BEHAVIOR-ALIGNED NEURAL SUPERVISION FOR VISION-LANGUAGE-ACTION MODELS

Anonymous authors
Paper under double-blind review

ABSTRACT

Vision-language-action (VLA) models learn embodied representations primarily from externally observable signals such as vision, language, and action trajectories. However, external observations may not fully capture behaviorally relevant structure and temporal dynamics during physical interaction. Biological organisms and embodied agents both operate through continuous perception–action loops under related physical constraints, motivating us to ask whether neural activity recorded during human interaction can provide complementary supervision for embodied learning. To answer this, we investigate EEG as a training-time supervisory signal for adapting pretrained embodied models. We first synchronously collect egocentric video and 128-channel EEG during natural manipulation tasks. We then introduce EEG-guided representation learning, with EEG used only during training and removed at inference. Using OpenVLA as the primary backbone, we evaluate the approach across multiple manipulation benchmarks and out-of-distribution settings involving visual, semantic, and execution shifts, and further validate its transfer to π_0 on LIBERO. EEG-guided training improves task success over the corresponding robot-only baselines across these settings and backbones. These gains depend on aligned EEG–behavior correspondence, as they are reduced when the alignment is disrupted or EEG supervision is ablated. Behavioral and representation analyses further associate the improvements with more effective task progression and completion, with the strongest representational changes occurring in late, action-proximal layers. These results suggest that behavior-aligned human neural dynamics can provide transferable training-time supervision for embodied representation learning without introducing an additional modality at deployment.

1 INTRODUCTION

Recent embodied models have evolved from task-specific policies toward more general systems that learn from vision, language, and large-scale interaction data. Vision-language-action models (VLAs), for example, combine semantic knowledge from vision-language pretraining with robot manipulation data to learn mappings from visual observations and language instructions to actions within a unified model. As model scale, training data, and cross-task experience grow, these systems have shown improved generalization across tasks, objects, and robot embodiments (Zitkovich et al., 2023; Open X-Embodiment Collaboration, 2024; Kim et al., 2024). However, the learning signals used by current embodied models still come primarily from externally observable information, including vision, language, robot state, and action trajectories. Physical interaction is inherently partially observable, and finite observations may not fully specify behaviorally relevant states or their temporal evolution (Kaelbling et al., 1998). Moreover, similar observations can correspond to different plausible actions or trajectories (Florence et al., 2022; Shafiullah et al., 2022; Chi et al., 2023). This raises a broader question: beyond collecting more visual and robot interaction data, can signals that evolve alongside behavior provide an additional source of supervision for embodied learning?

Human behavior emerges from continuous perception–action interactions among the nervous system, the body, and the environment. Processes such as movement preparation, sensorimotor transformation, and behavioral monitoring are accompanied by rapidly changing neural activity, which EEG can capture with high temporal resolution (Pfurtscheller & Lopes da Silva, 1999; Gehring et al., 1993). Although biological organisms and embodied agents differ in morphology and control mechanisms, both must continuously perceive, act, and adapt to feedback in the physical world under constraints

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

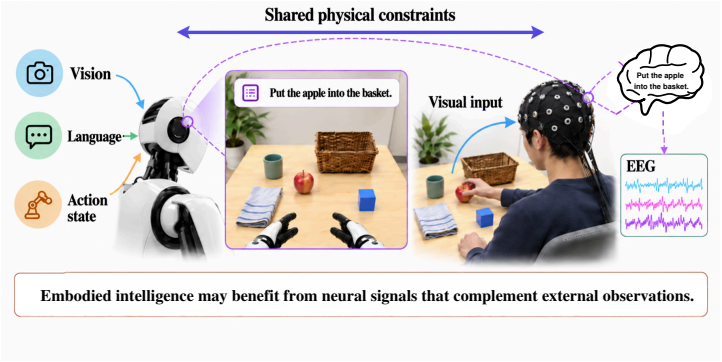

Figure 1: **Neural supervision for embodied learning.** Humans and embodied agents operate under related physical constraints. We investigate whether neural activity recorded during human interaction can provide complementary supervision for embodied representation learning.

imposed by dynamics, sensory uncertainty, and the environment (Pfeifer & Bongard, 2007; Todorov & Jordan, 2002; Cisek, 2007). Human neural dynamics recorded during natural behavior may therefore provide a training signal that differs from external observations while remaining closely tied to the ongoing embodied interaction (Figure 1). Egocentric video provides a natural interface for linking such neural activity to behavior: synchronized egocentric video and EEG place the actor's interaction process and neural dynamics on a shared temporal axis. Prior work has established the value of egocentric data for modeling natural human behavior and has begun to explore synchronized EEG–egocentric recordings for human action understanding (Grauman et al., 2022; 2024; Lin et al., 2026). This motivates extending synchronized neural–behavioral data from human behavior modeling to the training of embodied models.

Building on this idea, we synchronously collect egocentric video and 128-channel EEG during natural manipulation and introduce EEG-guided embodied representation learning, where neural activity temporally aligned with human behavior provides training-time supervision for pretrained VLA models. EEG is completely removed after training, and the robot performs inference using only its standard deployment inputs, requiring no online neural recording. We evaluate the approach across multiple manipulation benchmarks and out-of-distribution settings involving visual, semantic, and execution shifts using OpenVLA as the primary backbone, and further examine its transfer to π_0 on LIBERO. EEG-guided training improves task success over the corresponding robot-only baselines across the evaluated settings and backbones, with larger gains on more complex sequential manipulation tasks and under execution shifts. Further analyses associate these improvements with task progression and completion, while the strongest representational changes occur in late, action-proximal layers. Disrupting the temporal correspondence between EEG and behavior reduces the gains, indicating that the effect depends on behavior-aligned neural supervision. These findings broaden the role of human neural dynamics in embodied learning: rather than serving only as signals to decode or as an additional sensing modality, they can also provide training-time supervision for shaping representations in artificial embodied systems.

Our main contributions are as follows:

- We establish a synchronized human neural–behavioral data paradigm for embodied learning, aligning egocentric behavior during natural manipulation with EEG. This provides an experimental basis for incorporating real human neural dynamics into the training of embodied models.
- We introduce EEG-guided embodied representation learning, where neural activity aligned with behavior is used as a training-time supervisory signal to shape the internal representations of pretrained VLA models, while EEG is removed at inference. This allows neural information to contribute to model learning without changing the sensing and deployment interface of the robot.
- We evaluate the approach across multiple manipulation benchmarks and out-of-distribution settings involving visual, semantic, and execution shifts, and examine its transfer across OpenVLA and π_0. Further analyses show that the gains are consistent and depend on behavior-aligned neural supervision. These results provide empirical support for using human neural dynamics as a training signal for embodied representation learning.

2 Related Work

Vision-Language-Action Models. VLAs have progressed through larger robot datasets, improved action modeling, and more effective adaptation. RT-2 showed that large-scale vision–language

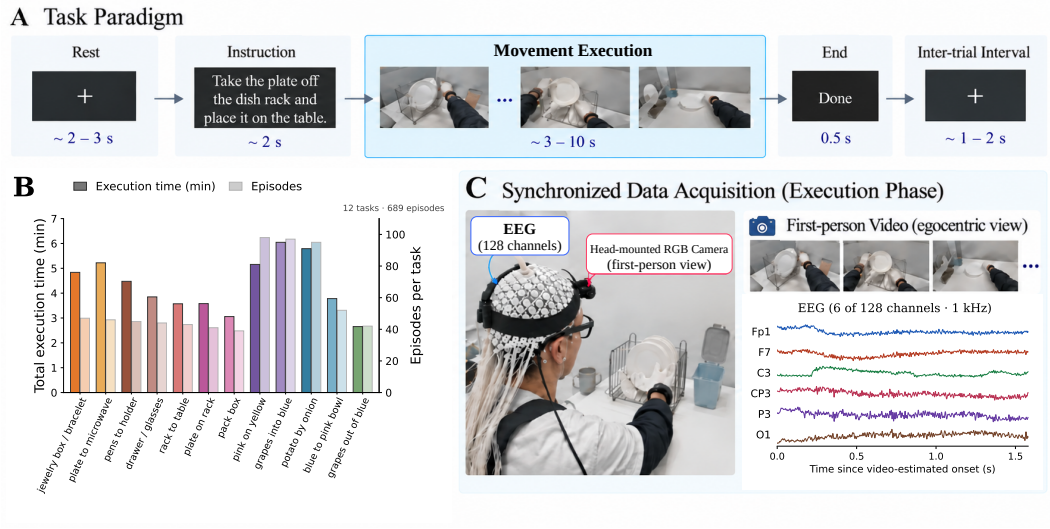

Figure 2: **Synchronized EEG–egocentric data collection. (A)** Task instruction, execution, and randomized inter-trial interval. **(B)** Execution duration and episode count across 12 tasks. **(C)** Synchronized egocentric video and 128-channel EEG during execution.

pretraining can transfer to end-to-end robotic control (Zitkovich et al., 2023). Open X-Embodiment aggregated robot experience across diverse platforms, while OpenVLA introduced an open pretrained VLA for downstream manipulation (Open X-Embodiment Collaboration, 2024; Kim et al., 2024). Recent work has further improved action representation and adaptation, as in OpenVLA-OFT (Kim et al., 2025), and expanded transfer across tasks and embodiments through models such as π_0, $\pi_{0.7}$ and GR00T N1 (Black et al., 2024; Physical Intelligence et al., 2026; Bjorck et al., 2025). These approaches primarily improve embodied policies through external behavioral data and model adaptation. In contrast, we study whether neural activity recorded during human behavior can provide complementary training-time supervision for VLA representations.

Egocentric Learning and Neural Data. Egocentric datasets provide natural human behavior for embodied learning. Ego4D introduced large-scale first-person video, and Ego-Exo4D extended this setting with synchronized multi-view and multimodal recordings (Grauman et al., 2022; 2024). Such data have also been used for robot learning; EgoMimic combines egocentric human demonstrations with robot data for manipulation policy learning (Kareer et al., 2025). More recently, EgoBrain synchronously records egocentric video and EEG for action understanding across participants and environments (Lin et al., 2026). Existing Ego–EEG work, however, mainly treats EEG as an input modality for human action understanding, whereas we use EEG only as training-time supervision for adapting a pretrained embodied model.

Neural Supervision for Representation Learning. Prior work has used measured brain activity to shape representations in pretrained models. Schwartz et al. fine-tuned BERT with fMRI- and MEG-based objectives, showing that neural supervision can bias language representations toward brain-relevant structure (Schwartz et al., 2019). Brain-tuning later extended this idea to pretrained speech models, improving model–brain alignment and semantic downstream performance (Moussa et al., 2025), with subsequent work demonstrating cross-participant generalization (Moussa & Toneva, 2025). These studies establish neural activity as a supervisory signal for representation learning, but focus mainly on language and speech. We extend this principle to embodied models, where representations must support perception and action during physical interaction.

3 METHODS

We investigate whether neural activity synchronized with human embodied behavior can provide auxiliary supervision for adapting VLA representations. Sec. 3.1 describes synchronized egocen-

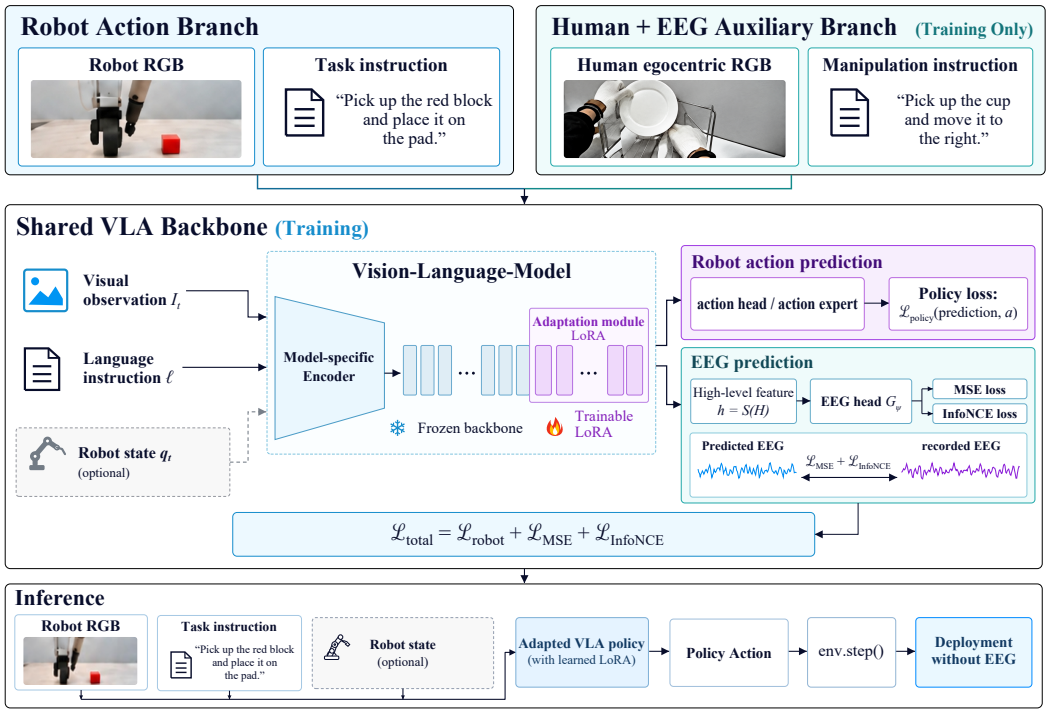

Figure 3: **EEG-guided training framework.** The robot branch learns the native policy objective, while the human branch provides temporally aligned EEG supervision. Both branches jointly optimize the shared adaptation parameters of the VLA backbone, with EEG used only during training and removed at inference.

tric video–EEG collection, Sec. 3.2 presents EEG-guided representation adaptation, and Sec. 3.3 introduces EEG-free inference.

3.1 Synchronized Egocentric Video and EEG Data

To obtain neural supervision temporally aligned with human behavior, we synchronously record egocentric video and EEG, as shown in Fig. 2. Each episode begins with a task instruction, followed by the corresponding object manipulation and a randomized inter-trial interval. The dataset contains 2 participants and 12 embodied manipulation tasks, with repeated trials for each task. Fig. 2(B) summarizes the execution duration and episode count across both participants, illustrating the distribution of data across tasks. EEG is recorded from 128 channels at 1 kHz and resampled to 200 Hz after preprocessing. Egocentric video is captured at a resolution of 1920×1080 and 30 fps, with temporal synchronization to the EEG stream, as illustrated in Fig. 2(C). We use only the task execution phase for subsequent analysis and training, preserving the alignment between EEG supervision and ongoing manipulation behavior.

3.2 Behavior-Aligned Neural Supervision

As illustrated in Figure 3, our framework builds two training branches on a pretrained VLA backbone. The robot branch learns action prediction from robot observations and task instructions, while the human branch uses synchronized egocentric video, task instructions, and EEG to provide neural supervision. The two branches share the same trainable adaptation parameters, allowing the EEG objective to shape representations used by the robot policy.

Given an image observation x and task instruction u, the VLA produces intermediate representations

$$H = F_{\Theta,\Phi}(x, u), \tag{1}$$

where Θ denotes the pretrained backbone parameters and Φ the trainable adaptation parameters. We use OpenVLA as the primary backbone and additionally evaluate π_0; backbone-specific adaptation and action-learning details are provided in the Appendix.

Robot Action Learning. The robot branch learns from manipulation demonstrations using the native action-learning objective of each VLA backbone. Given a robot observation x_i^r, task instruction u_i^r, and target action a_i, we write the robot objective as

$$\mathcal{L}_{\text{robot}} = \mathcal{L}_{\text{policy}} \left(F_{\Theta,\Phi}(x_i^r, u_i^r), a_i \right). \tag{2}$$

This formulation preserves the original action representation and prediction objective of each backbone while allowing the same EEG-guided supervision to act on its trainable representations.

Human-to-EEG Auxiliary Prediction. The human branch takes an egocentric image x_i^h and task instruction u_i^h and predicts the EEG segment aligned with the subsequent 1-s behavioral interval. From the intermediate backbone representations H_i, we select a high-level vision-language representation

$$h_i^h = S(H_i), \tag{3}$$

where $S(\cdot)$ denotes a backbone-specific representation-selection operation. This representation is passed to a lightweight EEG prediction head,

$$\hat{e}_i = G_\psi(h_i^h), \tag{4}$$

where G_ψ is a two-layer MLP with a hidden dimension of 128. The output corresponds to 128 EEG channels over 1 s at 200 Hz. We do not use a pretrained EEG encoder. EEG waveform prediction is supervised using mean squared error:

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N} \left\| \hat{e}_i - e_i \right\|_2^2, \tag{5}$$

where N denotes the number of EEG elements in the target segment.

Contrastive EEG Supervision. Waveform prediction encourages point-wise agreement but does not explicitly distinguish temporally aligned EEG from mismatched neural activity. We therefore introduce an InfoNCE objective (van den Oord et al., 2018), where the aligned EEG segment forms the positive pair and EEG segments from other samples serve as negatives. Predicted and target EEG are normalized along the temporal dimension and flattened in a fixed electrode order. Negatives are drawn from other valid samples in the distributed batch and from a first-in, first-out (FIFO) queue containing up to 256 target EEG representations from recent batches. The contrastive loss is

$$\mathcal{L}_{\text{InfoNCE}} = -\frac{1}{B} \sum_{i=1}^{B} \log \frac{\exp\left(\text{sim}(z_i^p, z_i^t)/\tau\right)}{\exp\left(\text{sim}(z_i^p, z_i^t)/\tau\right) + \sum_{z_j \in \mathcal{N}_i} \exp\left(\text{sim}(z_i^p, z_j)/\tau\right)}, \tag{6}$$

where $\text{sim}(\cdot,\cdot)$ denotes cosine similarity, $\mathcal{N}_i$ is the valid negative set, and $\tau = 0.1$.

Joint Optimization. The full training objective combines the native robot policy objective with EEG waveform prediction and contrastive supervision:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{robot}} + \lambda_{\text{MSE}} \mathcal{L}_{\text{MSE}} + \lambda_{\text{NCE}} \mathcal{L}_{\text{InfoNCE}}. \tag{7}$$

We set $\lambda_{\text{MSE}} = 0.1$ and $\lambda_{\text{NCE}} = 0.005$. The robot and EEG objectives jointly optimize the shared adaptation parameters Φ, allowing behavior-aligned neural supervision to shape the representations used by the robot policy.

3.3 EEG-Free Inference

EEG is used only during training. At inference, each adapted VLA operates with its standard deployment inputs, without any auxiliary EEG components or online neural recording. This preserves the original inference interface of each backbone. Additional backbone-specific training and hyperparameter details are provided in the Appendix.

4 EXPERIMENTS

4.1 EXPERIMENTAL SETUP

We evaluate our method on multiple robot manipulation settings, including LIBERO (Liu et al., 2023), as well as the RL4VLA benchmark (Liu et al., 2025a). We use OpenVLA-7B as the primary backbone for the main benchmark evaluation. To assess whether neural supervision remains effective across different VLA architectures, we additionally evaluate π_0 (Black et al., 2024) on all four LIBERO suites. Their action-generation mechanisms differ from the autoregressive action-token formulation of OpenVLA, providing a complementary test of cross-backbone transfer. Comparisons within each backbone use the same robot training data and evaluation protocol.

For LIBERO, we report task success rates over 50 episodes per task (Liu et al., 2023), while for RL4VLA, we evaluate each setting over 64 episodes (Liu et al., 2025a). During joint training, the robot branch uses a global batch size of 32, while the human EEG branch uses a global batch size of 64. Experiments are conducted on a compute cluster with 20 NVIDIA H100 GPUs. Additional training hyperparameters, benchmark-specific evaluation settings, and implementation details are provided in the Appendix.

As a diagnostic check, we examine whether execution-phase EEG contains task-related information. Fig. 4 shows above-chance 12-way task decoding across multiple models and EEG context lengths.

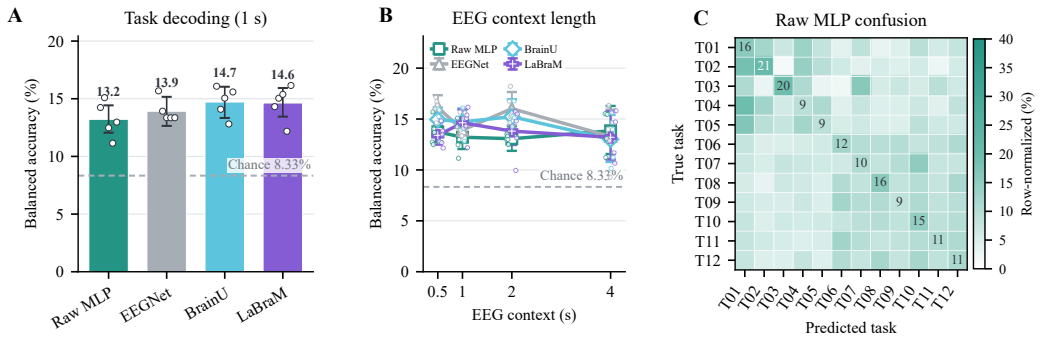

Figure 4: **Task-related information in execution-phase EEG. (A)** 12-way decoding from 1-s EEG windows using four models, including EEGNet (Lawhern et al., 2018), BrainMu (The Brainmu Authors, 2026), and LaBraM (Jiang et al., 2024). **(B)** 12-way task-decoding accuracy across EEG context lengths of 0.5–4 s. **(C)** Row-normalized confusion matrix for the 1-s Raw MLP. Dashed lines indicate the 8.33% chance level.

4.2 MAIN RESULTS

We evaluate EEG-guided training using OpenVLA as the primary backbone on four LIBERO suites—Spatial, Object, Goal, and Long—and three RL4VLA out-of-distribution settings corresponding to visual, semantic, and execution shifts.

Table 1 summarizes the OpenVLA results. EEG-guided training improves mean task success across all seven evaluated settings, suggesting that the benefit is not restricted to a single task category or evaluation protocol. The larger gains on LIBERO-Long and Execution OOD are particularly informative: both settings place greater demands on maintaining task progression under extended or perturbed execution, where failures often arise from incomplete subgoals, delayed transitions, or unsuccessful action completion. This pattern suggests that behavior-aligned neural supervision may be especially useful for shaping representations related to execution dynamics and task completion, which we examine further in the subsequent analyses.

To examine whether this effect is specific to OpenVLA, we further evaluate EEG-guided training with π_0 on the four LIBERO suites. As shown in Table 2, EEG-guided training also improves performance over the corresponding robot-only baseline, with relatively larger gains on LIBERO-Goal and LIBERO-Long. The smaller gains relative to OpenVLA may partly reflect the stronger π_0 baseline, which leaves less room for improvement. Its flow-matching action expert also introduces a

Table 1: **Main results with OpenVLA on LIBERO and RL4VLA.** Success rates are reported as mean ± standard deviation. Gain is the absolute improvement over the corresponding robot-only baseline.

Benchmark	Setting	Method	EEG Sup.	Episodes	Success Rate	Gain (pp)
LIBERO	Spatial	Robot-only	–	50	79.53 ± 1.03	–
		EEG-guided	InfoNCE	50	$\mathbf{80.93 \pm 1.42}$	$\mathbf{+1.40}$
	Object	Robot-only	–	50	66.93 ± 2.34	–
		EEG-guided	InfoNCE	50	$\mathbf{68.53 \pm 1.70}$	$\mathbf{+1.60}$
	Goal	Robot-only	–	50	77.87 ± 0.12	–
		EEG-guided	InfoNCE	50	$\mathbf{79.40 \pm 1.04}$	$\mathbf{+1.53}$
	Long	Robot-only	–	50	53.47 ± 0.42	–
		EEG-guided	InfoNCE	50	$\mathbf{56.73 \pm 0.70}$	$\mathbf{+3.26}$
RL4VLA	Vision OOD	Robot-only	–	64	93.33 ± 1.26	–
		EEG-guided	InfoNCE	64	$\mathbf{96.25 \pm 0.54}$	$\mathbf{+2.92}$
	Semantic OOD	Robot-only	–	64	71.50 ± 1.27	–
		EEG-guided	InfoNCE	64	$\mathbf{72.99 \pm 0.00}$	$\mathbf{+1.49}$
	Execution OOD	Robot-only	–	64	32.99 ± 2.10	–
		EEG-guided	InfoNCE	64	$\mathbf{39.00 \pm 1.86}$	$\mathbf{+6.01}$

Table 2: **Results with π_0 on LIBERO.** Success rates are reported as mean ± standard deviation. Gain is the absolute improvement over the corresponding robot-only baseline.

Benchmark	Setting	Method	EEG Sup.	Episodes	Success Rate	Gain (pp)
LIBERO	Spatial	Robot-only	–	50	93.67 ± 0.42	–
		EEG-guided	InfoNCE	50	$\mathbf{94.00 \pm 0.20}$	$\mathbf{+0.33}$
	Object	Robot-only	–	50	90.17 ± 0.37	–
		EEG-guided	InfoNCE	50	$\mathbf{91.13 \pm 1.09}$	$\mathbf{+0.96}$
	Goal	Robot-only	–	50	92.93 ± 1.89	–
		EEG-guided	InfoNCE	50	$\mathbf{94.20 \pm 0.87}$	$\mathbf{+1.27}$
	Long	Robot-only	–	50	80.20 ± 0.72	–
		EEG-guided	InfoNCE	50	$\mathbf{81.33 \pm 1.22}$	$\mathbf{+1.13}$

different action-generation pathway that may interact differently with representation-level supervision. Together, these results suggest that behavior-aligned neural supervision is not restricted to OpenVLA and can transfer across different VLA action-generation paradigms.

EEG Control Experiments. To further examine whether the gains from EEG-guided training depend on the correspondence between EEG and behavior, we conduct control experiments with OpenVLA on LIBERO-Long. In addition to the robot-only baseline and the full model, we evaluate a variant that removes InfoNCE while retaining EEG reconstruction, as well as an EEG-Shuffle condition using EEG from an incorrect trial.

Table 3: **Effects of EEG supervision, alignment, and InfoNCE on LIBERO-Long.**

Setting	EEG Sup.	Success Rate
Robot-only	None	53.47 ± 0.42
w/o InfoNCE	Aligned EEG	54.00 ± 1.39
EEG-Shuffle	Mismatched EEG	53.27 ± 1.80
EEG-guided	Aligned EEG	$\mathbf{56.73 \pm 0.70}$

As shown in Table 3, the full model achieves the highest success rate. Removing InfoNCE retains little of the improvement, whereas EEG-Shuffle reduces performance to a level close to the robot-only baseline. Since EEG-Shuffle preserves the same EEG branch and training objectives while disrupting the EEG–behavior pairing, the gain appears to depend on the correct correspondence between EEG and behavior rather than simply on additional EEG-related supervision.

Table 4: **Ablation studies on LIBERO-Long.** Each component is varied independently with all other settings fixed. † denotes the default configuration used in the main experiments.

LoRA Module		LoRA Depth		EEG Head Dim.		EEG Window	
Setting	**SR (%)**	**Setting**	**SR (%)**	**Setting**	**SR (%)**	**Setting**	**SR (%)**
Attention only†	$\mathbf{56.73 \pm 0.70}$	Last 4	54.33 ± 0.83	64	53.07 ± 1.70	1 s†	$\mathbf{56.73 \pm 0.70}$
FFN only	55.73 ± 2.73	Last 8†	$\mathbf{56.73 \pm 0.70}$	128†	$\mathbf{56.73 \pm 0.70}$	2 s	52.40 ± 0.92
Attention + FFN	56.40 ± 1.56	All layers	53.20 ± 0.92	256	56.07 ± 2.00	3 s	53.47 ± 2.58

4.3 Ablation Studies

We analyze several design choices in EEG-guided training with OpenVLA on LIBERO-Long, including the placement and depth of LoRA adaptation, the capacity of the EEG MLP, and the temporal extent of EEG supervision. Table 4 summarizes the results. Additional optimization-related hyperparameters are reported in the Appendix.

LoRA Placement and Depth. We examine where EEG-guided adaptation should be applied by varying both the adapted modules and the adaptation depth. At the module level, attention-only, FFN-only, and joint attention–FFN adaptation perform similarly, with attention-only achieving the highest mean success rate and relatively low variation, suggesting that extending adaptation to the FFN provides no clear additional benefit. At the depth level, adapting the last eight layers outperforms both the last-four-layer and all-layer settings, indicating that EEG supervision is most effective when applied to a sufficiently broad set of upper-layer representations without extending adaptation throughout the full language backbone.

EEG MLP Capacity. We vary the hidden dimension of the EEG MLP among 64, 128, and 256. The 128-dimensional head achieves the best mean performance, while both smaller and larger alternatives perform worse. This suggests a trade-off between mapping capacity and information bottleneck: a moderate auxiliary head may encourage the shared representation to retain compact, task-relevant neural information without relying on the MLP to model fine-grained EEG variations.

Temporal Extent of EEG Supervision. Finally, we vary the EEG target window from 1 to 3 s. The 1-s window achieves the highest and most stable performance, while longer windows reduce success rates. This suggests that EEG supervision benefits from a relatively tight temporal correspondence with the current behavioral state; extending the window may incorporate broader behavioral dynamics but also less directly relevant neural activity.

4.4 Interpretability Analysis

Behavior-level analysis. To understand where the gains from EEG-guided training arise, we analyze OpenVLA rollout behavior on LIBERO-Long. As shown in Fig. 5(A), the gains vary across tasks and are larger for tasks involving sequential subgoals, stage transitions, and explicit completion steps, such as T9 and T2. In contrast, Task 3 (*put the black bowl in the bottom drawer of the cabinet and close it*) and Task 4 (*put the white mug on the left plate and put the yellow and white mug on the right plate*) place greater demands on object identification and spatial target binding and show smaller or negative gains. The failure analysis further shows that robot-only failures typically reach the 520-step horizon (17.33 s), often with late-stage stalling or continued motion without completion. Representative rollouts in Fig. 5(B) show that EEG-guided models instead complete the tasks within 7.6–14.9 s. These patterns suggest that EEG-guided supervision is more closely associated with *task progression and action completion* than with fine-grained visual grounding or reduced motion jitter.

Representation-level analysis. We next examine whether the behavioral differences are reflected in the model's internal representations. As shown in Fig. 6(A), the representational difference increases in the upper Transformer layers and is largest at Layer 32, closest to the action-token output head. This suggests that EEG-guided supervision primarily affects *late, action-proximal representations*. Fig. 6(B) further shows that the Layer-32 difference is most pronounced around the grasp phase, consistent with the behavioral pattern associated with *task progression and action completion*.

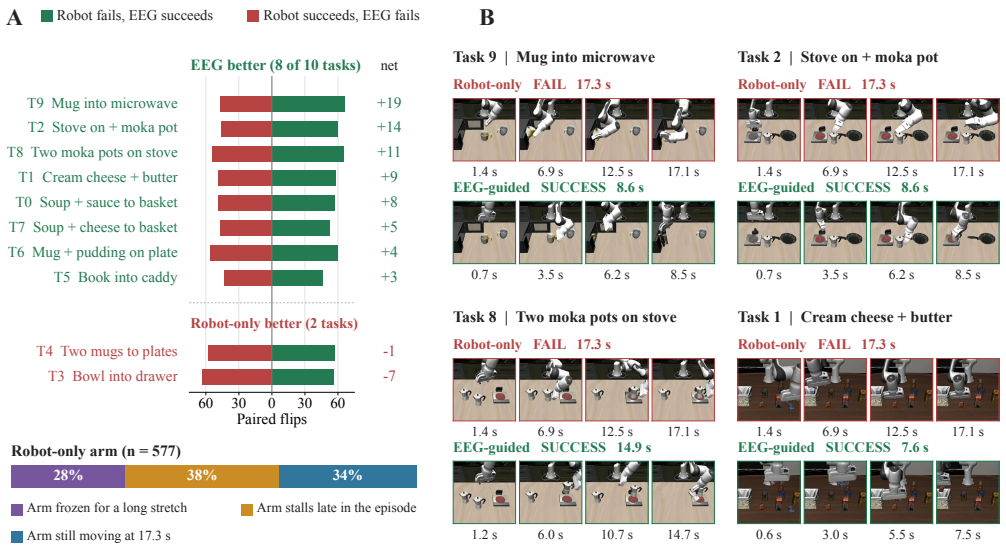

Figure 5: **Task-wise outcome changes and rollout analysis on LIBERO-Long. (A)** Paired outcome changes between robot-only and EEG-guided models; rightward bars indicate gains and leftward bars indicate regressions. Robot-only failure modes are summarized below. **(B)** Representative paired rollouts from the two models.

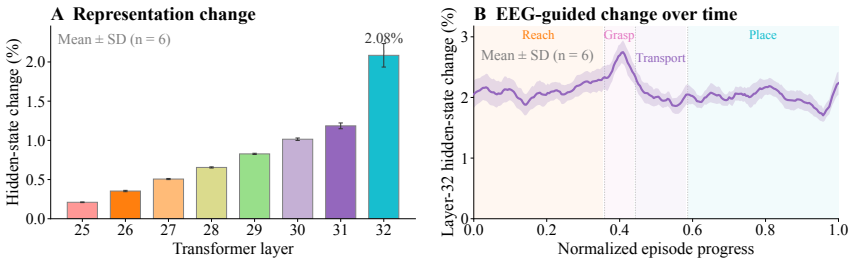

Figure 6: **Representation analysis. (A)** Hidden-state differences across layers. **(B)** Layer-32 differences across task phases.

5 Conclusion and Discussion

In this work, we examine whether human brain activity can serve as a training-time supervisory signal for embodied representation learning. We find that training with behavior-aligned EEG improves downstream task performance even when EEG is removed at inference, whereas disrupting the EEG–behavior correspondence reduces this gain. We further observe improvements over the corresponding robot-only baselines across VLA backbones, including OpenVLA and π_0, suggesting that the benefit is not specific to a particular model architecture. Overall, these results indicate that human neural dynamics can provide transferable training-time supervision for existing embodied models, offering an additional source of performance gains and a potential path toward leveraging internal behavioral-state information beyond external observations for embodied learning.

More broadly, neural supervision may contribute at different stages of the embodied learning pipeline. Current VLA development spans heterogeneous pre-training, task-specific post-training, and reinforcement-learning-based policy refinement (Bjorck et al., 2025; Black et al., 2025; Kim et al., 2025; Liu et al., 2025a). Beyond vision, force and audio have been incorporated into manipulation policies, while EEG-derived error or trust signals have been explored for reward shaping and closed-loop policy optimization (Yu et al., 2025; Liu et al., 2025b; Kim et al., 2017; Xu et al., 2025). Our study focuses on parameter-efficient adaptation, where EEG provides representation-level supervision. An important direction is therefore to determine where neural signals provide the greatest value: shaping general embodied representations during pre-training, adapting them during post-training, or providing online feedback for policy refinement. These findings add to growing evidence that neural activity can provide complementary information beyond conventional embodied data, suggesting that brain-derived signals may become an additional source of training gain for future embodied models.

ETHICS STATEMENT

All participants provided informed consent prior to data collection. The collected EEG and egocentric video data were de-identified and handled under appropriate privacy protections. The study involved non-invasive EEG recording and everyday object-manipulation tasks with minimal risk to participants.

REPRODUCIBILITY STATEMENT

Implementation details, preprocessing procedures, task definitions, model configurations, training settings, evaluation protocols, and random seeds are provided in the Appendix. The code used for training and evaluation is also included in the supplementary material. We will release the code and relevant documentation publicly upon acceptance.

AI USE STATEMENT

We used generative AI tools solely to assist with language editing and polishing of the manuscript, including improving grammar, clarity, and readability of author-written text. All AI-assisted revisions were carefully reviewed and verified by the authors. The authors take full responsibility for the accuracy, integrity, and final content of the manuscript.

REFERENCES

Johan Bjorck, Fernando Castañeda, Nikita Cherniadev, Xingye Da, Runyu Ding, Linxi Jim Fan, et al. GR00T N1: An open foundation model for generalist humanoid robots, 2025.

Kevin Black, Noah Brown, James Darpinian, Karan Dhabalia, Danny Driess, et al. $\pi_{0.5}$: A vision-language-action model with open-world generalization. In *Proceedings of the 9th Conference on Robot Learning*, volume 305 of *Proceedings of Machine Learning Research*, pp. 17–40. PMLR, 2025.

Kevin Black et al. π_0: A vision-language-action flow model for general robot control. *arXiv preprint arXiv:2410.24164*, 2024.

Cheng Chi, Siyuan Feng, Yilun Du, Zhenjia Xu, Eric Cousineau, Benjamin Burchfiel, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion. In *Robotics: Science and Systems*, 2023.

Paul Cisek. Cortical mechanisms of action selection: The affordance competition hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1485):1585–1599, 2007.

Pete Florence, Corey Lynch, Andy Zeng, Oscar A. Ramirez, Ayzaan Wahid, Laura Downs, Adrian Wong, Johnny Lee, Igor Mordatch, and Jonathan Tompson. Implicit behavioral cloning. In *Proceedings of the 5th Conference on Robot Learning*, volume 164 of *Proceedings of Machine Learning Research*, pp. 158–168, 2022.

William J. Gehring, Brian Goss, Michael G. H. Coles, David E. Meyer, and Emanuel Donchin. A neural system for error detection and compensation. *Psychological Science*, 4(6):385–390, 1993.

Kristen Grauman, Andrew Westbury, Eugene Byrne, Zachary Chavis, Antonino Furnari, Rohit Girdhar, et al. Ego4D: Around the world in 3,000 hours of egocentric video. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 18995–19012, 2022.

Kristen Grauman, Andrew Westbury, Lorenzo Torresani, Kris Kitani, Jitendra Malik, Triantafyllos Afouras, et al. Ego-Exo4D: Understanding skilled human activity from first- and third-person perspectives. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 19383–19400, 2024.

Wei-Bang Jiang, Li-Ming Zhao, and Bao-Liang Lu. Large brain model for learning generic representations with tremendous EEG data in BCI. In *International Conference on Learning Representations*, 2024.

Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1–2):99–134, 1998.

Simar Kareer, Dhruv Patel, Ryan Punamiya, Pranay Mathur, Shuo Cheng, Chen Wang, Judy Hoffman, and Danfei Xu. EgoMimic: Scaling imitation learning via egocentric video. In *2025 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 13226–13233, 2025.

Moo Jin Kim, Karl Pertsch, Siddharth Karamcheti, Ted Xiao, Ashwin Balakrishna, Suraj Nair, et al. OpenVLA: An open-source vision-language-action model. In *Proceedings of the 8th Conference on Robot Learning*, volume 270 of *Proceedings of Machine Learning Research*, pp. 2679–2713, 2024.

Moo Jin Kim, Chelsea Finn, and Percy Liang. Fine-tuning vision-language-action models: Optimizing speed and success. *arXiv preprint arXiv:2502.19645*, 2025.

Su Kyoung Kim, Elsa Andrea Kirchner, Arne Stefes, and Frank Kirchner. Intrinsic interactive reinforcement learning: Using error-related potentials for real world human-robot interaction. *Scientific Reports*, 7:17562, 2017. doi: 10.1038/s41598-017-17682-7.

Vernon J. Lawhern, Amelia J. Solon, Nicholas R. Waytowich, Stephen M. Gordon, Chou P. Hung, and Brent J. Lance. Eegnet: A compact convolutional neural network for eeg-based brain–computer interfaces. *Journal of Neural Engineering*, 15(5):056013, 2018. doi: 10.1088/1741-2552/aace8c.

Nie Lin, Yansen Wang, Dongqi Han, Wei-Bang Jiang, Jingyuan Li, Ryosuke Furuta, Yoichi Sato, and Dongsheng Li. EgoBrain: Synergizing minds and eyes for human action understanding. In *International Conference on Learning Representations*, 2026.

Bo Liu, Yifeng Zhu, Chongkai Gao, et al. LIBERO: Benchmarking knowledge transfer for lifelong robot learning. In *Advances in Neural Information Processing Systems*, 2023.

Jijia Liu, Feng Gao, Bingwen Wei, Xinlei Chen, Qingmin Liao, Yi Wu, Chao Yu, and Yu Wang. What can RL bring to VLA generalization? an empirical study. In *Advances in Neural Information Processing Systems*, volume 38, 2025a. doi: 10.52202/085713-3248.

Zeyi Liu, Cheng Chi, Eric Cousineau, Naveen Kuppuswamy, Benjamin Burchfiel, and Shuran Song. ManiWAV: Learning robot manipulation from in-the-wild audio-visual data. In *Proceedings of the 8th Conference on Robot Learning*, volume 270 of *Proceedings of Machine Learning Research*, pp. 947–962. PMLR, 2025b.

Guido Montúfar, Keyan Ghazi-Zahedi, and Nihat Ay. A theory of cheap control in embodied systems. *PLoS Computational Biology*, 11(9):e1004427, 2015. doi: 10.1371/journal.pcbi.1004427.

Omer Moussa and Mariya Toneva. Brain-tuning improves generalizability and efficiency of brain alignment in speech models. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

Omer Moussa, Dietrich Klakow, and Mariya Toneva. Improving semantic understanding in speech language models via brain-tuning. In *International Conference on Learning Representations*, 2025.

Open X-Embodiment Collaboration. Open X-embodiment: Robotic learning datasets and RT-X models. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 6892–6903, 2024.

Rolf Pfeifer and Josh Bongard. *How the Body Shapes the Way We Think: A New View of Intelligence*. MIT Press, 2007.

Gert Pfurtscheller and F. H. Lopes da Silva. Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology*, 110(11):1842–1857, 1999.

Physical Intelligence, Bo Ai, Ali Amin, Raichelle Aniceto, Ashwin Balakrishna, et al. $\pi_{0.7}$: A steerable generalist robotic foundation model with emergent capabilities, 2026.

Dan Schwartz, Mariya Toneva, and Leila Wehbe. Inducing brain-relevant bias in natural language processing models. In *Advances in Neural Information Processing Systems*, volume 32, pp. 14123–14133, 2019.

Nur Muhammad Mahi Shafiullah, Zichen Cui, Ariuntuya Altanzaya, and Lerrel Pinto. Behavior transformers: Cloning k modes with one stone. In *Advances in Neural Information Processing Systems*, volume 35, 2022.

The Brainmu Authors. Brainmu project, 2026. URL `https://github.com/BAAI-Brain-Inspired-Group/Brainmu`.

Emanuel Todorov and Michael I. Jordan. Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, 5(11):1226–1235, 2002.

Aaron van den Oord, Yazhe Li, and Oriol Vinyals. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748*, 2018.

Caiyue Xu, Hongrui Sang, Changming Zhang, Yanmin Zhou, Zhipeng Wang, Xin Li, and Bin He. NeuTRL: Neural trust-guided reinforcement learning for human-robot collaboration. *IEEE Robotics and Automation Letters*, 10(8):8155–8162, 2025. doi: 10.1109/LRA.2025.3583508.

Jiawen Yu, Hairuo Liu, Qiaojun Yu, Jieji Ren, Ce Hao, et al. ForceVLA: Enhancing VLA models with a force-aware MoE for contact-rich manipulation. In *Advances in Neural Information Processing Systems*, volume 38, 2025. doi: 10.52202/085713-3124.

Brianna Zitkovich, Tianhe Yu, Sichun Xu, Peng Xu, Ted Xiao, Fei Xia, et al. RT-2: Vision-language-action models transfer web knowledge to robotic control. In *Proceedings of the 7th Conference on Robot Learning*, volume 229 of *Proceedings of Machine Learning Research*, pp. 2165–2183, 2023.

A Additional Dataset Details

A.1 Data Acquisition

We adopt an instruction-guided natural object manipulation paradigm. At the start of each episode, a textual task instruction is presented on the screen, after which the participant performs the corresponding object manipulation. Once the manipulation is completed, the episode ends and is followed by an inter-trial interval of random duration. Although the complete experimental protocol includes additional phases, only the task execution phase is used for the analyses and model training in this work.

The dataset used in this study contains 2 participants and 12 embodied manipulation tasks. Each participant completes all 12 tasks, with multiple repeated trials per task. During data collection, EEG and egocentric video are recorded synchronously to establish temporal correspondence between neural activity and the ongoing behavioral process.

Figure 2(B) depicts an overview subset of 689 episodes across the 12 tasks, selecting one source dataset per task. The EEG task-informativeness analysis in Section A.4 instead includes both available source datasets for tasks represented in both, yielding 1,224 episodes.

EEG is recorded using a 128-channel saline electrode cap at an original sampling rate of approximately 1 kHz. Egocentric video is captured using a custom-built head-mounted camera at a resolution of 1920×1080 and 30 fps. The video covers the entire task execution period and is temporally synchronized with the EEG recordings.

A.2 EEG Preprocessing

EEG preprocessing is performed independently for each episode. We first apply notch filters at 50 Hz and 100 Hz to suppress power-line interference and its harmonic, followed by a 1–100 Hz band-pass filter. We then assess signal quality channel by channel and reconstruct channels identified as bad using spherical interpolation based on the spatial configuration of neighboring electrodes.

Next, we apply Infomax independent component analysis (ICA) to reduce non-neural artifacts. The EEG signals are decomposed into 64 independent components. Artifact-related components are identified using automatically adjusted thresholds, with particular emphasis on components dominated by electromyographic activity. These components are removed and the remaining components are projected back into sensor space to obtain cleaned EEG signals.

After preprocessing, the cleaned EEG signals are resampled to 200 Hz and stored in H5 format for each episode, together with the corresponding egocentric images and JSON metadata. The metadata records episode-level information and the correspondence between the EEG and video streams. All subsequent EEG feature extraction and EEG–visual representation alignment are performed using these preprocessed signals. We retain only EEG and egocentric video recorded during the actual execution period of the task. Continuous data are segmented at each episode's execution boundaries, covering the period from the beginning of task execution to task completion. Signals recorded during non-execution phases are excluded.

Consequently, the EEG and video samples used for model training correspond to the same behavioral process, preventing neural activity from unrelated phases, such as rest or motor imagery, from being introduced into the training data.

A.3 EEG Channel Distribution and Representative Examples

Figure 7 illustrates the spatial coverage of the 128-channel saline EEG system. Representative landmarks from the international 10–20 system orient the montage across the frontal, central, parietal, and occipital scalp regions. These landmarks indicate the anatomical coverage of the dense-array montage; they neither enumerate all 128 electrode locations nor depict task-related activation.

Figure 8 shows a real clothes-folding episode. Panels A–D contain egocentric frames at 37.0, 38.5, 40.0, and 41.5 s. The adjacent traces display stored EEG values from 34 to 45 s for six channels not marked as bad, spanning anterior to posterior scalp locations: Fp1, F3, C3, CP3, P3, and O1. The displayed 1-kHz waveforms are read from the stored preprocessed acquisition archive without

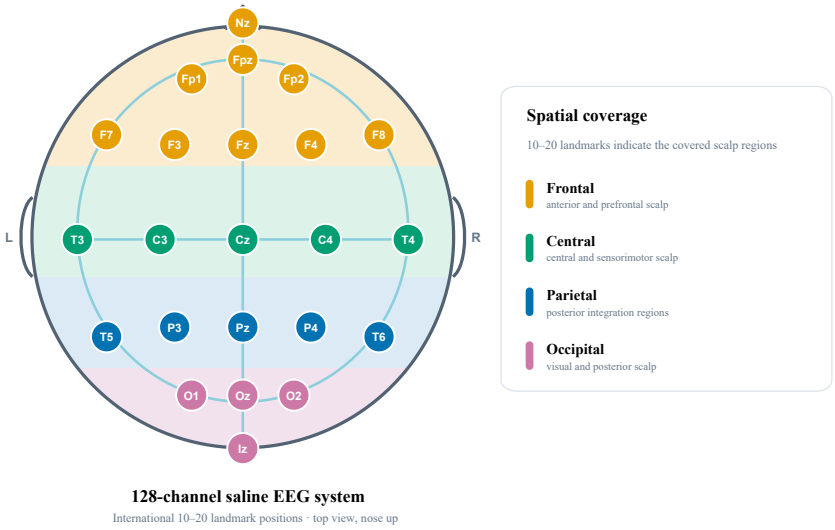

Figure 7: **Spatial coverage of the 128-channel saline EEG system.** Representative international 10–20 landmarks illustrate coverage over frontal, central, parietal, and occipital scalp regions.

further filtering or resampling for this visualization; fixed vertical offsets separate the channels. The waveforms vary in amplitude and structure across the episode, qualitatively illustrating how the recorded neural signals evolve during the behavior.

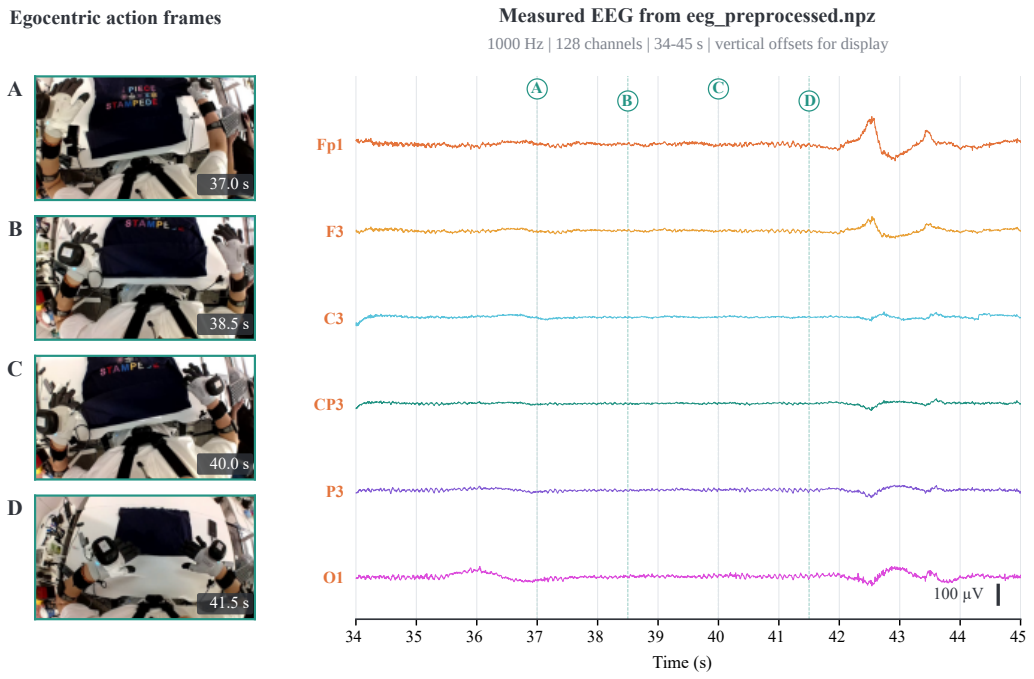

Figure 8: **Synchronized egocentric views and measured EEG from a clothes-folding episode.** Panels A–D show camera frames sampled every 1.5 s at 37.0, 38.5, 40.0, and 41.5 s. The right panel displays stored EEG values from 34 to 45 s for Fp1, F3, C3, CP3, P3, and O1 in the same acquisition archive. Dashed lines mark the frame times.

A.4 EEG Signal Informativeness

We conduct an independent diagnostic analysis to test whether execution-phase EEG contains information that distinguishes ongoing tasks. The analysis includes 6,291 non-overlapping 1-second windows from 1,224 episodes spanning 12 tasks. We use five-fold cross-validation grouped by episode, ensuring that windows from the same episode never appear in both the training and test sets of a fold. Chance accuracy for the 12-way classification task is 8.33%. Accuracy, balanced accuracy, and 95% confidence intervals are computed from out-of-fold predictions, with confidence intervals estimated by stratified bootstrap resampling at the episode level.

The raw-waveform MLP serves as an input-side diagnostic analogue of the waveform prediction head: it flattens the 128×200 input, applies LayerNorm, a 128-dimensional hidden layer, GELU, dropout of 0.1, and a 12-way output layer. Compact EEGNet is included as an additional task classifier. Brainμtokenizer and LaBraM are evaluated as frozen feature extractors followed by class-balanced linear probes. These models are used solely for signal diagnostics and are not interchangeable with the OpenVLA-to-EEG prediction head used for VLA training.

As shown in Figure 4(A), the balanced-accuracy intervals of all four diagnostic models lie above the 8.33% chance level. Because these models operate on complementary signal representations, their agreement provides a robustness check indicating that the execution-phase EEG retained after preprocessing, artifact-control, and quality-control pipeline contains detectable task-category information under episode-held-out evaluation. Figure 4(B) further reports the corresponding results across EEG context lengths from 0.5 to 4 seconds. The 4-second analysis is exploratory because the shortest-duration task class contributes only four eligible 4-second contexts. Figure 4(C) shows the row-normalized out-of-fold confusion matrix of the 1-second raw-waveform MLP.

The task indices in Figure 4(C) are: T01, move grapes from the blue bowl to the pink bowl; T02, open a drawer and put the glasses inside; T03, open a jewelry box, take out the bracelet, and close the box; T04, pack a box; T05, place the potato to the right of the onion; T06, put food from a plate into a microwave; T07, put grapes into the blue bowl; T08, put pens into the pen holder; T09, put a plate on a drying rack; T10, stack the pink block on the yellow block; T11, take grapes out of the blue bowl; and T12, move a plate from a drying rack to a table.

The analysis is intended to determine whether task-category information is present, not to rank the diagnostic models. It does not localize that information to a specific cortical region or frequency band, identify an underlying neural mechanism, or demonstrate generalization across subjects.

B EEG-Guided OpenVLA Training Details

B.1 Overall Training Pipeline

Algorithm 1 summarizes the joint training procedure.At each optimization step, a robot batch and a human EEG batch are processed in separate forward passes through the same OpenVLA model, whose base parameters remain frozen while the LoRA parameters are shared and trainable across both branches. The robot branch provides action supervision, whereas the EEG branch provides waveform reconstruction and cross-modal contrastive supervision. Gradients from both objectives update the shared LoRA parameters, while the EEG prediction head is optimized exclusively by the EEG objective. EEG therefore serves only as privileged supervision during training. At deployment, the entire EEG branch and all EEG-specific components are discarded, and robot rollouts rely solely on the OpenVLA backbone with the learned LoRA parameters.

Algorithm 1. Overall Training Pipeline

Require: Frozen OpenVLA base parameters Θ, shared trainable LoRA parameters Φ, EEG head Ψ
Require: Robot data $\mathcal{D}_r$, paired human image–EEG data $\mathcal{D}_e$
Require: EEG loss weight $\lambda = 0.1$, within-EEG contrastive weight $\beta = 0.05$
1: **for** each training step **do**
2: Sample robot batch (x^r, u^r, a) from $\mathcal{D}_r$
3: Sample human batch (x^e, u^e, e) from $\mathcal{D}_e$
4: $O^r \leftarrow F_{\Theta,\Phi}(x^r, u^r)$

$\mathcal{L}_{\text{robot}} \leftarrow \text{CE}(O^r, a)$
$H^e \leftarrow F_{\Theta,\Phi}(x^e, u^e)$
$h^e \leftarrow$ hidden state at the final prompt token
$\hat{e} \leftarrow G_\Psi(h^e)$
$\mathcal{L}_{\text{EEG}} \leftarrow \text{MSE}(\hat{e}, e) + \beta\,\text{InfoNCE}(\hat{e}, e)$
$\mathcal{L} \leftarrow \mathcal{L}_{\text{robot}} + \lambda\,\mathcal{L}_{\text{EEG}}$
Backpropagate $\mathcal{L}$ to Φ and Ψ
Update Φ and Ψ with AdamW
end for
Inference: discard G_Ψ and all EEG-related components
return robot policy $F_{\Theta,\Phi}(x^r, u^r)$

B.2 Training Configuration

Model initialization and batching. For each LIBERO suite, all compared settings are initialized from the same suite-specific OpenVLA checkpoint. For example, LIBERO-Object uses `openvla/openvla-7b-finetuned-libero-object`, while LIBERO-Spatial, LIBERO-Goal, and LIBERO-Long use their corresponding publicly released suite-specific checkpoints. The visual backbone, multimodal projector, and original language-model parameters remain frozen throughout training. LoRA adapters are inserted into the `q_proj` and `v_proj` modules of language-model layers 24–31, corresponding to the final eight layers, with a default rank of 16. At each optimization step, one robot batch and one human EEG batch are independently sampled and processed in separate forward passes. The global batch sizes are 32 for the robot branch and 64 for the EEG branch. Each run uses H100 GPUs.

Visual preprocessing. Robot images are resized to 224×224 and augmented online using random resized cropping with a crop ratio of 0.9, brightness perturbation of ± 0.2, contrast and saturation scaling in $[0.8, 1.2]$, and hue perturbation of ± 0.05. Egocentric human frames are first center-cropped to retain 80% of the image area and resized to 224×224 using bilinear interpolation. The same online augmentation is then applied.

EEG preprocessing. EEG targets are extracted from 1 s windows, $[t, t+1)$, during the motor execution period. For each window, the recorded waveform is linearly interpolated onto a uniform 1,000 Hz temporal grid and then downsampled to 200 Hz, yielding a target of size 128×200. A window is retained only if it contains 900–1,100 raw samples, has strictly increasing timestamps, has an absolute boundary deviation of at most 2 ms from the nominal window boundaries, and has a maximum inter-sample interval below 3 ms. All retained signals are reordered to a common 128-channel electrode configuration. EEG targets are standardized independently for each channel using statistics computed exclusively from the training split:

$$\tilde{x}_{n,c,t} = \frac{x_{n,c,t} - \mu_c}{\max(\sigma_c, 10^{-3})}, \tag{8}$$

where μ_c and σ_c are computed across all training windows and time points for channel c. The same statistics are applied unchanged to the validation split.

Optimization. We use AdamW with a weight decay of 0.01. The learning rates are 3×10^{-5} for the LoRA parameters and 3×10^{-4} for the EEG prediction head. Unless otherwise specified, the LoRA rank is 16. The EEG auxiliary loss is weighted by $\lambda = 0.1$, and the contrastive term within the EEG objective is weighted by $\beta = 0.05$, as defined in Algorithm 1. Training runs for 2,000 optimization steps, with a linear warmup over the first 5% of steps followed by cosine decay to 0.1 times the initial learning rate. Gradients are clipped to a global norm of 1.0, and training uses bf16 precision with FlashAttention-2.

Control variants. The control settings reported in the main text use the same suite-specific initialization, robot training data, and optimization schedule unless otherwise stated. **Robot-only** removes the human EEG branch entirely and optimizes only the robot action objective. **w/o InfoNCE** retains the aligned EEG waveform-reconstruction objective but removes the contrastive term by setting $\beta = 0$. For **EEG-Shuffle**, the full EEG branch and training objective are retained, but EEG targets

are permuted across training samples using a fixed, seed-specific random permutation within each $(\text{task}, \text{source-dataset})$ cell. This preserves the EEG target distribution and number of training samples while breaking the original image–EEG correspondence. The full **EEG-guided** setting uses the original aligned EEG targets together with both reconstruction and contrastive supervision.

B.3 Additional Ablations

The main text reports the control experiments used to isolate the contribution of aligned EEG supervision (Table 3). Here, we provide additional ablations on the optimization configuration and the form of EEG supervision. All experiments are conducted on LIBERO-Long, and results are reported as mean success rate $\pm$ standard deviation across repeated runs.

Table 5: **Additional ablations on LIBERO-Long.** The first three groups vary one optimization hyperparameter at a time while keeping the remaining settings fixed. The final group compares alternative training-time EEG supervision variants under matched robot-training and rollout-evaluation protocols. Results are reported as mean success rate $\pm$ standard deviation across repeated runs. Bold denotes the highest mean within each group, and † marks the default setting. The default result is repeated across groups for reference.

LoRA Learning Rate		**LoRA Rank**		**Batch Size**		**EEG Supervision**	
Setting	**SR (%)**	**Setting**	**SR (%)**	**Setting**	**SR (%)**	**Setting**	**SR (%)**
1.00×10^{-5}	52.00 ± 2.91	4	53.67 ± 1.92	32	51.00 ± 0.92	Raw EEG†	$\mathbf{56.73 \pm 0.70}$
3.00×10^{-5}†	$\mathbf{56.73 \pm 0.70}$	8	53.93 ± 3.00	64†	$\mathbf{56.73 \pm 0.70}$	BrainμTokenizer	54.30 ± 1.40
5.00×10^{-5}	54.93 ± 1.21	16†	$\mathbf{56.73 \pm 0.70}$	128	54.80 ± 2.91	LaBraM	53.50 ± 1.00
1.00×10^{-4}	54.00 ± 2.40	32	53.80 ± 0.80	256	52.73 ± 0.70		
5.00×10^{-4}	55.07 ± 1.17	54	53.53 ± 1.80	512	53.73 ± 1.29		

Optimization sensitivity. We first examine the sensitivity of EEG-guided training to the main optimization hyperparameters by varying one factor at a time while keeping the remaining settings fixed. As shown in Table 5, the default configuration—a LoRA learning rate of 3×10^{-5}, LoRA rank 16, and EEG batch size 64—achieves $56.73 \pm 0.70\%$ success rate and yields the highest observed mean in each of the three sweeps. Reducing the LoRA learning rate to 1×10^{-5} decreases performance to $52.00 \pm 2.91\%$, whereas increasing the learning rate does not improve over the default setting. A similar pattern is observed for LoRA rank and batch size; in particular, reducing the EEG batch size to 32 lowers the success rate to $51.00 \pm 0.92\%$. These results identify the default configuration as a favorable operating point, while also indicating non-negligible sensitivity to the optimization settings.

EEG supervision variants. We next investigate whether pretrained EEG representations provide more effective privileged supervision than the default raw-EEG formulation. Specifically, we replace the default EEG supervision branch with frozen Brainμtokenizer or LaBraM teachers. Each pretrained EEG encoder is used only during training to construct fixed target representations for the auxiliary contrastive objective and receives no gradient from VLA training. The VLA hidden representation is mapped to a normalized query and trained to retrieve the corresponding EEG-derived target using a one-way InfoNCE objective. All EEG-specific teacher and projection components are discarded at deployment, so robot rollout remains entirely EEG-free.

Under the corresponding matched robot-training and rollout-evaluation protocol, the raw-EEG formulation achieves $56.73 \pm 0.70\%$ success rate, compared with $54.30 \pm 1.40\%$ for Brainμand $53.50 \pm 1.00\%$ for LaBraM. Thus, pretrained EEG teachers do not improve downstream rollout performance over direct raw-EEG supervision under the present training protocol, and we retain the raw-EEG formulation as the default.

Loss-weight sensitivity. To further examine the sensitivity of EEG-guided training to loss weights, we conducted a loss weight ablation on LIBERO-Long. We vary the auxiliary EEG loss weight λ and the InfoNCE weight β while keeping the remaining training settings fixed. Here, λ weights the complete auxiliary objective of EEG, while β weights the InfoNCE term within that objective.

Table 6: **EEG loss-weight ablations on LIBERO-Long.** Results are reported as mean success rate $\pm$ sample standard deviation across repeated runs. The bold row denotes the default configuration.

Setting	λ	β	**Success Rate (%)**
Low λ	0.01	0.05	54.20 ± 0.72
High λ	0.20	0.05	53.80 ± 1.71
Low β	0.10	0.01	52.33 ± 0.50
Increased β	0.10	0.02	54.40 ± 1.73
Default	**0.10**	**0.05**	$\mathbf{56.73 \pm 0.70}$

As shown in Table 6, the default configuration ($\lambda = 0.10, \beta = 0.05$) achieves the highest success rate of $56.73 \pm 0.70\%$. With $\beta = 0.05$ fixed, setting λ to 0.01 or 0.20 reduces the success rate to $54.20 \pm 0.72\%$ and $53.80 \pm 1.71\%$, respectively. With $\lambda = 0.10$ fixed, setting β to 0.01 or 0.02 results in success rates of $52.33 \pm 0.50\%$ and $54.40 \pm 1.73\%$, respectively. We therefore retain $\lambda = 0.10$ and $\beta = 0.05$ as the default loss weights for the remaining experiments.

C Extended Discussion

An important distinction in our setting is that EEG is not introduced as an additional sensing modality for robot control. Rather, it provides a training signal that reflects the internal dynamics of a human performing a similar type of physical interaction. Neural activity is closely linked to sensorimotor processing, action selection, and behavioral monitoring (Pfurtscheller & Lopes da Silva, 1999; Gehring et al., 1993; Cisek, 2007), and previous work has shown that brain recordings can help shape representations learned by artificial models (Schwartz et al., 2019). The persistence of performance gain when EEG is removed at inference suggests that EEG serves as a supervision for representation learning, not as an additional input at deployment.

This perspective is particularly relevant to embodied systems, in which intelligence depends on interactions among computation, the body, and the physical environment (Pfeifer & Bongard, 2007). Physical agents must contend with constraints on sensing, latency, computation, and energy, as well as actuator dynamics and uncertain environmental feedback. Biological motor control is similarly organized around the continuous interaction between action and sensory feedback Todorov & Jordan (2002). Prior work on embodied control suggests that physical structure can reduce the complexity required of the controller by exploiting regularities in the body and environment (Montúfar et al., 2015). From this perspective, brain-inspired embodied learning need not reproduce specific neural architectures. Instead, it can investigate whether representations shaped by biological neural dynamics provide useful structure for artificial agents operating under similar physical constraints.

The present results should therefore be viewed as a proof of concept rather than evidence of a specific neural mechanism. A broader question is whether neural supervision captures structure that generalizes across subjects, tasks, and embodiments, and whether that structure can eventually be expressed as more general learning objectives rather than direct EEG prediction. Beyond task success, future work could examine whether such supervision improves data efficiency, robustness to partial observations, adaptation to changing dynamics, or performance under resource constraints. These directions may help clarify when neural signals provide useful information beyond behavioral observations and whether biological neural dynamics can inform representation learning in physically embodied agents.