Available Online: 23 October 2026 DOI: 10.54254/2977-3903/2026.38186

Beyond initial training: experience-driven continual adaptation in robotic manipulation

Yaohao Guo

School of Management, University College London, London, UK

gordonguo66@gmail.com

Abstract. Robots can learn useful manipulation skills from demonstrations and large datasets, but their initial training may not prepare them for every object or task they encounter. This review examines how robots can use subsequent physical experience to adapt to unfamiliar conditions and learn new manipulation skills. It draws on research in imitation learning, real-world reinforcement learning, offline-to-online learning and continual learning, including recent work on generalist robot policies. Silver and Sutton's Era of Experience provides a starting point for considering learning as an ongoing part of a robot's operation. The reviewed studies show several ways forward: learning from corrective guidance, updating an existing policy with new attempts, and learning to combine familiar skills to solve more complex tasks. These approaches depend on useful starting knowledge and feedback about action outcomes. Human corrections can guide learning during physical trials, and policies updated with earlier and newly collected experience can improve performance after deployment. Their wider use remains limited by the cost of physical practice, unsafe exploration, unreliable rewards and interference with earlier learning. Evaluations should retest earlier tasks after each learning stage and record human assistance. The review argues that post-deployment experience can extend a robot's capabilities, while identifying reduced human dependence and sustained learning across changing tasks as important research priorities.

Keywords: robotic manipulation, experience-based learning, post-deployment adaptation, reinforcement learning, continual learning

1. Introduction

Robotic manipulation involves acting on objects, from grasping a container to assembling a component. Learning from demonstration and imitation learning allow robots to acquire such behaviours from examples [1, 2]. Larger datasets have extended this approach. RT-1 learns many manipulation tasks within one model, while Open X-Embodiment investigates how experience collected across different robots can support shared learning [3, 4]. These developments provide useful starting capabilities, but also raise a question about what happens after initial training.

Consider a robot that can grasp individual objects but cannot yet use a tool to complete an unfamiliar task. Further demonstrations may help. The robot might also learn by trying actions, observing their effects and receiving guidance when it becomes stuck. The interest lies in how this experience can help it learn to perform

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

a task it cannot currently complete. Silver and Sutton's Era of Experience presents a broader vision of agents developing through continuing interaction with their environments [5].

This paper asks: how can physical experience after initial training help manipulation robots adapt to unfamiliar conditions and learn further skills, and what currently limits this process? A focused narrative review connects foundational work with recent studies available by September 2026. The analysis considers prior capabilities, learning procedures, human support and experimental outcomes. It focuses on learning after initial competence, particularly during deployment; controlled training studies are included where they explain relevant methods. It does not provide an exhaustive survey or report original experiments.

2. Pre-deployment robot learning and its limitations

2.1. The role of pre-deployment learning

Initial training gives a robot a policy, meaning a learned rule for choosing actions from observations. Demonstrations show useful ways of performing a task and reduce the need to discover every movement through trial and error [1, 2]. They can therefore prepare the robot for further learning: knowing how to reach and grasp makes later attempts more directed than starting without these skills. Initial training need not teach every future task to provide this useful starting point.

Broader training can also support generalisation. RT-1 studies the use of a large demonstration dataset to learn multiple tasks, and RT-X shows that data from other robotic platforms can improve performance [3, 4]. A robot may therefore succeed in a new situation without further learning. Initial training has two roles: providing abilities that can be used immediately and supplying knowledge on which later learning can build.

2.2. Distribution shift and incomplete training coverage

Distribution shift occurs when a robot encounters conditions that differ from those represented in its training data. An object may look familiar but have a different weight or surface friction. The robot must then deal with action outcomes that its earlier experience did not adequately cover. Training coverage and distribution shift are closely connected: one concerns what the dataset contains, while the other concerns how operating conditions may differ from the training data.

Some difficulties involve more than adjusting an existing movement. A robot may know how to pick up a tool without knowing how to use it. This example raises a question about both adapting existing skills and learning new task behaviours. Which kind of learning is possible depends on the robot's starting knowledge and the experiences it can collect.

3. The role of experience in physical robot learning

Physical experience includes what a robot observes, what it does and what happens next. For example, an attempted insertion may reveal that a component is misaligned. Reinforcement learning uses feedback expressed as rewards to improve a policy through experience [6]. Simply repeating a task is insufficient: the learning process must use the outcomes to change how future actions are selected.

Silver and Sutton argue that agents should learn through ongoing interaction and discover solutions beyond those directly supplied by human examples [5]. Their AlphaProof example illustrates learning beyond human-provided knowledge, although it concerns mathematical proofs rather than robots [5]. For manipulation, a child learning to use an unfamiliar utensil offers a useful analogy for practice, without implying identical

learning processes. A robot must test actions on physical objects and use feedback from their outcomes to adjust its behaviour. Human guidance can remain part of this experience.

The Robot Kindergarten initiative proposes a setting for safe exploration, repeated physical interaction and timely feedback [7]. Its cited source is a project announcement, not an experimental study. It is included as an emerging initiative rather than evidence that robots can already learn indefinitely. Practice in such a facility may also precede deployment.

4. Learning from experience after deployment

4.1. Reinforcement learning through interaction

Reinforcement learning uses rewards from attempted actions to improve how a robot performs a task [6]. However, feedback can be limited when success is rare. Consider an insertion task that gives a reward only when the part is fully inserted. If every attempt fails, the reward does not distinguish a nearly correct movement from a badly misaligned one. The robot may then struggle to discover which actions lead towards success.

HIL-SERL combines initial demonstrations with the robot's own attempts and human corrections. When the robot becomes stuck, a person can briefly take control and correct its movement. The system records these corrections alongside the robot's attempts and uses both to train the policy that controls its actions [8]. Human help therefore contributes to learning beyond the attempt in which it is given. The study reports that manipulation performance improves while intervention rates decline during training [8].

Bhatia et al. study how a robot can learn to complete a new task using movements it already knows. Their Semantic Action Reinforcement Learning (SARL) method uses a pretrained robot model that responds to language commands. A separate learning component chooses which command to give it at each stage. For example, the robot may fail to follow a complex instruction but respond successfully to a simpler command such as "move down and grasp" [9]. SARL tries commands on the robot and uses the outcomes to learn which ones help complete the task. In the real-robot experiments, humans also provide initial demonstrations of command sequences [9]. Together, these studies show how human corrections can guide learning during physical trials, while selecting effective commands allows existing movements to support new tasks [8, 9].

4.2. Offline-to-online learning and policy adaptation

Offline-to-online learning connects learning from an existing dataset with further learning through interaction. Earlier data provide useful examples; new attempts reveal what happens in the current setting. Lee et al. show that a policy can initially get worse during this transition. Their method reuses earlier data relevant to the current behaviour and avoids overly optimistic initial estimates of how well unfamiliar actions will work. In simulated control tasks, this supports more stable improvement during further interaction [10].

Julian et al. examine this connection on a real grasping robot. They introduce changes including unfamiliar transparent objects and altered gripper geometry, collect new grasp attempts, and fine-tune an existing policy using old and new data. For transparent bottles, the reported grasping success rate rises from 49% before adaptation to 66% after fine-tuning [11]. Using the same dataset of 800 new grasp attempts, a randomly initialised policy achieved only 27% success [11]. This comparison shows the value of starting from learned grasping behaviour. Updates occur after data collection, so adaptation here involves separate periods of practice and training.

Wang et al.'s 2026 Learning While Deploying preprint extends this approach to a shared vision-language-action policy, which maps images and language instructions to robot actions. Sixteen robots perform eight

types of manipulation task. The system records their attempts, including occasions when human operators take control to correct their movements. The policy is updated using earlier and newly collected experience, then returned to the robots. Online learning improves the reported task scores over the offline version [12]. This makes deployment itself a source of further training, although the study still uses a specified set of tasks. Both studies use prior learning as a starting point for adaptation. Julian et al. update a policy after collecting grasp attempts; Wang et al. repeatedly combine new fleet experience with earlier data to update a shared policy [11, 12].

4.3. Continual and lifelong robot learning

Continual learning aims to let a robot learn new tasks over time while retaining useful earlier skills [13]. Earlier learning may also help it learn later tasks. For example, experience of opening one drawer might help a robot learn to open a different drawer. After learning the second task, it should still be able to open the first drawer.

Julian et al. fine-tune the same grasping policy across several changed conditions and report that the original grasping ability is retained [11]. In these experiments, the robot continues learning from its existing policy as conditions change.

LIBERO examines knowledge transfer and forgetting in sequences of simulated manipulation tasks learned from demonstrations [14]. In these experiments, methods that reduced forgetting generally performed worse at learning later tasks than sequential fine-tuning, which trains the same model on one task after another [14]. This does not imply an unavoidable trade-off. It shows why reduced forgetting alone is insufficient to judge a learning method. For robots learning through physical interaction, evaluations should therefore measure performance on each new task and retest previously learned tasks. This would reveal whether better retention is accompanied by successful new learning.

5. Practical limits of learning from physical experience

5.1. Interaction costs and sample efficiency

Physical practice requires robot time, and failed attempts may require someone to restore the scene. Tang et al. identify sample efficiency and stable real-world learning as continuing priorities [15]. A method's usefulness therefore depends on both the improvement achieved and the effort needed to obtain it.

The 2026 PARTS preprint investigates concentrating practice on difficult subtasks. A pretrained policy handles the parts it already performs, while reinforcement learning trains corrections at identified failure points. The authors report improved completion of the full tasks [16]. This offers a way to avoid repeatedly practising every step. Humans still identify the difficult subtasks during setup and reset the scene when the robot cannot do so. Fewer robot attempts do not necessarily mean less human work. Evaluations should record both robot practice time and the time people spend helping or putting objects back in place.

5.2. Safety, feedback and retention

Exploration can damage objects or the robot. Safe reinforcement learning research therefore considers both learning objectives and restrictions on exploratory behaviour [17]. A penalty received after a collision cannot undo that collision. Physical limits and human intervention may be needed while a new behaviour is being learned.

The robot also needs a dependable indication of success. HIL-SERL uses trained classifiers to recognise task completion [8]. If such a classifier misjudges an unfamiliar object, the robot may learn from incorrect feedback. New tasks can therefore require changes to the learning setup as well as to the policy.

Finally, updates can cause catastrophic forgetting, where learning new behaviours reduces performance on earlier ones. Replaying earlier examples or limiting changes to learned parameters can help, although these approaches introduce storage and adaptation trade-offs [13]. Checks on earlier tasks are needed to establish what the robot has retained.

6. Discussion: combining initial training with later experience

Initial training should be judged partly by how well it prepares a robot to keep learning. Useful starting abilities can make exploration more directed and reduce the amount of new physical data needed [9, 11]. The reviewed studies do not identify a set of basic skills required by every robot. The practical aim is to provide enough task-relevant ability for further attempts to produce useful feedback.

The need for further learning depends on what the robot cannot yet do. Changed object properties may require adjusting a familiar action. A new task may require combining existing actions differently or learning a missing movement. This distinction helps determine what experience to collect.

Reducing human involvement requires more than reducing demonstrations. Robots may still need help recovering from failures, resetting objects and recognising success [8, 16]. A useful research goal is to identify when assistance is necessary and whether that need declines as learning progresses.

Evaluation should track new-task performance and retest earlier tasks after each learning stage. It should also record practice time and human assistance. Together, these measures would show whether a robot gains usable abilities without losing earlier skills or requiring increasing supervision.

7. Conclusion

Physical experience can help manipulation robots extend what they learned before deployment. New attempts reveal how actions work in the current environment, and their outcomes can guide changes in behaviour. The reviewed research supports several routes: refining an existing policy, learning with occasional human corrections, and learning to combine available skills into more complex task behaviour. These routes address different gaps in initial competence and should be selected accordingly.

The recent research direction is increasingly concerned with connecting broad pretrained capabilities to further learning from real interaction. Shared experience across robot fleets and practice focused on difficult subtasks offer ways to make this process more useful. They also show why initial training remains important: it supplies behaviour from which further exploration can begin. Human guidance can support this process without specifying every action throughout learning.

Sustained adaptation still depends on reliable feedback, manageable physical costs and ways to preserve useful knowledge. The ambition of a robot that keeps acquiring skills throughout its working life therefore remains broader than the capabilities established by any single experiment discussed here. Future studies should examine sequences of changing tasks and report how learning performance and human assistance evolve together.

This is a selective narrative review, and several recent contributions are preprints. Results should be interpreted in the context of each study's hardware, tasks and evaluation procedures. Nevertheless, the evidence supports treating deployment as a potential source of continued learning, with research focused on how experience can produce lasting and practically useful additions to a robot's abilities.

References

[1] Argall, B. D., Chernova, S., Veloso, M., & Browning, B. (2009). A survey of robot learning from demonstration. *Robotics and Autonomous Systems, 57*(5), 469–483. https://doi.org/10.1016/j.robot.2008.10.024

[2] Osa, T., Pajarinen, J., Neumann, G., Bagnell, J. A., Abbeel, P., & Peters, J. (2018). An algorithmic perspective on imitation learning. *Foundations and Trends in Robotics, 7*(1–2), 1–179. https://doi.org/10.1561/2300000053

[3] Brohan, A., Brown, N., Carbajal, J., Chebotar, Y., Dabis, J., Finn, C., Gopalakrishnan, K., Hausman, K., Herzog, A., Hsu, J., Ibarz, J., Ichter, B., Irpan, A., Jackson, T., Jesmonth, S., Joshi, N. J., Julian, R., Kalashnikov, D., Kuang, Y., ... Zitkovich, B. (2022). *RT-1: Robotics transformer for real-world control at scale*. arXiv. https://doi.org/10.48550/arXiv.2212.06817

[4] Open X-Embodiment Collaboration, O'Neill, A., Rehman, A., Gupta, A., Maddukuri, A., Gupta, A., Padalkar, A., Lee, A., Pooley, A., Gupta, A., Mandlekar, A., Jain, A., Tung, A., Bewley, A., Herzog, A., Irpan, A., Khazatsky, A., Rai, A., Gupta, A., Wang, A., ... Vakil, J. (2023). *Open X-Embodiment: Robotic learning datasets and RT-X models.* arXiv. https://doi.org/10.48550/arXiv.2310.08864

[5] Silver, D., & Sutton, R. S. (2025). *Welcome to the era of experience* [Preprint of a chapter]. https://storage.googleapis.com/deepmind-media/Era-of-Experience%20/The%20Era%20of%20Experience%20Paper.pdf

[6] Kober, J., Bagnell, J. A., & Peters, J. (2013). Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research, 32*(11), 1238–1274. https://doi.org/10.1177/0278364913495721

[7] Shougang Group. (2026, July 6). *Robot kindergarten inaugurated at Shougang Park*. https: //www.shougang.com.cn/m1/sgyw/20260706/14728.html

[8] Luo, J., Xu, C., Wu, J., & Levine, S. (2025). Precise and dexterous robotic manipulation via human-in-the-loop reinforcement learning. *Science Robotics, 10*(105), eads5033. https: //doi.org/10.1126/scirobotics.ads5033

[9] Bhatia, J. S., Wagenmaker, A., Chen, W., & Levine, S. (2026). *Adapting generalist robot policies with semantic reinforcement learning*. arXiv. https: //doi.org/10.48550/arXiv.2606.31958

[10] Lee, S., Seo, Y., Lee, K., Abbeel, P., & Shin, J. (2022). Offline-to-online reinforcement learning via balanced replay and pessimistic Q-ensemble. In *Proceedings of Machine Learning Research* (Vol. 164, pp. 1702–1712). PMLR. https: //proceedings.mlr.press/v164/lee22d.html

[11] Julian, R., Swanson, B., Sukhatme, G. S., Levine, S., Finn, C., & Hausman, K. (2021). Never stop learning: The effectiveness of fine-tuning in robotic reinforcement learning. In *Proceedings of Machine Learning Research* (Vol. 155, pp. 2120–2136). PMLR. https://doi.org/10.48550/arXiv.2004.10190

[12] Wang, Y., Li, X., Xie, P., Yang, P., Nie, B., Cai, Y., Zhang, Q., Qu, C., Wu, J., Song, J., Ren, X., Huang, J., Pan, M., Feng, S., Chen, Z., & Luo, J. (2026). *Learning while deploying: Fleet-scale reinforcement learning for generalist robot policies*. arXiv. https://doi.org/10.48550/arXiv.2605.00416

[13] Lesort, T., Lomonaco, V., Stoian, A., Maltoni, D., Filliat, D., & Díaz-Rodríguez, N. (2020). Continual learning for robotics: Definition, framework, learning strategies, opportunities and challenges. *Information Fusion, 58*, 52–68. https://doi.org/10.1016/j.inffus.2019.12.004

[14] Liu, B., Zhu, Y., Gao, C., Feng, Y., Liu, Q., Zhu, Y., & Stone, P. (2023). *LIBERO: Benchmarking knowledge transfer for lifelong robot learning*. arXiv. https://doi.org/10.48550/arXiv.2306.03310

[15] Tang, C., Abbatematteo, B., Hu, J., Chandra, R., Martín-Martín, R., & Stone, P. (2025). Deep reinforcement learning for robotics: A survey of real-world successes. *Annual Review of Control, Robotics, and Autonomous Systems, 8*, 153–188. https://doi.org/10.1146/annurev-control-030323-022510

[16] Su, S., Yang, B., Deng, Z., Liang, B., Yeung, Y. F., Wang, Z., & Sun, L. (2026). *From pretraining to proficiency: Real-world subtask RL for long-horizon manipulation with minimal human intervention*. arXiv. https://doi.org/10.48550/arXiv.2609.21788

[17] García, J. and Fernández, F. (2015). A comprehensive survey on safe reinforcement learning. *Journal of Machine Learning Research, 16*(42), 1437–1480. https://jmlr.org/papers/v16/garcia15a.html