Implicit Communication in Human-Machine Collaboration

FA9550-17-1-0308

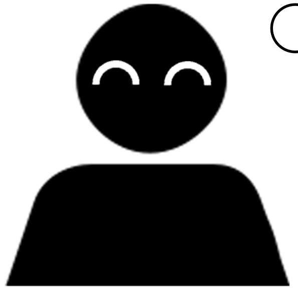
Anca Dragan

BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

InterACT Laboratory

Inferring internal state from human behavior.


Speed: 0.80



Reward

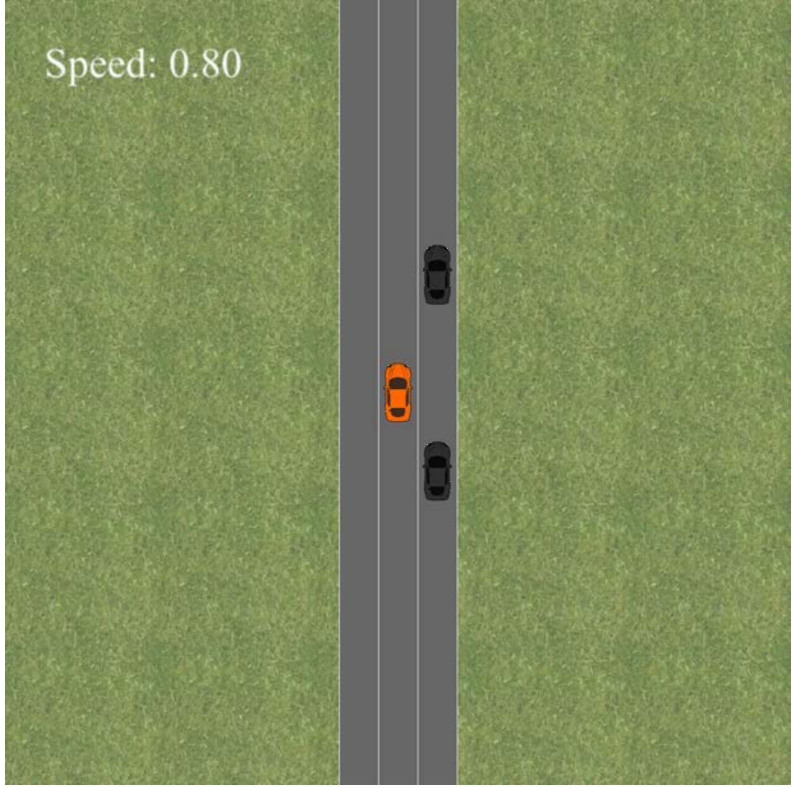
Features

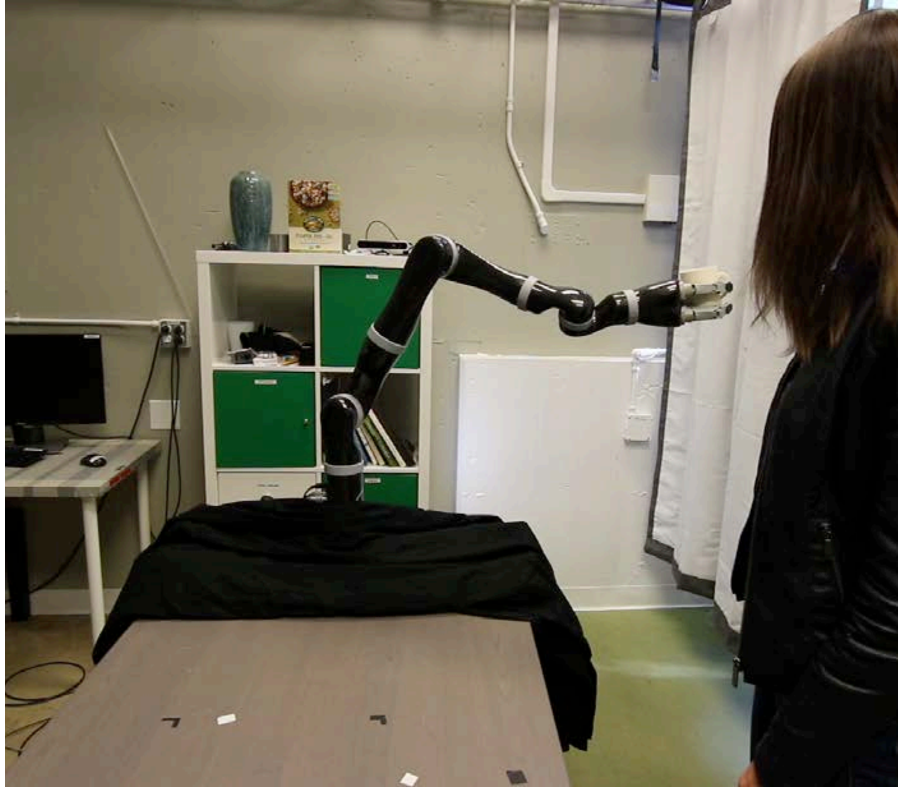
$$R_{\theta}(\Phi(\xi))$$

Weights

Trajectory

$$p(\xi_H|\theta) = \frac{e^{R_{\theta}(\Phi(\xi_H))}}{\int_{\Xi} e^{R_{\theta}(\Phi(\bar{\xi}))} d\bar{\xi}}$$

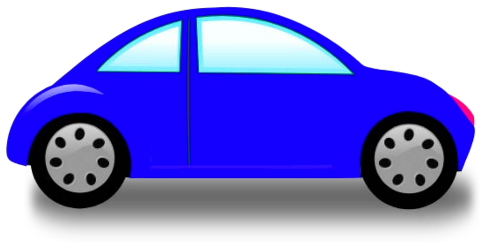


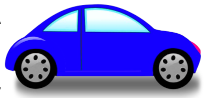


[Ziebart et al., 2009]
[Henry et al., 2010]
[Vasquez et al., 2014]
[Pfeiffer et al., 2016]
[Kretzschmar et al., 2018]
[Fisac, Bajcsy* et al., 2018]*

[Ziebart et al., 2008]
[Kitani et al., 2012]
[Wulfmeier et al., 2015]
[Sadigh et al., 2016]

[Kalakrishnan et al., 2013]
[Mainprice et al., 2013]
[Mainprice et al., 2015]
[Finn et al., 2016]
[Bobu et al., 2018]



$v($  $)$

$v($  $)$

Luce: $$P(o) = \frac{v(o)}{\sum_{\overline{o} \in O} v(\overline{o})}$$

Luce-Shepard: $$P(o) = \frac{e^{R_\theta(o)}}{\sum_{\overline{o} \in O} e^{R_\theta(\overline{o})}}$$

Inference: $$P(\theta | o) \propto \frac{e^{R_\theta(o)}}{\sum_{\overline{o} \in O} e^{R_\theta(\overline{o})}} P(\theta)$$

What if we have the wrong human model?

we don't have the right hypothesis space

Inference: $P(\theta|o) \propto \frac{e^{R_\theta(o)}}{\sum_{\bar{o} \in O} e^{R_\theta(\bar{o})}} P(\theta)$

this is not how people choose

1. **Fix the model.**
2. **Detect that the model is wrong.**

What if we have the wrong human model?

we don't have the right hypothesis space

$$\text{Inference:}\quad P(\theta|o) \propto \frac{e^{R_\theta(o)}}{\sum_{\bar{o}\in O} e^{R_\theta(\bar{o})}} P(\theta)$$

this is not how people choose

1. **Fix the model.**
2. **Detect that the model is wrong.**

Are there fundamental, general-purpose ways in which the model is wrong?

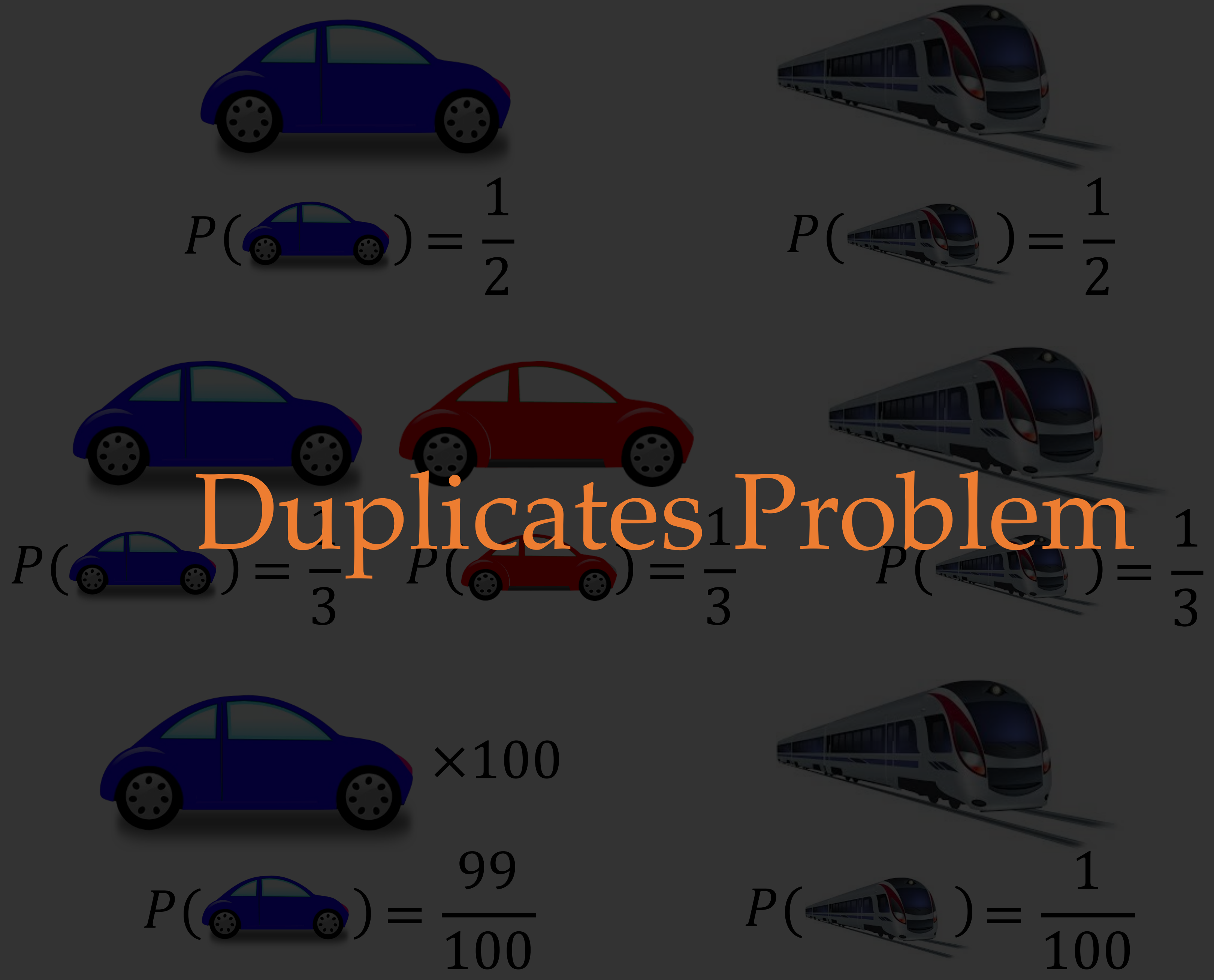
we don't have the right hypothesis space

$$\text{Inference: } P(\theta|o) \propto \frac{e^{R_\theta(o)}}{\sum_{\overline{o}\in O} e^{R_\theta(\overline{o})}} P(\theta)$$

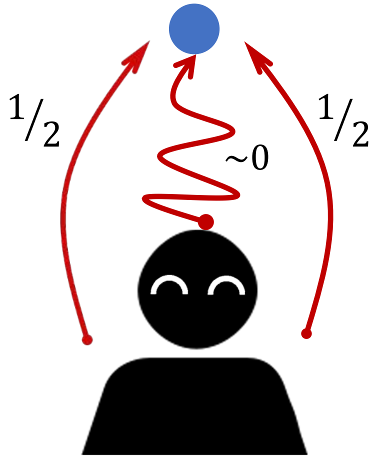
this is not how people choose

1. **Fix the model.**
2. **Detect that the model is wrong.**

Duplicates Problem



Boltzmann: $p(\xi_H) = \dfrac{e^{R(\Phi(\xi_H))}}{\int_{\Xi} e^{R(\Phi(\bar{\xi}))} d\bar{\xi}}$



Key Insight

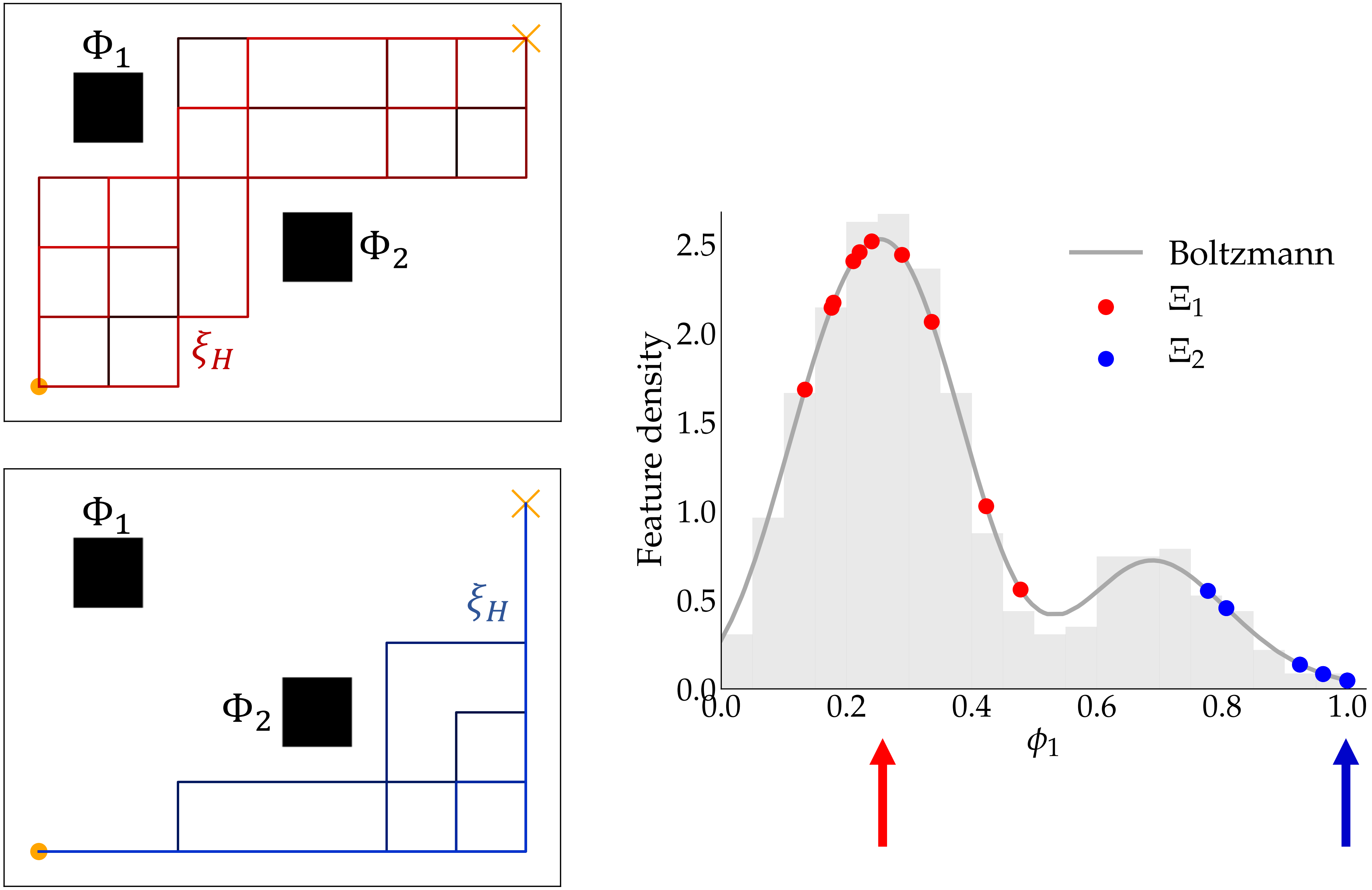
We need to account for how similarity in trajectories should influence their probability.

Boltzmann: $P(\xi) \propto e^{R(\phi(\xi))}$

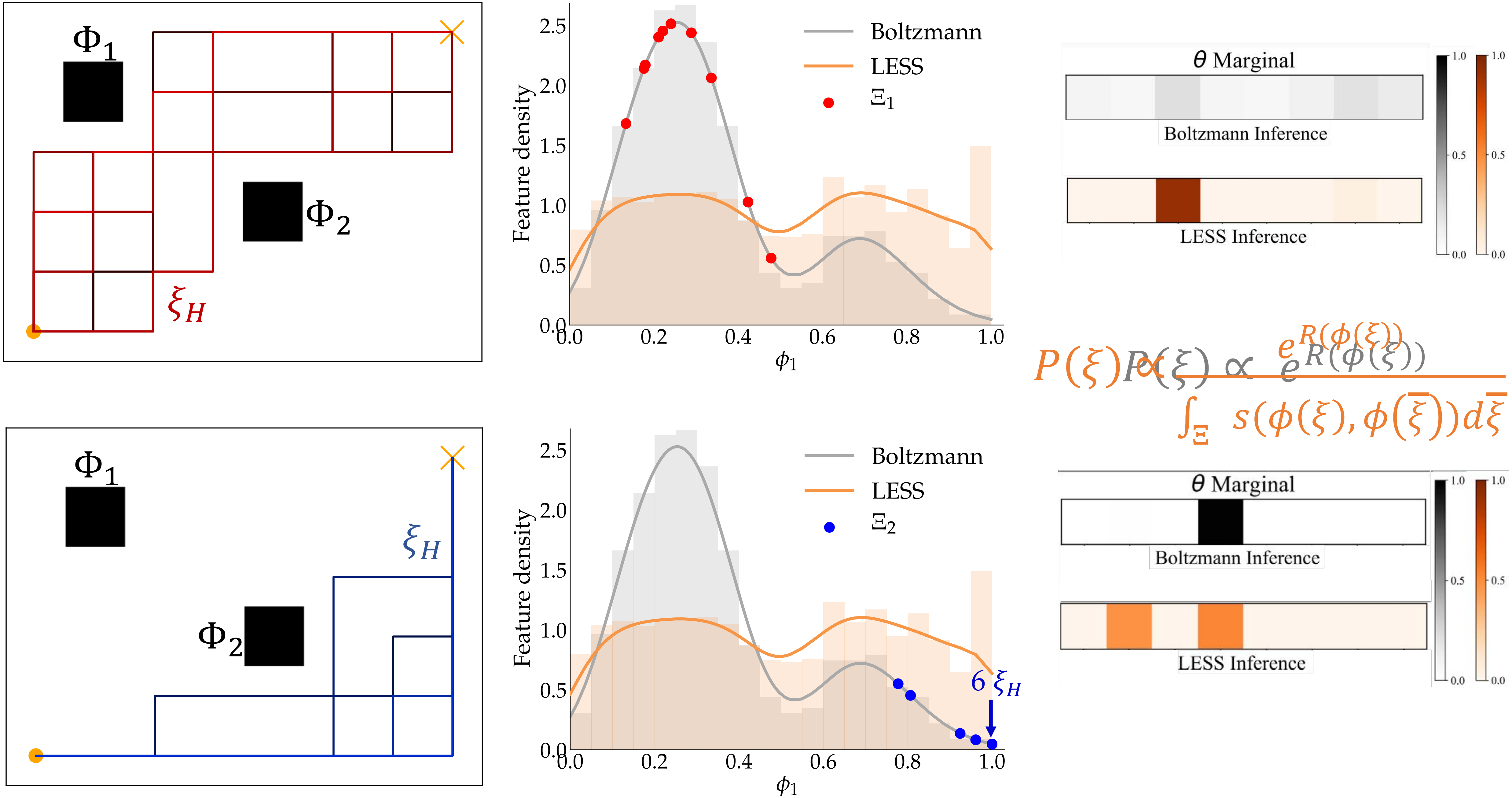
Limit Errors due to Similar Selections

LESS: $P(\xi) \propto \dfrac{e^{R(\phi(\xi))}}{\int_{\Xi} s(\phi(\xi), \phi(\bar{\xi}))d\bar{\xi}}$

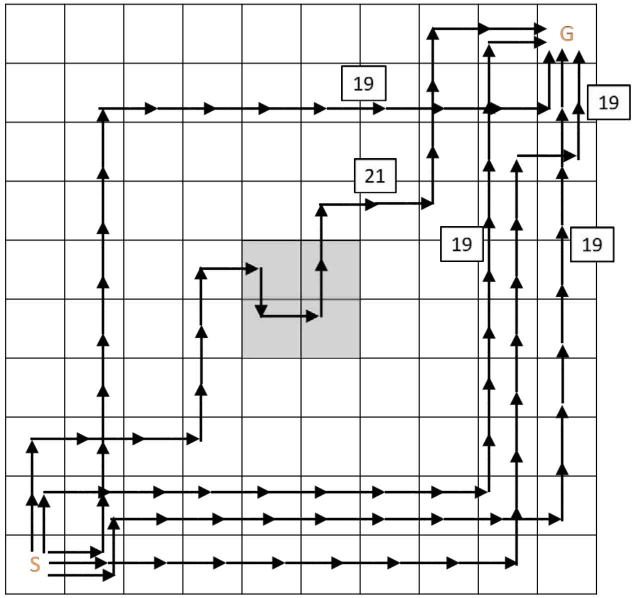
Similarity Metric

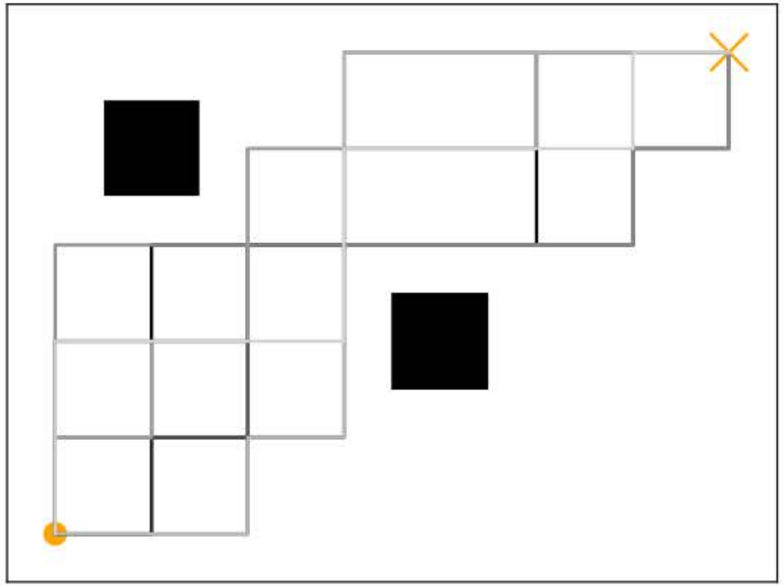


Φ_1
Φ_2
ξ_H
Φ_1
Φ_2
ξ_H
Feature density
2.5
2.0
1.5
1.0
0.5
0.0
0.0
0.2
0.4
0.6
0.8
1.0
ϕ_1
Boltzmann
Ξ_1
Ξ_2

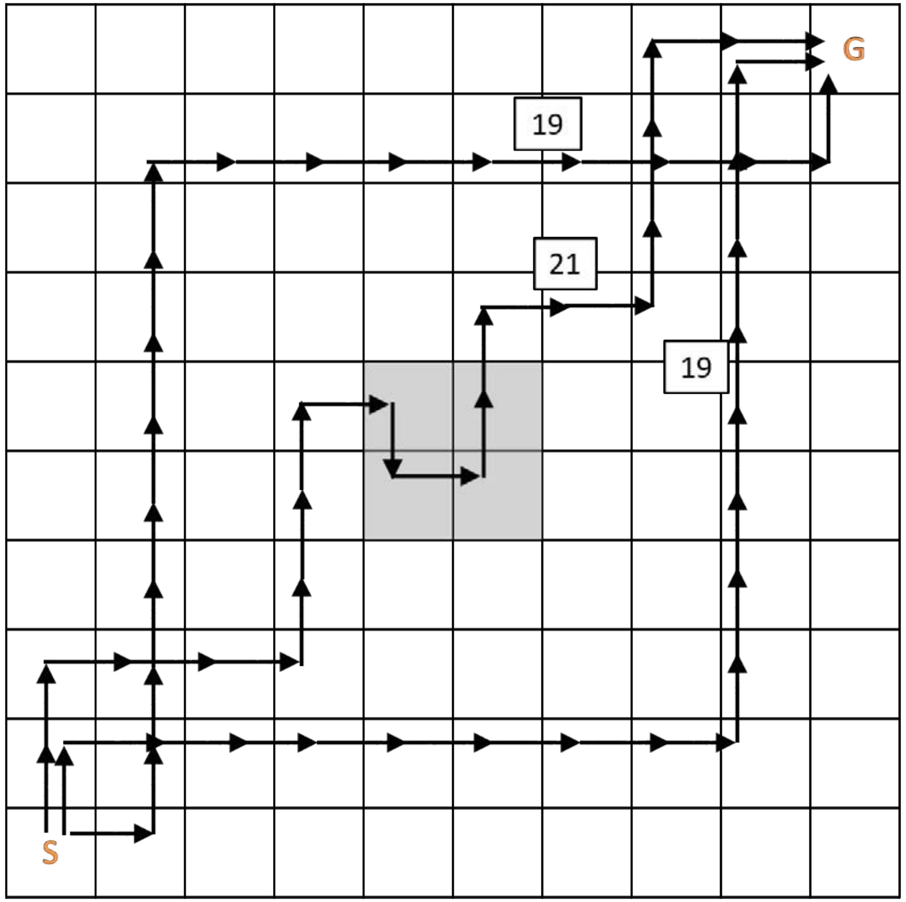


Φ1
Φ2
ξH
Boltzmann
LESS
Ξ1
Ξ2
Feature density
ϕ1
0.0
0.5
1.0
1.5
2.0
2.5
0.2
0.4
0.6
0.8
θ Marginal
Boltzmann Inference
LESS Inference
6 ξH
P(ξ) ∝ e^{R(ϕ(ξ))} / ∫Ξ s(ϕ(ξ), ϕ(ξ̄)) dξ̄
P(ξ) ∝ e^{R(ϕ(ξ))}

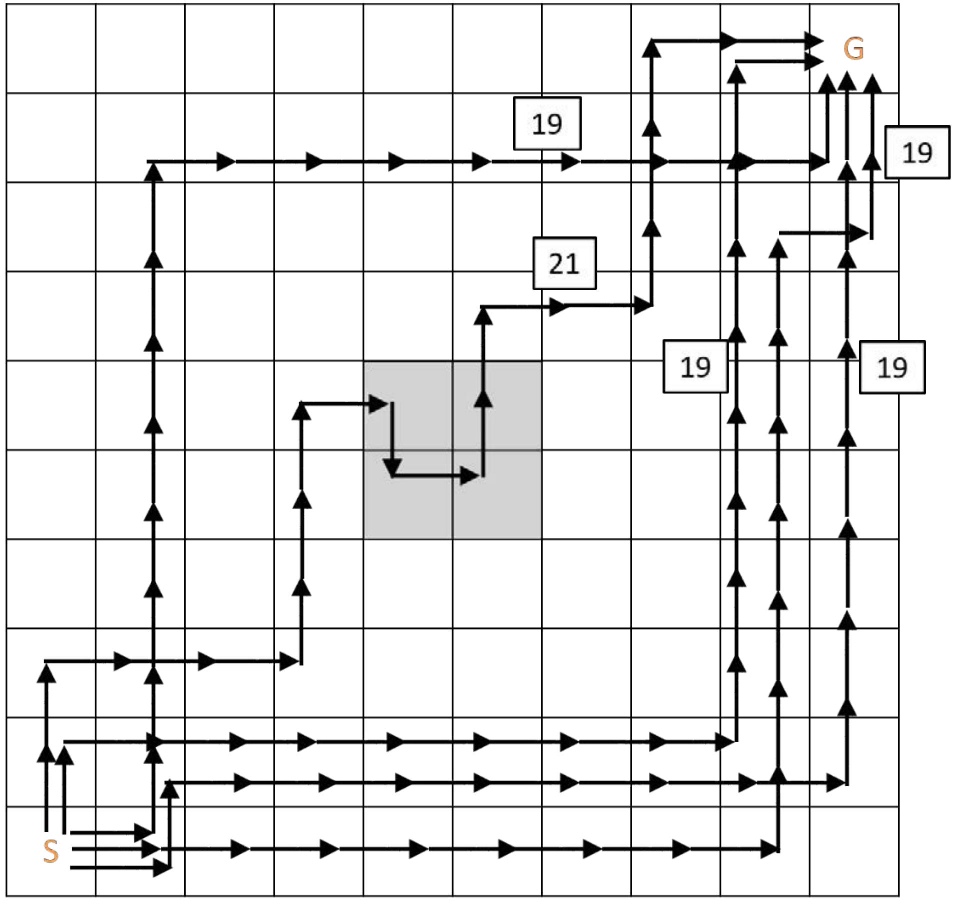

G
19
19
21
19
19
S



LESS as a Human Decision Model

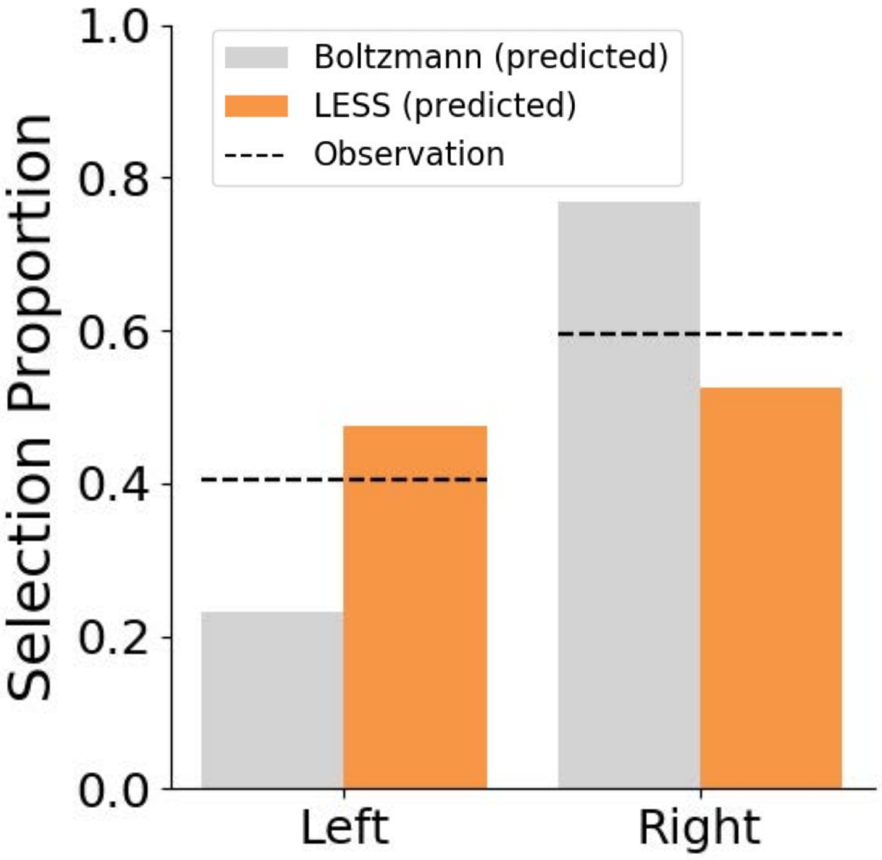


Control trial



Experimental trial

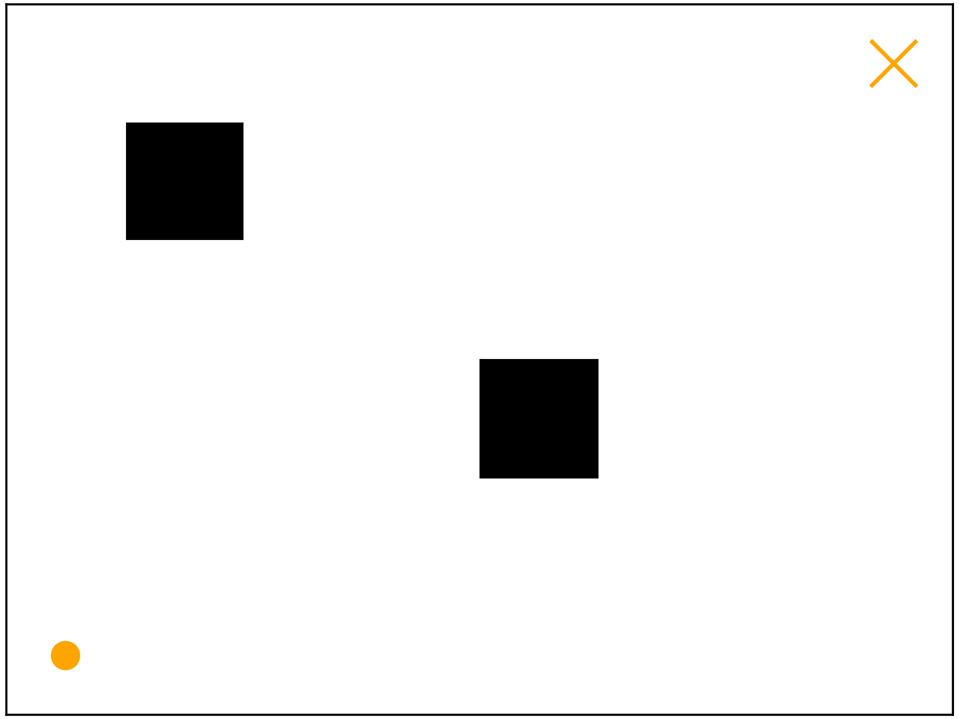
LESS as a Human Decision Model



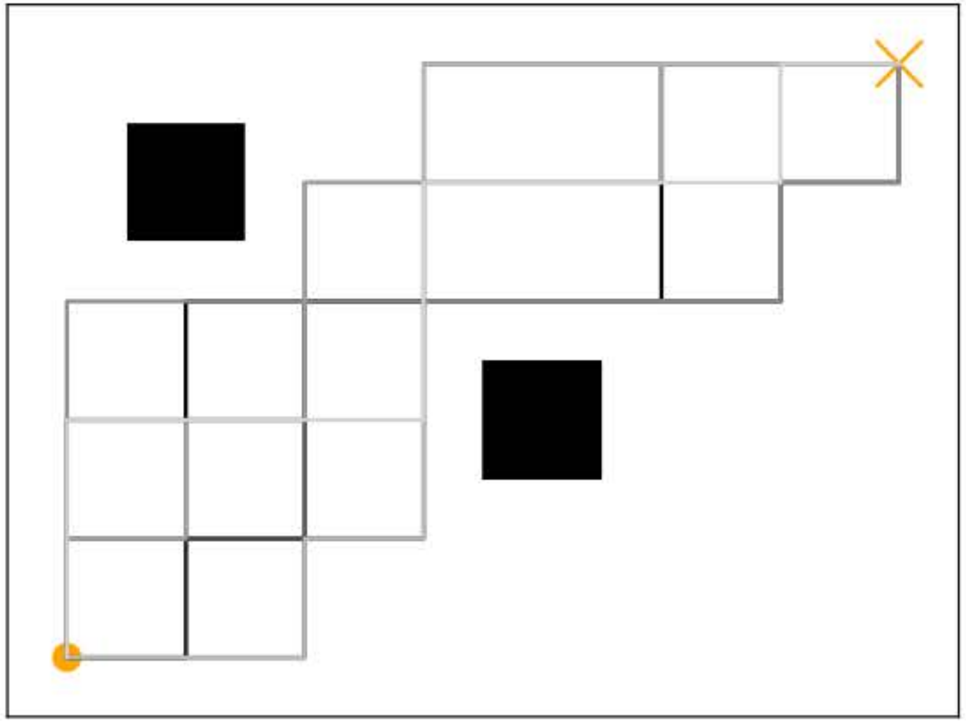
O1: Boltzmann predictions are significantly different from observed proportions.

O2: LESS predictions have a tighter equivalence bound to the observed proportions than Boltzmann predictions.

Using LESS for Robot Inference

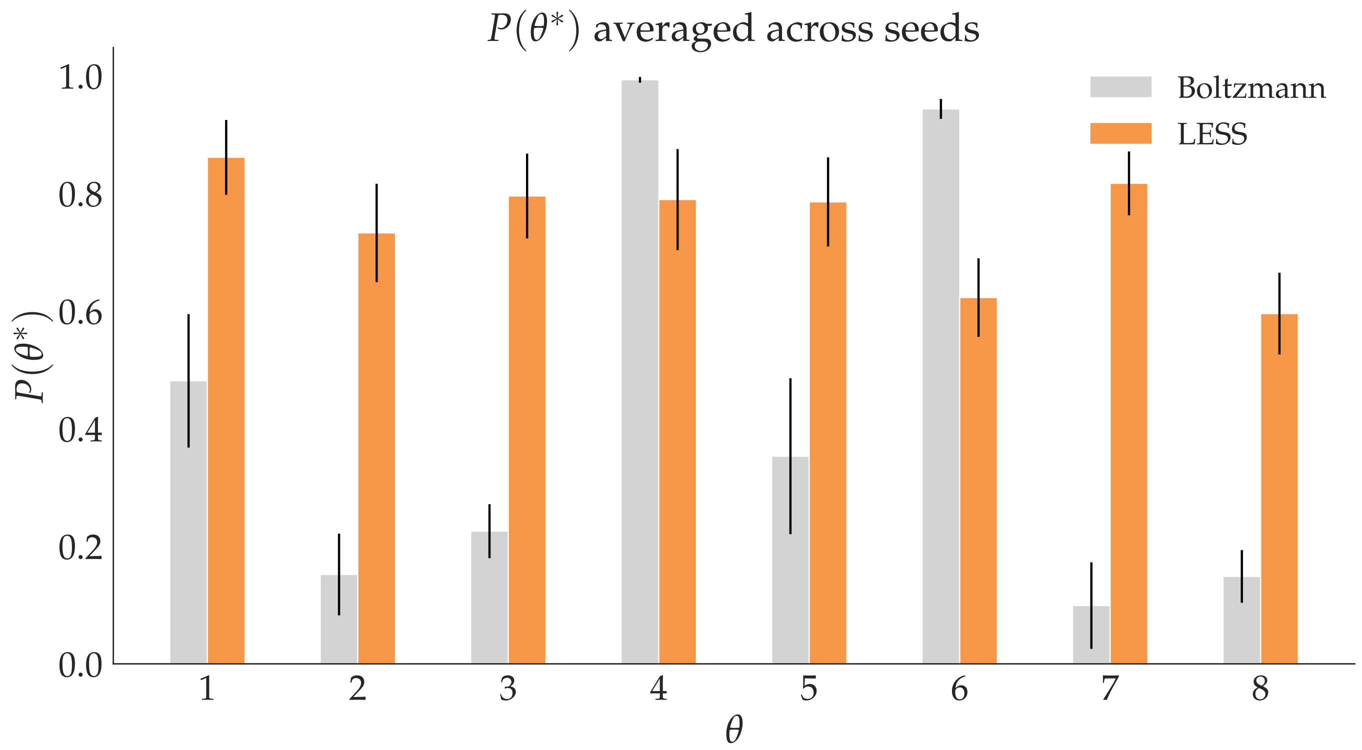


Gridworld



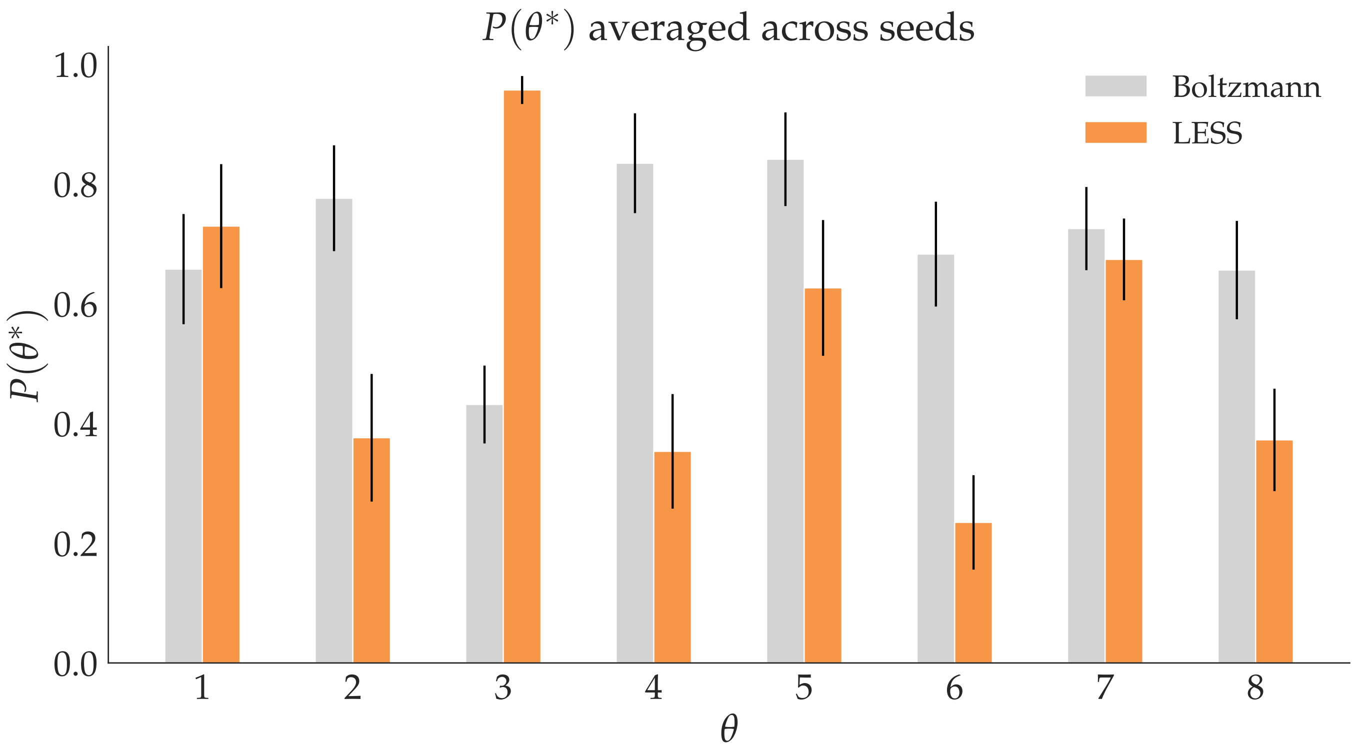
Sampling under θ

LESS Sampling



When human input is generated using LESS, inference quality is significantly higher with LESS than with Boltzmann.

Boltzmann Sampling



When human input is generated using Boltzmann, inference quality is significantly higher with Boltzmann than with LESS.

Robust Inference for Robot Arms

Boltzmann: $P(\xi) = \dfrac{e^{R(\phi(\xi))}}{\int_{\Xi} e^{R(\phi(\bar{\xi}))} d\bar{\xi}}$

Computing the denominator is intractable!

LESS: $P(\xi) \propto \dfrac{e^{R(\phi(\xi))}}{\int_{\Xi} s(\phi(\xi), \bar{\xi}) d\bar{\xi}} = \dfrac{\dfrac{e^{R(\phi(\xi))}}{\int_{\Xi} s(\phi(\xi), \bar{\xi}) d\bar{\xi}}}{\int_{\Xi} \dfrac{e^{R(\phi(\hat{\xi}))}}{\int_{\Xi} s(\phi(\hat{\xi}), \bar{\xi}) d\bar{\xi}} d\hat{\xi}}$

Need to sample Ξ!

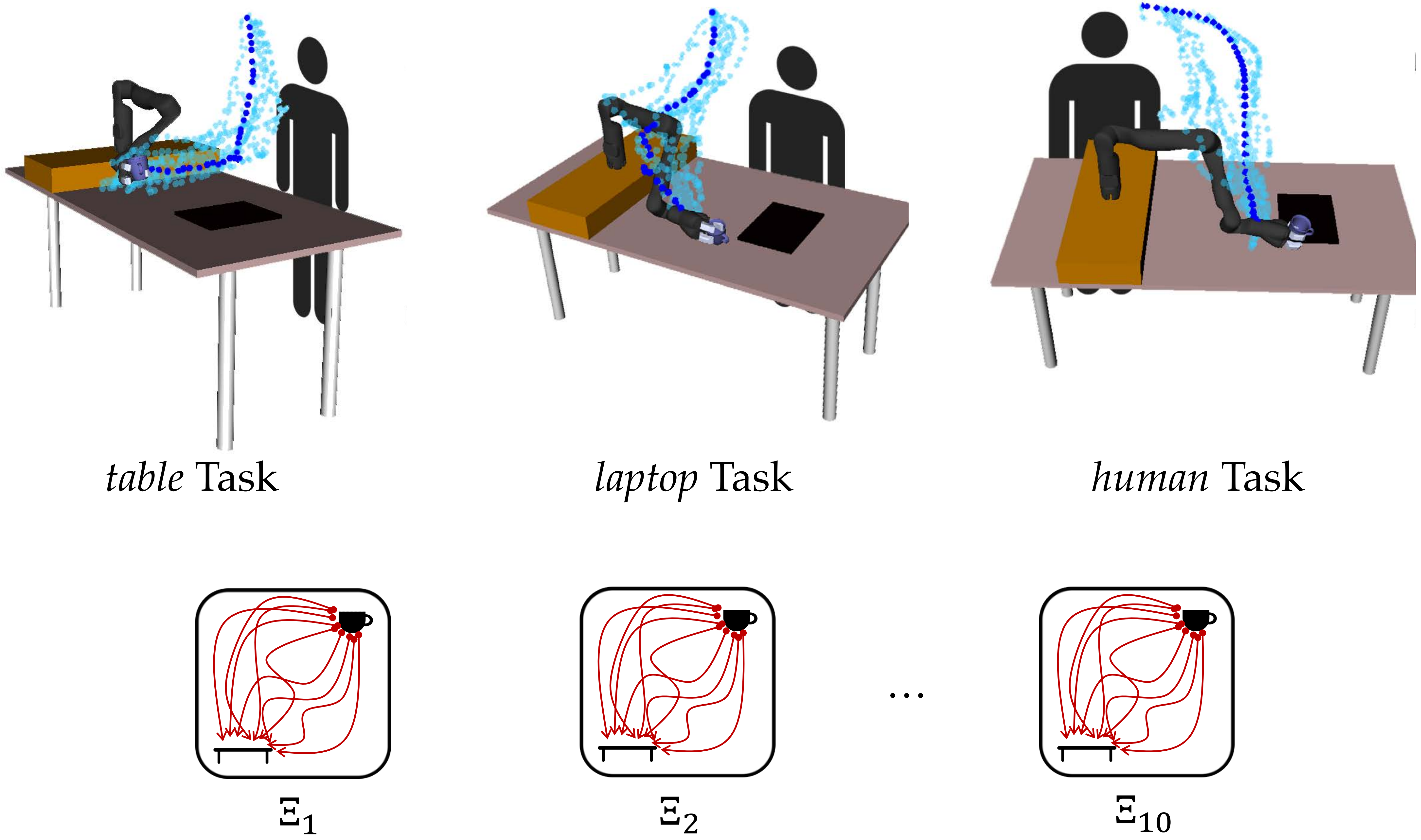
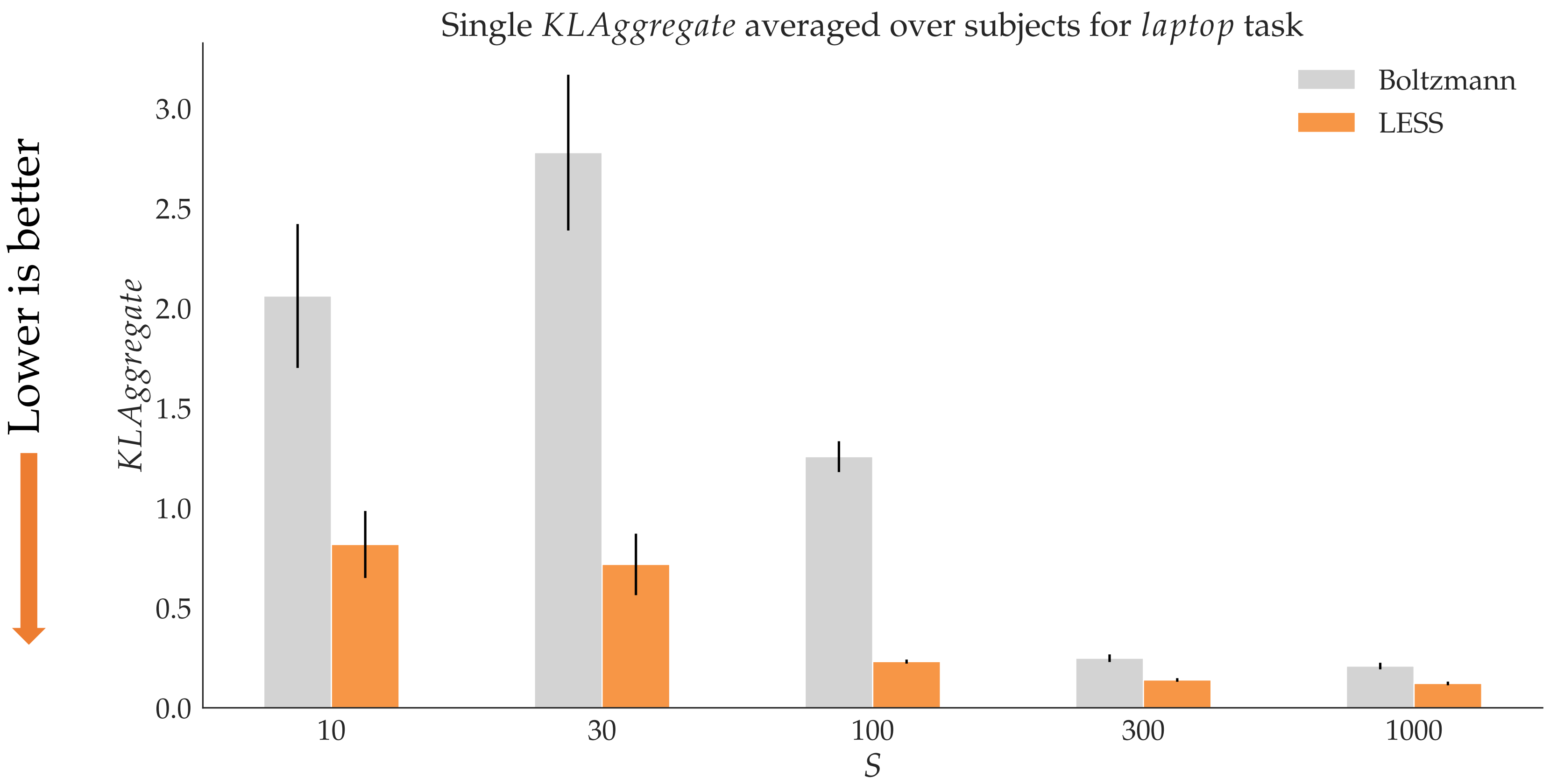


table Task
laptop Task
human Task
Ξ_1
Ξ_2
$\dots$
Ξ_{10}



Performing inference with LESS across multiple trajectory sets results in higher robustness than inference with Boltzmann.

What if we have the wrong human model?

we don't have the right hypothesis space

$$\text{Inference:}\quad P(\theta|o) \propto \frac{e^{R_\theta(o)}}{\sum_{\overline{o}\in O} e^{R_\theta(\overline{o})}} P(\theta)$$

this is not how people choose

1. **Fix the model.**
2. **Detect that the model is wrong.**

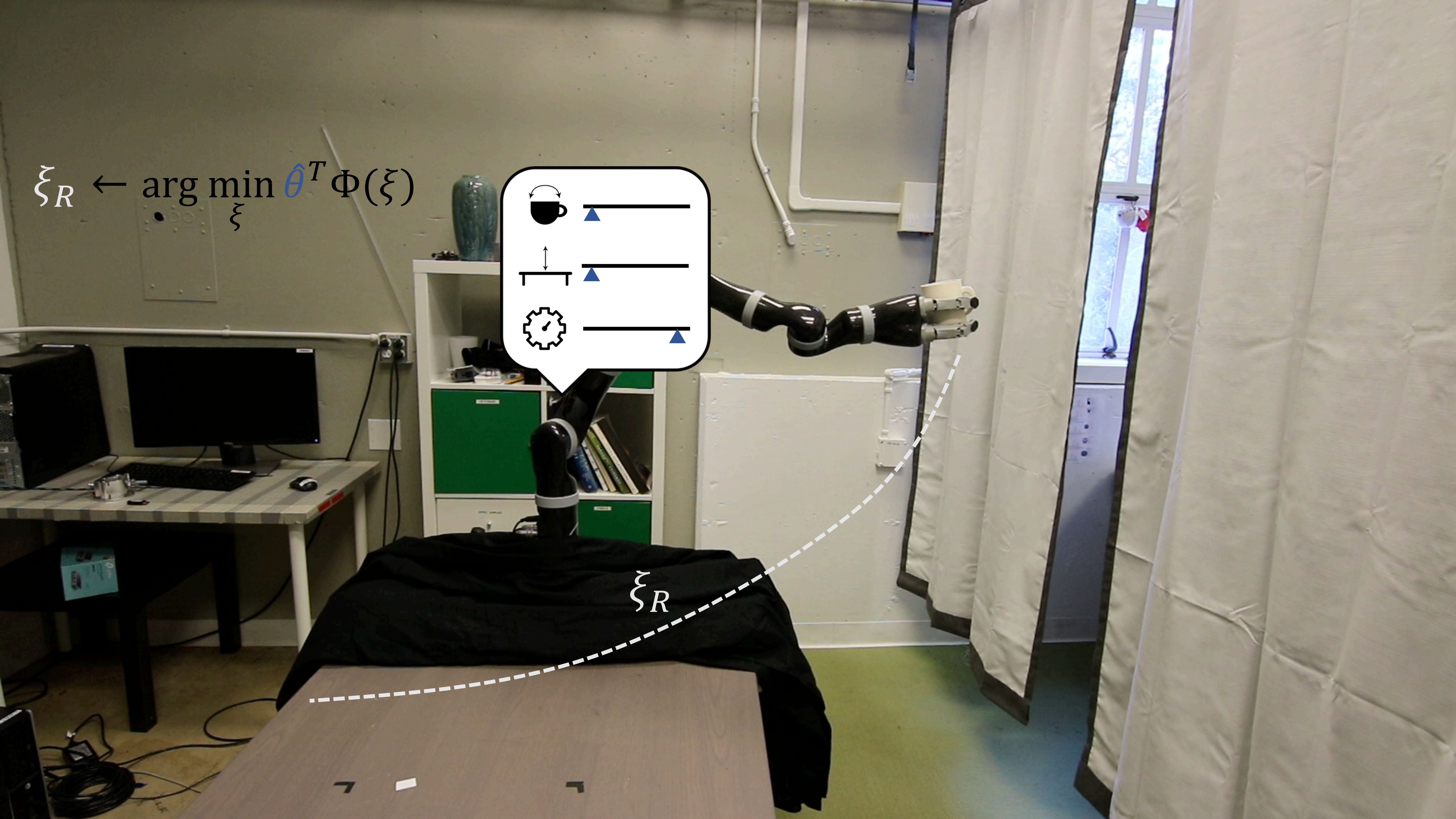
What if we have the wrong human model?

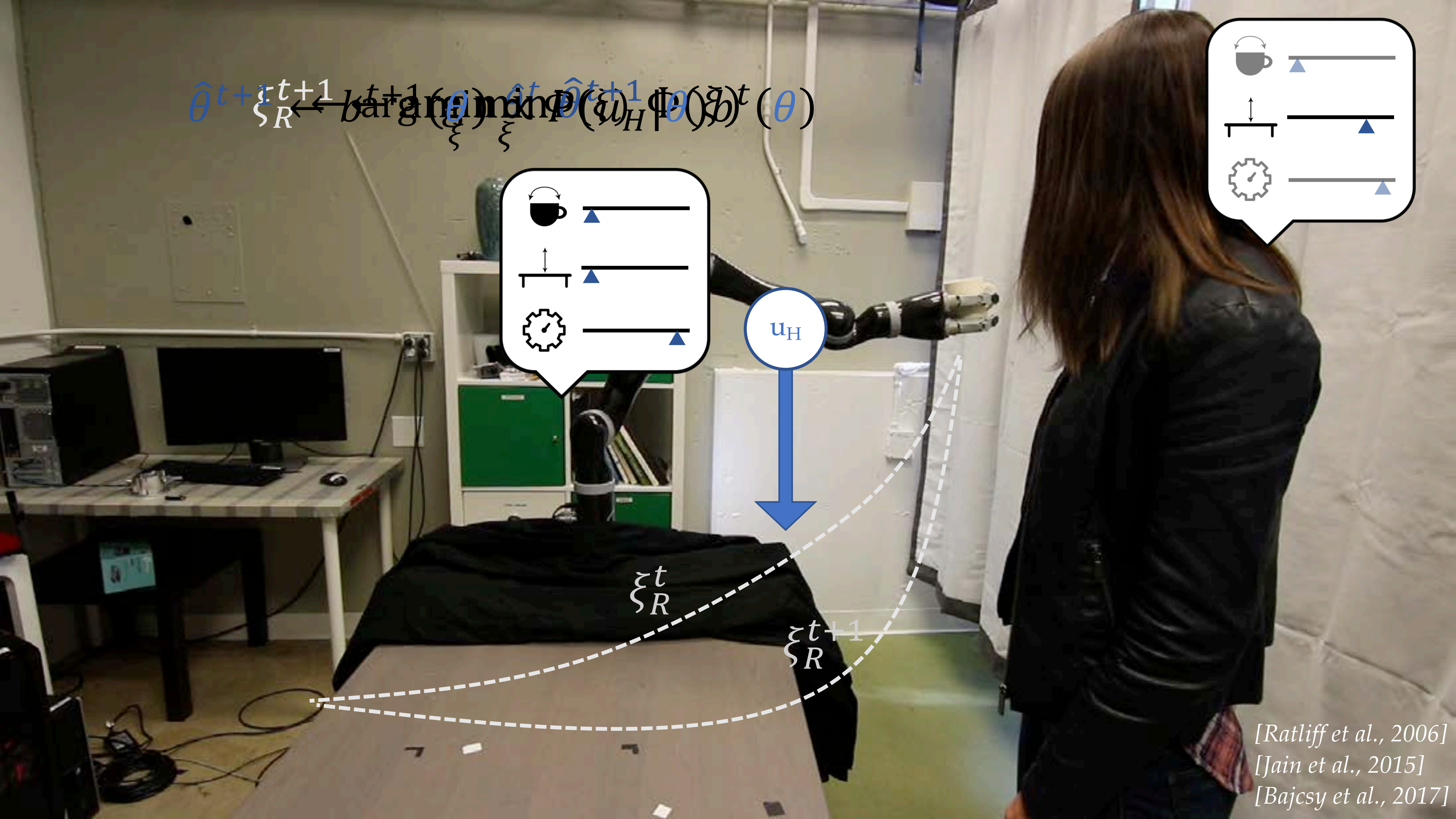
we don't have the right hypothesis space

$$\text{Inference:}\quad P(\theta|o) \propto \frac{e^{R_\theta(o)}}{\sum_{\overline{o}\in O} e^{R_\theta(\overline{o})}} P(\theta)$$

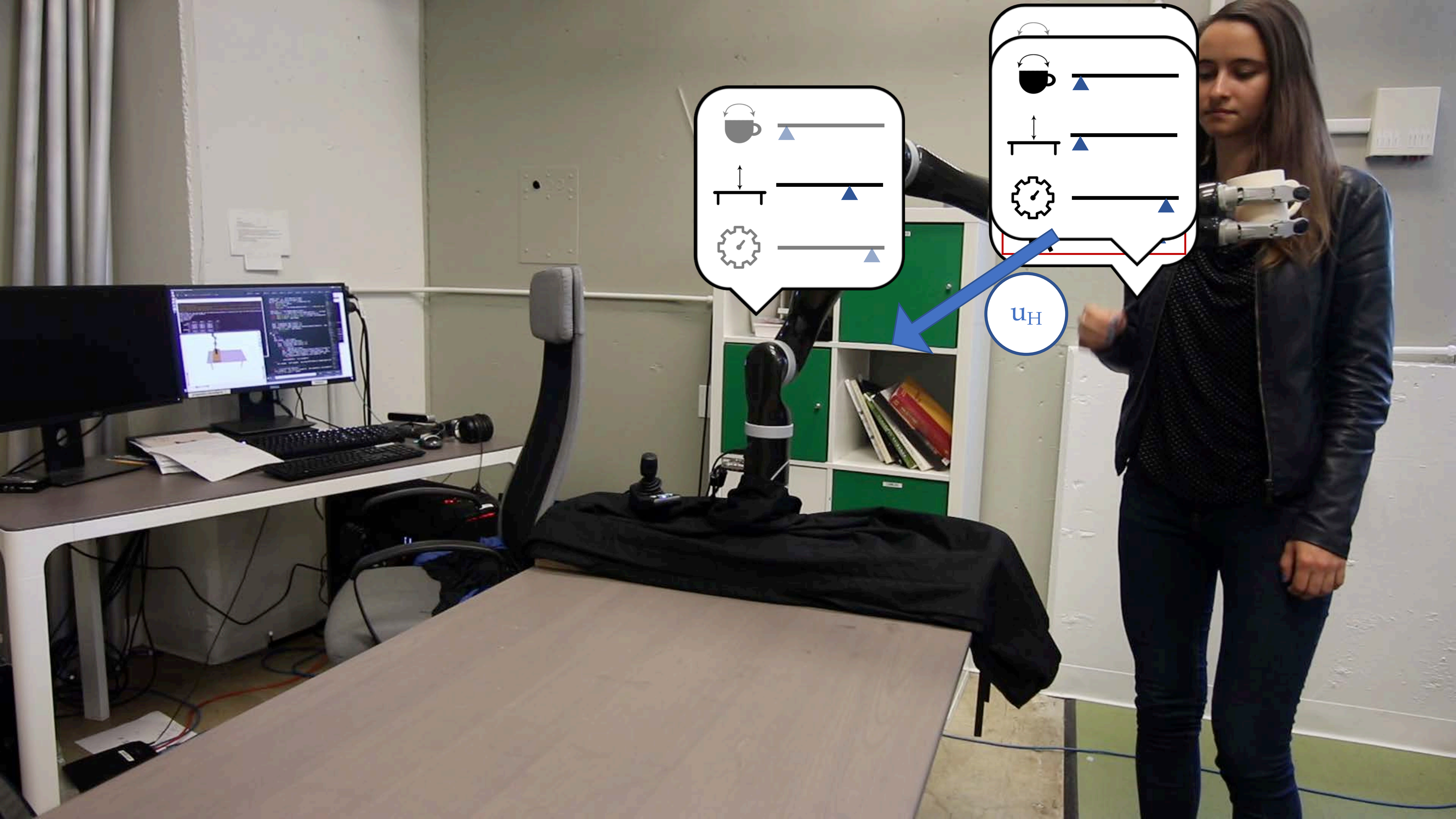
this is not how people choose

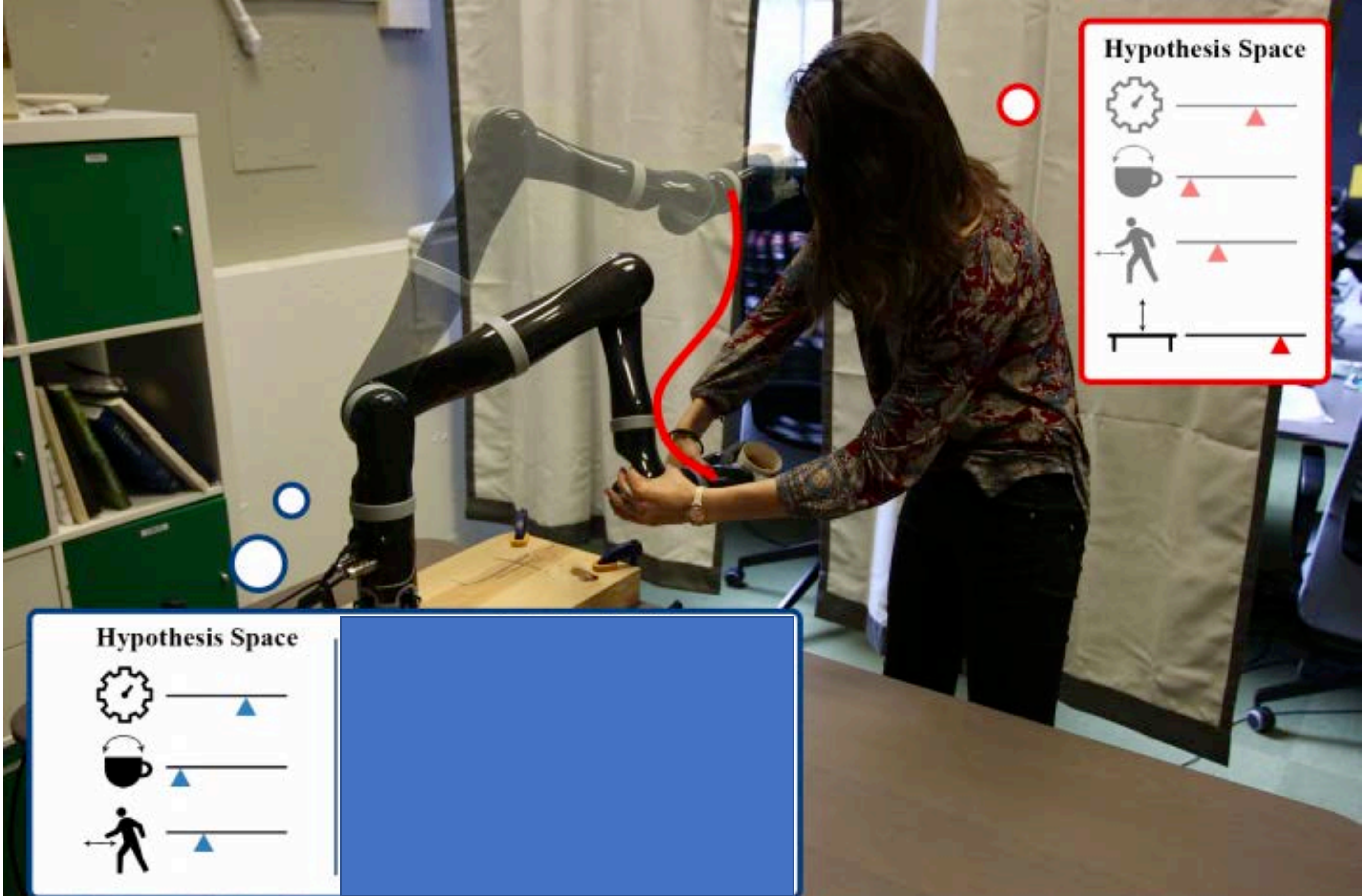
1. **Fix the model.**
2. **Detect that the model is wrong.**


$\xi_R \leftarrow \arg\min_{\xi} \hat{\theta}^T \Phi(\xi)$
ξ_R



[Ratliff et al., 2006]
[Jain et al., 2015]
[Bajcsy et al., 2017]


u_H


Hypothesis Space
Hypothesis Space

$$P(\theta|o) \propto \frac{e^{R_\theta(o)}}{\sum_{\overline{o}\in O} e^{R_\theta(\overline{o})}} P(\theta)$$

Rationality coefficient

$$P(\theta|o) \propto \frac{e^{\beta R_\theta(o)}}{\sum_{\overline{o} \in O} e^{\beta R_\theta(\overline{o})}} P(\theta)$$

Key Insight

Estimate apparent rationality.

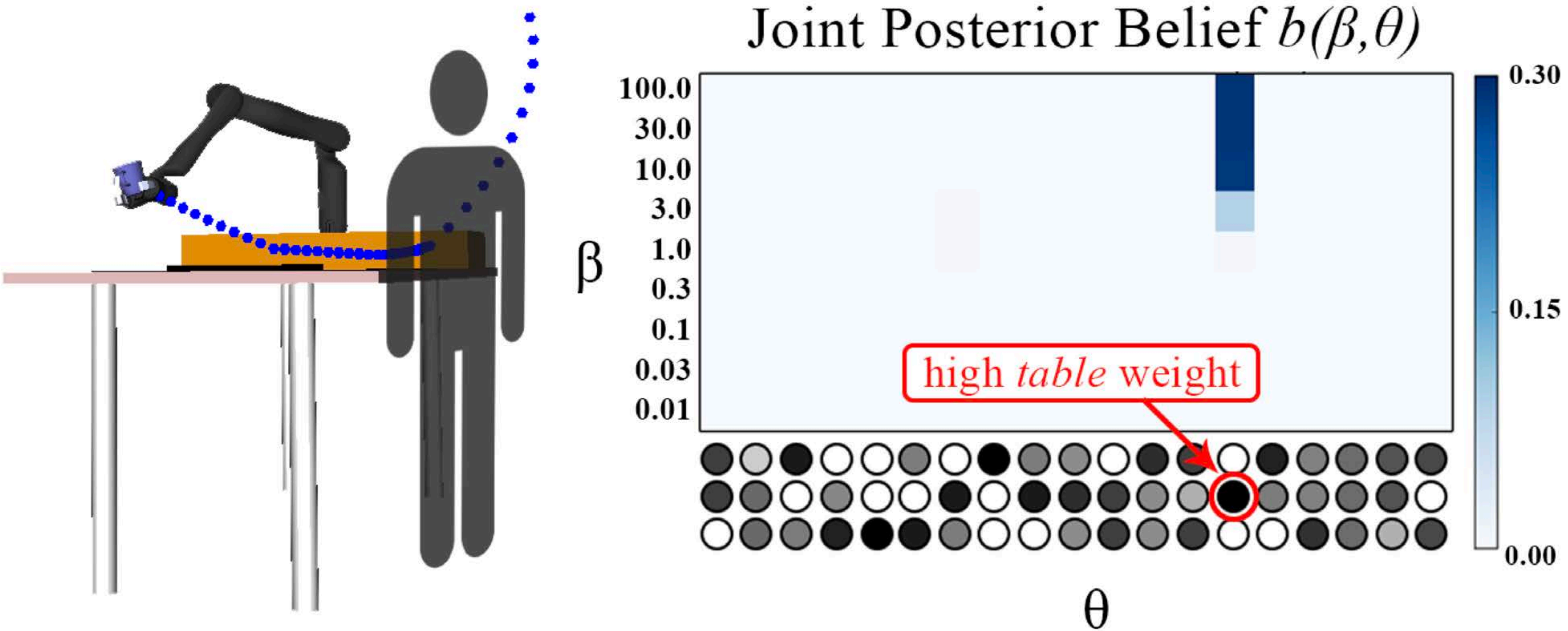
If the human appears irrational to the robot,

the robot has the wrong model of the human.

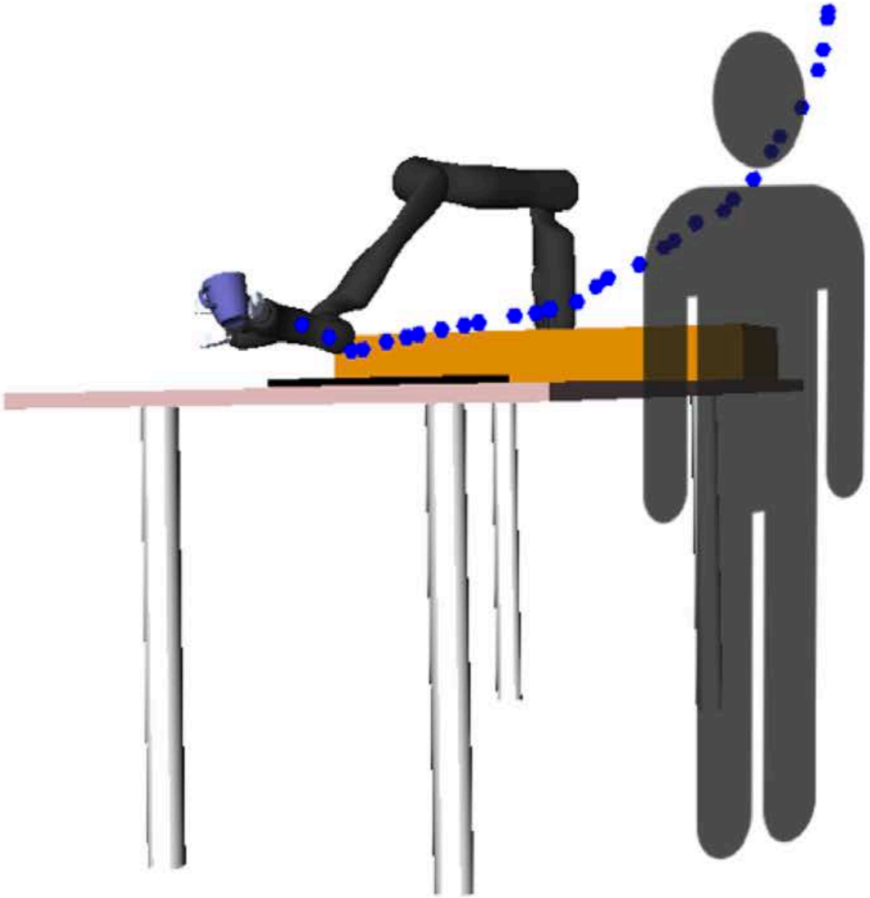
Joint estimation

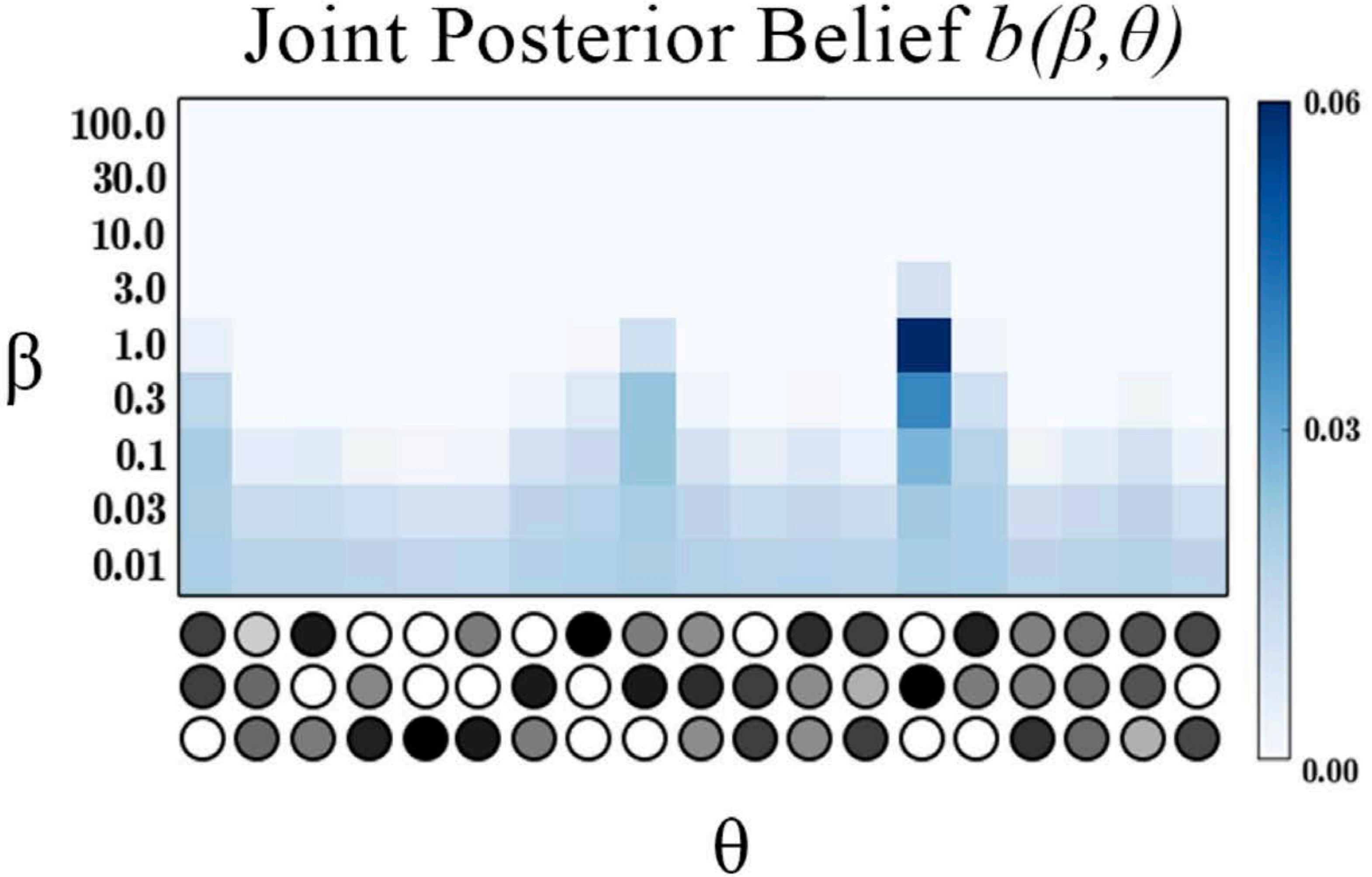
$$P(\theta, \beta | o) \propto \frac{e^{\beta R_\theta(o)}}{\sum_{\bar{o} \in O} e^{\beta R_\theta(\bar{o})}} P(\theta)$$

Demonstration, correctly specified

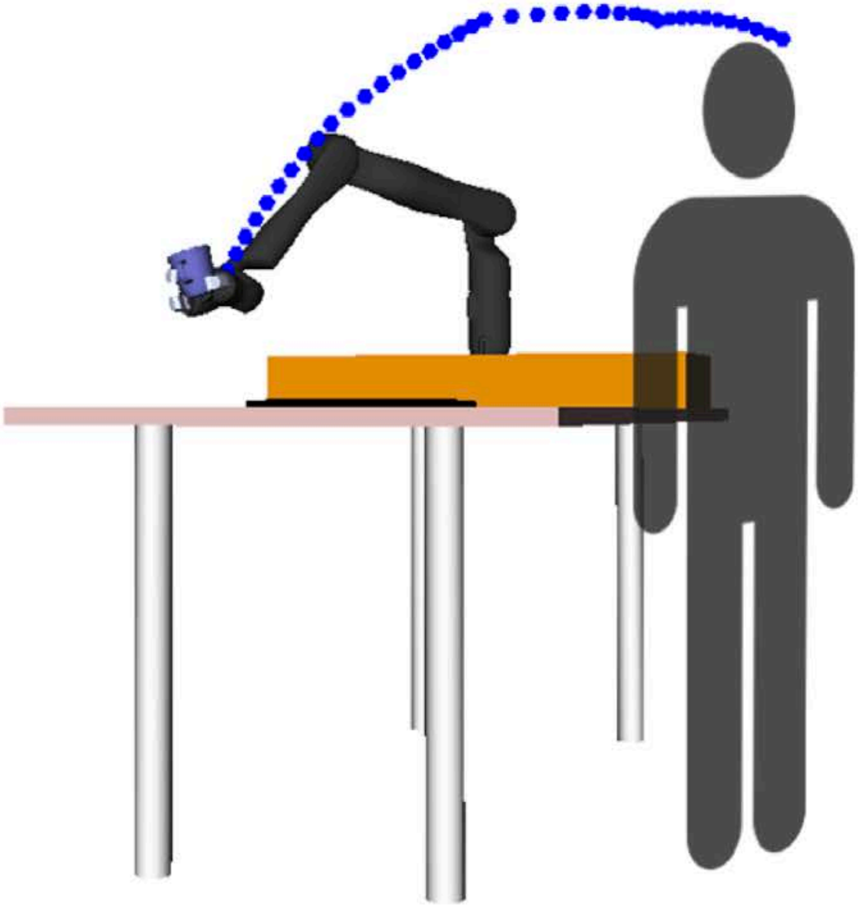


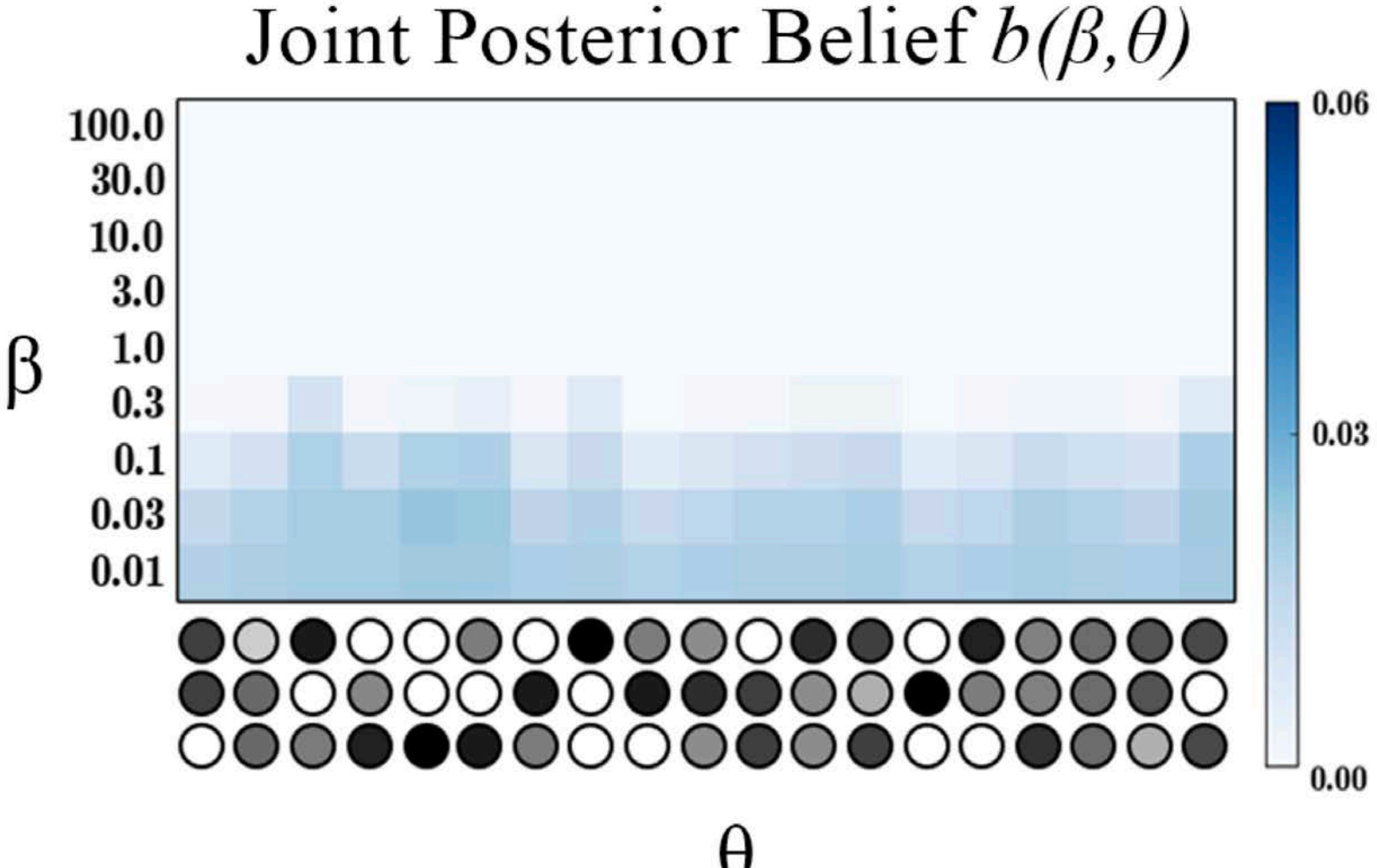
Demonstration, noisy

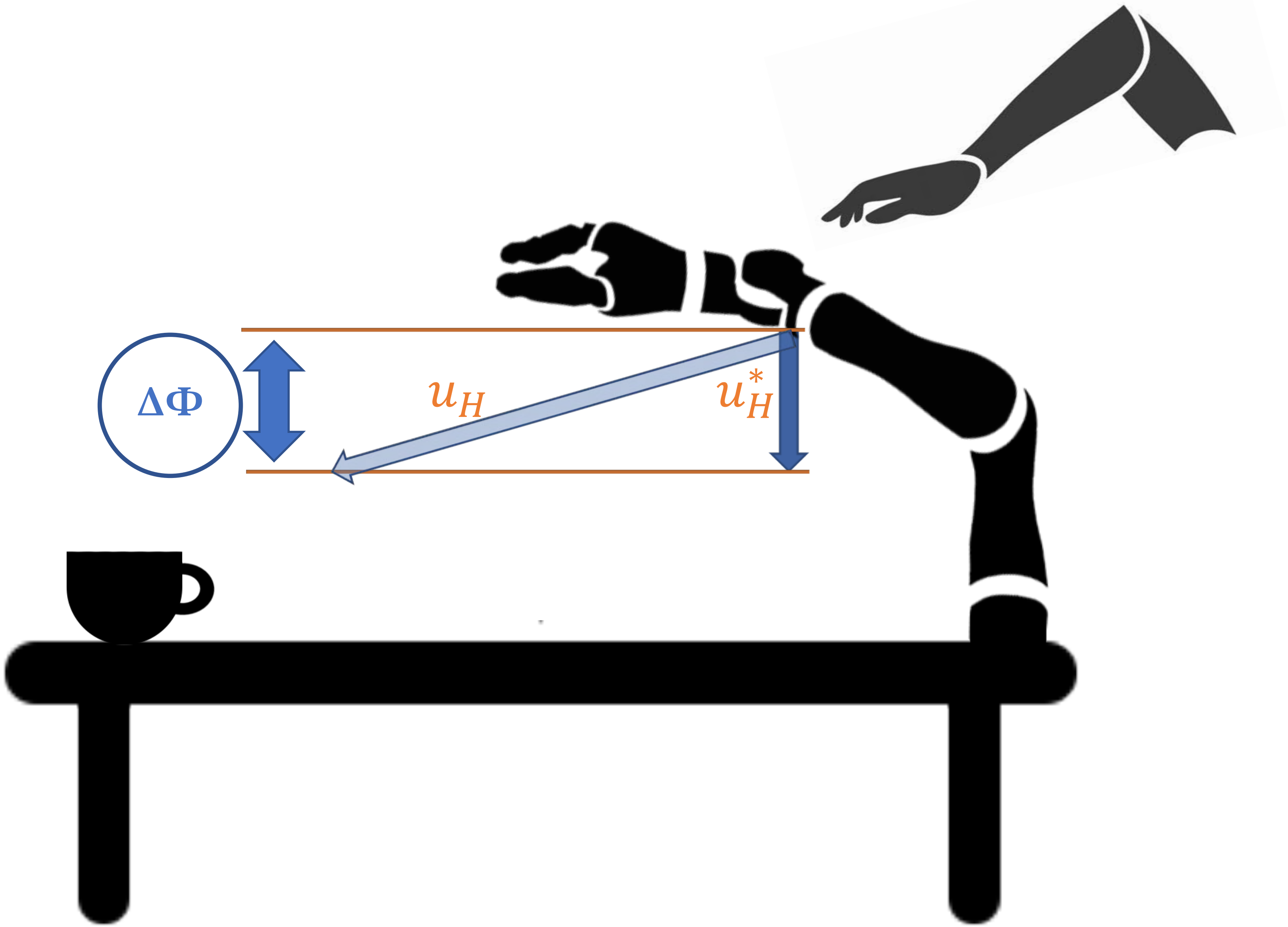




Demonstration, misspecified



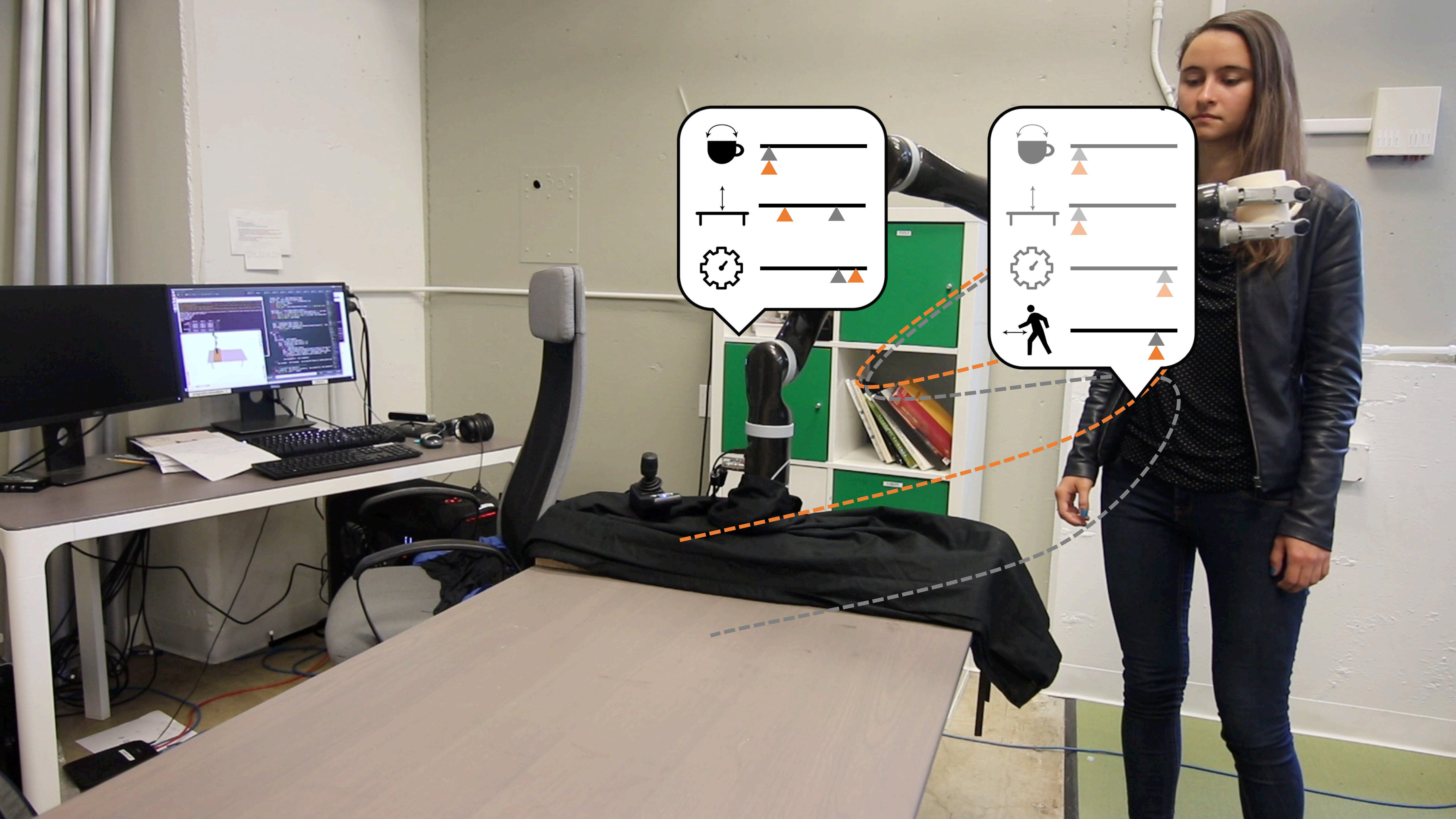


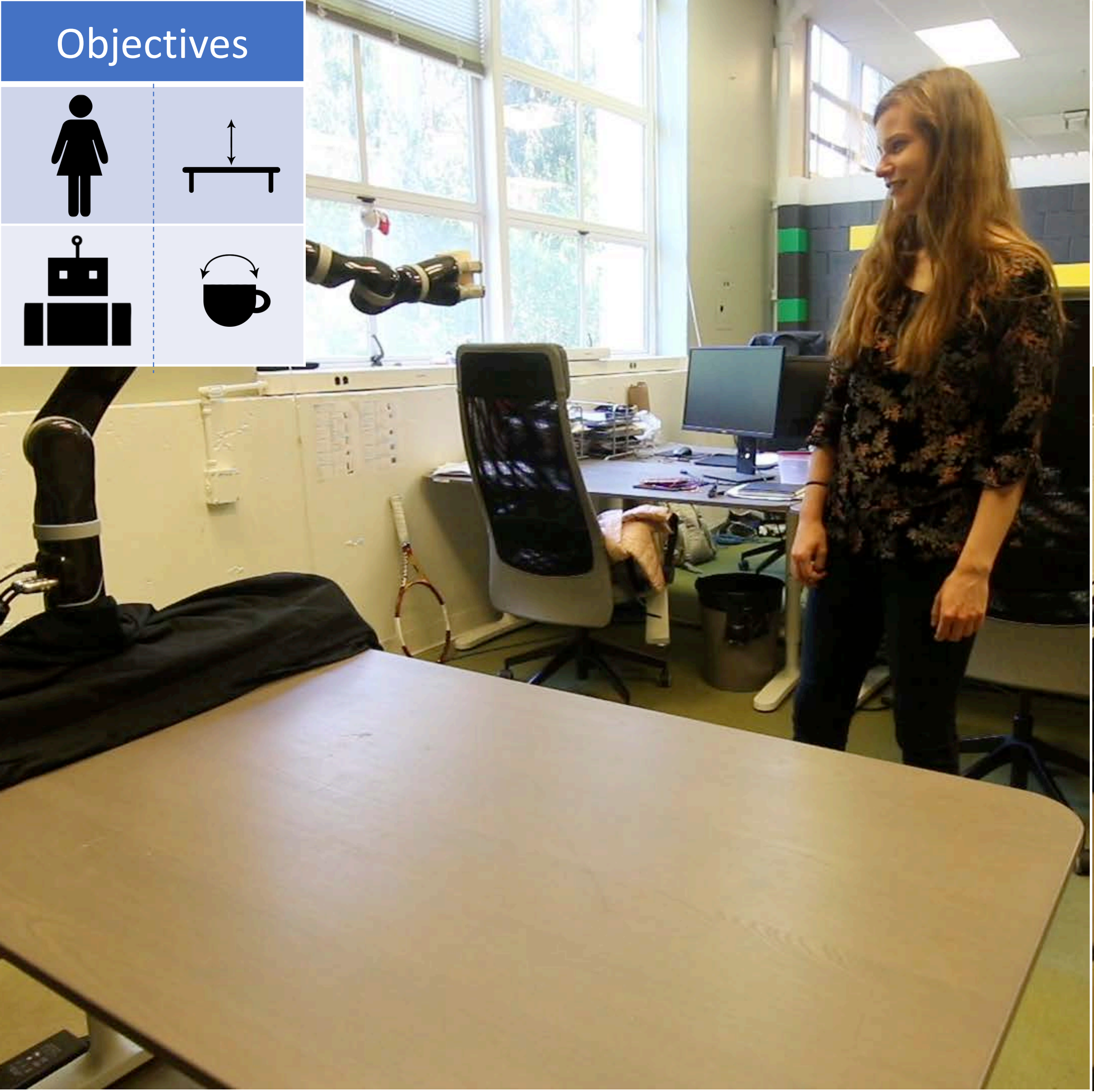


$\Delta\Phi$
u_H
u_H^*

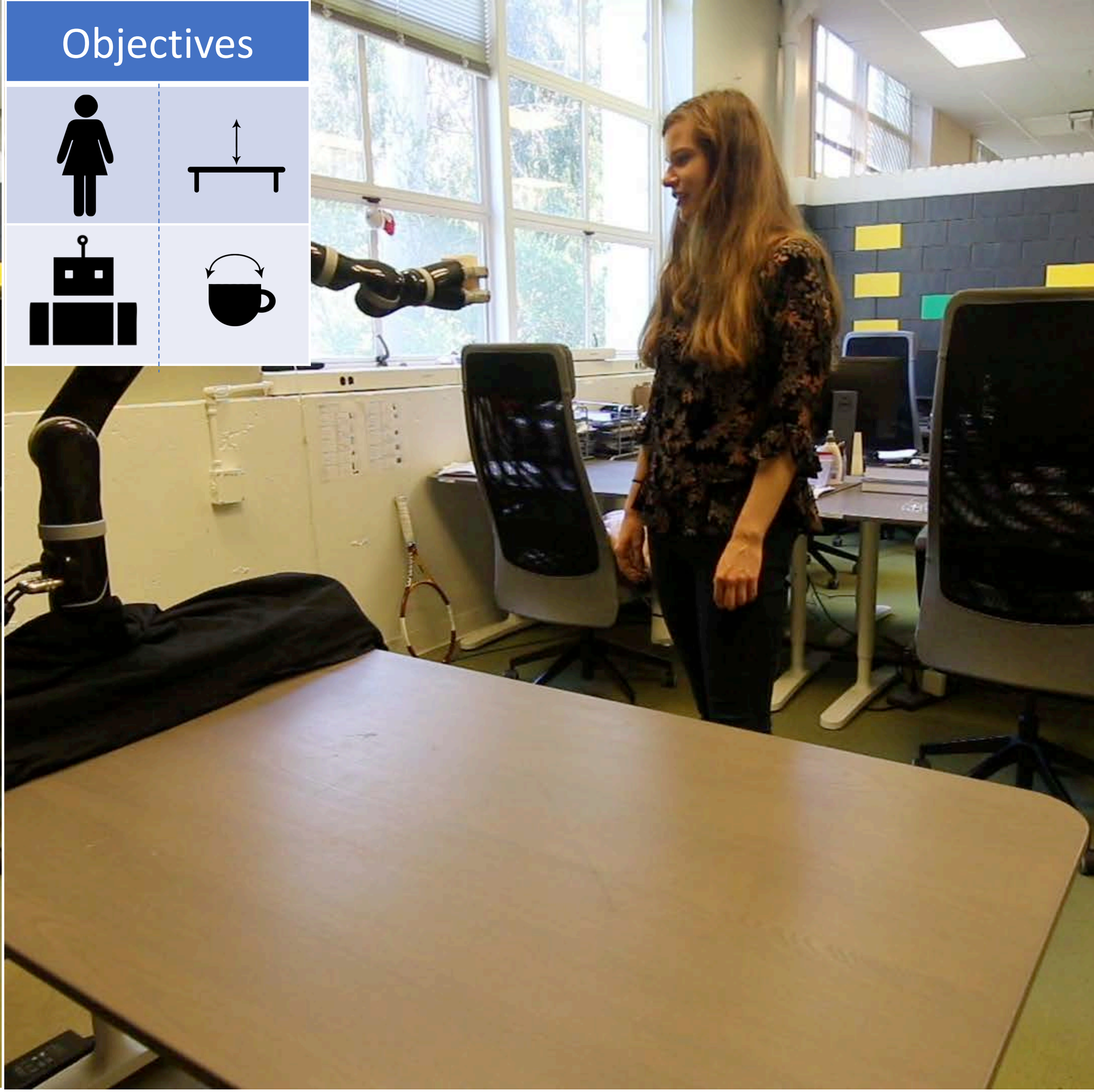
Key Insight
(Updated for efficiency)

When the human input appears to waste effort,
the robot has the wrong model of the human.

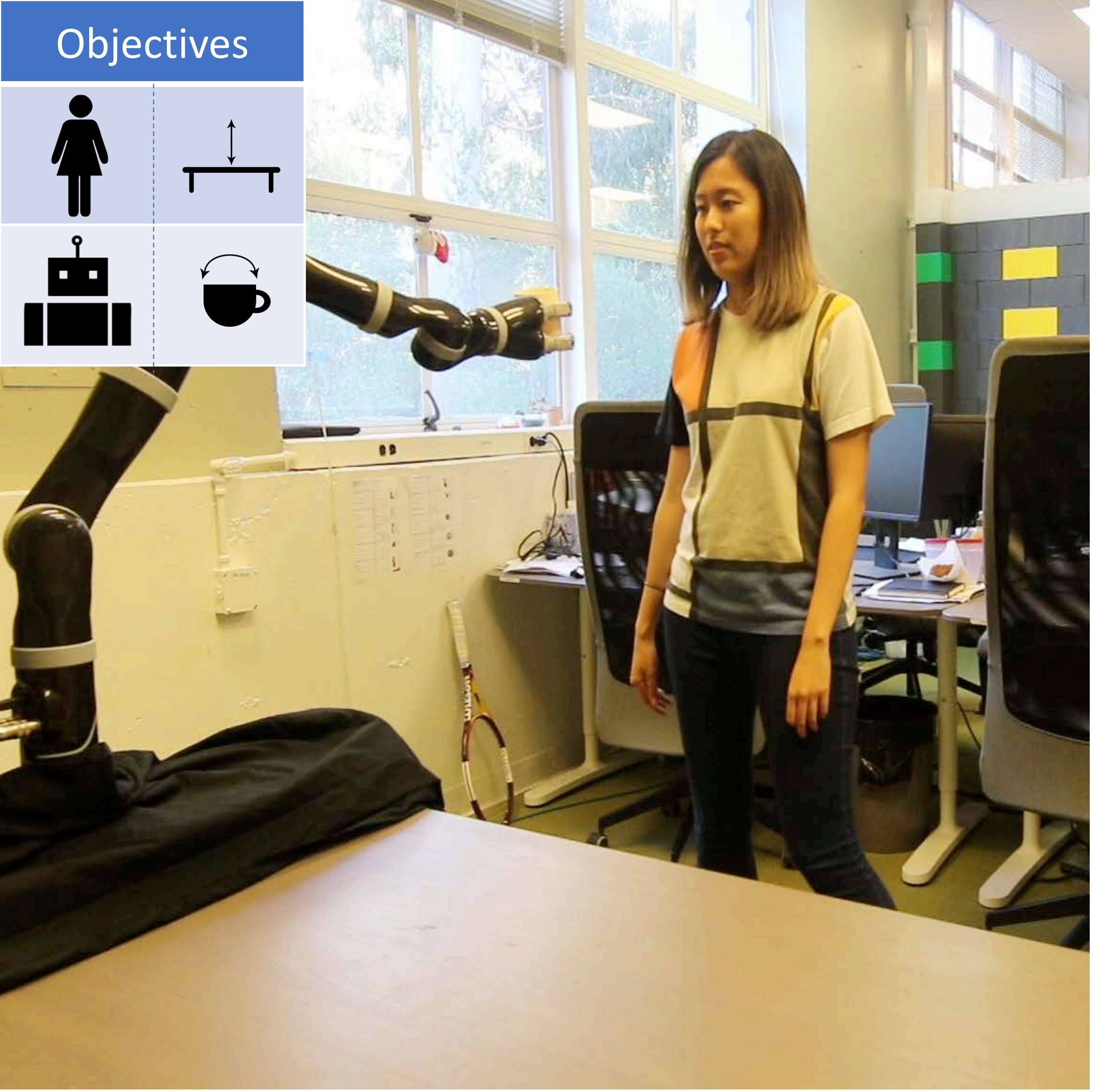




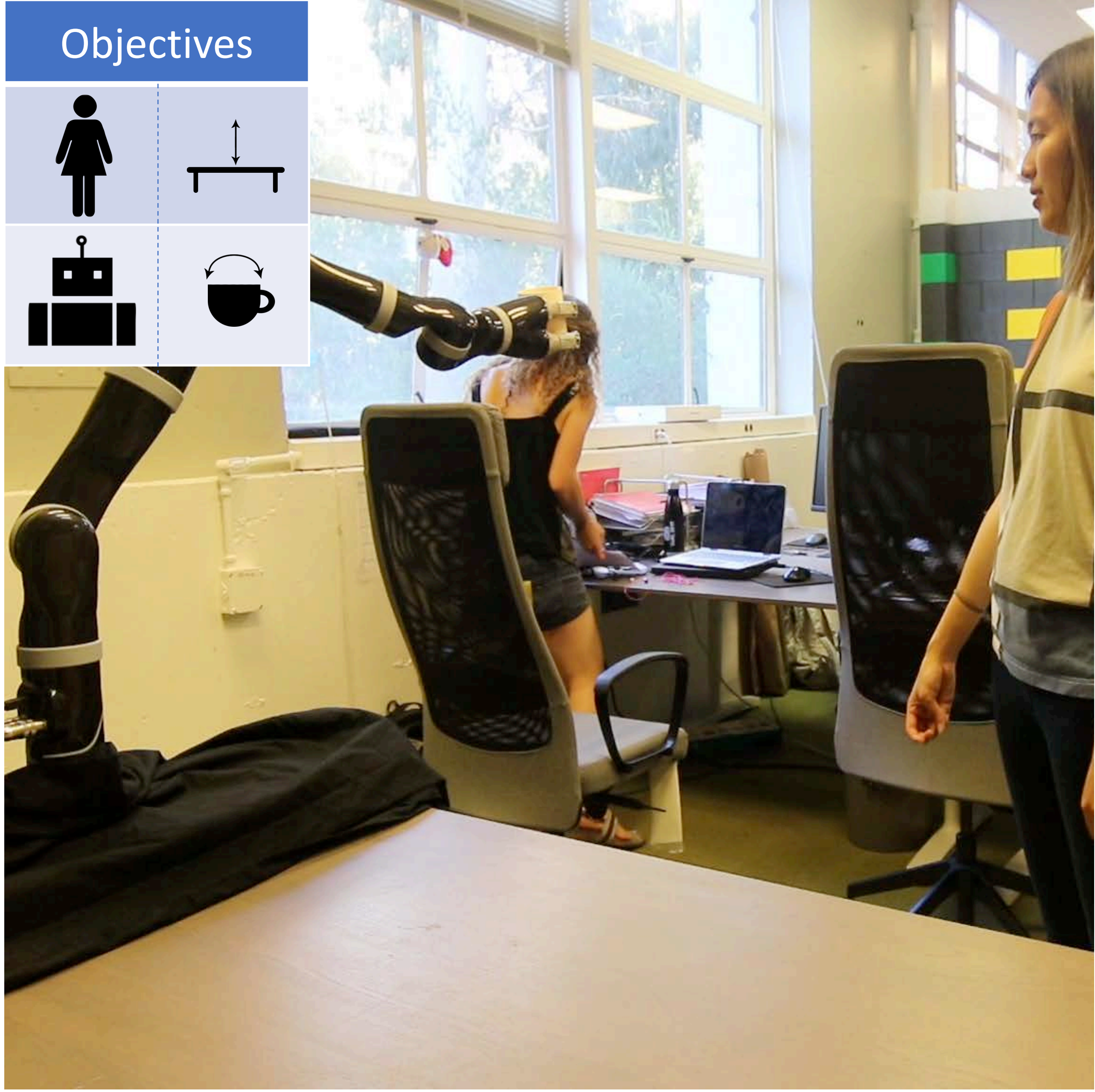
Naïve Learning



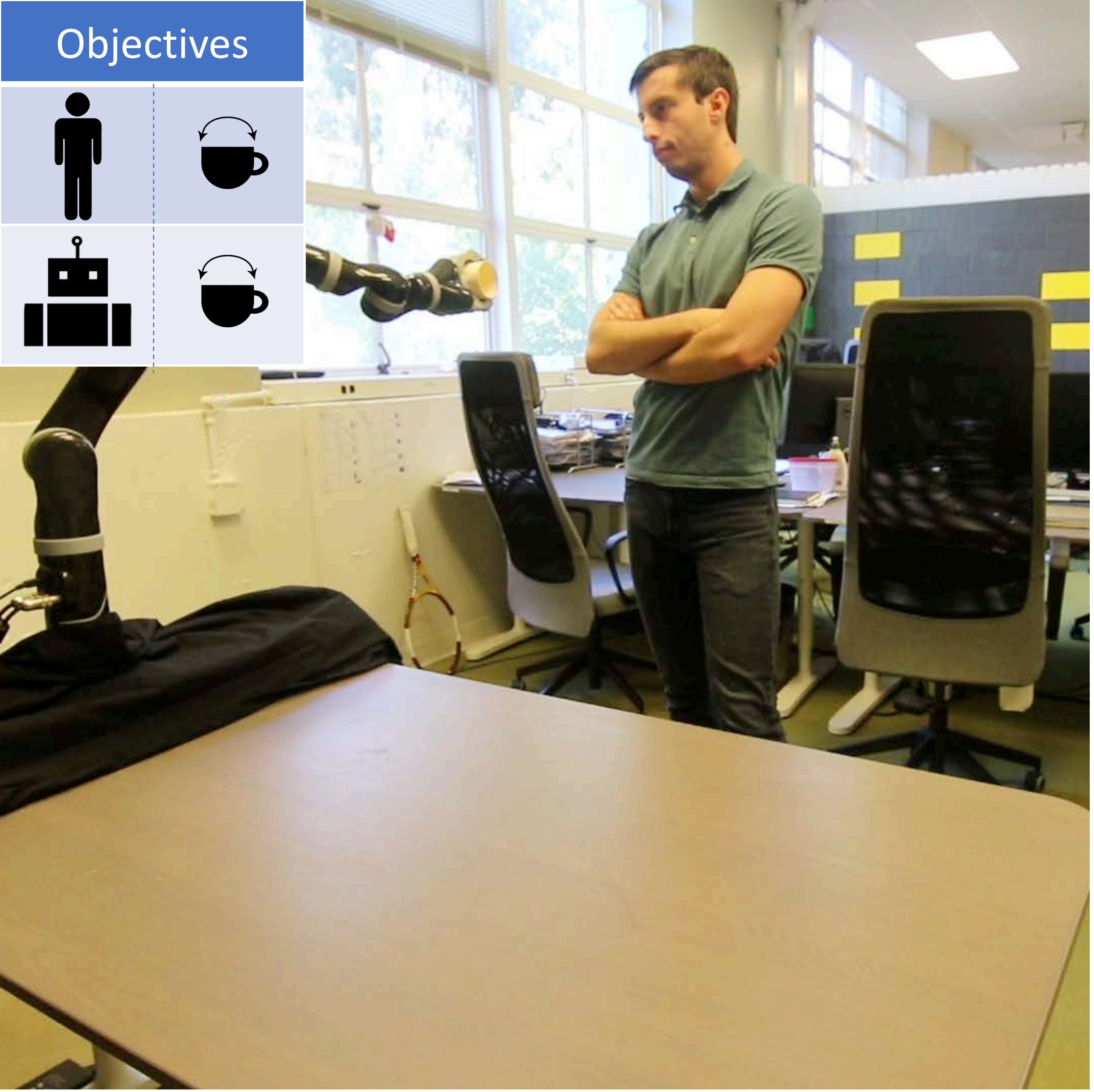
Relevance-Aware Learning



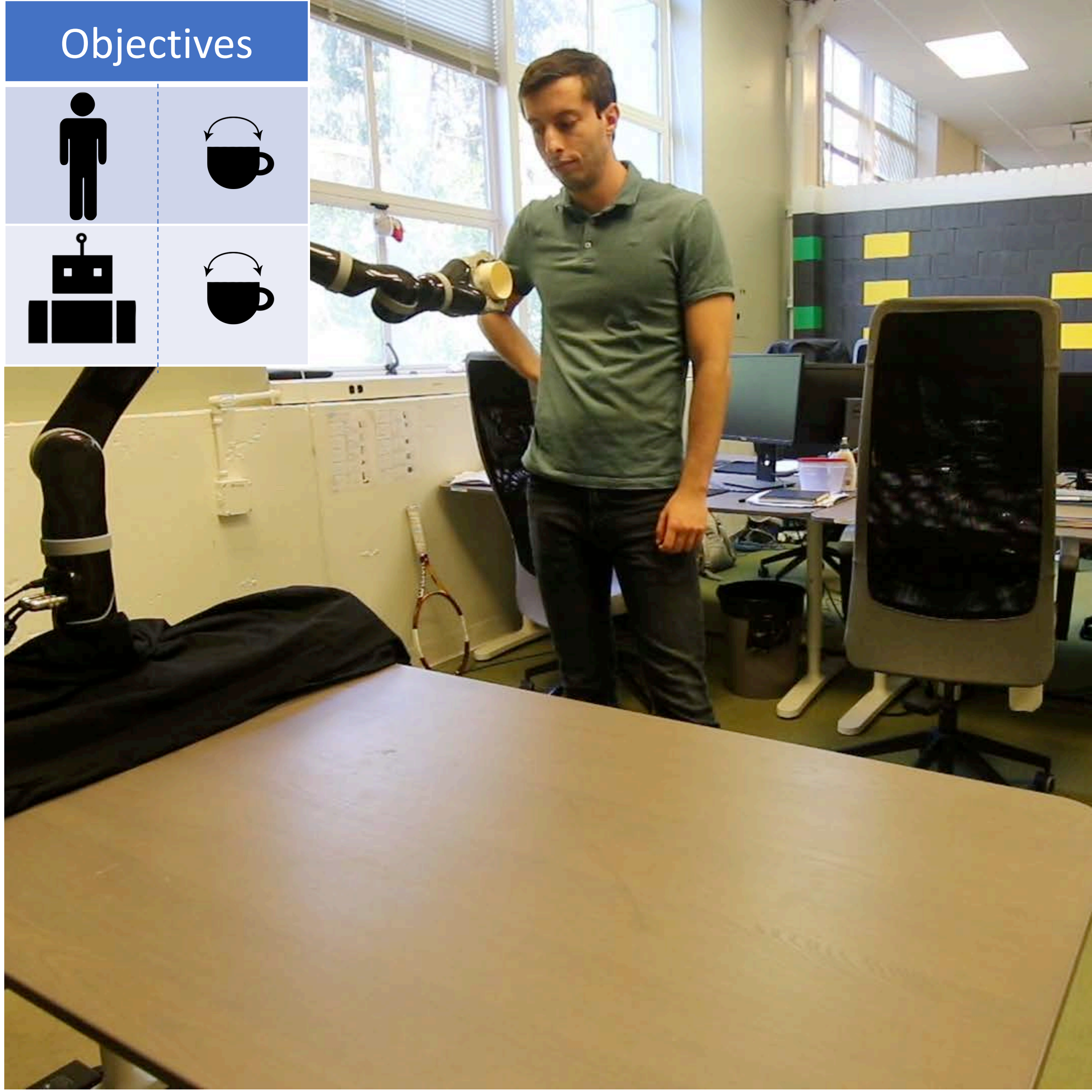
Naïve Learning



Relevance-Aware Learning



Naïve Learning



Relevance-Aware Learning

Implicit Communication in HMC (Dragan)

Research Objectives:

1. *Enable robots to infer implicit meaning in human actions.*
2. *Enable humans to infer implicit meaning in robot actions.*

Technical Approach:

1. *Run (active) Bayesian estimation over human internal state.*
2. *Augment physical state with human belief, and optimize KL-divergence of human belief from true robot state.*

Key Scientific Contributions:

- *observation models for human preferences, goals, etc.*
- *dynamics models for human belief*
- *tractable optimization with human state*
- *theoretical analysis of collaboration*

DoD Benefits:

- *better human-machine coordination*
- *machines that adapt to human preferences, capabilities, moods*
- *machines that are robust to under- or mis-specification of objectives*

Implicit Communication in Human-Machine Collaboration

FA9550-17-1-0308

Anca Dragan

BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

InterACT Laboratory