SLIDE 1 OF 8 / 0:00-0:30

Teaching a household robot in VR

Main idea: We used a VR headset to record corrections to a robot's mistakes.

EXPLANATION CUES

Introduce this as the project talk: a simulated household robot tries a task, a person takes over through VR, and the system saves the correction.

Define teleoperation as controlling a robot from another place or interface. Here the person uses a headset and hand controllers to control a robot inside a simulator.

The research question is whether recordings made around mistakes can teach behavior missing from successful demonstrations.

POINT OR DEMONSTRATION

Keep the setting concrete: a robot with two arms, a virtual home, and a person holding the controllers.

TRANSITION

First, what did the robot already know?

IF ASKED

This work adapts human-gated DAgger. The automatic-expert variants later in the talk do not use the same human gate.

Source: Harvesting Failures for Imitation, team report, 3 November 2026

SLIDE 2 OF 8 / 0:30-1:45

Our robot learned from demonstrations

Main idea: Successful demonstrations give a starting point, but mistakes require additional examples.

EXPLANATION CUES

BEHAVIOR is a challenge made of household tasks in a simulator. It includes 100 tasks and 20,000 recorded demonstrations by people.

Define controller, also called policy: a learned model that chooses robot movements using camera images and information about the robot. Avoid requiring students to know the model architecture.

Use picking up a can as the example. If the grasp fails, save that moment. An expert can return to it and demonstrate a correction instead of repeating the whole task.

A simulator snapshot is a saved game state: it stores enough about the scene to resume near that point. Simulation can still vary, so replay is not guaranteed to unfold identically.

POINT OR DEMONSTRATION

Walk the four steps on the right. “An expert” can mean a person or an automatic controller; explain those alternatives later.

TRANSITION

Here is what the human correction looks like in VR.

IF ASKED

The environment is OmniGibson, the robot is R1Pro, and the learned controller is based on pi-0.5. These implementation names are optional answers, not prerequisites for following the talk.

Source: Harvesting Failures for Imitation, team report, 3 November 2026 (Sections 3, 4.1-4.2); https://behavior.stanford.edu/challenge/

SLIDE 3 OF 8 / 1:45-3:25

The person takes over with one button

Main idea: The person can switch between watching and controlling with one button.

EXPLANATION CUES

Click the poster and play the 13-second recording. Point out the red border: it marks the moments a person controls the robot.

The headset shows the virtual scene. While the robot moves on its own, the person watches. Pressing the right controller’s B button transfers control to the person.

The robot’s arms follow movements of the hand controllers. After the correction, pressing B again gives control back. The software records which side controlled each moment.

Explain why the clear switch matters: a useful teaching example needs a clear beginning, a correct movement, and a clear hand-back.

POINT OR DEMONSTRATION

Pause the clip on a red-border frame and identify the robot’s camera views. Allow the playback time within this slide’s timing.

TRANSITION

How does that interaction become something a robot can learn from?

IF ASKED

This policy-plus-human loop was used on older tasks. On new tasks, people corrected saved failures directly without a running policy. Control updated 30 times per second; that is not a measurement of end-to-end latency.

Source: Harvesting Failures for Imitation, team report, 3 November 2026 (Section 4.3, Fig. 2a)

SLIDE 4 OF 8 / 3:25-4:35

What the robot sees and what we teach

Main idea: A training example pairs what the robot saw with the action a teacher took.

EXPLANATION CUES

Give one concrete pair: the camera shows a can beside an empty hand, and the teacher moves the hand into a position that grasps it.

The recording saves images and movements together. It also marks whether a person or the robot controlled that moment, so later processing can keep the human correction.

Explain a training label as the action associated with an image. It is an instruction example, not only a written description of the scene.

The software converts hand-controller motion into the robot's movement format. Recording a good-looking video alone would not provide those action examples.

POINT OR DEMONSTRATION

Point from "Camera image" to "Correct movement". This is the essential data pair.

TRANSITION

We also learned that the moment where replay begins changes the lesson.

IF ASKED

Technical details: LeRobot v2.1 episodes, per-frame intervention mask, next reached arm/torso positions, base velocity and gripper command. The collection and processing were built, but the planned movement-training test on human takeovers was not run.

Source: Harvesting Failures for Imitation, team report, 3 November 2026 (Sections 4.3, 4.6)

SLIDE 5 OF 8 / 4:35-6:00

The replay starting point changed the lesson

Main idea: Starting before a mistake encouraged people to prevent it rather than recover from it.

EXPLANATION CUES

Use a game-replay analogy: restart before a bad move and you may simply avoid making it. Restart after it and you have to solve the resulting problem.

Our early replay resumed five seconds before the marked failure. In 29 of 33 place recordings, the can was already in the hand. Those recordings often did not demonstrate recovering a missed grasp.

The default later moved to three seconds after the mark, aiming to collect recovery behavior. The change came after 78 of the initial 82 recordings.

Avoid calling prevention useless. It teaches a different behavior. The collection goal determines which starting point is useful.

POINT OR DEMONSTRATION

Read 29 out of 33 aloud, explain the denominator, and return to the empty-hand example.

TRANSITION

Choosing good replay moments helped, but recording enough corrections was still slow.

IF ASKED

This observation concerns one human-recorded batch, not a controlled study of replay timing. The notebook reports only one human re-grasp among those place recordings.

Source: Harvesting Failures for Imitation, team report, 3 November 2026 (Tables 1-2, Sections 4.3, 7.1-7.2)

SLIDE 6 OF 8 / 6:00-7:15

People could not record every correction

Main idea: Human teaching time was limited, so we tried automatic experts too.

EXPLANATION CUES

The robot could generate failures faster than people could fix them. One batch produced 78 human recordings over four days on one task.

We added scripts that use simulator information to perform particular actions, and an older, stronger learned controller that could sometimes fix the newer controller's mistakes.

Both started from saved failed situations. An automatic checker tested the result. Selected script recordings also went through an AI video review followed by a human decision.

Explain why a checker is not enough: a task might finish by accident, or the successful movement may be a poor example for the learner. Success and useful teaching data are different properties.

POINT OR DEMONSTRATION

Separate the human VR column from the automatic-help column. Do not imply all recordings were human takeovers.

TRANSITION

Did any of these correction examples improve the robot?

IF ASKED

Scripts produced 4,166 recordings on 48 new tasks. The stronger controller produced 754 successful attempts out of 3,961, about 19 percent. The first training round kept 500 recordings; the run recipe lists 499. These counts are different stages of the pipeline.

Source: Harvesting Failures for Imitation, team report, 3 November 2026 (Sections 4.4-4.5, Table 1)

SLIDE 7 OF 8 / 7:15-8:50

A better grasp still left the full task unsolved

Main idea: Some corrections helped individual pick-ups, while complete new tasks remained unsolved.

EXPLANATION CUES

Explain the percentages intuitively: the measured rates are roughly 38 versus 45 successful pick-ups per 100 attempts. These are rates scaled to 100 for explanation, not a claim that the experiment used only 100 attempts.

The actual comparison uses 383 held-out pick-up steps. One controller received extra training on stronger-controller corrections; the comparison controller received the same extra training using demonstrations alone.

The gain comes from the stronger automatic controller's data. Our planned test of human VR corrections for movement learning did not run, so we cannot attribute this gain to the people in headsets.

A household task contains many steps. In the submitted evaluation of new tasks, zero of 490 runs completed a whole task. Improving one grasp is useful but not sufficient.

POINT OR DEMONSTRATION

Point to 37.6 and 44.6, then to the whole-task result. Explain the two different things being counted.

TRANSITION

The project's clearest lessons are about both learning and interface design.

IF ASKED

The gain is 7.0 percentage points, with a 95 percent interval of +1.0 to +12.7 over tasks. Separately, human corrections supplied images to a step-finished detector: 66 encoded recordings, 45 in training and 21 validation. The overall self-reported score includes older tasks, where the 18 successes occurred.

Source: Harvesting Failures for Imitation, team report, 3 November 2026 (Sections 6.1, 6.3, 7.5, Table 3)

SLIDE 8 OF 8 / 8:50-10:00

What VR design changes

Main idea: VR interface choices shape the examples a robot receives.

EXPLANATION CUES

Revisit the workflow: the person needs a clear indication of who has control and a clear way to hand it back.

The saved starting point determines whether the person demonstrates prevention or recovery. Choose it deliberately and check the resulting recordings.

Look at both the video and recorded movements. In our automatic data, an unfamiliar waist-joint movement was an important source of mismatch; a visually successful recording can still teach an incompatible action.

The next study should ask whether people can produce more useful corrections with clearer takeover controls or less delay. Afterward, test learning from those corrections with a matched full-task evaluation.

POINT OR DEMONSTRATION

Close by connecting a familiar VR interaction to a concrete consequence: the interaction changes the training examples.

TRANSITION

Invite questions about the headset workflow, replay timing, or what the result does and does not show.

IF ASKED

We ran at most one training round per correction dataset. A full DAgger-style study would repeat collection with the updated controller. AI tools helped build the slides; project AI coding and video-review tools are separate uses described in the report.

Source: Harvesting Failures for Imitation, team report, 3 November 2026 (Sections 7.1-7.5); https://arxiv.org/pdf/1810.02890

Plain-language glossary

Use the explanation before introducing the technical term.

Teleoperation

A person controls a robot through a remote interface, here a VR headset and hand controllers.

Policy / controller

The learned model that chooses robot movements.

Simulator snapshot

A saved scene state that allows replay near a chosen moment, like a saved game.

Training example / label

A situation paired with the action the teacher supplies.

Recovery

Correcting the situation after something has already gone wrong.

Individual step vs whole task

Picking up the can is a step. Completing the household instruction requires a sequence of steps.

Rehearsal

Speak from the cues in your own words. Use the running-time markers to stay near ten minutes. In the project talk, the VR clip is included in slide 3's time.

The same slide notes are available in the Google Slides speaker-notes pane and Presenter view.

Open the Google Slides presentation

Questions and short answers

Was the robot physical?

These BEHAVIOR recordings use a simulated robot in OmniGibson, viewed through a real VR headset.

Did the headset recordings cause the seven-point gain?

That was not tested. The measured gain uses a stronger controller's corrections. Some human recordings instead contributed to a detector of completed steps.

Why resume after the mistake?

It encourages recording a recovery. Resuming before the mistake often encourages prevention. Both can be useful for different collection goals.

Why did scripts sometimes hurt learning?

They could reach the goal using actions unlike the demonstrations, including waist motion. The negative result also has training-duration and fumbling confounders.

Does 45 percent mean the robot finishes 45 percent of tasks?

No. It is success at individual pick-up steps. Zero of the 490 evaluated new-task runs completed a whole task.