al., 2017). Neurofeedback training with fMRI benefits from fMRI's high spatial resolution and has allowed researchers to probe the plasticity of neural substrates underlying cognitive processes including visual perception (Shibata, Watanabe, Sasaki, & Kawato, 2011), attention (deBettencourt, Cohen, Lee, Norman, & Turk-Browne, 2015), and memory (deBettencourt, Turk-Browne, & Norman, 2019). Despite its potential for answering ages-old questions in cognitive neuroscience, neurofeedback studies have had limited impact thus far, due to the extensive training they require (5–10 fMRI sessions) and the variable outcomes they have yielded.

Prior studies have not considered how the computational principles underlying neural activity (i.e., neural manifolds) may relate to learning in the human brain. As such, neurofeedback training agnostic to a participant's neural manifold may actually encourage brain states that are computationally inefficient or impossible to generate, hindering the generation of novel behaviors. Here, we trained humans to perform a virtual reality (VR) navigation task with their brains using real-time fMRI neurofeedback. Participants successfully learned to modulate activity in a network of navigation-related regions to self-directed navigation within one neurofeedback training session when administered feedback on-manifold but not off-manifold. Participants adjusted and re-learned BCI mappings still within the manifold, but not off the manifold. In sum, non-invasive BCI learning in humans shows low-dimensional constraints similar to non-human primate studies, which extends beyond motor learning to higher-order cognition.

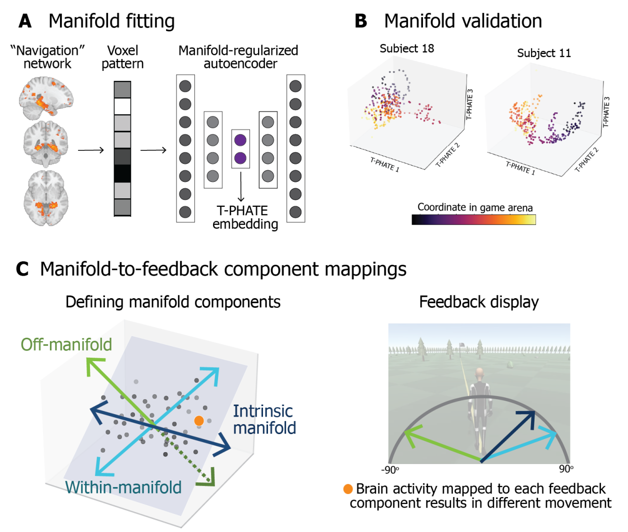

Figure 1: (A) Calibration data were preprocessed and masked. Then, they were used to fit a 20-D T-PHATE manifol embedding and an autoencoder used to extend the manifold. (B) Visualizations of activation during the task, where a point is a brain volume in T-PHATE space colored by the avatar's location in the game arena at that timepoint. (C) Defining intrinsic, within, and off-manifold mappings such that the same brain activity in T-PHATE space (orange point) could result in 3 different angles.

Materials and methods

Experimental overview We enrolled 20 participants (9 female; 25.8 ± 5.5 years) in a 4–5 session real-time fMRI experiment. In each session, participants completed 4 runs (approx. 10 min/run) of a game where they had to direct an avatar through a virtual environment to a target. Runs contained trials of approximately 20s (dependent upon performance) followed by 6s rest. In session 1, participants played the game using a MR-compatible joystick in the scanner. Using these data, we estimated a neural manifold of this task from a network of regions (defined with the search term "navigation" on neurosynth.org). Manifold learning used an algorithm optimized for the dimensionality, noise, and dynamics of fMRI data (Busch et al., 2023) (Fig 1A,B). In subsequent sessions, we used *rt-cloud* (Wallace et al., 2022) and a custom procedure to map brain activity to the avatar's turning angle in real-time (Fig. 2A) (Busch et al., 2022; Huang et al., 2022).

Neurofeedback training fMRI volumes were acquired every 2s and transmitted from the scanner to a remote high-performance computing cluster for processing and embedding into the pre-fit manifold (Fig. 2A) (Wallace et al., 2022). Embedded data were then used to decode the avatar's next turning angle via one of three manifold components (i.e., intrinsic-, within-, and off-manifold [IM, WM, OM]), which captured the most, second most, and least variance in the manifold.

Participants were instructed to navigate the avatar to the target and were rewarded for minimizing excess distance traveled within each trial. Neurofeedback training used a staircasing procedure to adapt the control a participant's brain state exerted over the avatar's movement. The feedback signal was a weighted combination of the decoded angle and an adaptive parameter (*Brain Control*). After each trial, *Brain Control* increased, decreased, or remained the same depending upon performance. In the first feedback session (session 2), all participants received the IM mapping, and all participants' *Brain Control* staircasing began at 20%. For each of the final two sessions, participants received either the WM or OM component (e.g., session 3 = WM and session 4 = OM, or vice-versa), with session order counterbalanced across participants. This procedure allows us to consider each subject as their own control.

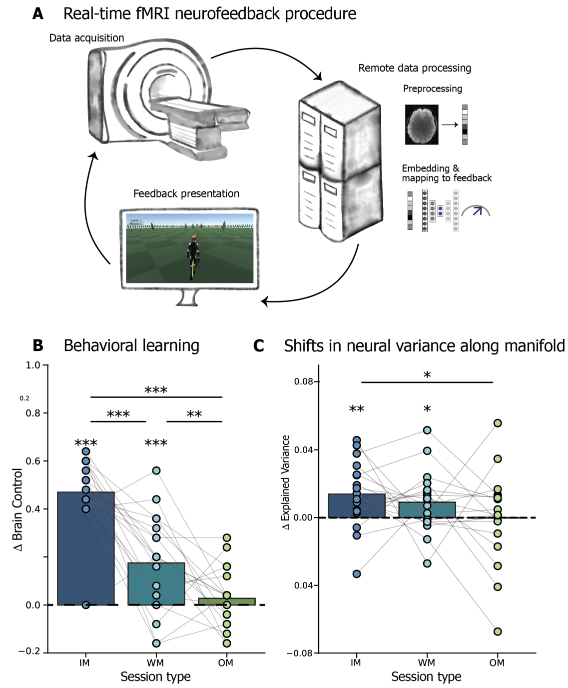

Figure 2: (A) The closed-loop procedure collects whole-brain data and transmits them to a remote HPC cluster for motion-correction, masking, normalization, and analysis, returning a new angle to the display computer within 2s. (B) Behavioral learning was quantified by subtracting the ending *Brain Control* from the starting *Brain Control* during each session. (C) Shifts in neural variance were computed as the explained variance of a manifold component at the start vs. the end of neurofeedback training, within a session. $*** = p < 0.001$, $** = p < 0.01$, $* = p < 0.05$, $\sim = p < 0.1$

Results

First, we validated that the T-PHATE embeddings could capture task structure represented in the brain. Visually, embeddings of brain activity during this task reflect the location in the game arena (shown for two sample subjects in Fig. 1B). This lends confidence that our region of interest is implicated in this task and the T-PHATE manifold highlights this.

As *Brain Control* is scaled linearly with task performance, higher *Brain Control* serves as a behavioral metric of learning to control the BCI. We quantified behavioral learning as the change in *Brain Control* across all trials in a session (Fig. 2B). Learning was significant in IM sessions, significant in WM sessions, but not in OM sessions.

Behavioral learning effects are hypothesized to be supported by neural changes, specifically an increase in the alignment of neural activity to the component of the manifold mapped to the feedback display (i.e., the one being trained). We calculated alignment of new neural activity with the manifold components as the change in the neural variance explained by the component at the start vs. end of a neurofeedback session. During IM and WM sessions, the proportion of explained variance increased significantly from the start to end of the session. The IM session showed a greater increase in variance explained than the OM session, which did not significantly change in the variance it explained (Fig. 2C).

In sum, we introduced a framework for administering neurofeedback that respects that computational constraints on learning in the brain, and we do so in a way optimized for the challenges of non-invasive human neuroimaging. By processing brain activity as it lies along a low-dimensional manifold, behavioral metrics of learning in our experiment were stronger, more robust across brains, and faster relative to other neurofeedback experiments. Further, we show that manifold constraints on learning exist not just in motor cortex, as investigated in the non-human primate literature (Sadtler et al., 2014; Hennig et al., 2018; Golub et al., 2018; Oby et al., 2019), but also in brain regions related to higher-order cognition. Our results suggest important applications for brain-based cognitive training and neuroprothetics.

Acknowledgments

E.L.B was supported by a NSF GRFP. N.B.T-B. was supported by CIFAR. G.L. was supported by a Canada CIFAR AI Chair and a Canada Research Chair in Neural Computation and Interfacing.

References

Busch, E. L., Huang, J., Benz, A., Wallenstein, T., Lajoie, G., Wolf, G., ... Turk-Browne, N. B. (2023, March). Multi-view manifold learning of human brain-state trajectories. *Nature Computational Science*, 1–14. (Publisher: Nature Publishing Group) doi: 10.1038/s43588-023-00419-0

Busch, E. L., Letrou, A., Huang, J., Wolf, G., Lajoie, G., Krishnaswamy, S., & Turk-Browne, N. B. (2022). *A neural manifold learning framework for real-time fmri neurofeedback.* Real-Time Functional Imaging and Neurofeedback meeting.

deBettencourt, M. T., Cohen, J. D., Lee, R. F., Norman, K. A., & Turk-Browne, N. B. (2015, March). Closed-loop training of attention with real-time brain imaging. *Nature Neuroscience*, *18*(3), 470–475. doi: 10.1038/nn.3940

deBettencourt, M. T., Turk-Browne, N. B., & Norman, K. A. (2019, October). Neurofeedback helps to reveal a relationship between context reinstatement and memory retrieval. *NeuroImage*, *200*, 292–301. doi: 10.1016/j.neuroimage.2019.06.001

Golub, M. D., Sadtler, P. T., Oby, E. R., Quick, K. M., Ryu, S. I., Tyler-Kabara, E. C., ... Yu, B. M. (2018, April). Learning by neural reassociation. *Nature Neuroscience*, *21*(4), 607–616. (Number: 4 Publisher: Nature Publishing Group) doi: 10.1038/s41593-018-0095-3

Hennig, J. A., Golub, M. D., Lund, P. J., Sadtler, P. T., Oby, E. R., Quick, K. M., ... Chase, S. M. (2018, August). Constraints on neural redundancy. *eLife*, *7*, e36774. (Publisher: eLife Sciences Publications, Ltd) doi: 10.7554/eLife.36774

Huang, J., Busch, E., Wallenstein, T., Gerasimiuk, M., Benz, A., Lajoie, G., ... Krishnaswamy, S. (2022, August). Learning Shared Neural Manifolds from Multi-Subject FMRI Data. *2022 IEEE 32nd International Workshop on Machine Learning for Signal Processing (MLSP)*, 01–06. (Conference Name: 2022 IEEE 32nd International Workshop on Machine Learning for Signal Processing (MLSP)) doi: 10.1109/MLSP55214.2022.9943383

Oby, E. R., Golub, M. D., Hennig, J. A., Degenhart, A. D., Tyler-Kabara, E. C., Yu, B. M., ... Batista, A. P. (2019, July). New neural activity patterns emerge with long-term learning. *Proceedings of the National Academy of Sciences*, *116*(30), 15210–15215. doi: 10.1073/pnas.1820296116

Sadtler, P. T., Quick, K. M., Golub, M. D., Chase, S. M., Ryu, S. I., Tyler-Kabara, E. C., ... Batista, A. P. (2014, August). Neural constraints on learning. *Nature*, *512*(7515), 423–426. (Number: 7515 Publisher: Nature Publishing Group) doi: 10.1038/nature13665

Shibata, K., Watanabe, T., Sasaki, Y., & Kawato, M. (2011, December). Perceptual Learning Incepted by Decoded fMRI Neurofeedback Without Stimulus Presentation. *Science*, *334*(6061), 1413–1415. doi: 10.1126/science.1212003

Sitaram, R., Ros, T., Stoeckel, L., Haller, S., Scharnowski, F., Lewis-Peacock, J., ... Sulzer, J. (2017, February). Closed-loop brain training: the science of neurofeedback. *Nature Reviews Neuroscience*, *18*(2), 86–100. doi: 10.1038/nrn.2016.164

Wallace, G., Polcyn, S., Brooks, P. P., Mennen, A. C., Zhao, K., Scotti, P. S., ... Norman, K. A. (2022, August). RT-Cloud: A cloud-based software framework to simplify and standardize real-time fMRI. *NeuroImage*, *257*, 119295. doi: 10.1016/j.neuroimage.2022.119295