

Whole-Brain Breath-State real-time fMRI Neurofeedback Links Decoder-Guided Regulation to Salience-Network

Jihun Kim ¹, Jae-eon Kang ¹, Hyojun Choi ², Yuri Kim ², Juyoung Kim ², Hackjin Kim ², Jong-Hwan Lee ¹

¹ *Dept. of Brain and Cognitive Engineering, Korea University, Seoul, Korea*

² *School of Psychology, Korea University, Seoul, Korea*

Introduction

Interoception depends strongly on the anterior insular cortex (AIC) and is implicated in psychiatric disorders (Craig, 2002; Khalsa et al., 2018). Breathing is trainable because the AIC and anterior cingulate cortex (ACC) encode anxiety-modulated respiratory predictions (Harrison et al., 2021). Conventional real-time fMRI neurofeedback prespecifies an insular ROI and may miss distributed breath-related brain-state signatures (Sitaram et al., 2017). We therefore fine-tuned a participant-specific whole-brain decoder to feed back breath-attentive-state probability and tested learning, specificity, necessity, and network change.

Methods

Primary analyses focused on seven protocol-complete EXP-SHAM yoked pairs who completed Day 5. Across five days, participants completed two shuffled 15-min runs of Wisconsin Card Sorting (WCST), Emotion Recognition (ERT), 3-back (NBT), BreathCounting (BC), NumberCounting (NC), and Rest; structural MRI and decoder fine-tuning; three 6-min NF runs; and a 5-min feedback-free transfer run. Seven-minute resting scans were acquired on Day 1 and 5. EXP viewed their own session-adapted decoded BC probability; SHAM viewed yoked non-contingent feedback (Fig. 1A). A GPT-style decoder with four multi-head-attention blocks mapped 10-TR DiFuMo-1024 time series (Dadi et al., 2020) to six probabilities. The deployment decoder was session-adapted before NF and fixed during each NF run.

Signed gradient $\times$ input attribution was summed over 59 interoceptive modes for the BC-vs-rest margin. We evaluated model reliance by zero-ablation of anatomical mode sets and compared BC-margin loss with size-matched nulls. Significant modes were projected to MNI space for visualization. Exploratory RVT/HR-regressed resting-state FC changes (Day 5 minus Day 1) were summarized as EXP-minus-SHAM Fisher-z differences across Yeo's seven networks.

Results

Decoded BC probability rose in EXP but stayed near baseline in SHAM (Fig. 1B); EXP rose from Day 1 to Day 5 ($p = 0.017$) and exceeded SHAM on Day 5 ($p = 0.022$; paired Wilcoxon). Within the 59-mode interoceptive ROI, signed attribution was positive only for BC ($+0.052$; $p < 0.05$, $g = 1.45$) and negative for every other task, which it significantly exceeded (all FDR $q < 0.001$; $g = 0.92$ - 1.67); interoceptive features therefore supported the BC decision specifically, not overall signal magnitude (Fig. 1C). Within these interoceptive modes, 40 of 59 were individually significant ($1.73\times$ enrichment, $p < 0.05$), led by the posterior circular insular sulcus ($g = 1.44$). Fig. 1D showed that zero-ablation of thalamic and bilateral insular mode sets produced significantly larger BC-margin losses relative to size-matched nulls. Similar sensitivity patterns were observed across groups, suggesting that this reflected decoder structure rather than an EXP-specific mechanism. When projected onto MNI space, these modes were primarily localized to the insula, thalamus, and cingulate cortex (Fig. 1E). Exploratory RVT/HR-regressed resting-state Day5-Day1 Functional Connectivity (FC) showed consistent EXP>SHAM increases in mPFC-dIPFC and Salience-FPN summaries, although network-level correction was trend-level. (Fig. 1F).

Conclusions

Whole-brain real-time fMRI NF produced progressive regulation of a breath-attentive state under decoder-contingent feedback. Interpretability analyses indicated BC-specific model reliance on interoceptive-mode sets, with the strongest zero-ablation sensitivity in thalamic and insular modes, suggesting model reliance rather than regional brain causality. Exploratory resting-state FC changes provided convergent but non-confirmatory network-level support, although larger samples, behavioral validation, confirmatory connectivity statistics, and detailed assessment of breathing patterns and end-tidal CO₂ with synchronized physiological monitoring remain needed.

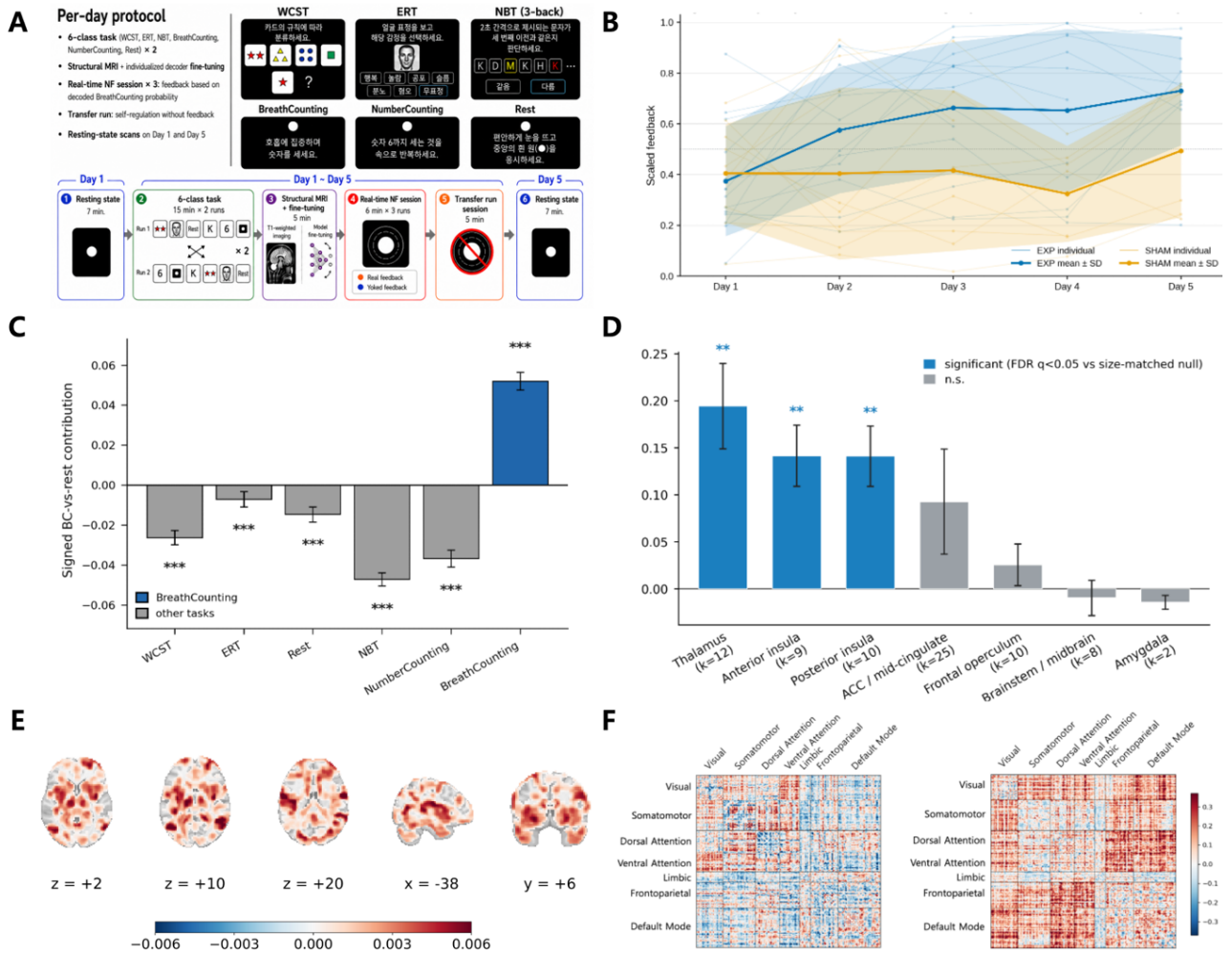


Figure 1. (A) Five-day whole-brain NF protocol; EXP received decoder-contingent BC probability and SHAM received yoked feedback. (B) BC-state learning by group. (C) Signed interoceptive contribution by task. (D) Zero-ablation sensitivity of anatomical mode sets relative to size-matched nulls; blue denotes FDR $q < 0.05$. (E) Significant DiFuMo effects on MNI, localized to insula, thalamus, and cingulate. (F) Exploratory Day5-Day1 resting-state FC changes; RVT/HR-regressed mPFC-dIPFC and Salience-FPN coupling favored EXP over SHAM.

References

1. Craig, Arthur D. "How do you feel? Interoception: the sense of the physiological condition of the body." *Nature reviews neuroscience* 3.8 (2002): 655-666.
2. Khalsa, Sahib S., et al. "Interoception and mental health: a roadmap." *Biological psychiatry: cognitive neuroscience and neuroimaging* 3.6 (2018): 501-513.
3. Harrison, Olivia K., et al. "Interoception of breathing and its relationship with anxiety." *Neuron* 109.24 (2021): 4080-4093.
4. Sitaram, Ranganatha, et al. "Closed-loop brain training: the science of neurofeedback." *Nature Reviews Neuroscience* 18.2 (2017): 86-100.
5. Dadi, Kamalaker, et al. "Fine-grain atlases of functional modes for fMRI analysis." *NeuroImage* 221 (2020): 117126.

Acknowledgements

This work was supported by the National Research Foundation (NRF) grant funded by the Korea government (MSIT) (No. RS-2023-00218987; RS-2025-00562405; RS-2026-25518987), and in part by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the MSIT (No. RS-2026-25507282) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. RS-2023-00241694).