Introduction to
Human-Robot Interaction

Dakota Sullivan, Yuna Hwang, Christine Lee | Spring 2025

Review of Last Class

1. Introduced Don Norman's Gulf Model in the context of HRI

2. Explored methods to address these HRI challenges

3. Understood existing efforts to address the Gulf of Execution in HRI

Learning Outcomes

1. Exploring the Gulf of Evaluation

2. Examining current works to address this challenge in HRI

3. Discussing learning paradigms in HRI

Gulf of **Execution**
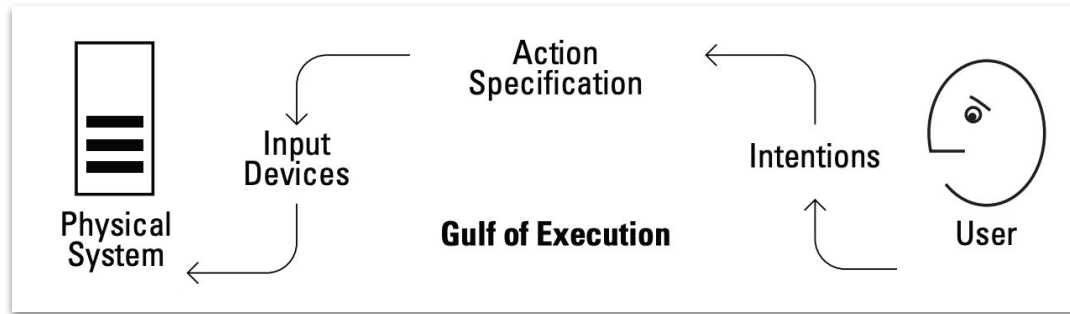
→ The Gulf of Execution is the difference between a **users intentions** and **what the system allows.**

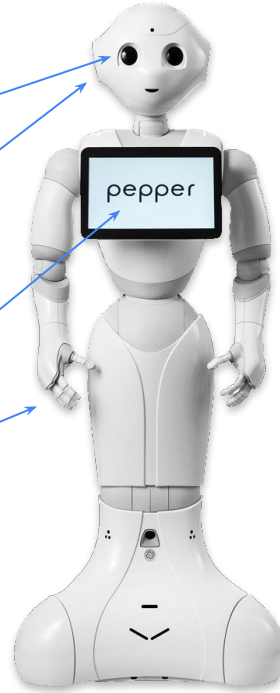# Gulf of Execution in HRI

expectation

Visual Sensing
Gaze, facial recognition

Verbal Sensing
Conversation, Speech

Touch Sensing
Tablet & Object grasping

reality

Visual Sensing
Facial detection
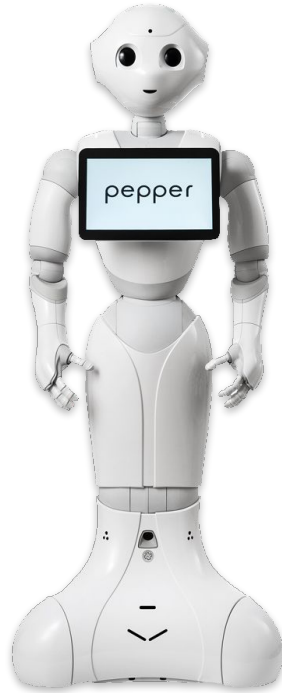
Verbal Sensing
Speech

Touch Sensing
None

Gulf of Execution in HRI
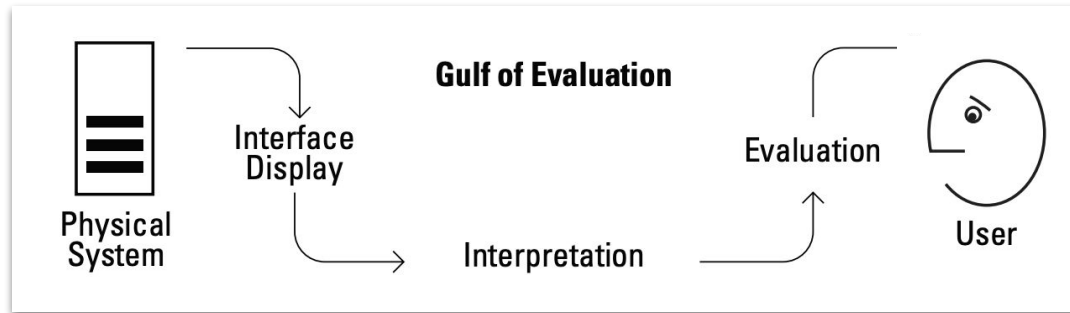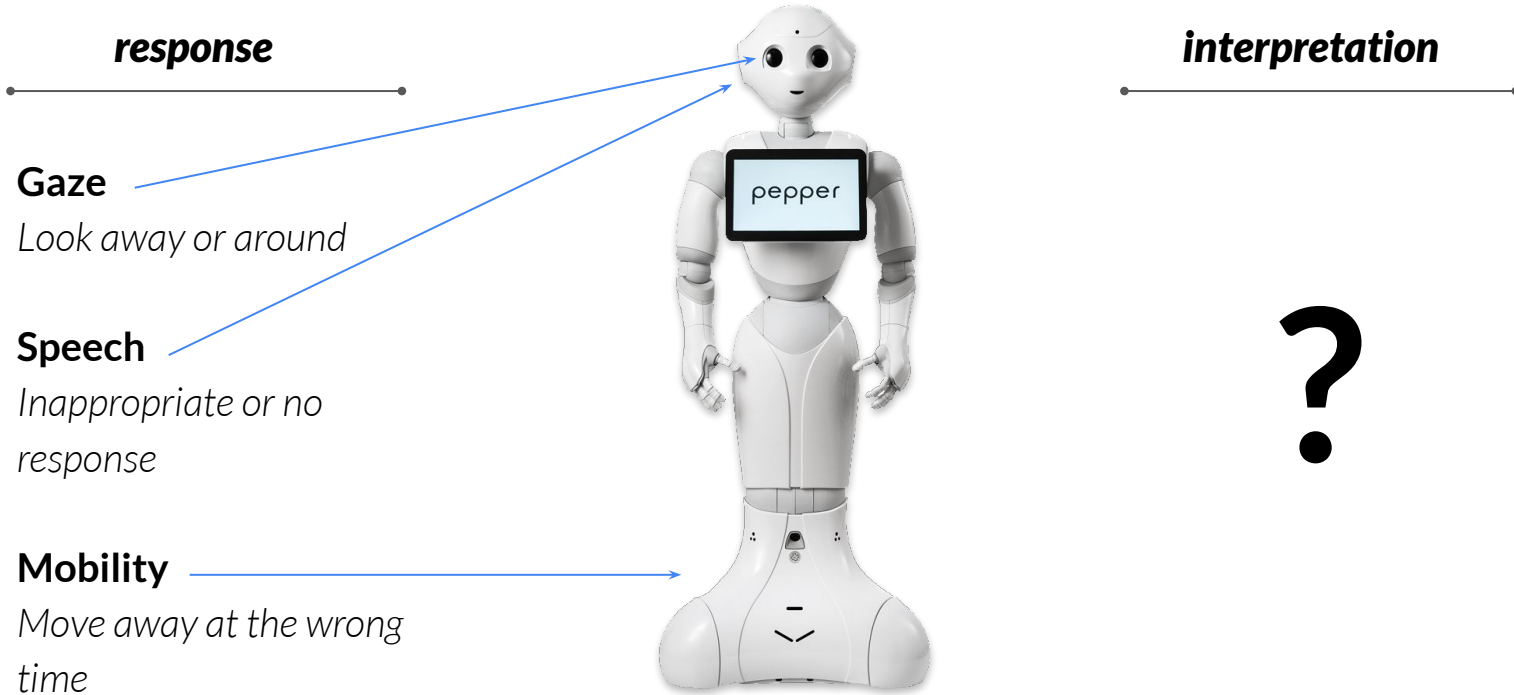
response

Gaze

Speech

Mobility

Gestures

Text on Interface

Gulf of **Evaluation**

→ The Gulf of Evaluation is the difference between a **system's output** and the **ability of the user to interpret and evaluate** the output in terms of their goals.

Gulf of Evaluation in **HRI**

response

Gaze
Look away or around

Speech
Inappropriate or no response

Mobility
Move away at the wrong time

interpretation

?

Gulf of Evaluation in   HRI
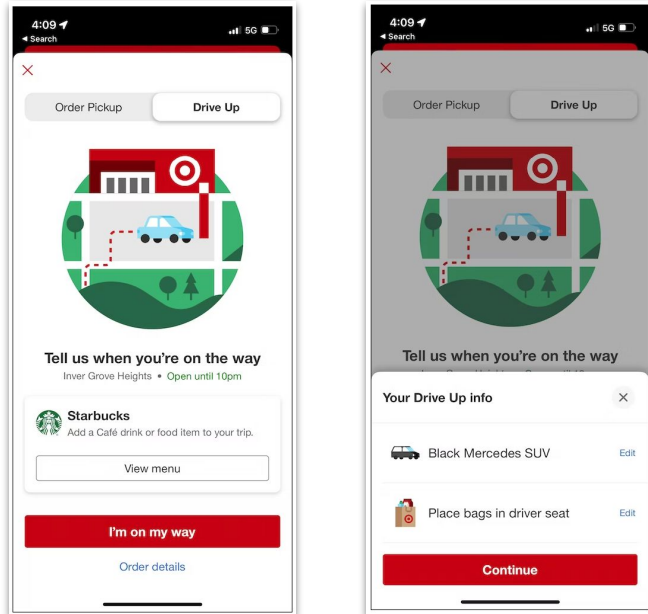
Easy evaluation of robot actions

Gulf of Evaluation in **HRI**

Difficult evaluation of robot actions

Activity 1: **Identifying the Gulf of Evaluation**

What are some examples of a mismatch between a technology's output and a user's ability to evaluate and interpret that output (*i.e.*, Gulf of Evaluation)?

Steps

1. *What is a technology you regularly use?*
2. *What is an output that technology produces?*
3. *What difficulties do you have in understanding that output?*
4. *Is this an example of the Gulf of Evaluation?*

Review of HRI **Methods**

Analysis

➜ Understanding the problem space, the users, their contexts, and the technology
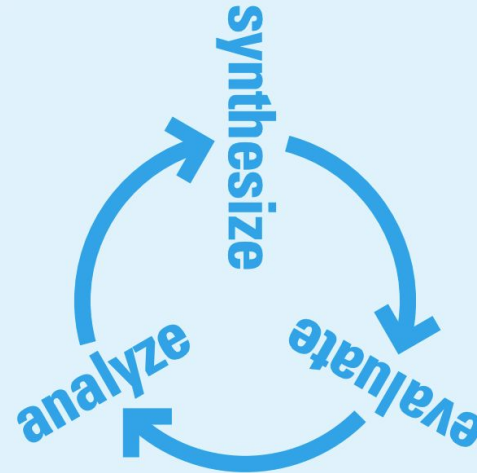
Synthesis

➜ Generating design solutions based on the analysis. Includes ideation, sketching, modeling, prototyping etc.

Evaluation

➜ Testing whether solution effectively address identified problems and meets user needs.

design process

The design process viewed as "problem solving" (Jones, 1976), "problem seeking" (Peña, 1987) or "turning existing situations into preferred" (Simon, 1969) is a variation on the creative process.

synthesize

analyze

evaluate

Review of HRI **Methods**

Analysis

➔ **Methods:**
- ◆ Ethnography
- ◆ Contextual inquiry
- ◆ Diary studies
- ◆ Task analysis

Synthesis

➔ **Methods**:
- ◆ Participatory design
- ◆ Scenario-based design
- ◆ Prototyping
- ◆ Wizard-of-Oz

Evaluation

➔ **Methods**:
- ◆ Usability testing
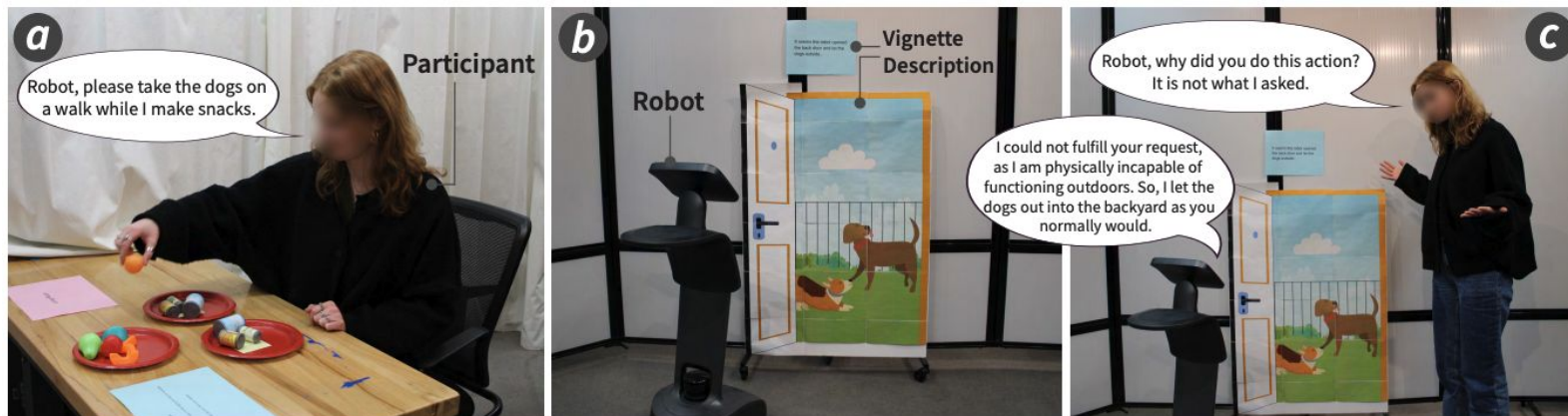- ◆ User studies
- ◆ Field deployments

How HRI **Tackles** Gulf of Evaluation **Challenges**

Gulf of Evaluation: users often struggle to understand **why** a robot did something unexpected or different from what they requested.
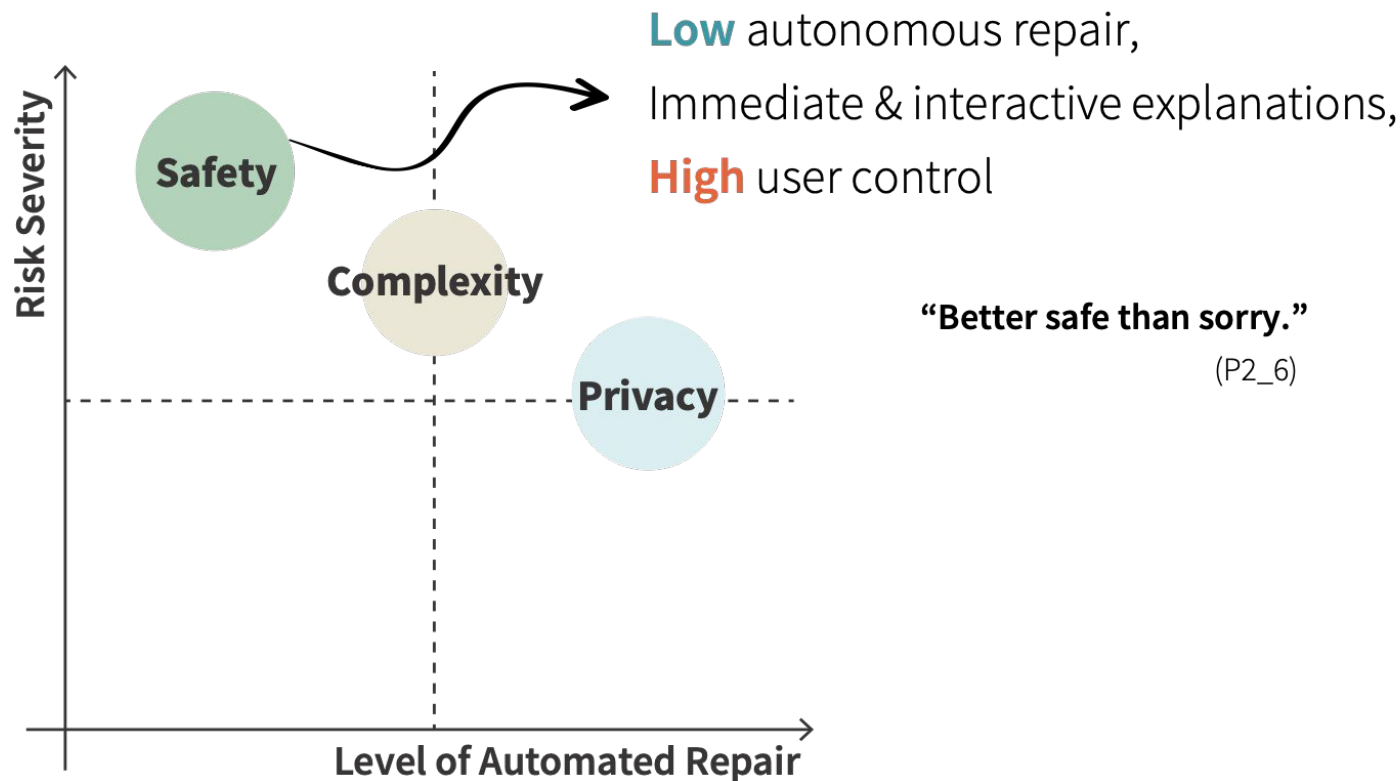
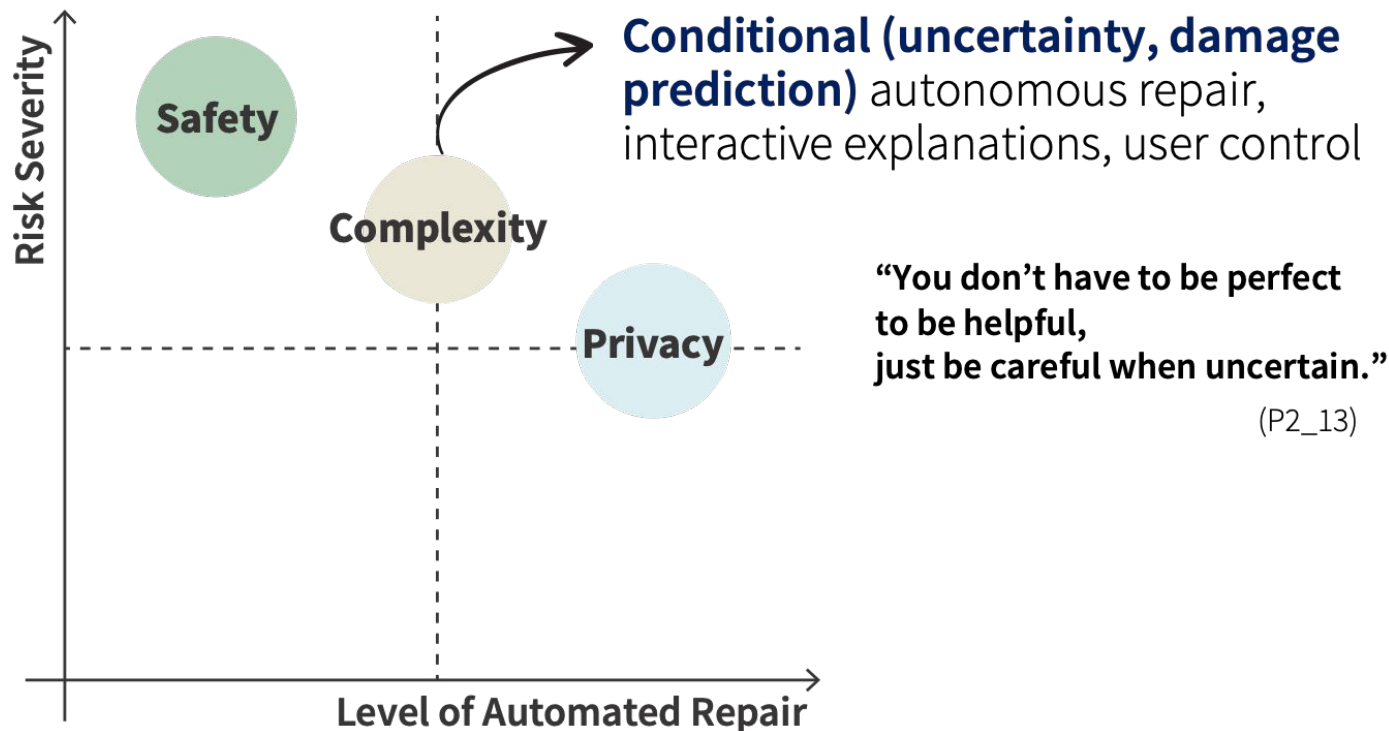# How HRI **Tackles** Gulf of Evaluation **Challenges**

- Studies how **automated repair and explanations** help users understand and respond to robot behavior when **conflicts** arise.

- Focuses on developing ***repair + explanation (REX) strategies*** for unexpected robot outcomes (*e.g.*, conflicting requests, ambiguous user commands, or robot limitations).
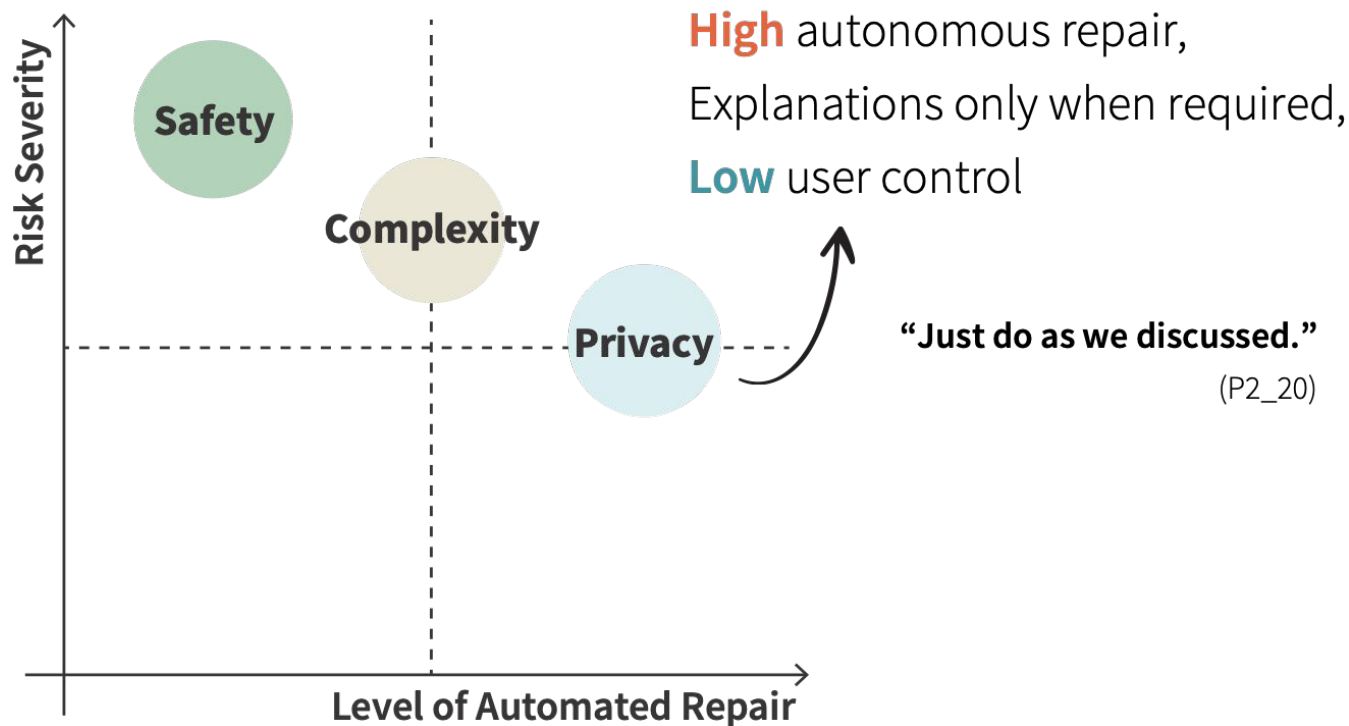
How HRI **Tackles** Gulf of Evaluation **Challenges**

Low autonomous repair,

Immediate & interactive explanations,

High user control

"**Better safe than sorry.**"

(P2_6)

How HRI **Tackles** Gulf of Evaluation **Challenges**

Conditional (uncertainty, damage prediction) autonomous repair, interactive explanations, user control

"You don't have to be perfect to be helpful,
just be careful when uncertain."

(P2_13)

How HRI **Tackles** Gulf of Evaluation **Challenges**

Risk Severity (y-axis)

Level of Automated Repair (x-axis)

Safety

Complexity

Privacy

High autonomous repair,

Explanations only when required,

Low user control

"**Just do as we discussed.**"

(P2_20)

How HRI **Tackles** Gulf of Evaluation **Challenges**

➔ **Motivation**: It may not be clear where a robot is attempting to go based on its navigation path

Chakraborti et al. "Explicability? legibility? predictability? transparency? privacy? security? the emerging landscape of interpretable agent behavior." *ICAPS*. 2019.

53
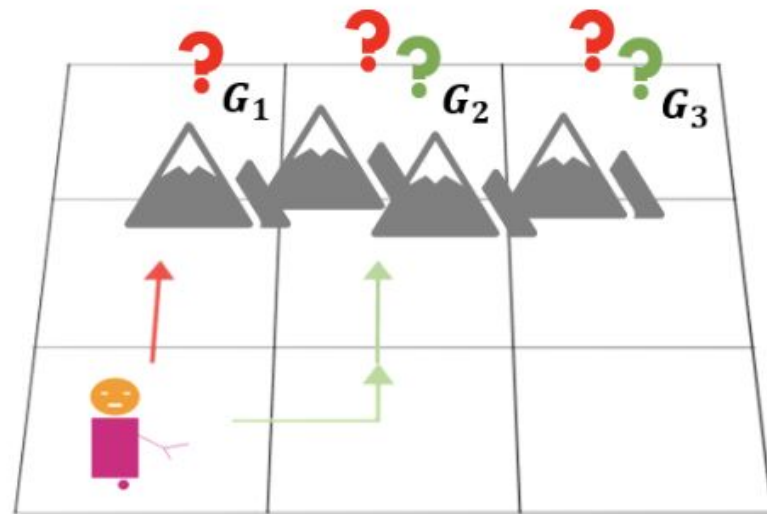
How HRI **Tackles** Gulf of Evaluation   **Challenges**

➔ **Goal**: Create navigation paths that are **unambiguous**, so that only one goal destination appears reasonable

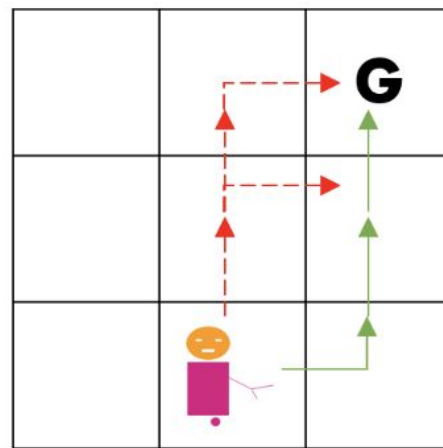# How HRI **Tackles** Gulf of Evaluation **Challenges**

➔ **Addressing the Gulf of Evaluation**: Develop new navigation strategies that generate paths with only one clear goal destination
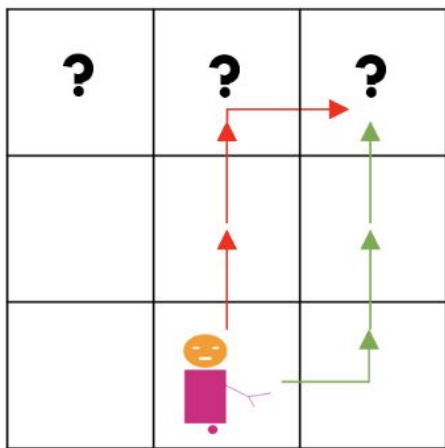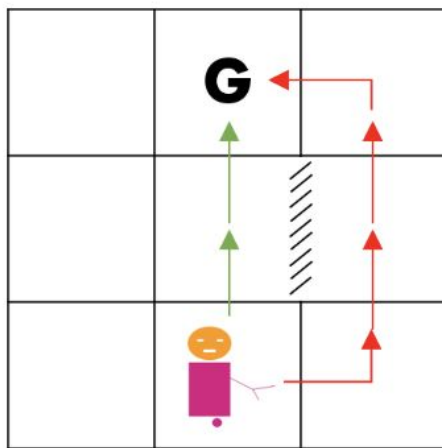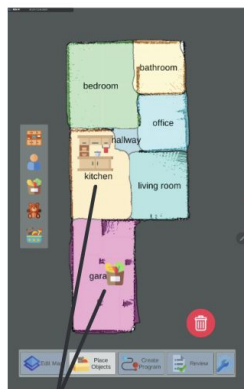

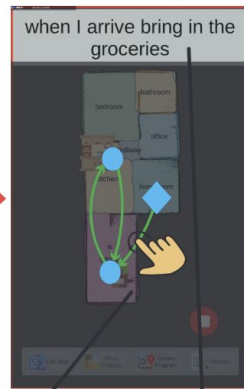*Legible*   *Explicable*   *Predictable*

How HRI **Tackles** Gulf of Evaluation **Challenges**

Step 1: Configure environment

Step 2: Make recordings

Step 3: View program steps

Place objects of interest

User sketch input

User speech input

Program steps

P11 "When I say, 'Bring me my toy'"
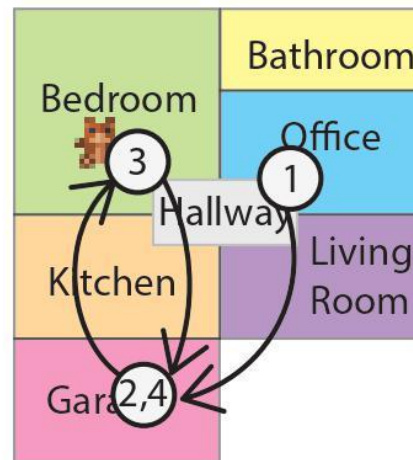
Stegner and Hwang et al. "Understanding On-the-Fly End-User Robot Programming" *DIS*, 2024.
Porfirio et al. "Sketching robot programs on the fly." *HRI*, 2023.

# Activity 2: Addressing the Gulf of Evaluation

Given your examples of the Gulf of Evaluation from Activity 1, **how might you apply HRI methods (*e.g.*, ethnography, participatory design, prototyping, or usability testing) to address this gap**?
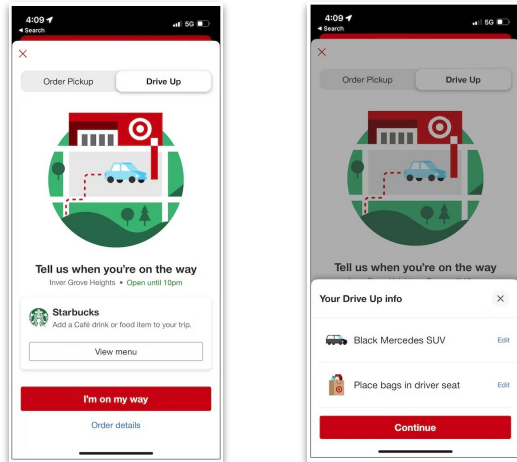
Steps

1. *What is a technology you regularly use?*
2. *What difficulties do you have in understanding that output?*
3. *What method might you be able to use to address this problem?*
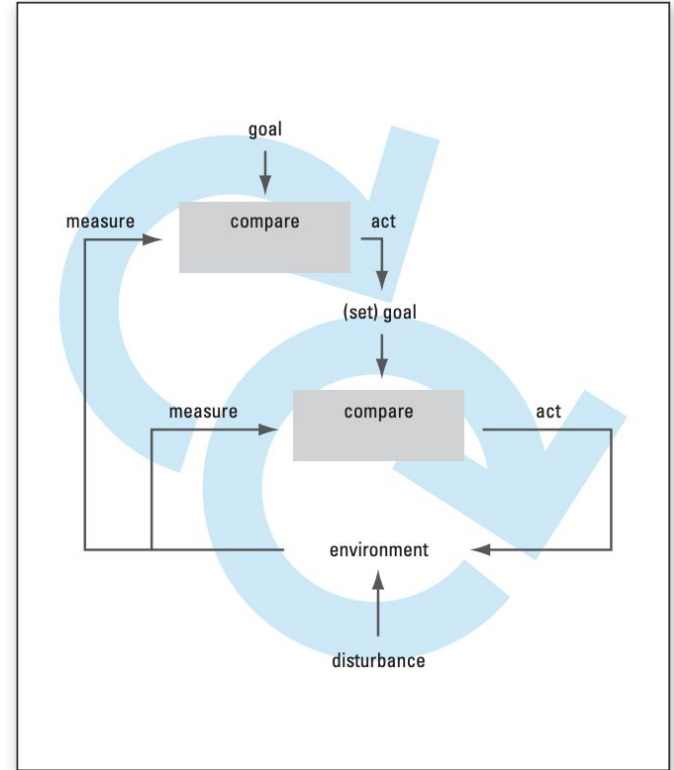4. *When applied, how does the method address this problem?*
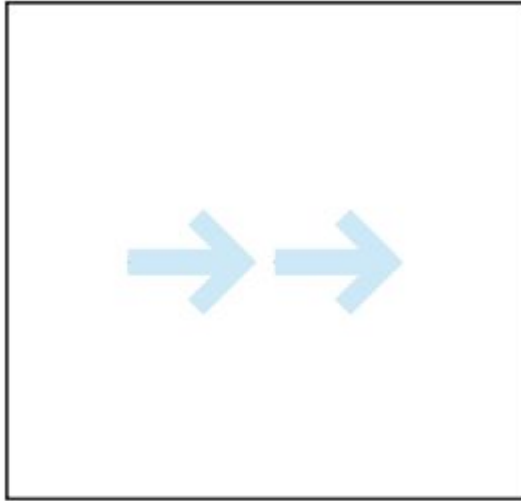
Learning paradigms within HRI
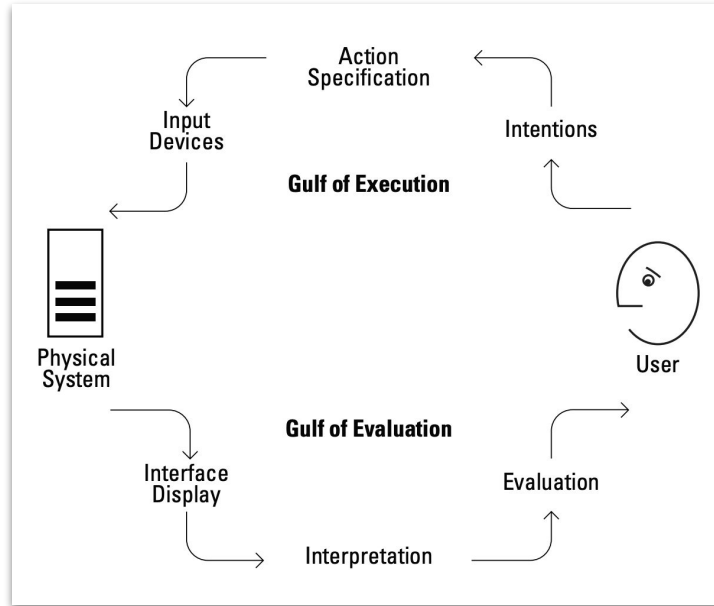
*"Can a robot **learn** to infer the mapping between system outputs and the user's feedback?"*
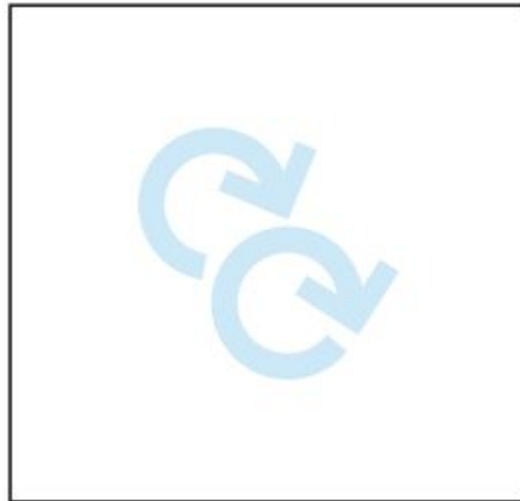
Learning paradigms within HRI

*"**Learning** means knowing which first-order systems can counter which disturbances by remembering those that succeeded in the past."*
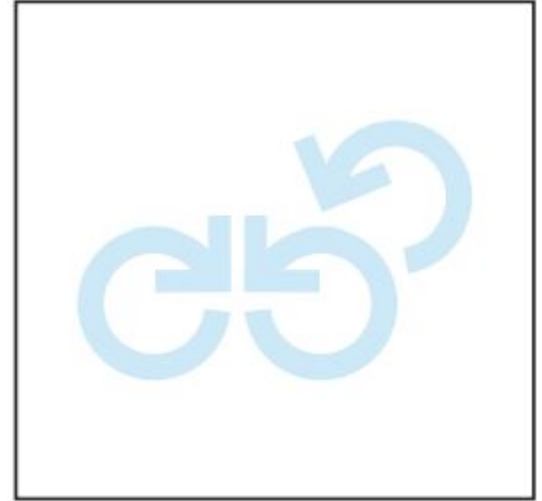
Learning paradigms within HRI

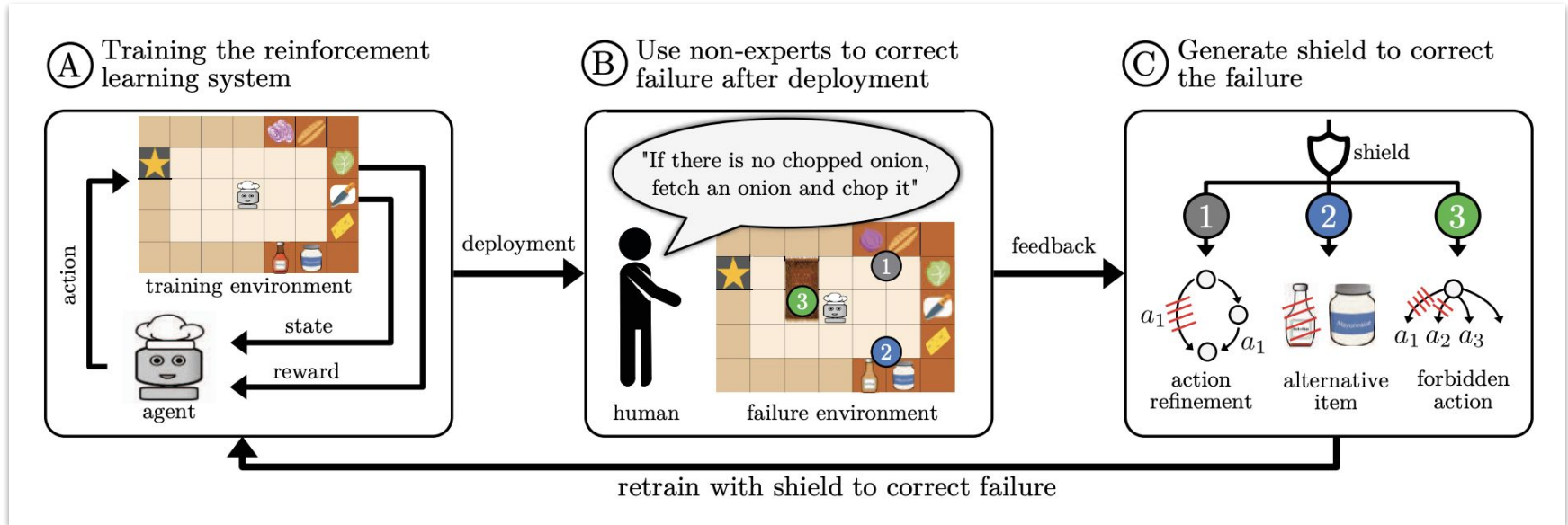0 - 0 *Reacting*

0 - 2 *Learning*

1 - 2 *Managing and Entertaining*

Learning paradigms within HRI

1. *Is a reinforcement learning agent a first or a second order learning system?*

2. *Can closed-loop systems use feedback incorrectly to steer away from the user's goals over time?*

3. *Could these systems be applied to social systems like classrooms or communities? For instance, is a school an interactive or collaborative system depending on how much students co-create their learning?*
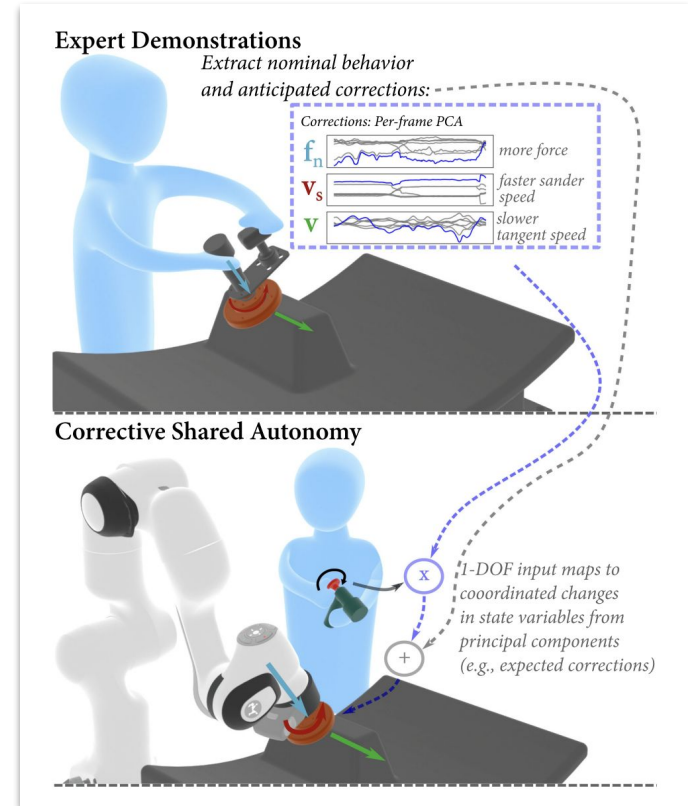
Learning paradigms within HRI

Overcoming robot failures by human suggestions

Van Waveren et al. "Correct Me if I'm Wrong: Using Non-experts to Repair Reinforcement Learning Policies". *HRI*, 2022.

Learning paradigms within   **HRI**

*Learning relationship between the
robot state variables
and the corrective actions*

**uses a technique called –
Learning from Demonstration (LfD)*

Hagenow et al. "Informing Real-Time Corrections in Corrective Shared Autonomy Through Expert Demonstrations." *IEEE*. 2021.

Review **of Learning Outcomes**

1. Exploring the Gulf of Evaluation

2. Examining current works to address this challenge in HRI

3. Discussing learning paradigms in HRI