A Human-Centered Perspective on Interactive Robot Learning

Helen Beierling, Anna-Lisa Vollmer

UNIVERSITÄT BIELEFELD

Interactive Robotics in Medicine and Care

TRR 318 CONSTRUCTING EXPLAINABILITY

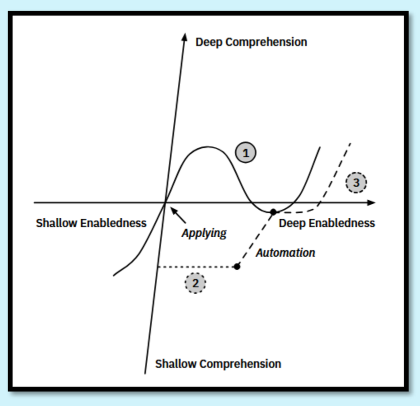
Recommendations

1. HRI & HIRL
   - Explicitly integrate the **Understanding Loop** into system development and evaluation
   - Treat learning as a **co-constructive process** between human and system
   - Consider human understanding as a **core component** of the learning process
2. Explainable AI (XAI)
   - Align explanations with the **intended form of user understanding**
   - Systematically identify **relevant explanatory concepts**
3. Didactic Perspective
   - Integrate **didactic methodologies** more strongly into XAI
4. AR for XAI
   - Visualization deepens **system understanding**
   - Potential challenge:
     - Increased understanding may shift attention away from the training process itself

Forms of understanding for XAI-Explanations

Buschmeier, Hendrik, et al. "Forms of understanding for XAI-Explanations." *Cognitive Systems Research* (2025): 101419

Dimensions:
- Comprehension / Enabledness
- Deep / Shallow



What you need to know about a learning robot: Identifying the enabling architecture of complex systems

Beierling, Helen, et al. "What you need to know about a learning robot: Identifying the enabling architecture of complex systems." *Cognitive Systems Research* 88 (2024): 101286.

Concept Extraction Process (Didactic Reconstruction)
1. **Collect expert explanations** of the system architecture.
2. **Collect lay user explanations** before, during, and after interaction.
3. **Reduce and structure concepts**
4. **Compare both maps** to identify known concepts, misunderstandings, and missing knowledge.

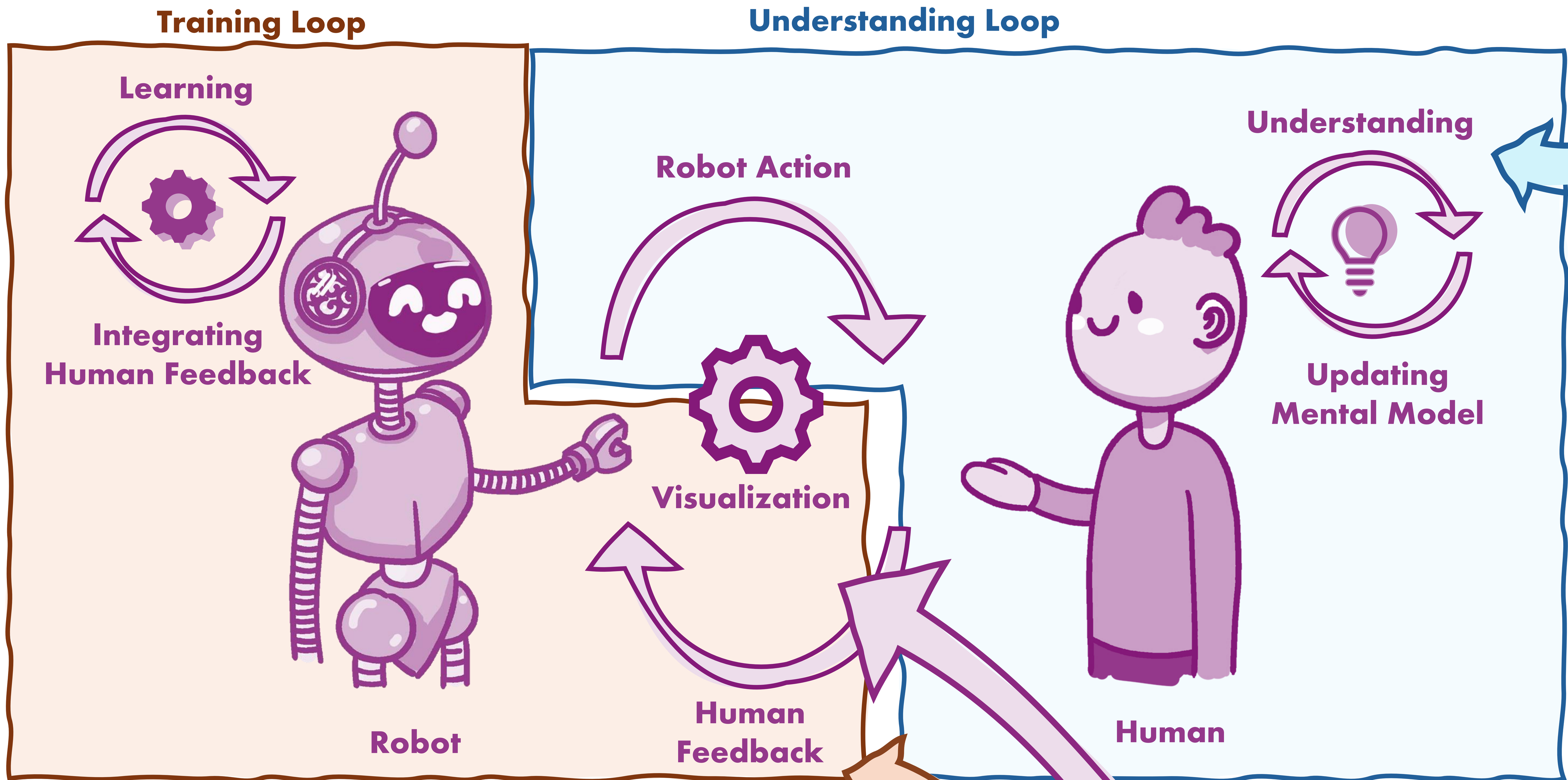
Method tested
- **Expert study:** 9 robotics/ML experts
- **Lay user study:** 10 lay users interacting with a
- **Task:** Teaching a robot mini-golf via preference feedback

Compare **expert explanations users' mental models**
→ Identify **missing concepts** and **misconceptions**

Preprocessing
- Remove function words

Quantitative Reduction
- Calculate concept frequencies
  - Per expert & Across all experts
- Remove concepts below threshold
- Syntax Merge
  - Merge different grammatical
- Semantic Merge
  - Merge concepts by meaning
  - Merge linguistic synonyms (keep the more abstract term)
  - Merge contextual synonyms (keep the more precise term)

Output
- Refined concepts Basis for concept maps

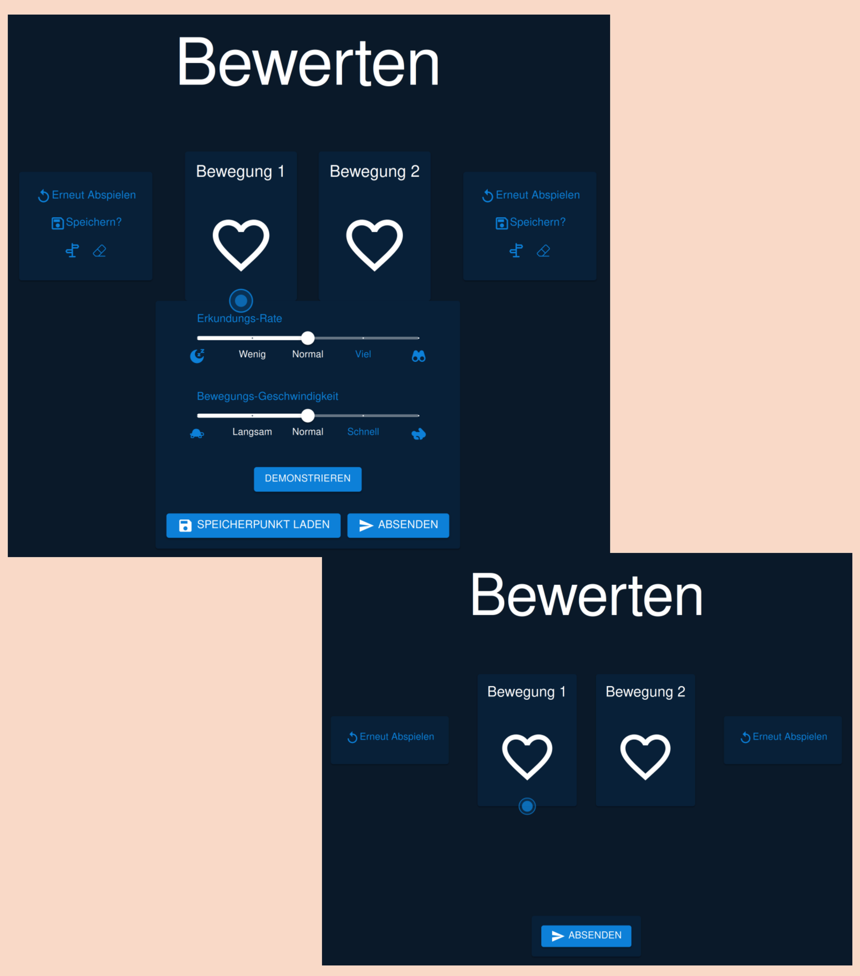


The power of combined modalities in interactive robot learning

Beierling, Helen, Robin Beierling, and Anna-Lisa Vollmer. "The power of combined modalities in interactive robot learning." *Frontiers in Robotics and AI* 12 (2025): 1598968.

Idea:
- **Lay users want to scaffold**
- **Multiple feedback modalities** enables **more flexible** improve robot learning

Method
- **Between-subject user study** (N = 33 lay participants)
- **Task**: Teach robot to minigolf 40 trials (80 trajectory evaluations) per participant
- **Learning framework:** Probabilistic Movement Primitives (ProMP) + Policy Improvement with Black Box (PIBB)
- **Group 1:** Baseline (Preference-only feedback)
- **Group 2:** Multimodal interface (Preference. Guidance, Correction, Demonstration, Exploration, Speed, Fallback)
- **Measures**:
  - Objective learning success (hits, first success)
  - Modality usage frequency
  - Perceived usefulness & usability (SUS, rankings)



Results
- The multimodal group (**N = 18**) significantly more task successes vs baseline group (**N = 15**) (Median hits: 26.5 vs. 5.0, p = .0013)
- The multimodal group also achieved their first successful hit earlier on average (17.6 vs. 34.6 trials)
- Users showed clear modality preferences, favoring Guidance and Speed, while Demonstration was rarely used

Key Insight
- Some modalities with **measurable learning impact** were **underestimated** by users:
  - Decreasing Exploration and increasing Speed correlated with higher success rates
  - In contrast, Guidance was highly preferred but showed no measurable correlation with task success

PRELIMINARY

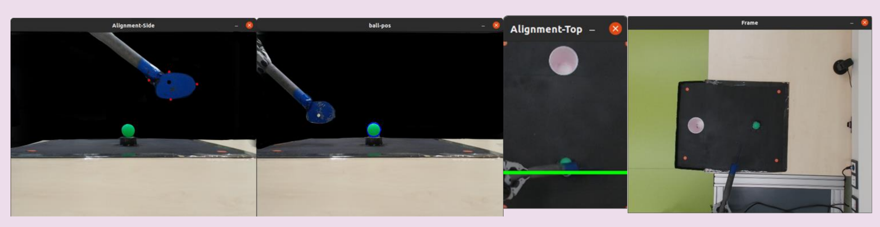
Understanding through seeing: the effect of AR-based transparency visualizations on interactive robot training

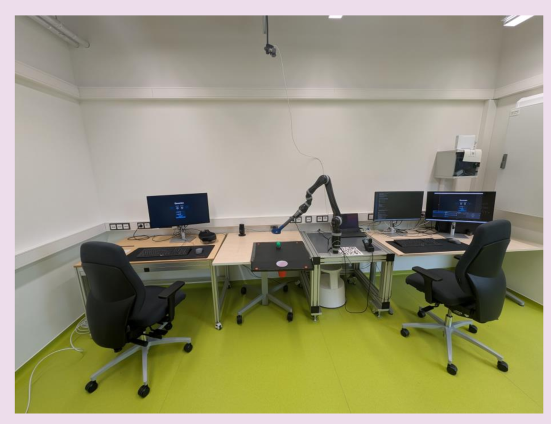
Helen Beierling, Robin Beierling, Anna-Lisa Vollmer

Idea:
- More Co-Construction leads to more understanding → better robot learning
- **Adaptive AR visualizations** explain hidden learning concepts identified in enabling architecture
- Visualizations responded **dynamically** to user feedback

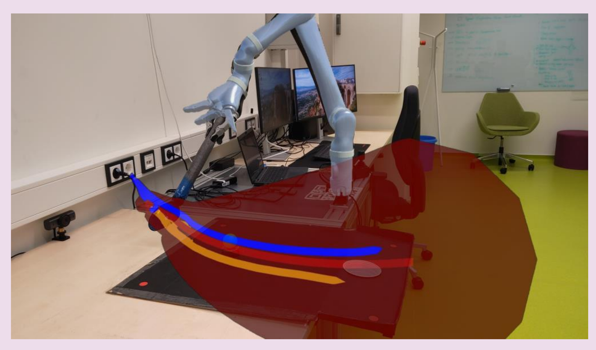
Method

User study (N = 48) comparing three levels of co-construction
- **AI**: Participants observe learning without input





- **Multi-Modal:** Users scaffold learning with multiple feedback modalities
- **Multi-Modal + AR:** Multimodal interaction plus AR transparency visualizations of hidden learning dynamics



Results

Understanding:
- Multi-Modal + AR showed the strongest gain in understanding

First Hit:
- Fastest learning tendency in Multi-Modal + AR (median 14.5 trials vs. 24 Multi-Modal vs. 39.5 Only AI)

Hit Rate:
- Multi-Modal showed the highest median performance (13 hits vs. 9 Multi-Modal + AR vs. 2 Only AI)

Key Insight

Higher co-construction increases user **understanding**, but greater transparency does **not** necessarily improve **training performance**