2026 IEEE/SICE International Symposium on System Integration (SII)
January 11-14, 2026. Cancún, México

Accelerating Teleoperation Skill Acquisition through Visuo-Haptic Replay

Francisco Ribeiro[1,2], Jihoon Park[1,3], Minoru Asada[1,3,4], Erhan Oztop[1,5]

Abstract— **Training humans to control dynamic systems such as prosthetic limbs or teleoperated robots typically requires extensive practice. In this study, we investigate whether passive exposure to skilled control behavior—delivered via synchronized haptic and visual playback—can accelerate visuomotor learning in a non-trivial control task. Using a simulated cart-pole environment, we compare performance between participants who received an initial session of passive visuo-haptic replay and those who directly began with active control. Our results show that passive exposure yields two clear benefits: improved initial performance and a steeper learning curve across subsequent sessions. These findings suggest that passive sensorimotor experience, even without haptic guidance during active control, can support the acquisition of motor skills necessary for dynamic control. This approach may provide a low-effort and scalable training paradigm for enhancing skill acquisition in robotic and assistive technologies.**

I. Introduction

Training users to control complex dynamic systems such as robotic arms, prosthetic limbs, or industrial machinery requires significant practice [1], [2], which may lead to frustration, inefficient use of time and resources. Prior work in *haptic guidance* [3], [4] has shown that physically guiding users along expert trajectories during task execution can accelerate learning by enforcing motor patterns. These approaches typically operate online, blending user input with autonomous controller assistance or constraining movement within a reference trajectory. A general classification of haptic training methodologies are given in [5] which include haptic guidance. Most of these methods find application in rehabilitation and sport science, where the primary goal is to restore or refinement the motor functions. In contrast, the application of haptic training for more general cognitive motor learning tasks, such as visuomotor control, remains limited. This study aims to explore the potential of *offline visuo-haptic demonstration* for improving learning in dynamic control scenarios where the mapping between perception and action must be acquired for novel tasks learning.

[1] All authors are with the Symbiotic Intelligent Systems Research Center, Institute for Open and Transdisciplinary Research Initiatives, Osaka University, Japan. {jihoon.park, erhan.oztop, asada}@otri.osaka-u.ac.jp

[2] Francisco Ribeiro is also affiliated with Instituto Superior Técnico – University of Lisbon, Lisbon, Portugal. {francisco.p.ribeiro}@tecnico.ulisboa.pt

[3] Jihoon Park and Minoru Asada are also affiliated with Center for Information and Neural Networks, National Institute of Information and Communications Technology, Osaka, Japan.

[4] Minoru Asada is also affiliated with International Professional University of Technology in Osaka, Japan.

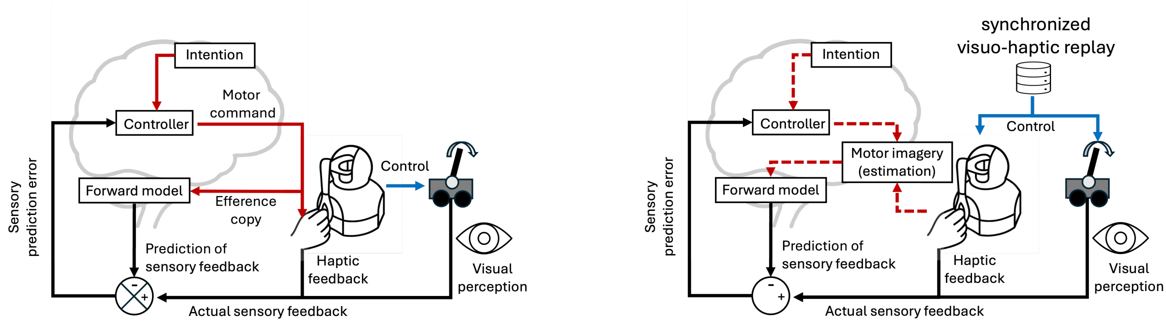
[5] Erhan Oztop is also affiliated with Ozyegin University, Istanbul, Turkey.

In a recent report [6], it has been shown that the so called ceiling effect for expert piano players can be overcome by haptic practice enforced to the piano players by a finger exoskeleton. While this study seems to cover our research interest, there is a critical difference. In piano playing what is learned is a complex sequence of motor activation patterns, an example of a new *motor memory* formation, which involves storing, and later recalling motor patterns as a function of time. On the other hand, *visuomotor learning*, is characterized by associating the visually sensed current state of the environment (e.g., pose of an object) with appropriate motor commands. Visuomotor learning is critical for tasks that require online adaptation or correction, which may involve complex dynamics such as driving a car.

The existing works on the effect of passive visuo-haptic feedback prior to visuomotor learning are limited, and focus on simple tasks such as rotation adaptation [7] or do not strictly use passive visuo-haptic playback, but rather have human participants learn with haptic guidance [3]. To the best of our knowledge, our study is the first to use passive haptic and visual playback of skilled execution performance to speed up the acquisition of visuomotor control skills in non-trivial dynamical control tasks. It is worthwhile to emphasize that our approach differs from the well established approaches that intervene during task execution such as haptic guidance [3], [4] and shared control [8], [9], since we deliver expert control experience as a synchronized visual and haptic information before the task learning begins with no intervention during human learning.

In addition to the scientific motivation, this study aims to develop practical training methods to allow users to efficiently acquire control skills for artificial devices, including prosthetic limbs, teleoperated robots, and industrial machinery. As a proof of concept, we use the cart-pole task in simulation where the users must move the cart in one axis so as to keep the pole close to vertical. By comparing the learning curves of participants who receive passive replay of expert control commands to those who do not, we assess whether passive visuo-haptic exposure, can boost subsequent learning of the cart-pole control task. We hypothesize that passive exposure facilitates skill acquisition by engaging *motor imagery* [12], [13] through the haptic sensation and movement induced by the control interface. Such engagement can enable forward model learning [14]–[16]. Fig. 1 illustrates the usual skill acquisition (Active Learning) on the left (a) together with proposed Passive Learning mechanism on the right (b).

978-1-6654-5784-2/26/$31.00 ©2026 IEEE



(a) Active visuomotor learning.

(b) Passive visuo-haptic replay.

Fig. 1: How passive visuo-haptic replay may facilitate visuo-motor skill acquisition is illustrated. During Active visuo-motor learning, the central nervous system (CNS) uses several mechanisms to learn to generate motor commands to succeed in the target task [10], [11]. One mechanism relies on a forward model that predicts sensory consequences of motor commands and updates them to minimize the prediction error between predicted and actual sensory feedback (a). In the Passive visuo-haptic replay, the hand of the user and the cart-pole are moved externally (blue arrows) rather than by CNS-generated motor commands. Yet, the haptic feedback and the passive hand movements may induce ‘motor imagery’ (red dashed arrows) and enable forward model learning thereby accelerating skill acquisition in the subsequent active learning trials (b).

In the rest of the paper, we first give the details of the learning task, control interface and the general experiment design in Methods. Then, in Experiments and Results we detail the experimental procedure and present the results obtained. Finally, in Conclusion we point out the limitations and future directions.

Methods

Apparatus and Task

In the experiments, participants interacted with a simulated cart-pole environment rendered in real time on a computer display (see Fig. 2). The task required maintaining the pole within 45 degrees offset from the vertical while preventing the cart from leaving the bounded horizontal area. If either the pole is passed 45 degrees or the cart pole leaves the play area the trail is ended. The control input was provided via a commercially available haptic interface (*Touch* from 3D Systems). In the teleoperation experiments, the velocity of the haptic pointer held by the participants was used to generate force commands on the cart-pole system.

During visuo-haptic replay the device replays pre-recorded force trajectories corresponding to successful task executions whereas during teleoperation, the hand velocity of the participants are scaled to control actions in the range of $[-1, 1]$, which are then converted into forces applied to the simulated cart (see the *Experimental Design* subsection). The simulation provides real-time visual feedback on the computer screen in both visuo-haptic replay and active control conditions. To compute the force applied to the cart, we extract the velocity in the table plane, i.e., x and y axes communicated by the haptic device at 100 Hz. We include the y component as users often use wrist articulation yielding radial motion in the table plane leading to displacement in y axis in addition to the main axis of x. The main axis movement is also used to determine the sign of the control command. Finally, the z component, measuring the motion perpendicular to the interaction plane is not used in control command computation. Thus we have:

Fig. 2: The experimental interface is shown. The setup allows delivering passive replay of prerecorded control trajectories through the haptic device held by the users. Alternatively the users can control the cart-pole in real-time.

$$v_{\text{hand}}(t) = sign(v_x(t))(v_x^2(t) + v_y^2(t))^{0.5} \tag{1}$$

where t indicates the time step. $v_{hand}(t)$ is then mapped to an action signal, $a(t)$ that is used by the cart-pole simulator for controlling the cart-pole. We use the following empirically tuned equation to make the mapping.

$$a(t) = Cv_{\text{hand}}^{1.5}(t) \tag{2}$$

where C is a tuned constant which is kept fixed through out the experiments for aligning the velocity range of the haptic device with the action range allowed by the simulation ($C \approx 7.78 \times 10^{-5}$). The exponentiation in Equation 2 introduces a desirable nonlinear sensitivity that facilitates an intuitive mapping: finer control at low velocities and amplified responses for higher velocities.

To suppress the effects of high-frequency noise in the control signal, a moving average filter is applied with a window size of N ($N = 3$ in the experiments reported). In addition, to assist users in keeping the cart in the game area we inject a small damping term, compensating for overshoots. With these, the net action driving the cart-pole simulation is giving as follows:

$$a_{\text{net}}(t) = \frac{1}{N}\sum_{i=0}^{N} a(t-i) - 0.05 v_{\text{cart}}(t) \tag{3}$$

Overall, the combination of hand velocity mapping (Equation 2), action filtering and cart-velocity based dampening (Equation 3) helps ensure responsive but stable control during teleoperation.

Cart-Pole Simulation and Dynamics

The cart-pole environment is based on the classical control problem originally described by Barto, Sutton, and Anderson, and implemented following reference implementation given by Sutton and Barto [17]. This system consists of a cart of mass M and a pole of mass m and length $2l$, hinged to the cart. The objective is to balance the pole upright by applying horizontal forces to the cart. The dynamics are governed by:

$$\ddot{x} = \frac{F + m\sin\theta\left(l\dot{\theta}^2 + g\cos\theta\right)}{M + m(1-\cos^2\theta)}$$

$$\ddot{\theta} = \frac{-F\cos\theta - ml\dot{\theta}^2\cos\theta\sin\theta - (M+m)g\sin\theta}{l\big(M + m(1-\cos^2\theta)\big)}$$

where M and m are taken as 1.0kg and 0.07kg respectively, and the half-pole length l is set to 1.0m. To increase the novelty of the dynamics, g, the gravitational acceleration is set to half-earth level. Finally, F is the force applied to the cart which is computed by $F = 50 \times a$ with a being the simulator action input in the range of $[-1, 1]$.

The simulation was implemented in Python based on a continuous-action version of the cart-pole task included in the Gymnasium toolkit [18], which was originally published by Matt Alan Wright [19], and modified by us to work with the current Gymnasium libraries.

Experimental Design

The experiment involved passive and active experience sessions as described in the following:

Passive Visuo-Haptic Experience Session: Participants passively hold a haptic joystick while it replayed pre-recorded force trajectories corresponding to successful task executions. Each participant goes through 30 trials, each lasting 15 seconds. All trials begin with the cart centered at zero linear and angular velocities, and the pole angle, θ is initialized to one of the angles from $\mathcal{A} = \{-3°, -2°, -1°, 1°, 2°, 3°\}$ depending on the shuffled order they receive for experiencing expert control playback.

Active Control Session: Participants actively control the cart-pole system without haptic feedback. Hand velocity, measured via the Touch device, is mapped to control actions in the range [-1, 1], which are then mapped into forces applied to the cart. Each participant is asked to complete 60 trials each of which may last up to a maximum of 20 seconds. As in the visuo-haptic experience sessions, the trials begin with the cart centered, and the pole angle initialized with $\theta \in \mathcal{A}$, while ensuring each initial condition is presented exactly ten times per participant.

Participant Groups

Participants were randomly assigned to one of two experimental groups. *Group 1* participants went through two consecutive active control sessions without prior passive visuo-haptic exposure (*Active* + *Active* group), whereas *Group 2* participants were given a passive visuo-haptic experience session prior to active control sessions (*Passive* + *Active* + *Active* group). The participants were allowed to rest between sessions for about one minute.

With this design, we can make the following comparisons. (i) By focusing on the first two session of each group, we can perform *Active* + *Active* vs. *Passive* + *Active* comparison as one the the main results of this study. In addition, (ii) we can consider having passive exposure or not as an independent variable and check its effects on the upcoming active learning sessions. This can generate useful information as to whether a cognitively less demanding passive session can boost active learning of the participants.

II. Experiment and Results

At the start of the experiment, all participants were given approximately one minute to familiarize themselves with the haptic joystick and the visual rendering of the cart-pole system. Participants were instructed to maximize the duration the pole remained within the allowed range and the cart stayed within the game area. Each experimental group consisted of five participants, for a total of ten participants in the study (9 men, 1 woman with mean age, 25.7 ± 3.5). These participants were recruited from university students and research staff who were not informed of the experimental details beforehand. To analyze the effect of visuo-haptic playback on task performance, all the haptic device movements of the participants and the task performances were recorded.

State Space Trajectories. The visual inspection of the state space trajectories of the participants confirm that, even after learning, their control performance did not reach the level of the expert demonstration. In Fig. 3, we show representative trajectories in the 3D state space of the cart-pole system, defined by pole angle, angular velocity, and cart position. The green and cyan curves correspond to typical trials from the expert demonstration data. Expert trajectories stay clustered near the origin, particularly along the cart position (Z-axis), indicating well stabilization of the pole with minimal cart displacement signature of a good control. In contrast, typical participants trajectories at the end of *Active* + *Active* (blue), *Passive* + *Active* (orange), and *Passive* + *Active* + *Active* (red) training exhibit broader excursions across the state space indicating less efficient control and more frequent deviations from the upright equilibrium. Especially high deviation in the Cart position (Z) axis shows that the subjects had difficulty keeping the cart in the center region and thus failed. The expert trajectories span 15 seconds whereas Active + Active and Passive + Active condition trajectories last 5.10 and 5.75 seconds respectively. The sample trajectory from the Passive + Active + Active condition shows 8.89 second survival performance, a marked increase compared to the other conditions. These times correspond to the path lengths shown in the state space plot (Fig. 3). The trajectories for the passive visuo-haptic exposure conditions (red and orange curves) exhibit more cyclic patterns resembling the expert trajectories, corroborating the longer survival times and indicating better control compared to the Active + Active condition.

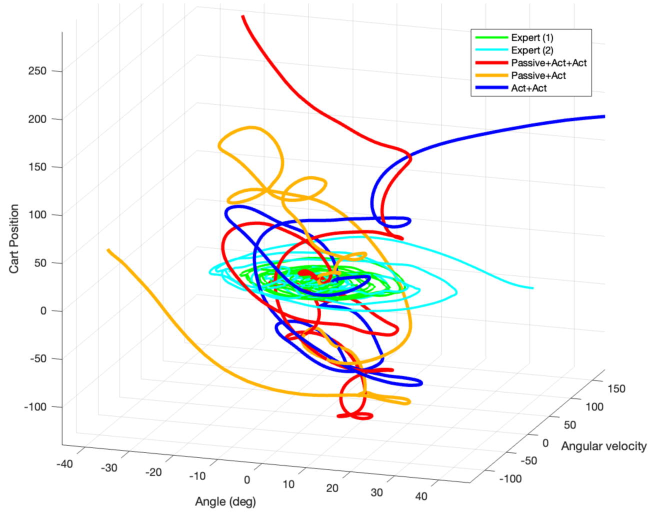


Fig. 3: Example trajectories in the 3D state space of the cart-pole system, plotted as a function of pole angle (deg), angular velocity, and cart position. Two typical expert trajectories are shown in green and cyan, alongside example trajectories from participant from the Passive + Active + Active group (orange and red curves) and Active + Active group (blue curve). The expert trajectories are tightly clustered near the origin, reflecting stable control around the upright equilibrium. In contrast, trajectories from the experimental groups show broader excursions, indicating greater variability and less consistent stabilization of the system.

Average survival times. To evaluate the impact of haptic playback on task performance, we compared average survival times across different sessions (see Fig. 4). Participants in the Passive + Active group outperformed those in the Active + Active group with small margin (Fig. 4c), despite having engaged in only one active control session. However, statistical significance cannot be determined due to small number of participants. This finding suggests that prior exposure to successful force trajectories through passive haptic playback may facilitate more efficient learning than repeated active control alone. Notably, both groups experienced the cart-pole environment in two sessions, controlling for total visual exposure. Thus, the performance gain observed in the Passive + Active group can be more confidently attributed to the effect of synchronized visuo-haptic playback.

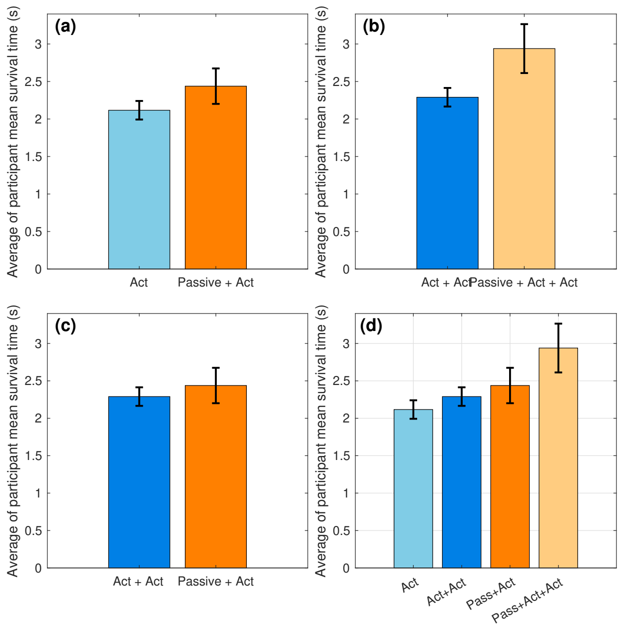


Fig. 4: As learning performance *mean survival times* are shown for different comparisons. (a) The effect of passive exposure to visuo-haptic playback. (b) The continued effect of the passive exposure in the next learning session. (c) The potential of substituting an active learning session with passive visuo-haptic exposure. (4) All experiment conditions together. Error lines show standard error of the mean (SEM).

In the comparisons of *Active* vs. *Passive + Active* and *Active + Active* vs. *Passive + Active + Active*, the performance-enhancing effect of having a passive exposure before going through active learning can be seen in Fig. 4a-b. It is interesting to note that the increment in the second active learning session is more pronounced (see also Fig. 5).

It is worth noting that participants in *Passive + Active + Active* group experienced the simulated environment across three sessions, whereas the *Active + Active* group went through only two. While the initial session of the former group did not involve active control, the additional observational exposure to the task could have supported performance by increasing familiarity or aiding in the development of mental models. We discuss this issue in the Discussion section and argue that this might not be a major contributor to the higher performance observed.

Improvement across sessions. To assess the second-order effect of learning rate change, we compared the difference in task performance between two active sessions across participant groups. Fig 5 shows that the group receiving initial passive exposure demonstrated a substantially greater improvement in mean survival time compared to the group that completed two consecutive active sessions. This extends earlier findings of enhanced performance in the Passive + Active group, and now reveals that initial exposure also facilitates a steeper learning trajectory. While, current result show that passive synchronized visuo-haptic observation can enhance subsequent active learning, the weights of the individual contributions from the components of haptics, vision and their contingency cannot be determined in this comparison.

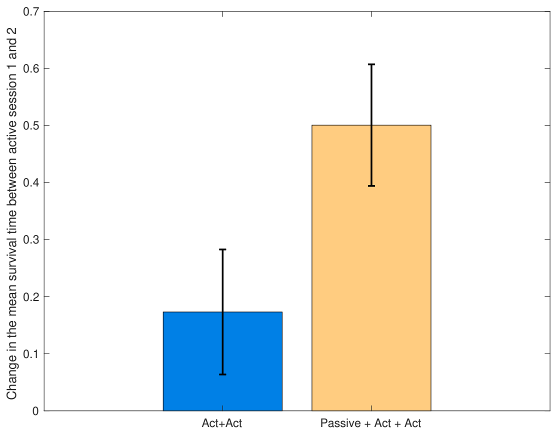


Fig. 5: The effect of passive experience on the learning rate of active learning is illustrated. Having a passive experience prior to engaging in active learning not only improves the learning level (see Fig.4) but also increases the learning rate. Bars represent mean survival time change between the groups of *Active + Active* and group *Passive + Active + Active*. The error lines indicate standard error of the mean (SEM).

Post-experiment questionnaire. We also administered a post-experiment questionnaire (see Table I) that probed the participants' perception of the task to see whether there is any perceived difference between the experimental groups. The answers to questionnaire mostly showed similar patterns

TABLE I: Post-Experiment Questionnaire administered (compacted for the sake of space). Participants were given A4 paper and asked to mark one of the circles indicating the rating between strongly disagree (1) and strongly agree (7).

Statement	1–7
Q1. I felt that I was in control of the cart-pole system during the active control stage.	ooooooo
Q2. The movements of the cart-pole system felt like they were a direct result of my own actions.	ooooooo
Q3. The cart-pole system felt as if it had its own intentions or was alive.	ooooooo
Q4. The movements of the cart-pole system seemed autonomous or self-driven at times.	ooooooo
Q5. I felt that the cart-pole system responded reliably to my actions.	ooooooo
Q6. It felt possible to influence the outcome of each trial through my own movements.	ooooooo
Q7. Over the course of the trials, I felt that my ability to control the cart-pole system improved.	ooooooo
Q8. I could notice specific strategies or adjustments I made to improve my performance.	ooooooo
Q9. I believe that, given more time or practice, I could successfully perform the task for longer durations.	ooooooo
Q10. I would be interested in participating in a follow-up experiment of this kind in the future.	ooooooo

among the groups (see Fig 6). One exception is Q3, where we asked "The cart-pole system felt as if it had its own intentions or was alive.". Intriguingly, the group who have experienced the cart-pole moving itself and feeling it through haptic feedback, i.e., Passive + Active + Active group, responded with a (almost) significantly lower rating (p-value t-test $<$ 0.073). This suggests that passive experience enhanced the perception of the system as mechanical one with predictable responses. Corroborating with this, the first two questions probing agency and controllability show a higher average score for the Passive + Active + Active group. Finally, the slightly higher score of the same group (p-value t-test $<$ 0.21) in Q10 that asks whether the participant would like to join a follow-up experiment may suggest that the passive visuo-motor experience delivery increases user engagement and motivation for learning.

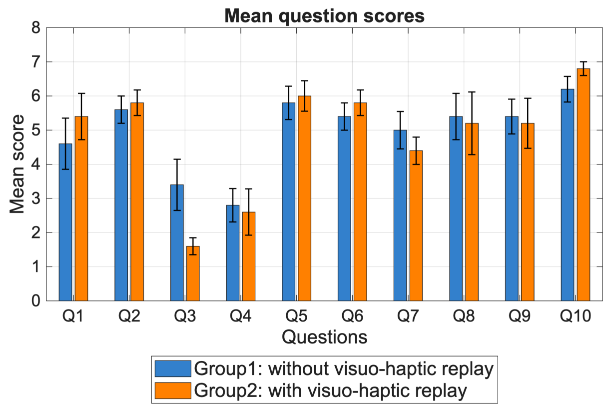


Fig. 6: The average response of participants to the post-experiment questionnaire (see Table I is shown). The error lines indicate standard error of the mean (SEM).

Overall results. These findings provide preliminary support for the use of synchronous visuo-haptic playback as a learning aid for learning the control of complex dynamic tasks, particularly in contexts where reducing human effort or training time is critical. This approach could be especially valuable in domains where control policies are difficult to verbalize or demonstrate visually, such as in surgical robotics, assistive devices, or high-DOF manipulators.

III. Discussion

The comparison between the Passive + Active and Active + Active conditions revealed a small but consistent performance advantage for participants who received passive visuo-haptic playback prior to active control (Fig. 4). While the difference was modest, it suggests that exposure to successful force trajectories may support more efficient initial learning in teleoperation tasks. Given the limited number of participants (five per group), further studies with larger sample sizes are needed to better characterize the robustness of this effect.

A more pronounced improvement was observed in the Passive + Active + Active condition, which achieved the highest average performance overall. Furthermore, the between-session improvement was also strikingly higher for the participants who received initial exposure to passive feedback

(Fig. 5). These improvements may indicate that combining passive visuo-haptic experience with active practice enhances learning beyond what repeated active control alone can offer. While the additional passive session introduces the possibility of increased visual task familiarity, we argue that the nature of the task limits the likelihood that visual exposure alone accounts for the observed gains. Research on the mirror neuron system suggests that action observation can facilitate motor learning, particularly for tasks within the natural human action repertoire—such as reaching, grasping, or tool use [20], [21]. However, the cart-pole balancing task involves abstract, non-anthropomorphic dynamics that are far removed from everyday motor behavior. As such, it is unlikely to engage the mirror system to a meaningful extent [22], reducing the plausibility that visual input alone drove the performance improvements.

To further disentangle the contributions of visual and haptic modalities during passive exposure, future studies could include a control group receiving visual-only passive exposure without haptic feedback. Furthermore, incorporating eye-tracking measurements during passive exposure could provide insights into how participants allocate their visual attention, revealing which features they focus on while observing the task and the extent to which they maintain attention on the task. This would help clarify the extent to which passive haptic playback contributes uniquely to performance improvements. Future studies incorporating a visual-only passive exposure condition would help clarify the specific contribution of haptic feedback in facilitating performance improvements, and better establish the role of passive experience in learning to control novel complex dynamical systems.

IV. CONCLUSIONS

This study suggests that passive visuo-haptic exposure to expert demonstrations has the potential to enhance the acquisition of dynamic control skills in teleoperation tasks. Participants who experienced this form of offline replay before active control exhibited both higher initial performance and a steeper learning trajectory compared to those trained through active control alone. These benefits were elicited even though active sessions involved no haptic guidance, highlighting the effect of prior passive exposure to subsequent visuomotor learning~~.~~ and suggesting potential implications for tasks in which active control for training is less desirable due to the risk of higher costs, or safety concerns (e.g., damage of teleoperated robots).

Beyond potential applications in training for assistive robotics or prosthetics, these findings encourage further exploration of passive experience as a low-effort, scalable strategy for skill acquisition. Future work should examine the relative contributions of visual and haptic modalities and assess generalizability to more complex and real-world tasks.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number JP23K24926, and Grant Number JP25H01236.

REFERENCES

[1] C. W. Antuvan, M. Ison, and P. Artemiadis, "Embedded human control of robots using myoelectric interfaces," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 4, pp. 820–827, 2014.

[2] S. Pooladvand, H. Taghaddos, A. Eslami, A. N. Tak *et al.*, "Evaluating mobile crane lift operations using an interactive virtual reality system," *Journal of Construction Engineering and Management*, vol. 147, no. 11, p. 04021117, 2021.

[3] J. Bluteau, S. Coquillart, Y. Payan, and E. Gentaz, "Haptic guidance improves the visuo-manual tracking of trajectories," *PLOS ONE*, vol. 3, no. 3, p. e1775, Mar. 2008.

[4] L. Marchal-Crespo and D. J. Reinkensmeyer, "Review of control strategies for robotic movement training after neurologic injury," *Journal of NeuroEngineering and Rehabilitation*, 2009.

[5] C. K. Williams and H. Carnahan, "Motor learning perspectives on haptic training for the upper extremities," *IEEE Transactions on Haptics*, vol. 7, no. 2, p. 240–250, Apr. 2014.

[6] S. Furuya, T. Oku, H. Nishioka, and M. Hirano, "Surmounting the ceiling effect of motor expertise by novel sensory experience with a hand exoskeleton," *Science Robotics*, vol. 10, no. 98, p. eadn3802, Jan. 2025.

[7] E. H. F. van Asseldonk, M. Wessels, A. H. A. Stienen, F. C. T. van der Helm, and H. van der Kooij, "Influence of haptic guidance in learning a novel visuomotor task," *Journal of Physiology, Paris*, vol. 103, no. 3–5, p. 276–285, 2009.

[8] A. D. Dragan and S. S. Srinivasa, "A policy-blending formalism for shared control," *The International Journal of Robotics Research*, 2013.

[9] N. Amirshirzad, A. Kumru, and E. Oztop, "Human adaptation to human–robot shared control," *IEEE Transactions on Human-Machine Systems*, vol. 49, no. 2, p. 126–136, Apr. 2019.

[10] J. W. Krakauer and P. Mazzoni, "Human sensorimotor learning: adaptation, skill, and beyond," *Current Opinion in Neurobiology*, vol. 21, no. 4, p. 636–644, Aug. 2011.

[11] K. Doya, "Complementary roles of basal ganglia and cerebellum in learning and motor control," *Current Opinion in Neurobiology*, vol. 10, no. 6, p. 732–739, Dec. 2000.

[12] M. Jeannerod, "The representing brain: Neural correlates of motor intention and imagery," *Behavioral and Brain Sciences*, vol. 17, no. 2, p. 187–202, 1994.

[13] M. Lotze, P. Montoya, M. Erb, E. Hülsmann, H. Flor, U. Klose, N. Birbaumer, and W. Grodd, "Activation of cortical and cerebellar motor areas during executed and imagined hand movements: an fmri study," *Journal of Cognitive Neuroscience*, vol. 11, no. 5, p. 491–501, 1999.

[14] E. Oztop, D. Wolpert, and M. Kawato, "Mental state inference using visual control parameters," *Brain Research. Cognitive Brain Research*, vol. 22, no. 2, p. 129–151, Feb. 2005.

[15] M. Kawato, "Internal models for motor control and trajectory planning," *Current Opinion in Neurobiology*, vol. 9, no. 6, p. 718–727, Dec. 1999.

[16] D. M. Wolpert, Z. Ghahramani, and M. I. Jordan, "An internal model for sensorimotor integration," *Science (New York, N.Y.)*, vol. 269, no. 5232, p. 1880–1882, 1995.

[17] R. S. Sutton and A. G. Barto, "pole.c: Classic cart-pole control implementation," 1998, http://incompleteideas.net/sutton/book/code/pole.c.

[18] F. Foundation, "Gymnasium: A toolkit for developing and comparing reinforcement learning algorithms," 2023, https://github.com/Farama-Foundation/Gymnasium.

[19] M. A. Wright, "gym-continuous-cartpole," 2018, adapted by the authors in 2025 to work with Gymnasium and updated Python libraries. Available at https://github.com/MattAlanWright/gym-continuous-cartpole.

[20] G. Rizzolatti and L. Craighero, "The mirror-neuron system," *Annual Review of Neuroscience*, vol. 27, p. 169–192, 2004.

[21] G. Buccino, S. Vogt, A. Ritzl, G. R. Fink, K. Zilles, H.-J. Freund, and G. Rizzolatti, "Neural circuits underlying imitation learning of hand actions: An event-related fmri study," *Neuron*, vol. 42, no. 2, p. 323–334, Apr. 2004.

[22] J. M. Kilner and C. D. Frith, "Action observation: Inferring intentions without mirror neurons," *Current Biology*, vol. 18, no. 1, p. R32–R33, Jan. 2008.