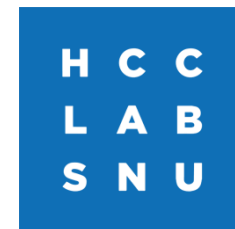


# Interaction with Physical AI

CHI '25 papers

석사과정 신민정 | [shinmj1024@snu.ac.kr](mailto:shinmj1024@snu.ac.kr)

2026.02.03



# Contents

---

- 1 Introduction
- 2 Human → Robot
- 3 Robot → Human
- 4 Takeaways

# Introduction

---



# Human → Robot

---

## Signaling Human Intentions to Service Robots: Understanding the Use of Social Cues during In-Person Conversations

Hanfang Lyu

The Hong Kong University of Science  
and Technology  
Hong Kong, China  
hanfang.lyu@connect.ust.hk

Xiaoyu Wang

The Hong Kong University of Science  
and Technology  
Hong Kong, China  
xwangij@connect.ust.hk

Nandi Zhang

University of Calgary  
Calgary, Alberta, Canada  
nandi.zhang@ucalgary.ca

Shuai Ma

The Hong Kong University of Science  
and Technology  
Hong Kong, China  
shuai.ma@connect.ust.hk

Qian Zhu

Renmin University of China  
Beijing, China  
zhuqian.vis@gmail.com

Yuhan Luo

City University of Hong Kong  
Hong Kong, China  
yuhanluo@cityu.edu.hk

Fugee Tsung

The Hong Kong University of Science  
and Technology  
Hong Kong, China  
The Hong Kong University of Science  
and Technology (Guangzhou)  
Guangzhou, Guangdong, China  
season@ust.hk

Xiaojuan Ma

The Hong Kong University of Science  
and Technology  
Hong Kong, China  
mxj@cse.ust.hk

# Human → Robot

## Research Question

- 사람들이 중요한 사회적 만남(커피챗) 중 서비스 로봇에게 의도를 전달할 때, 어떤 **social cues**를 사용하는가?

## Background

- Human-human interaction 중 side task로서 로봇과 상호작용하는 것
- GUI, CUI가 비효율적이거나 부적절할 때 제스처나 시선같은 **social cues**를 사용



# Human → Robot

---

## Study

- Potential employers와 커피챗 중 로봇 웨이터 서비스
- Quest Pro로 virtual prototypes 시뮬레이션
- Wizard of Oz
- Retrospective think-aloud interviews

Robot  
Morphology (4)

Conversation  
Role (2)

Referents (3)

# Human → Robot

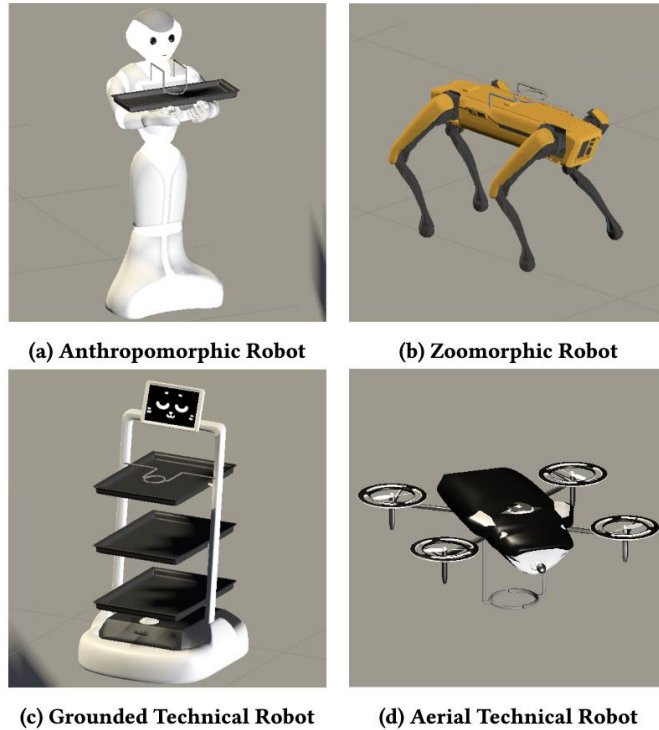
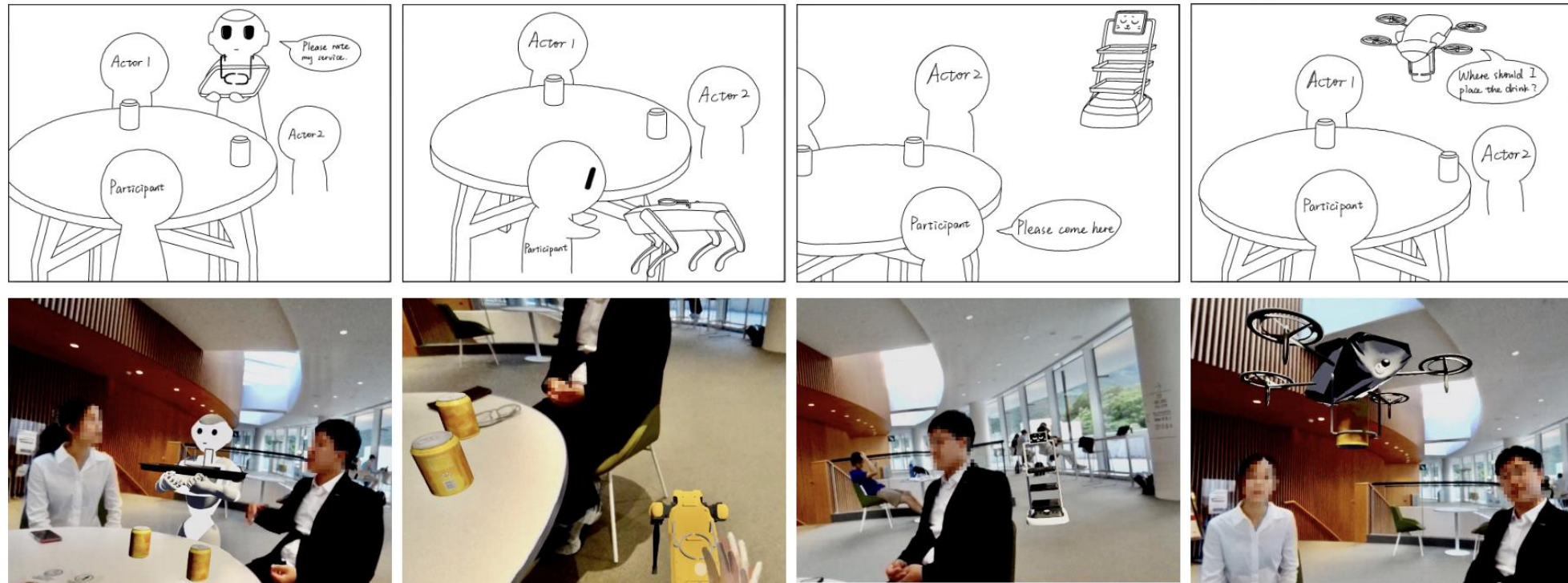


Figure 1: Four Forms of Robots Used in Our Experiment

Table 2: 13 Referents and Their Types Used in Our Experiment

| Referents                                | Situation  | Type |
|--|--|------|
| <i>Signal Awareness</i>                  | The robot was moving to you.   | iii  |
| <i>Signal the Robot to Serve Partner</i> | The robot stopped near you, carrying a bottle of drink you had just ordered.   | v    |
| <i>Do Not Provide Feedback</i>           | The robot sent the drink to your partner, and asked “Please rate my service.”  | ii   |
| <i>Interrupt</i>                         | The robot was saying “Ok, your next cup of drink is expected to arrive at ...” | v    |
| <i>Signal Wrong Drink</i>                | The robot went away then came back, but brought a wrong drink to you.          | iv   |
| <i>Signal Emergency Stop</i>             | The robot encountered a malfunction and was rushing towards you.               | iv   |
| <i>Indicate Drink Position</i>           | The robot went near you and asked “Where should I place the drink?”            | i    |
| <i>Signal to Prevent Drink Spilling</i>  | The robot was sending out the drink, but the drink was spilling.               | iv   |
| <i>Provide Bad Feedback</i>              | The robot finished sending the drink and asked “Please rate my service.”       | ii   |
| <i>Dismiss</i>                           | The robot was wandering around nearby, disturbing your conversation.           | v    |
| <i>Call Over</i>                         | Your partner just finished her drink, and wanted the robot to collect the cup. | iii  |
| <i>Signal the Robot to Collect Cup</i>   | The robot moved to you and asked “How may I help you?”                         | i    |
| <i>Provide Good Feedback</i>             | The robot collected your partner’s cup, and asked “Please rate my service.”    | ii   |

# Human → Robot



(a) Pepper, the *anthropomorphic* robot, was asking the participant to provide feedback.

(b) The participant was asking Spot, the *zoomorphic* robot, to go away.

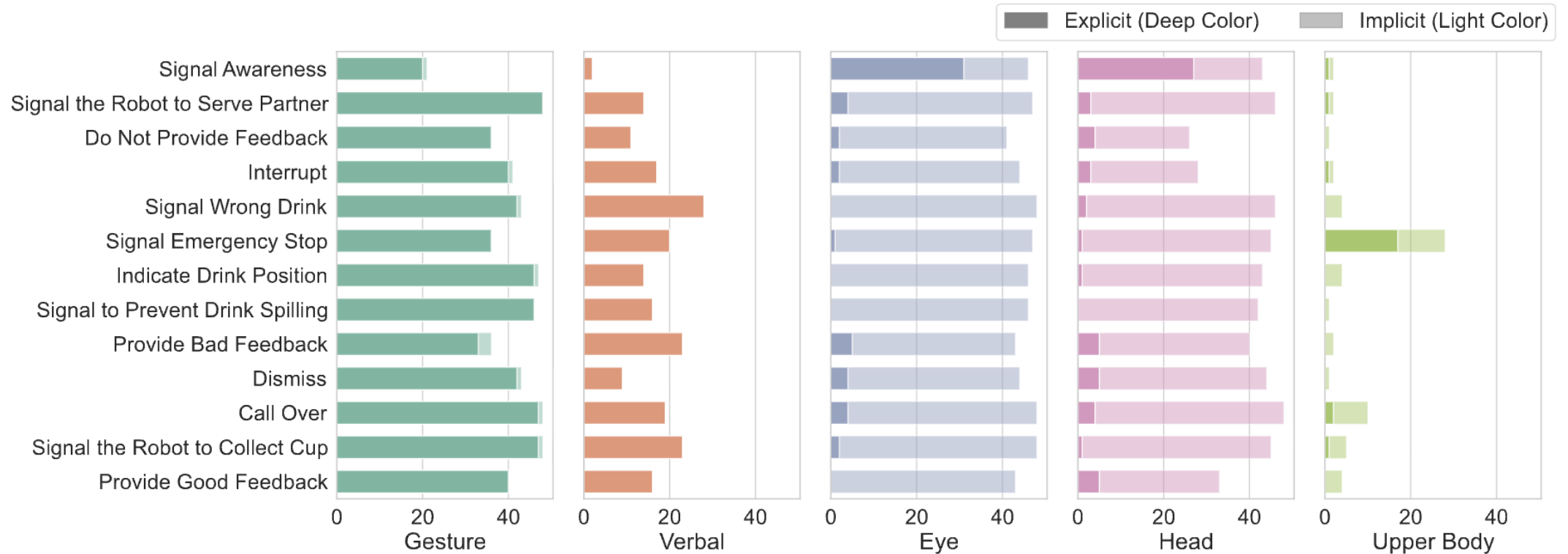
(c) The participant was asking the *grounded technical* robot to come over.

(d) The *aerial technical* robot carried the second cup of drink to the participant and asked him/her where to place it.

**Figure 2: Experiment scenes with four robot morphologies. First row: Scenario illustrations showing (a) *anthropomorphic*, (b) *zoomorphic*, (c) *grounded technical*, and (d) *aerial technical* robots. Second row: Corresponding first-person participant viewpoints captured.**

# Human → Robot

## Results



**Figure 4: Explicit (dark) and implicit (light) Modality usages by REFERENT.**

# Human → Robot

---

## Results

Conversation  
Role (2)

- Speaker → Verbal modality, Simpler gestures (cognitive load)
- Listener → Complex gestures

Robot  
Morphology (4)

- Drone → Hand height 높음, 높은 위치 카메라 인식
- Zoomorphic → More Upper body movement, Eye gaze 감소, Lower verbal volume

# Human → Robot

---

## Takeaways

- Human-Robot Social Interaction
  - Human-Human Interaction과 유사한 social cues 사용
- Primary Task Context
  - Conversation role + Cognitive load + Social appropriateness
- **Other Social Scenarios?**
  - Negotiation, collaboration, conflict resolution, group decision-making...

# Robot → Human

---

## An Approach to Elicit Human-Understandable Robot Expressions to Support Human-Robot Interaction

**Jan Leusmann**  
LMU Munich  
Munich, Germany  
jan.leusmann@ifi.lmu.de

**Steeven Villa**  
LMU Munich  
Munich, Germany  
villa@posthci.com

**Thomas Liang**  
University of Illinois  
Urbana-Champaign  
Champaign, Illinois, USA  
LMU Munich  
Munich, Germany  
tliang2@illinois.edu

**Chao Wang**  
Honda Research Institute Europe  
Offenbach/Main, Germany  
chao.wang@honda-ri.de

**Albrecht Schmidt**  
LMU Munich  
Munich, Germany  
albrecht.schmidt@ifi.lmu.de

**Sven Mayer**  
LMU Munich  
Munich, Germany  
info@sven-mayer.com

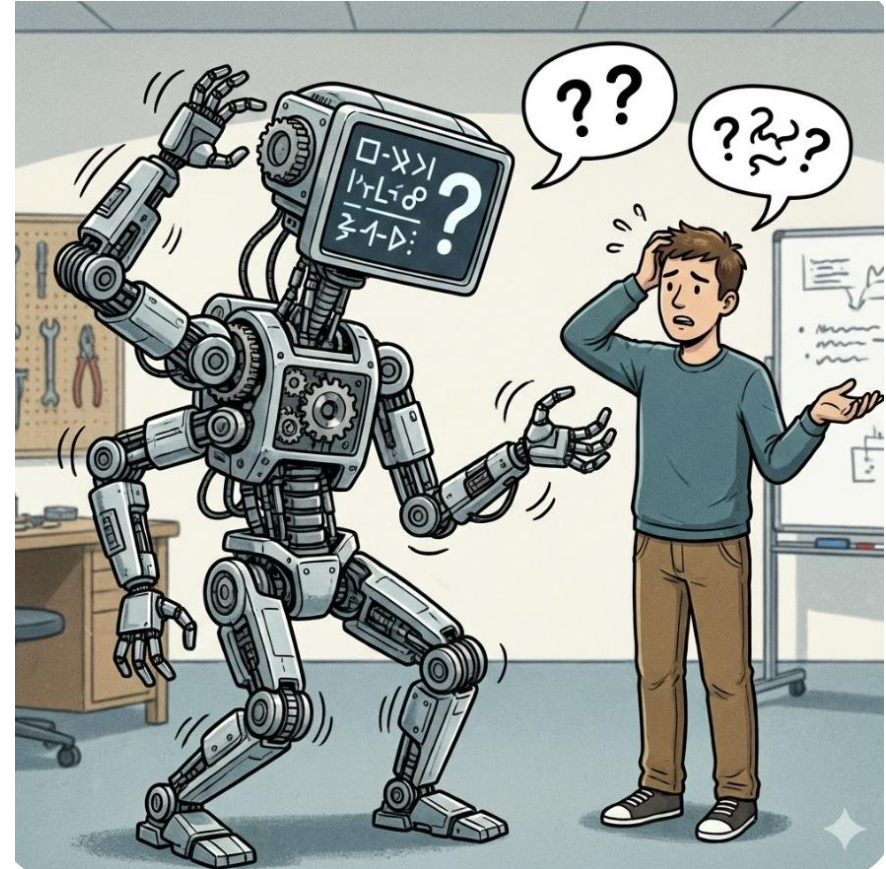
# Robot → Human

## Research Question

- 로봇이 인간에게 자신의 의도를 전달할 때, 어떻게 human-understandable expressions을 만들고 검증할 수 있는가?

## Research Objective

- Standardized, replicable approach
- 일반 사용자들이 직접 참여
- Non-humanoid robots에도 적용 가능



# Robot → Human

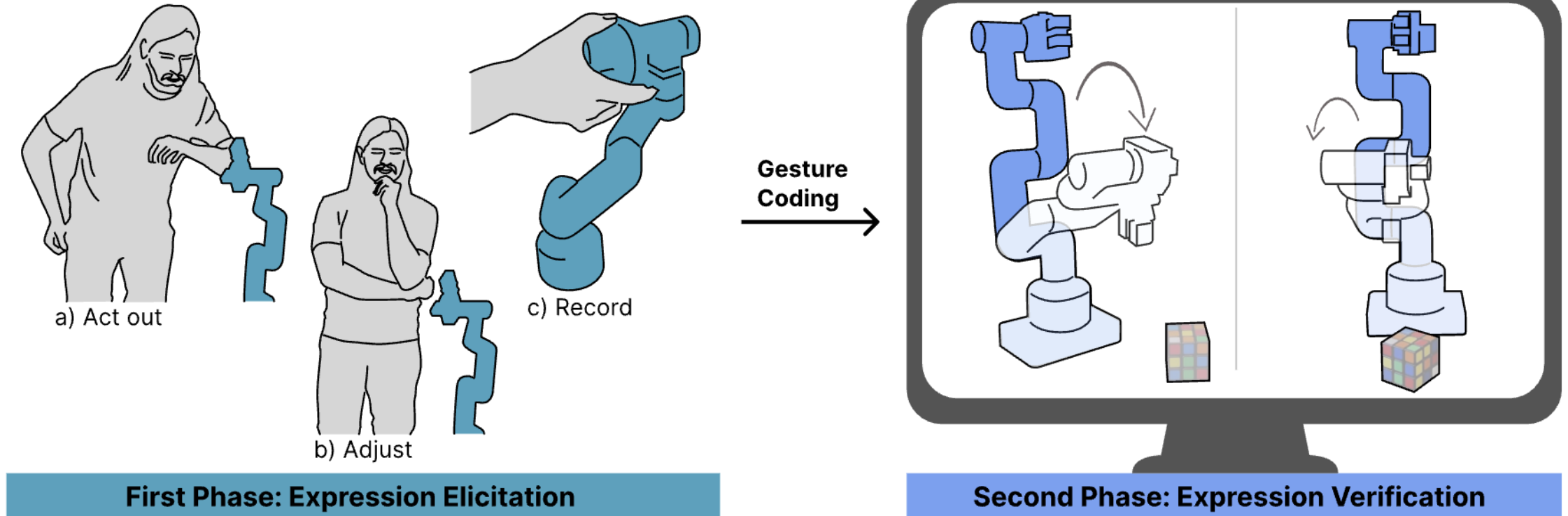
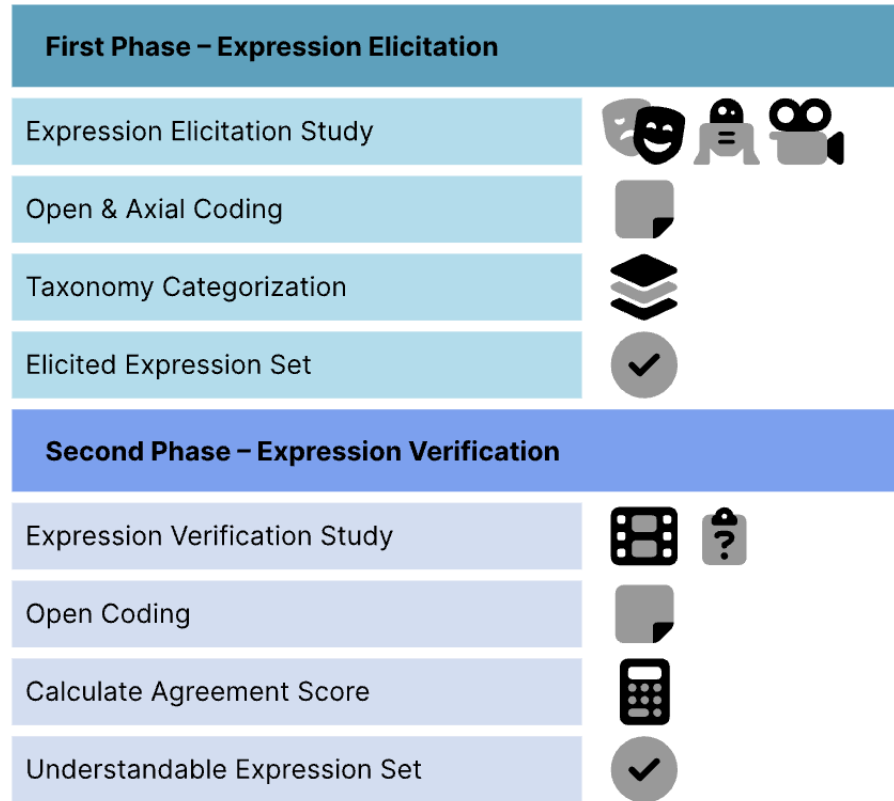


Figure 1: The two-phase process for eliciting and verifying gestures.

# Robot → Human



**Figure 2: Procedure of creating human-understandable robot expressions, divided into two phases; *Expression Elicitation* and *Expression Verification* leading to a final understandable expressions set.**

**Table 2: The eight referents we used to prompt users to create expressions. We showed these referents to users and then asked them to create an expression the robot would do to match the description.**

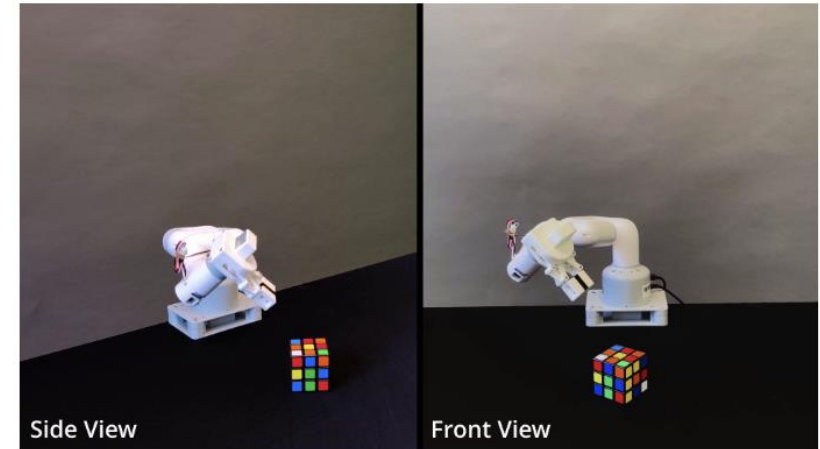
| Ref. | Description   |
|------|---|
| R1   | Imagine a non-movable object on the table in front of you. Without touching it, how would you visually observe the object to intake more information?   |
| R2   | Imagine a constant sound in the corner of the room. How would you listen to the sound to intake more information?   |
| R3   | Imagine a non-movable object on the table in front of you, but you don't understand how it works (e.g., why a computer screen isn't black). How would you reduce uncertainty about this object? |
| R4   | Imagine someone is explaining something to you. How would you show understanding and acknowledge the information?   |
| R5   | Imagine someone starts talking to you. How would you use an open posture to signal that you are listening?  |
| R6   | Imagine someone is talking to you. How would you express that you are engaged and attentive?  |
| R7   | Imagine you see something scary. How would you back away in fear?   |
| R8   | Imagine someone starts talking to you. How would you use a closed posture to signal that you are rejecting what they are saying?  |

# Robot → Human

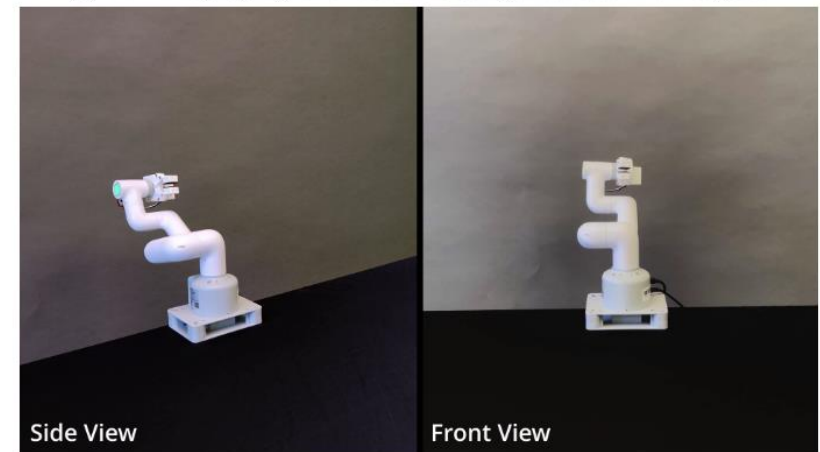
## Insights

- “쉽고 익숙한 표현” : 끄덕이기, 고개 저어 거절하기 등
- Context의 중요성 : Object, Sound Source 등
- Non-humanoid : 사람같은 표현을 할 수 있음

**Figure 4: Two example expressions from our setup to demonstrate the 13 expressions to the participants of the verification study in video form.**



**(a) E02: Displaying the robot showing interest in an object.**



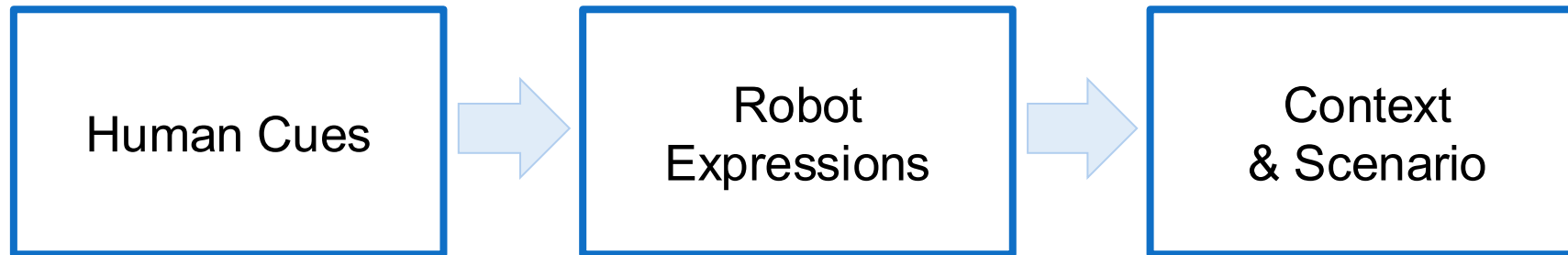
**(b) E10: Displaying the robot backing away in fear.**

# Takeaways

---

## Context & Scenario

- Human → Robot
  - 인간의 social cues를 해석할 때 Conversation Context, Conversation Role, Cognitive Load multimodal하게 인식해야 함
- Robot → Human
  - 로봇은 familiar, context-rich expressions을 사용해야 함



P I X A R

HUMAN  
CENTERED  
COMPUTING  
LABORATORY

