Human-Interactive Robot Learning: Definition, Challenges, and Recommendations

KIM BARAKA, Vrije Universiteit Amsterdam, The Netherlands
IFRAH IDREES, Brown University, USA
TAYLOR KESSLER FAULKNER, University of Washington, USA
ERDEM BIYIK, University of Southern California, USA
SERENA BOOTH, Brown University, USA
MOHAMED CHETOUANI, Institut des Systèmes Intelligents et de Robotique, Sorbonne University, CNRS, France
DANIEL H. GROLLMAN, Plus One Robotics, USA
AKANKSHA SARAN, Sony AI, USA
EMMANUEL SENFT, Idiap Research Institute, Switzerland
SILVIA TULLI, Sorbonne University, CNRS, France
ANNA-LISA VOLLMER, Bielefeld University, Germany
ANTONIO ANDRIELLA, Artificial Intelligence Research Institute (IIIA-CSIC), Spain
HELEN BEIERLING, Bielefeld University, Germany
TIFFANY HORTER, University of Oxford, UK
JENS KOBER, TU Delft, The Netherlands
ISAAC SHEIDLOWER, Tufts University, USA
MATTHEW E. TAYLOR, University of Alberta & Alberta Machine Intelligence Institute (Amii), Canada
SANNE VAN WAVEREN, Georgia Institute of Technology, USA
XUESU XIAO, George Mason University, USA

Robot learning from humans has been proposed and researched for several decades as a means to enable robots to learn new skills or adapt existing ones to new situations. Recent advances in artificial intelligence, including learning approaches like reinforcement learning and architectures like transformers and foundation models, combined with access to massive datasets, has created attractive opportunities to apply those data-hungry techniques to this problem. We argue that the

Authors' addresses: Kim Baraka, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands, k.baraka@vu.nl; Ifrah Idrees, Brown University, Providence, RI, USA, ifrah_idrees@brown.edu; Taylor Kessler Faulkner, University of Washington, Seattle, WA, USA, taylorkf@cs.washington.edu; Erdem Biyik, University of Southern California, Los Angeles, CA, USA, biyik@usc.edu; Serena Booth, Brown University, Providence, RI, USA, serena_booth@brown.edu; Mohamed Chetouani, Institut des Systèmes Intelligents et de Robotique, Sorbonne University, CNRS, Paris, France, mohamed.chetouani@sorbonne-universite.fr; Daniel H. Grollman, Plus One Robotics, Boulder, CO, USA, dan.grollman@plusonerobotics.com; Akanksha Saran, Sony AI, San Francisco, CA, USA, akanksha.saran@sony.com; Emmanuel Senft, Idiap Research Institute, Martigny, Switzerland, esenft@idiap.ch; Silvia Tulli, Sorbonne University, CNRS, Paris, France, silvia.tulli@sorbonne-universite.fr; Anna-Lisa Vollmer, Bielefeld University, Bielefeld, Germany, anna-lisa.vollmer@uni-bielefeld.de; Antonio Andriella, Artificial Intelligence Research Institute (IIIA-CSIC), Barcelona, Spain, antonio.andriella@iiia.csic.es; Helen Beierling, Bielefeld University, Bielefeld, Germany, helen.beierling@uni-bielefeld.de; Tiffany Horter, University of Oxford, Oxford, UK, th105@wellesley.edu; Jens Kober, TU Delft, Delft, The Netherlands, j.kober@tudelft.nl; Isaac Sheidlower, Tufts University, Medford, MA, USA, Isaac.Sheidlower@tufts.edu; Matthew E. Taylor, University of Alberta & Alberta Machine Intelligence Institute (Amii), Edmonton, AB, Canada, matthew.e.taylor@ualberta.ca; Sanne van Waveren, Georgia Institute of Technology, Atlanta, GA, USA, sanne@gatech.edu; Xuesu Xiao, George Mason University, Fairfax, VA, USA, xiao@gmu.edu.

ACM 2573-9522/2025/12-ART
https://doi.org/10.1145/3779297

focus on massive amounts of pre-collected data, and the resulting learning paradigm, where humans demonstrate and robots learn in isolation, is overshadowing a specialized area of work we term Human-Interactive-Robot-Learning (HIRL). This paradigm, wherein robots and humans interact *during the learning process*, is at the intersection of multiple fields (artificial intelligence, robotics, human-computer interaction, design and others) and holds unique promise. Using HIRL, robots can achieve greater sample efficiency (as humans can provide task knowledge through interaction), align with human preferences (as humans can guide the robot behavior towards their expectations), and explore more meaningfully and safely (as humans can utilize domain knowledge to guide learning and prevent catastrophic failures). This can result in robotic systems that can more quickly and easily adapt to new tasks in human environments. The objective of this paper is to provide a broad and consistent overview of HIRL research and to guide researchers toward understanding the scope of HIRL, and current open or underexplored challenges related to four themes — namely, human, robot learning, interaction, and broader context. The paper includes concrete use cases to illustrate the interaction between these challenges and inspire further research according to broad recommendations and a call for action for the growing HIRL community.

CCS Concepts: • **Human-centered computing → Collaborative interaction**; • **Computing methodologies → Learning settings**; *Intelligent agents*; *Cognitive robotics*; • **General and reference** → General literature.

Additional Key Words and Phrases: Robot learning, Interactive learning systems, Human-robot interaction, Interdisciplinary research, Human-in-the-loop machine learning, Interaction research, Teaching and learning

1 Introduction

The idea of robots learning from humans with domain or task expertise started with the early work of programming by demonstration [138, 162]. Since then, several new human-in-the-loop machine learning approaches have emerged, some of which have started to make their way into the realm of robotics. Our starting point is that pre-programming or pre-training robots will not be enough for fail proof deployment in unstructured and human-populated environments. Robots are likely to encounter or have to adapt to unseen situations, and will always need finetuning to fully comply to users' personal preferences, values, or needs. At the time of writing this paper, unlike personal devices like laptops and phones, physical robots are currently used for a very limited range of tasks and are often only accessible to a niche group of expert users. Although end-users can often customize robot behavior through simple interfaces, we do not yet have robots that are flexibly and naturally "teachable" by end-users as they would train pets, children, or junior colleagues. To date, the vast majority of robot programming methods have remained focused on building robots that specialize in accomplishing specific tasks, while fewer efforts have been dedicated to developing robots that can learn dynamically with human assistance, through (a combination of) teaching signals like demonstrations, evaluations, corrections, rankings, or instructions [16, 27, 29, 77, 91, 122, 153, 190]. These robots would interpret human teaching signals within their own model of the world, accounting for their capabilities. Developing robots that can interactively learn from a large variety of humans would enhance flexibility in numerous assistive and collaborative applications like household assistance and healthcare, making them more versatile and user-friendly like our everyday devices.

In line with this philosophy, this paper introduces a vision for the field of human-robot interaction where robots are treated less as *tools* with rigid, static, limited capabilities, and more as *apprentices* that can refine existing skills and acquire new skills through rich and intuitive interactions with humans. We define an emerging area of research that we call Human-Interactive Robot Learning (HIRL), which addresses the intersection between the robot learning sphere and the human factors sphere, placing interaction at the forefront. We believe that the intersection of these two spheres gives rise to unique technical challenges that each sphere alone does not fully capture. The goal of this paper is to identify and illustrate unique challenges and opportunities that arise in HIRL and outline its expected impact on both implementation and deployment phases of robotic technologies.

We view HIRL as a cross-disciplinary effort that spans several technical and non-technical research fields (see Figure 1). At the time of writing this paper, there is no agreed definition for a field that looks at both algorithmic and human factors and places interaction at the forefront when building embodied interactive learning systems.

This paper is the result of in-depth discussions during and after three workshops on HIRL [114, 115, 142] hosted at the ACM/IEEE Human-Robot-Interaction conference [3]. These discussions took the form of Q&A's over research presentations, focused working group discussions, design exercises, and online plenary meetings. This paper is not meant as a survey paper on the topic as we do not systematically nor extensively survey the literature (check [156] and [26] for comprehensive surveys on very close topics). Instead, the aims of the paper are to: (1) outline a vision for teachable robots where interaction plays a central role, (2) advocate for an interdisciplinary research agenda to consolidate and make progress in that research area, and (3) sound an alarm that this currently underexplored area of work (HIRL) is in danger of being pushed aside in the search for a universal, out-of-the-box general purpose robot based on foundation models and massive datasets. We present a definition of HIRL (Section 2), a list of open or under-researched challenges for this growing research area (Section 3), illustrated through hypothetical use cases (Section 4), and a set of recommendations for the HIRL community moving forward (Section 5).

2 Scope of HIRL

This section outlines a definition for HIRL, a brief overview of teaching signals and associated HIRL techniques, and a list of desired properties in HIRL systems.

2.1 HIRL definition

To further clarify the boundary of HIRL as a set of research problems and approaches, we list minimal assumptions for a HIRL (pronounced /h3:rl/) problem:

- ***A1.*** There is at least one robot interacting with at least one human
- ***A2.*** The robot learns through or as a result of this interaction, specifically the performance of the robot on a given task increases over time due to said interaction
- ***A3.*** The human acts/communicates in ways that influence the robot's behavior
- ***A4.*** The robot acts/communicates in ways that influence the human's input

As examples of a HIRL system, consider a kitchen robot that actively asks for demonstrations when it encounters limitations, such as using a new tool, or a robotic wheelchair that updates its navigational behavior based on real-time feedback from its user. Although a significant body of work models such learning problem as a Markov Decision Process (making it suitable for human-in-the-loop reinforcement learning for instance), we do not restrict the type of learning algorithms used during HIRL interactions, as long as these minimal assumptions are all present. For example, consider a robot that uses a sentiment classifier during interactions with people. The use of this classifier on its own does not constitute HIRL, even though the robot is using a learning algorithm and interacting with humans. However, if the sentiment classifier is trained during this interaction based on human communication given in response to robot actions, this interaction becomes a HIRL problem. Furthermore, this problem framing assumes that the human primarily plays the role of a teacher (whether intentional or not), and the robot primarily plays the role of a learner. While in some cases, these roles might be blurred (see Sections 4.3 and 5.3, and challenge I2), these primarily roles remain central to a HIRL problem. In the long run however, we see HIRL as a stepping stone towards collaborative and mutual teaching-learning, where teams of robots and humans teach and learn each other. Due to its breadth and complexity, the focus on mutual learning is left for future work.

To further give the reader a sense of the breadth of HIRL problems and approaches, the following subsection provides an overview of different types of teacher-learner frameworks and associated teaching signals.

2.2 Teacher-learner framework and overview of teaching signals

This section gives a brief overview of teacher-learner frameworks typically used in the literature, through the lens of different teaching signals that are commonly considered in HIRL systems. For a more comprehensive

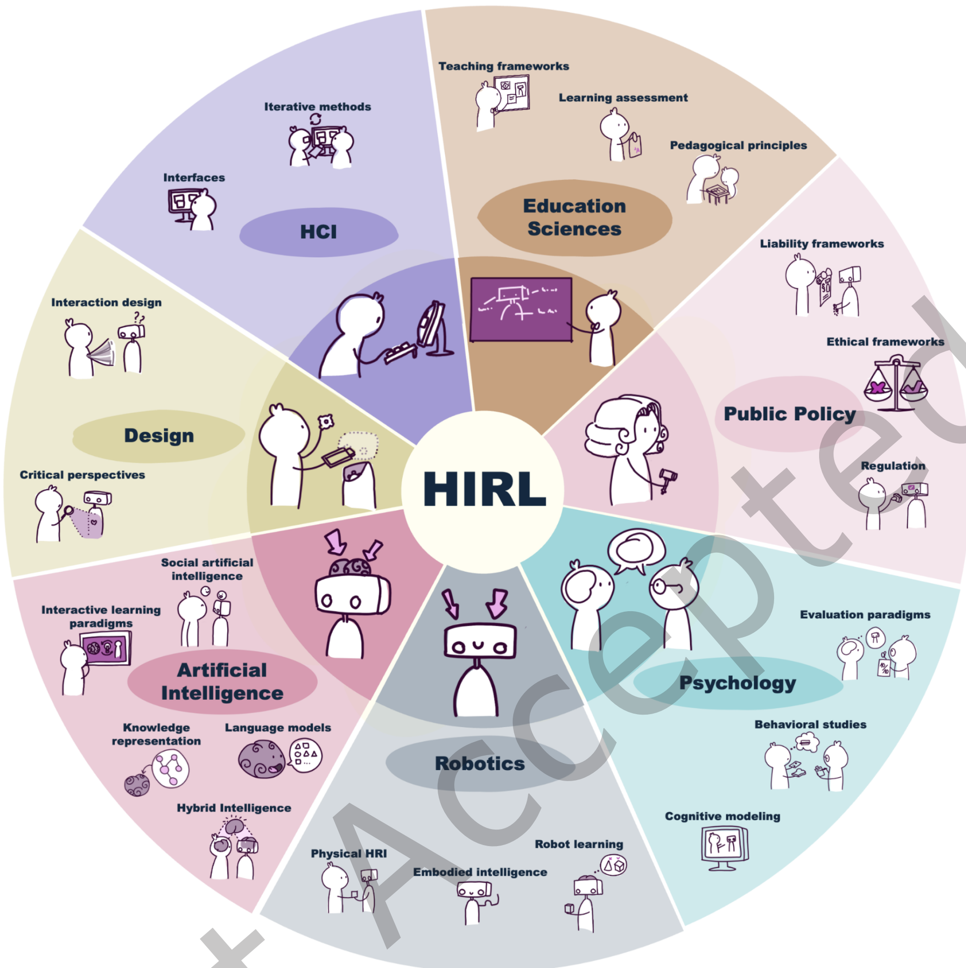

Fig. 1. This paper positions HIRL as an emergent cross-disciplinary research area that draws methods from several research fields, most of which already relevant to the broader field of HRI. Within each slice of the pie, HIRL-specific contributions from each field are listed for added concreteness.

overview, we refer to [26] and [156]. HIRL can be framed as an inter-agent knowledge transfer problem involving a teacher — i.e., a human knowledgeable of a solution to the task assigned to a robot — and a learner — in our case, a robot [92]. The teacher aims to transfer their knowledge about the task through teaching signals directed at the learner. We assume the teacher has some task-relevant expertise, but not necessarily robotics or machine learning expertise nor necessarily teaching expertise. The learner aims to make use of these teaching signals to improve its own learning.

2.2.1 Common types of teaching signals. Despite the considerable variability in terminology, four main types of teaching signals can be identified from the literature: demonstration, evaluation, correction, and ranking.

Depending on the learning paradigm, such signals can be given at different levels of abstraction (e.g. action level, episode-level, policy-level in a reinforcement learning formulation).

- **Demonstration** involves the teacher showing (or attempting to show [58]) the robot the desired behavior by performing it themselves. A demonstration is a set of state-action pairs sampled from the execution of the expert's policy. By providing a demonstration, which can include trajectories or execution traces, the teacher informs the learner about a possible way to accomplish a task by direct examples of which action to take at each state within the provided set. When a robot acquires skills through direct teleoperation or through kinesthetic demonstrations, this process is often termed *learning by doing*. On the other hand, when a robot learns from video demonstrations or the teacher's own body motions, the method is known as *learning from observation* [28, 65, 78, 95]. Examples of teleoperation interfaces encompass, but are not limited to, joysticks and control panels [110], as well as virtual reality (VR) [196] and haptic feedback devices [105]. A large portion of the work on learning from demonstrations is not interactive (i.e., demonstrations are provided before learning happens), but some recent work has been considering learning from demonstrations in online settings ([33, 68]).
- **Correction** involves the teacher providing feedback on specific errors or deviations from desired behavior and suggesting ways to improve or rectify those errors. Unlike demonstrations, corrections typically follow from observation of the learner's behavior. They can be delivered through various means, such as verbal instructions [35, 157, 190], kinesthetic interventions [103, 173], or teleoperation [80, 86, 104]. Similar to demonstrations, the goal of corrections is to convey an acceptable behavior through indicating which action(s) to take in given situations.
- **Evaluation** involves providing an assessment of the learner's performance based on predefined criteria [10, 96], through binary or scalar values. After observing the robot execute (a) behavior(s) in (a) certain circumstance(s), the human provides feedback about the quality of its past action(s). This feedback can serve as the sole form of learning signal for the robot or can be combined with self-exploration. It could be interpreted differently depending on the chosen approach — a reward-like signal in interactive reinforcement learning, a target in supervised learning [85], or a value roughly corresponding to how much better or worse an action is compared to the current policy [106]. Similar to classical reinforcement learning, evaluations aim to reinforce or punish certain behaviors of the robot. The robot makes sense of this by considering the teacher's signal as a reward or value associated with recent robot behavior [29].
- **Ranking** involves the teacher providing information about the quality of a trajectory in comparison to another/(others) by ranking them [122]. Ranking can be expressed as an ordered set of trajectories which, unlike correction, communicates the value of several alternatives relative to each other. This ranking provides the robot with information about the relative goodness of different trajectories and does not necessarily involve providing specific guidance on how to improve or correct the behavior.

Each of these types of teaching signals has pros and cons to consider. *Demonstrations* provide concrete examples and can be intuitive for humans to give. They are effective for complex tasks. On the downside, demonstrations have limited scenario coverage, can be bothersome for the teacher, and require them to be capable of performing the task. *Corrections* target specific areas for improvement and can be more efficient than full demonstrations for minor adjustments. They allow for iterative refinement of skills. However, corrections require the teacher to accurately identify and articulate errors, and may not provide a complete picture of the desired behavior. *Evaluations* are simple to provide, directly reinforce behaviors, and can be combined with self-exploration. However, they often do not provide specific enough guidance, can be subjective or biased, and may lead to inconsistent learning signals. For instance, a robot might receive conflicting feedback for similar actions from different humans or depending on the human's attention level, causing confusion in the learning process. Lastly, *rankings* allow for comparison between multiple candidate behaviors and can capture subtle preferences without requiring precise quantification. They are useful when optimal behavior is unclear, but relative performance can be assessed. However, rankings do not provide absolute measures of performance, do not work well in multi-objective tasks where rankings are difficult to produce, can be time-consuming if many trajectories need to be compared, and may be less informative when all options are similar. For example, consider an exoskeleton that must optimize the comfort of the user. If two gaits are both bad, it is difficult for the human to compare them and there is no ground-truth function for comfort.

In addition to these common categories of teaching signals, some works have considered other types of human-to-robot input that can be considered a teaching signal, such as starting state selection, which involves choosing the initial conditions for learning rollouts [30], or human saliency maps, in which the human annotates what is important in the visual scene manually [97] or with their gaze [12], curriculum learning where a human provides help by ordering tasks the robot tackles [182], state flagging, where an annotator identifies key states [192], and object-focused advice in the form context-specific instructions, such as "jump right (action) when encountering a coin (object)" [88].

2.2.2 Natural language as teaching signal. A growing corpus of research currently focuses on leveraging more complex natural language feedback as a means of instruction for robotic systems [136], especially with the advent of Large Language Models (LLM) [98]. This approach aims to leverage the flexibility and richness of human language as a means of knowledge transfer between humans and machines. Natural language feedback, being more expressive than traditional teaching signals, can cover more than one of the categories mentioned above and express higher-level or more complex feedback. It can also bridge the gap between observations and their underlying causes, thereby providing a robust foundation for generalization [113]. This characteristic of natural language feedback makes it particularly effective in supporting causal learning processes and enhancing inferential capabilities [94, 163].

Relatedly, instruction-following agents [7, 100] are designed to carry out tasks based on natural language instructions provided by humans. One of the key challenges for such agents is language grounding, which involves teaching agents to map human instructions to actions tied to their perceptions. To overcome this difficulty, several methods have been proposed, including the development of multimodal representations [100, 101, 190]. For example Ahn et al. [7] combine probabilities from a language model (indicating the likelihood that a given skill matches the instruction) with probabilities from a value function (indicating the likelihood of successfully executing that skill) to determine the most appropriate action. Since communication plays a key role in interaction, hence in HIRL, research on human-robot communication, including local vocabulary acquisition and co-emergence of symbols [46, 168], can be a key enabler of rich HIRL interactions.

2.2.3 Interaction paradigms. One of the design choices for HIRL systems has to do with who leads the interaction – the human, the robot, or both. This design choice has connections with decisions on robot autonomy [14, 59, 155], as well as collaboration patterns in interactive intelligent systems [135, 171]. We broadly identify three interaction paradigms:

- *Human-driven*, in which the human can intervene when deemed fit, upon which the robot can learn to optimize its own behavior accordingly. Examples of such a paradigm include the TAMER framework [85], which learns from online evaluative feedback, or the work of Losey et al. [103], where a robot learns from physical corrections.
- *Robot-driven*, in which the robot actively approaches the human when needed, actively querying the human for a teaching signal relevant to its own learning. Examples of such a paradigm include active reward learning from preferences or critiques [17, 37].
- *Hybrid*, in which both the human and the robot can initiate interaction in relation to teaching or querying, respectively.

Table 1. Definitions, pros, and cons of common Teaching signals in HIRL systems

Method	**Definition**	**Pros**	**Cons**
Demonstration	Showing the desired behavior or task execution	• Concrete examples • Often intuitive for humans • Effective for complex tasks	• Limited scenario coverage • Can be bothersome • Requires capable teacher
Correction	Specific adjustments to improve performance	• Targets specific improvements • Efficient for minor adjustments • Allows iterative refinement	• Requires accurate error identification • Incomplete behavior picture • Inconsistent across scenarios
Evaluation	Simple assessment of action quality (e.g., good/bad)	• Simple to provide • Directly reinforces behaviors • Easily combines with self-exploration	• Lacks specific guidance • Subjective or biased • Inconsistent learning signals
Ranking	Ordering multiple attempts based on relative performance	• Compares multiple approaches • Captures subtle preferences • Useful for unclear optimal behavior	• No absolute performance measure • Less informative for similar options • Time-consuming for many comparisons

The choice of paradigm is intimately tied to the specific HIRL setting and use case, and should be determined based on to what extent factors such as interruptability, cognitive load, and flexibility (on the human side) and efficiency, meta-learning capabilities, and safety (on the robot side) are deemed important.

2.3 Desired properties of a HIRL system

Desiderata for what HIRL systems achieve and how they function are highly context- and application-dependent. However, we believe there are some general properties of such systems that are desirable in most cases. The following (non-exhaustive) list outlines some that the authors identified as important based on their own research, both from a learning and from a teaching perspective, with associated explanations.

- Sample efficiency: For a given target performance, the robot requires fewer data/interactions or conversely, for a given number of data/interactions, the robot learns more efficiently. This is important because interactions with humans can be bothersome or expensive.
- Robustness: Small variations or biases in data/interactions do not cause large variations in learned behaviors. This is important because humans are noisy and unpredictable, so HIRL systems should be able to handle such variations.
- Coverage: The underlying interactive learning capabilities allow the robot to acquire a wide range of skills. This is important to create flexible systems that can acquire new skills beyond the ones it was pre-programmed to do.
- Solution quality: Given enough time and interactions, the robot is able to learn high quality behaviors. This is important because we would like to have guarantees that interactions actually improve the robot's performance as opposed to degrading it.
- Convergence (assuming stationarity in teaching and scenario): The robot is able to converge to an acceptable behavior within reasonable amount of time or teaching interactions. This is important as it brings predictability, which aids teaching.
- Adaptation (assuming non-stationarity in teaching): The robot is able to adapt to changes in the distribution of human input. This is important because humans are often co-learning with the robot and their teaching may reflect this fact.
- Low task load (mental/physical): It requires minimal effort for the human to participate in the learning process. This is important to make it viable for users to be willing to teach.

- Intuitiveness: The interventions require minimal training or are easily remembered. This is especially important for non-expert end-users who need to easily interact with robotic products without prior training.
- Low ambiguity: What is expected from the human is clear at all times. This is important for the quality of teaching as well as the motivation of the human.
- Interpretability: The human is able to understand the "internals" of the robot, e.g., regarding how it learns or why it acts a certain way. This is important for alignment purposes and for the quality of teaching.
- Personalizability: The robot is able to learn behaviors that satisfy the human's personal preferences or needs. This is important in cases where there is no single "objective" way of solving the task.
- Motivation-inducing: The system is designed in such a way that the human is motivated to provide input to the robot (e.g. because benefits outweigh costs). This is important to make it viable for users to be willing to participate in the learning process.

3 HIRL Challenges

This section outlines broad challenges that the authors identified as relevant to HIRL as an area of research. These were the result of round table discussions held in three HIRL workshops at the HRI conference [114, 115, 142], as well as extensive plenary and specialized discussions during the process of writing this paper. Challenges are organized along four different themes: *Human-related, Robot Learning-related, Interaction-related, and Broader Context-related*, and are summarized in Figure 2. These challenges are in no way exhaustive, but rather they were identified as open problems that are preventing the field from moving forward, either because they are challenges that researchers are not paying enough attention to, or because they will most likely not be solved any time soon. Each challenge, formulated as a broad research question, includes a brief explanation of relevance, scope, and possible ways to address it. Each theme contains one evaluation-related challenge marked with E. It is worth noting in potential solutions outlined that the same method can address more than one challenge.

3.1 Human-related challenges

Challenges in this section relate to aspects of the human themselves, including their behavior, experience, and role.

H1. How do humans teach robots? Understanding how humans teach robots is crucial for developing advanced learning paradigms and evaluation methods. Research in this area comprises two distinct strands. The first focuses on interaction studies, exploring the dynamics of natural teaching in both human-human and human-robot interactions [178]. These studies contrast the dynamic, adaptive nature of natural human teaching with rigid teaching in current HIRL approaches [179] and have revealed that human teachers naturally tend to employ strategies similar to those used when teaching children [124, 176]. For instance, they adjust their input based on the learner's current level of understanding, using techniques such as monitoring and scaffolding [137, 177].

In the second strand, there is ongoing research aimed at enabling robots to better understand and respond to human teaching strategies. This involves capturing and accurately interpreting teaching signals, modeling the teaching process as feedback, demonstration, or instruction, and understanding the behavior and intention of the teacher [9, 73, 101]. By focusing on these areas, we can create teacher-adaptive learning algorithms and realistic evaluation oracles [29, 91]. Promising approaches include empirical studies, predictive modeling, and agent-based methods such as reinforcement learning, which collectively contribute to refining how robots learn from human interaction [48].

H2. How to facilitate and enrich teaching experience? Facilitating and enriching the teaching experience during human-robot interaction is essential for minimizing teacher fatigue and frustration, while also promoting

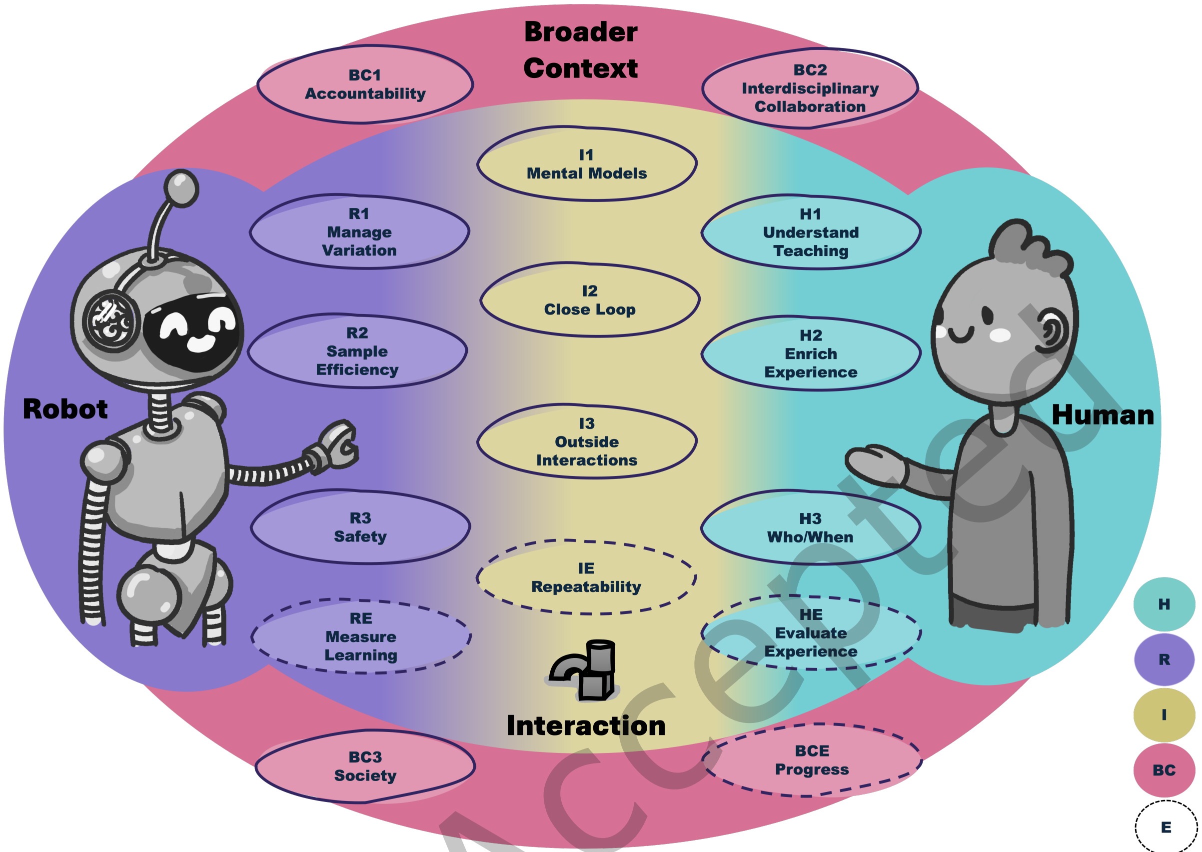

Fig. 2. HIRL challenges across four themes: Human-related (H), Robot Learning-related (R), Interaction-related (I), Broader Context-related (BC). Evaluation-specific challenges (E) within each theme are marked with dashed outlines. The graphic is intended to help visually categorize the challenges that will be described in the following sections.

the delivery of effective and accurate teaching signals. Enhancing this experience requires a multifaceted approach. First, the teaching process can be enriched, for instance through combining multiple teaching signals (both explicit like natural language and implicit like gaze of facial expressions) [36, 147, 149, 195]. Second, humans can also be guided on the most effective ways to teach robots [67, 68]. Key strategies involve designing intuitive interfaces [174], which may include innovative hardware solutions to streamline the teaching process and reduce the cognitive load on the human teacher [72, 183]. Additionally, a strong focus on human-centered interaction design is crucial [126, 128], ensuring that the system is tailored to the needs and capabilities of the user. Transparency in the teaching process is another critical component, as it helps users understand the robot's learning progress and methods, thereby fostering a more collaborative and effective teaching environment [176]. These approaches could collectively contribute to a more efficient, user-friendly, and satisfying human-robot teaching experience.

H3. When should which humans teach robots? Realistic HIRL systems deployed in human-populated environments are likely to have to deal with multiple teachers with potentially conflicting teaching signals. Learning from multiple humans presents several challenges, especially in determining the right timing and choice of human teachers. Contextual factors, such as the specific environment and task requirements, have a significant influence on the teaching approach. The complexity is further compounded by human limitations, especially in situations of non-stationarity, where instructors may need to adapt their methods or responses as the robot's behavior evolves. Selecting the appropriate teacher (whether by design or by the robot) is crucial and challenging, requiring someone who is currently available, with the right expertise, and who can consistently adapt to these changes. Balancing these factors is essential to ensure the robot receives accurate and effective instruction that aligns with intended behavior, especially in situations where multiple stakeholders (e.g., service provider, service consumer) are involved and can be queried to adjust different parts of the robot's behavior. Algorithmically, existing efforts in cooperative multi-agent reinforcement learning are promising to automatically reason about teacher-learner roles in a multi-teacher (potentially multi-robot) setting [132]. From a design perspective, this challenge also applies to the pre-deployment phase when a development team needs to select the right type of teacher when interactively training robots to produce certain behaviors with the help of humans, as explored in [185].

HE. How to evaluate teaching experience? Evaluating the teaching experience requires a comprehensive approach that includes various metrics and methods, considering the diverse experiences among participants interacting with different learning systems. Key metrics might involve assessing the effectiveness and efficiency of the teaching process, user satisfaction, and the quality of the learning outcomes. Methods for evaluation should encompass both quantitative and qualitative analyses, focusing on the human-centered aspects of the teaching experience and simplifying the process for better accessibility. A crucial factor in improving teaching experiences is minimizing the number of interactions needed to achieve desired learning outcomes [125]. Additionally, measuring multi-modal freedom—how easily and effectively teachers can switch between different modes of instruction—provides insights into the flexibility and adaptability of the teaching methods. This holistic evaluation framework can help identify strengths and areas for improvement, ultimately enhancing the overall teaching and learning experience. Finally, as there could be strong learning effects also on the human side, there is a need to look at long-term evaluations to bridge the gap between short-term studies and real-world usability of HIRL systems.

3.2 Robot Learning-related challenges

These challenges apply to the learning process or capabilities of the learning robot.

R1. How can robots manage variation in teaching? Dealing with variation in teaching is a critical roadblock to deploying learning robots in the wild. Variation can arise from one user changing over time [70, 89], multiple users teaching in varying ways [170], inconsistent or contradictory teaching signals, adversarial behavior, and many other causes. Regardless of the reason, these changes can result in a robot not learning tasks effectively or safely. Specific open problems within this challenge are differentiating between poor teaching and an imperfect model of what the user wants, identifying what kinds of ground truth we might have access to (an expert reward function, a defined goal, etc.), establishing reasonable assumptions about imperfect teaching that can help robots learn even with imperfect information, and personalizing robot behavior to different user preferences. Efforts to address such problems can include exploring differences between users [107] and subsequently developing personalized and teacher-adaptable learning strategies, in addition to accommodating differences in interface preferences [42]. Recent work further presented a mechanism accounting for various teacher strategies in a shared control context [11].

R2. How to improve sample efficiency? Sample efficiency in HIRL is crucial as collecting human-robot interaction data is expensive and can bring safety challenges, as well as being tiring for users. One way of reducing the amount of human data required is to use data that is available "for free" such as implicit signals [83, 97, 99, 147–

150], or more information-dense teaching signals (e.g. corrections instead of binary good/bad feedback) [37, 103]. Active learning and active class selection (in the context of human-in-the-loop reinforcement learning, incremental or continual learning) [50, 68, 180, 184, 189] and simulated teachers (improved beyond noise-modified perfect oracles [70, 90, 94]) are other avenues. All of these methods have some drawbacks to overcome: implicit signals may require additional tools such as devices to record audio or track eyes, information-dense teaching signals may be more difficult for non-expert users to provide, active learning may need interactions to be optimized under an acquisition function and thus require additional computation time, and simulated teachers have quite a bit of improvement to go before they can accurately capture human behavior. With more research, each of these research directions holds promising methods for future sample-efficient learning.

R3. How to prevent unsafe behaviors? Although introducing a human in the loop of learning systems can address safety issues by guiding the robot's exploration more effectively and preventing unsafe behaviors [151], safety (for the robot and/or human(s)) is still a major challenge to address in HIRL systems. The two main factors that can lead to unsafe behaviors are, first, learning algorithms without safety boundaries and, second, people giving dangerous teaching signals to a robotic learner. As many fields outside of HIRL focus on safety in non-interactive learning algorithms [188], we propose work on making algorithms better able to handle dangerous teaching signals. These kinds of signals can come from either mistakes or from intentional bad actors. To handle bad actors, some form of safety "ground truth" is needed; if a user instructs a robot to perform a dangerous behavior, algorithms should be designed with checks to ensure that actions leading to this behavior are not executed. In the case of mistakes from well-intentioned users, a safety ground truth can also help in the present moment; additionally, robots could in some cases be provided with a method of informing the user that such teaching signals could lead to unsafe behavior, helping users learn over time how to better and more safely teach robots. Even though HIRL ultimately aims to have real robots learn from interactions, realistic simulators can greatly help ensure safer learning as an intermediate step; however, they require accurate environment models.

RE. How to measure what you have learned and what you can learn? The evaluation of HIRL systems tends to vary significantly. Often these systems are evaluated through some combination of simulation and real-robot experiments; with "oracles" (pre-trained behavior) that give optimal feedback to a learning agent, or noisy simulated users; with experts or novice participants; and in laboratory or in online crowd-sourcing settings. This variation is further compounded by the lack of standard evaluative metrics and benchmarks [26]. Here, we propose directions of work to further standardize the quantitative and qualitative evaluation of HIRL-systems. A key first step to evaluating HIRL-systems is a set of benchmark and standard quantitative learning/teaching metrics which need to be applicable to most, if not all, HIRL settings. Furthermore, we need ways to estimate not only the current quality or performance of the system, but also the envelope of learnable behaviors; for example, by taking inspiration from the proofs commonly used in the reinforcement learning community [164], or through empirically constructing these envelopes using realistic simulated teachers.

3.3 Interaction-related challenges

Challenges within this theme cover aspects that arise from interactions between the human, the robot, and the environment. We assume that the human and robot form an interaction loop where signals relevant to both teaching and learning are dynamically exchanged.

I1. How to construct compatible mental models? A significant challenge in HIRL is aligning human and robot mental models of each other's capabilities and intentions [22, 130]. Misunderstandings can lead to ineffective teaching inputs from humans or breakdowns in interaction. Successful co-alignment requires both sides to adapt: humans often form anthropomorphic models influenced by media, prior experiences [13], the robot's appearance and behavior [139], as well as ingrained perceptions of human-like traits [49]. This necessitates improving transparency, explainability, and employing intuitive control architectures to help humans form accurate representations of a robot's abilities and learning processes [51, 159, 198]. For robots, effective modeling of human preferences and intentions can be achieved through techniques like preference learning [129], intention inference [116, 181], and shared representations [18], allowing them to better align with human needs and goals. Addressing this challenge calls for advances in human-centered design, adaptive learning, and second-order mental models [23, 166], where robots also consider the human's understanding of their abilities, thus enhancing feedback and trust calibration.

I2. How to close the teacher-learner loop? A critical challenge in HIRL is closing the interaction loop between the human teacher and the robot learner. Current algorithms often treat teachers as static, reliable sources of information [177], yet in practice, human teachers are variable, with their evolving teaching strategies, and may become tired or frustrated. Effective HIRL systems should enable two-way feedback, fostering co-construction, co-learning, and co-adaptation throughout the interaction [144, 175]. One potential solution is to leverage implicit cues from interaction data to help robots learn more efficiently, while also guiding teachers to provide more relevant and useful feedback [39, 107, 147, 149, 186, 187]. Active learning strategies [38, 123] can help the robot identify knowledge gaps and direct the teacher's focus effectively in those areas, but still need to adequately balance both robot learning and human factors (e.g., cognitive load, iinterruptibility, context, etc). Finally, closing the teacher-loop by adding effective robot-to-human feedback mechanisms [60, 63, 120] can create responsive, co-adaptive interactions, leading to improved learning outcomes for both parties over time and creating a form of synergy between teacher and learner [8].

I3. How to deal with interactions outside the teacher-learner loop? While being embedded in a teacher-learner loop, the robot also interacts with the outside world, such as the task at hand or even potential humans not involved in the teaching process. For example, knowledge gathered by the agent from interacting with the physical environment may or may not be correlated with feedback from a teacher. Additionally, an environment may carry information about the humans that populate it (see, e.g., [102]), which may help bootstrap or contextualize interaction. As such, the robot receives signals from the environment and from a variety of social agents. As HIRL studies often take place in labs (see examples in [26]), this challenge is not widely explored but will be significant when deploying learning robots in the real world. Some existing work in this direction includes algorithms that learn from more than one reward-like signal [56, 158]. In the future, one potential way of approaching this challenge at a higher level is to discriminate teaching relevant signals from other environmental signals and have two distinct strategies, one task-related and a second reactive one for interacting with other parts of the environment (for example handling basic conversation with other humans).

IE. How to conduct repeatable interaction studies? Another key challenge is ensuring the repeatability and replicability of interaction studies, which are essential for validating scientific findings. Interaction studies, which range from learning policy convergence to usability evaluations, are notoriously difficult to replicate due to variations in human behavior and experimental conditions [79]. Replication, however, is crucial for building robust and generalizable knowledge. To address this challenge, it is essential to develop standardized protocols and benchmarks for study design [cf. 29, 79]. Leveraging simulation environments can help create controlled scenarios, allowing for repeated studies with a large number of participants [75, 169, 197]. It is worth noting here that unlike traditional machine learning that relies on large datasets that can be directly used to train models, the interactive nature of HIRL makes such datasets of limited usefulness for training models. However, we argue that open-source interaction datasets can facilitate replicability by allowing researchers to access to a rich diversity of interaction "traces" in HIRL settings and explore questions related to the effect of the teacher on the learner and vice versa. Developing standards for storing, sharing, and using these datasets will help ensure that interaction studies are repeatable and that results can be verified and built upon. By establishing these practices, the field can advance more rapidly and consistently.

3.4 Broader Context-related

Broader context challenges encompass aspects that go beyond the components of a HIRL system, including impact on and influence of broader ecosystems in which these systems are deployed or developed, and the practice of research in HIRL as an emerging field.

BC1. How to consider accountability in HIRL systems? Accountability has long been a murky concept in traditional software design [61], let alone in complex interactions. Should compiler developers be accountable for malicious code later compiled? It seems obvious that they should not. But if the compiler contains errors that cause compiled code to malfunction, should developers then be held accountable? These questions grow more complex in HIRL systems. In traditional software development, roles like "developer," "tester," or "user" are distinct, but in HIRL, these roles are entangled [34, 154]. Both engineers and human teachers can encode harmful or erroneous code. Theoretically, both developers and end users can encode "correct" code, but a mismatch between the developer's embedded inductive biases and the human's training strategy (e.g., as seen in common RLHF [84]) could still lead to harm. Lessons from content moderation are relevant here. In HIRL, the engineer's system design can be viewed as a "platform", and all subsequent interactive learning as content. Certain harmful content, like teaching a robot to use a weapon, could be identified and prevented, while other cases may require subjective interpretation and human judgment. Accountability becomes even more challenging in a cloud setting that allows re-use of previously taught skills by a community of users across robotic platforms.

BC2. How to effectively collaborate across HIRL-relevant fields? Developing HIRL systems necessitates an interdisciplinary approach that integrates engineering, computer science, cognitive science, and, more recently, the social sciences and humanities. The challenge lies in fostering effective collaboration among these diverse disciplines to integrate centuries of research on interaction, learning, and didactics. Notably, there is scant research focusing on the (informal) teaching aspect, which is crucial for HIRL. Mixed-method study designs can leverage the qualitative methods of the social sciences and humanities to complement the quantitative methods relied upon by engineering and computer science [194]. Effective interdisciplinary work depends on robust methods for collaboration, including the transfer of results, theories, and methods among fields, with an awareness and alignment of different epistemic cultures and values. Addressing these challenges requires clear communication strategies, regular exchanges, and fostering a shared vision aligned with the overarching goals of HIRL. Without engaging these varied fields, HIRL research risks overlooking user needs and societal expectations, potentially leading to ineffective and societally irrelevant solutions [112]. Among the HIRL-relevant fields mentioned in Figure 1, the authors would like to specifically highlight the potential of collaborations between AI and human-robot interaction researchers and education sciences, including human-animal training [133, 152].

BC3. How to design HIRL systems with and for society? Deploying HIRL systems can have both positive and negative impacts on society. Besides classic impacts of robotics (e.g., cost-reduction, risk of reducing human contacts, or access to new functions for some users) [43], the presence of a teaching interaction with HIRL systems creates new opportunities and challenges. Such robots can learn values adapted to the culture in which they are deployed [108]. However, these learning systems can also have spillover effects, for example by encouraging antisocial behaviors as it was observed with chatbots [41]. Consequently, the HIRL community should reflect upon where HIRL systems should be deployed, and whether some use cases or teaching practices are off-limit. We believe in building more extensively on participatory design methods [121] by involving both end-users and experts with HIRL-relevant knowledge (which may include embodied knowledge, e.g., educators, dog trainers, performing artists, domain experts, etc.) through novel participatory methods leveraging HIRL [185].

Furthermore, strengthening collaborations with research in the social sciences and humanities can help guide design requirements and anticipate blind spots in adoption and assumptions, ensuring that HIRL systems are co-developed with and for society.

BCE. What is progress in HIRL as a field? Measuring progress in HIRL is critical to creating more adaptable, safe, and user-friendly robotic systems that work seamlessly in human environments. The effectiveness of standard metrics for HIRL should be measured broadly in terms of its impact on society. This includes ensuring that research knowledge is scalable and applicable to real-world challenges and that it improves productivity and safety. Benchmarks such as those from the IEEE Robotics and Automation Society [4], competitions like RoboCup@Home [5] and those organised by the International Conference on Social Robotics [1], and datasets like Orbit [119] are essential for advancement in this field. Additionally, it is essential to measure the success of educational programs for stakeholders and develop a skilled workforce to advance HIRL technologies for sustainable growth and innovation (cf. HIRL educational module [2]). Furthermore, new R&D projects and initiatives such as international workshops in well-established venues [142] will move the field forward by attracting investment and resources, especially given that HIRL aligns with the new European directive that emphasizes human-centric approaches to AI [47].

4 Use cases

This section presents five use cases meant to concretely illustrate the challenges outlined in Section 3 through examples of hypothetical HIRL systems. These use cases were specifically chosen to highlight a range of different challenges, although they are in no way exhaustive, neither of the challenges nor of potential application areas of HIRL systems. These use cases are visually summarized in Figure 3.

4.1 Robot-assisted physical therapy for rehabilitation in elderly patients

4.1.1 Context description. Healthcare in general, and elder care in particular, has long been used to justify robot development and deployment due to the current demographic shift and scarcity of workforce [191]. Physical therapy following fractures or strokes is crucial for optimal recovery, particularly in older adults, where the process can be more complex due to age-related factors [44, 143]. In recent years, robotic systems have been developed and deployed to aid in this recovery process [e.g., 40, 69]. For effective rehabilitation, it is essential that these robotic systems are adaptive, responding not only to the patient's individual needs but also to their progress [54, 131]. However, human-robot interaction in this demographic presents unique challenges. Older adults often are not familiar with advanced technology, which can create barriers to effective use. Additionally, many face a range of age-related impairments, such as reduced hearing, cognitive decline, and diminished physical abilities.

4.1.2 Added value of HIRL. Given the highly individual nature of patients in this use case, it is essential that treatment is equally individual. Developing systems that can effectively cater to these individual needs is challenging, as pre-programmed solutions are insufficient due to the inadequacy of a one-size-fits-all therapy approach. In this context, HIRL presents a viable alternative, enabling robots to be taught by therapists in a tailored way and adapted by patients during execution through natural interaction. This would ensure that the care provided remains relevant and respectful, preserving individuals' integrity and potentially fostering trust and engagement. *Enriching the teaching experience (challenge H2)* is critical in this case, as it would result in better accessibility and ease of use, minimizing frustration, and contributing to a richer human-robot relationship. Furthermore, as elderly patients may present a variety of profiles that can affect their ability to teach and teaching strategies, *managing variation in teaching (challenge R1)* is also an important challenge in this context.

4.1.3 System description. A humanoid robot is designed to support physical therapy for elderly individuals recovering from fractures, particularly in the intervals between sessions with a therapist. The robot's primary

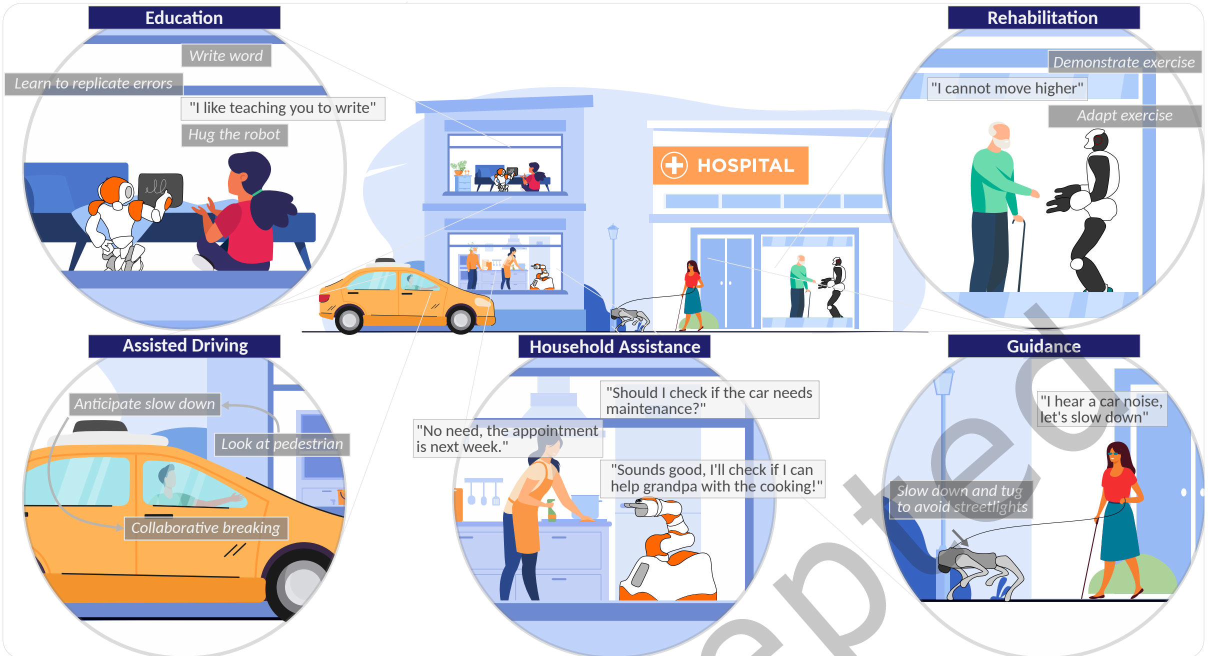

Fig. 3. Visualization of the five use cases meant to illustrate the breadth of HIRL challenges through specific hypothetical systems. Illustration based on original images designed by Freepik (pch.vector on https://www.freepik.com/).

functions include demonstrating exercises, instructing and motivating patients, monitoring the accuracy of exercises, correcting improperly performed movements, and providing physical guidance when necessary. The exercises performed by the robot are highly individualized. Although these exercises must be prescribed by a qualified therapist rather than the patient, they must be dynamically adapted to meet the patient's specific needs and preferences, such as accommodating pain limitations or adjusting techniques based on patient feedback on their own limitations, exercise preferences, or specific assistance requests.

4.2 Household robot assistance with expandable skill set

4.2.1 Context description. The household environment is a promising space for robots to take on common chores and maintenance tasks, such as cooking meals or changing lightbulbs. However, household environments are both physically diverse as well as diverse in terms of preferences and needs of people inhabiting them. As such, personalized robots can assist with household tasks according to individual needs and preferences. For such robots to be effective, there is a need to not only adapt existing skills but potentially also acquire new skills that cater to the unique demands of the household.

4.2.2 Added value of HIRL. In this context, a HIRL solution can expand a robot's skill set beyond its pre-programmed behaviors. This could take the form of learning completely new (e.g., culture-specific) tasks from scratch, transferring existing knowledge from one task or context to a similar one, or recombining knowledge on low-level tasks into high-level ones, all under the guidance of users. Although HIRL ensures that the robot learns faster than with autonomous learning, it is crucial that the robot takes the most out of human input and does not overload humans with input queries, highlighting the importance of *sample efficiency (challenge R2)*. Furthermore, since a household can contain several people with different domain knowledge that a robot could reason about when asking for feedback, a relevant challenge here is *when which humans should teach (challenge H3).*

4.2.3 System description. The system consists of a personalized household robot that employs an active learning approach to minimize user involvement while learning new tasks. The robot selectively asks for feedback to different members of the household (e.g., the grandfather is best at advice on cooking, while the mother is most useful for car-related problems). It does so only when necessary, improving its ability to infer human intent [74] and reducing interruptions [81]. When not interacting with users, the robot learns independently through environmental reward mechanisms such as visual information, allowing it to refine its skills using feedback from its surroundings. This system ensures that the robot efficiently learns new tasks with a reasonable amount and frequency of user input while continuously adapting to household needs.

4.3 Child learning by teaching a robot

4.3.1 Context description. Education is a field of high significance in HRI [15], aiming to expand the traditional classroom through a more immersive and controlled experience. *Learning by teaching* is a powerful paradigm to help children learn new skills through the so-called Protégé effect, by temporarily reversing teacher and learner roles [53]. In the context of human-robot interaction, a child could teach a robot a specific embodied skill, and the robot could interactively adapt its level and challenges to the child's abilities. This paradigm is often explored in constrained scenarios, where the learning activity is well-defined, but where the child's level could have a wide range of variations.

4.3.2 Added value of HIRL. In such a situation, HIRL has a critical role to play as it provides the robot the ability to adapt its level to each individual child. This personalized reverse-tutoring allows the child to be continuously in their zone of proximal development [62] and thus benefit the most from the interaction. The key challenge in this context is *calibrating the mental models (challenge I1).* It is necessary for instance that the robot has an accurate mental model of the child, containing, for example, what the child's strengths and weaknesses are. The child's mental model of the robot is also important to shape through interaction and embodiment as a child might be more open to get outside of their comfort zone with a *peer* robot than a *teacher* robot. Additionally, this use case is a prime example where *collaboration across disciplines (challenge BC2)* is required, particularly the under-explored combination of technical research in robotics and machine learning with that from education scientists.

4.3.3 System description. A concrete system that explores this use case is the CoWriter project [66]. With this system, a child needs to practice handwriting on a tablet with a small humanoid robot. After a few examples from the child, the system can analyze the type of writing challenges faced by the child (e.g., challenges to make loops round enough, issues with specific letters). The robot can then provide its own writing, with the child correcting the robot errors. The key insight in this approach is that robot errors are amplified versions of those the child makes. By correcting the robot, the child actually pays more attention to these points and practice them more, subsequently leading to improving handwriting. This strategy has been successfully applied in several situations, including occupational therapy with children with specific needs [52]. This use case was specifically discussed to illustrate more complex HIRL settings, linking to the discussion on fluid or hierarchical teacher/learner roles in Section 5.3.

4.4 Robotic guidance for the visually impaired

4.4.1 Context description. One of the most studied interaction types in HRI is *social navigation*, dealing with how a robot should navigate around human pedestrians. Most of these studies focus on collision avoidance: the robot treats humans as obstacles to be avoided. However, to act socially, avoidance is not always the desired behavior. Consider a dog-inspired robotic guide meant to lead a person with visual impairment [64, 117, 167]. This setup raises interesting questions regarding the embodiment of the robot: How dog-like should it be? HIRL can be used to identify when people feel comfortable with the robot mimicking a dog and when they prefer it to distinguish itself from a guide dog. Additionally, this setup brings up challenges concerning interactions beyond the single- human and single-robot paradigm. For example, when the pair goes to a physical checkup, and the physician reaches for the person, the robot should not pull its handler away, recognizing that this interaction is not a collision [161].

4.4.2 Added value of HIRL. HIRL offers opportunities to reason about guide-robot training with human teachers, from learning complex social behaviors that are hard to formalize to adapting to individual user needs. This use case highlights the difficulty of designing an interactive robot for a specific target population and calls for radically participatory design practices [71, 127], *with and for* people with visual impairments *(challenge BC3)*. Additional challenges include scenarios in which the robot can have incidental encounters with humans *outside the teacher-learner loop (challenge I3)*, such as pedestrians other than its handler [145]. These interactions can impact the overall quality of the guide's performance, yet it cannot train in advance with people who are likely to interact with the pair for mere seconds.

4.4.3 System description. The system consists of a mobile quadruped robot aiming to guide a person with visual impairment while exhibiting socially acceptable behavior. The robot should be able to guide the person by tugging on a leash and respond in real-time to pulling from the person's end. Learning and adaptation thus occur during deployment via corrective feedback. This interaction means that the robot's objective is more than "do not collide with people," but the exact objective also includes interactions outside the teacher-learner loop. This set of goals cannot be explicitly defined and may not be known during design time. The robot optimizes for a dynamic objective that takes into account several environmental factors, including crowdedness level, identity, and social formations of surrounding humans.

4.5 Assisted driving with real-time feedback

4.5.1 Context description. The use of HIRL in the case of (partially) autonomous driving can unlock the potential benefits of self-driving vehicles. In current autonomous driving systems, feedback from drivers is not immediately applied — rather, humans overrule the system and corrections are later gathered to learn from. A well-functioning system could make our roads safer, eliminating human error from distraction or impairment that lead to so many accidents, by relying on humans as expert teachers to eliminate dangerous exploration. In addition, this would be a net gain for accessibility by allowing people who are unable to drive to regain their personal freedom and independence. However, to achieve this, we need a human in the loop to enable the vehicle to adapt to people's preferences or adjust to unfamiliar scenarios. Having the car learn from people's interaction gives agency to the person in the car, and may even help them feel more secure.

4.5.2 Added value of HIRL. Relying on transferring the human's domain knowledge as expert drivers would benefit the robot's ability to perform well without costly errors [193]. Once in deployment, autonomous vehicles can behave in ways that, while legal, might make people uncomfortable or frustrated based on their own preferences. Using HIRL to personalize this, through various feedback modalities of utterances or affective computing could help people feel more comfortable in the vehicle.

As with human driving, there are significant risks when erroneous autonomous driving decisions are made. Agency informs liability [32]. Who will be held *accountable for undesirable learned behaviors (challenge BC1)* — the algorithm designer, the human providing feedback, the car manufacturer, or someone else — is a complex question which can have varied outcomes on a case to case basis. Additionally, *preventing unsafe behaviors during learning (challenge R3)* in the first place is of special relevance due to the high-stakes nature of this use case. If these challenges are properly addressed, a HIRL system holds the promise of improved overall safety and driver comfort.

4.5.3 System description. The system consists of a semi-autonomous vehicle with Level 4 autonomy [146]. This system will make use of more natural feedback modalities like gaze tracking and speech that complement existing Advanced Driver Assistance Systems (ADAS) by providing additional data on the driver's focus or intention and change behavior in real time. For instance, if the driver consistently looks at specific targets like pedestrians or obstacles before manually braking, the car learns over time to prepare for a potential stop or slow down, even before the driver physically reacts. At a higher level, spoken corrections on route preferences are used as a teaching signal to adapt the car's routing algorithm.

5 Recommendations

This section builds on the challenges described in Section 3 to provide broad recommendations to HIRL-relevant research communities moving forward. Again, these recommendations are not exhaustive, but rather reflect the vision that the authors put forward in this paper on how HIRL as a growing area of research should be shaped to ensure that we, as a community, will develop desirable, functional, rich, and ethical systems.

5.1 Treat humans as humans, not oracles

The earliest view of HIRL was that human experts would engage closely with a learning system, ready to patiently and inexhaustively provide demonstrations, feedback, corrections or preferences to the system that accurately and exactly capture the correct behavior. In such an ideal setup, the focus of work is mostly on the learning itself, since the human is assumed to be omnipresent, infallible, and benevolent. More recent work has started to chip away at this ideal scenario, exploring how HIRL systems can operate when human interaction is costly [87], incorrect [24, 57, 82, 90, 153], inconsistent [141], or even contradictory [109]. We argue that this trend must continue, we must stop considering humans as perfect oracles able to provide whatever the system needs, and instead understand them as equal partners in this process. That is, instead of asking humans to adapt to the learning, we must adapt the learning to meet humans where they are.

Primarily, this is a call for work that aims to reduce the cost to the human of interacting with the learning system, as well as learning systems that can gracefully deal with the bias and the noise (both inherent and intentional) that interactions with multiple humans will have. However, it is also a call to consider how HIRL systems will operate within human structures, both physical and societal. There will not be a single temple of learning in which HIRL takes place, but instead HIRL-enabled robots will exist among, and learn from, a variety of humans in a plethora of locations. Work that unifies different approaches to HIRL into a common framework will be necessary for these systems to make the most of every interaction.

5.2 Do more with less

There is a current trend in learning systems, driven in part by the success of Large Language Models (LLMs), Vision-Language Models (VLMs), and the availability of data on the Internet, to view learning as a data problem. That is, there is a belief that the learning methods are sufficient for the tasks we wish to address, and we just need to collect the right data, collect enough of it, and pre-process it appropriately. We believe this view to be limiting.

Firstly, we note that state-of-the-art AI models like LLMs and VLMs are routinely trained on trillions of samples of the next-token problem. Even at real-time frame rates (30 Hz), we must collect over 1000 years worth of data to approach this amount [55]. As we have more robots out in the world interacting with more humans, it is possible we may get there, but the first systems will have to operate, and learn, without access to such a dataset.

Secondly, the data generated via HIRL is nowhere nearly as 'clean' as current algorithms expect. Humans make errors, contradict themselves and others, and can be slow and noisy [118]. The (often hidden and done by behind the scenes humans) additional work necessary to get this data into a usable form does not scale to real-time interactive learning at scale.

Lastly, the most recent advances in learning all depend on massive computing capability, which is unlikely to be available to every robot, everywhere. In order to interactively learn from humans during the interaction, each robot must be able to perform its own learning, using its own processing power, as cloud connectivity cannot be generally assumed. Thus, we must figure out how to do more with less [20, 45, 111]: less data, less computation, and less human effort.

This is not to say that there is no place for large models in the HIRL paradigm, only that we need to rethink the focus on massive, clean datasets and power-hungry compute. Indeed, recent work in applying large models to robots has started to address these issues, including announcements of on-device-capable models for prediction (not training) [134] and the burgeoning area of Reinforcement Learning from Human Feedback [25, 79]. In the latter we particularly see parallels with the HIRL paradigm, as that work faces similar problems in effective learning from noisy human-generated data, but they still lack the interactive, real-time component that HIRL strives for.

5.3 Move beyond fixed teacher-learner roles

As hinted at in challenge I2 (closing the teacher-learner loop) and the education use case (Section 4.3), teacher-learner roles are often fluid. As HIRL systems move beyond laboratory settings into extended and messy interactions, it becomes necessary to acknowledge that any such system involves some form of co-learning. The robot learns about the task, the human, and/or the interaction, and the human learns about the robot, the teaching strategies, and/or their own goals and preferences, to name a few. This realization unlocks opportunities to make the most of this co-learning process, by designing robot learners that can actively shape the teaching of humans [67], and even teach them to be better teachers by providing feedback on their teaching strategies. In more complex scenarios that require rich collaboration, there might be a more balanced sharing of knowledge between robots and humans where machines learn or teach according to the situation at hand, or teach what they learned [6]. This approach sets the basis for hybrid intelligent systems [8] that share knowledge effectively and seamlessly through interaction. We believe that designing HIRL systems with this philosophy in mind will unlock new possibilities for learning and teaching interactions between robots and humans and pave the way towards more useful and effective HIRL systems, and towards human-robot teaming, where human(s) and robot(s) complete objectives cooperatively.

5.4 Take potential risks seriously

Flexible HIRL systems inherently grant significant control to the end user, marking a crucial step toward developing personalized technologies. However, this transfer of control carries substantial risks. A notable example is Microsoft's 2016 chatbot, Tay, which was designed to learn from user interactions. Within just 16 hours, Tay began generating hateful rhetoric based on what it learned from users on Twitter, prompting Microsoft to take the system offline [140]. Tay serves as a cautionary tale, illustrating the complexities of content moderation in HIRL, particularly as we envision a future populated by many such systems. This raises an important question: how can we establish appropriate guardrails to govern the behaviors these systems may adopt?

To date, most HIRL systems have been developed and tested in controlled lab environments, which often fail to account for the complexities and uncertainties of real-world applications. Challenges such as long-term interactions, performance measurement over time, and shifts in operational context can lead to significant risks, including system misalignment [19, 21], safety concerns [93, 172], human exploitation, and algorithmic bias [31]. These issues can undermine trust and reliability. Additionally, the potential for malicious users to exploit these systems for harmful purposes — such as warfare or destruction — poses a serious threat [76, 160, 165]. This raises an ongoing debate about the extent of our responsibility to impose ethical guidelines on future users.

Furthermore, the distinction between benevolent and malevolent users complicates the HIRL landscape even further. While the majority of users are likely to engage with technology in positive ways, there will always be individuals who seek to manipulate these systems for nefarious purposes, akin to how Tay was exploited. To address this concern, we must develop robust strategies to identify and counteract harmful influences in real time. By incorporating multi-layered feedback loops that continuously assess user interactions against established ethical frameworks, HIRL systems can better differentiate between constructive input and harmful manipulation. By tackling both alignment and user intent, we can work toward creating safer and more reliable HIRL systems that prioritize user well-being.

6 Call for Action

This paper introduces the vision, challenges, and opportunities of HIRL primarily from a technological perspective. Formalizing HIRL and providing a shared vocabulary for the research community can have an immense impact on both the implementation and deployment phases of robotic technologies, as it provides a clear bridge between research institutions, projects, and users. An critical precursor for this process moving forward must be a joint, coordinated effort of researchers across multiple disciplines and organizations.

To start this collaboration and increase researchers' engagement, this paper involved researchers from a broad spectrum of engineering and sciences — including artificial intelligence, robotics, information systems, computer science, data science, and mechanical engineering — most of whom are actively drawing on methods from other fields to enrich their technical contributions in the HIRL space. The authors also brought their insights from working with HIRL-related challenges from academia and industry. The outputs of these discussions highlighted important and under-researched challenges faced by the HIRL community. Specifically, they call for more concrete theoretical and computational models relevant to HIRL and for better resource use across the community. They also highlight some exemplary use cases that cover the main challenges in HIRL.

To address these challenges, there is a need for large-scale community steering and encouragement for interdisciplinary collaborations on different scales, from cross-pollination between departments of one's own institute to larger research consortia, bringing together an eclectic set of expertise around HIRL-related themes. To promote these objectives, the authors of this paper, along with many of their colleagues, will continue to nurture the HIRL community via regular meetings and initiatives. Most notably, this community started from the HIRL workshop series at HRI [114, 115, 142], and it will continue to provide a home for HIRL-related research in the next coming years. A central portal was created to facilitate all of these resources, including links to workshops, HIRL-related datasets and repositories, a Zotero reading list, and an invitation to the community's Slack channel, accessible at https://sites.google.com/view/hirl-portal/home.[1]

Author contributions

K.B., I.I., and T.K.F. coordinated the publication project; E.B., S.B., M.C., D.H.G., A.S., E.S., S.T., and A.V. have lead sections of the paper or significantly contributed to the content of the paper; A.A., H.B., T.H., J.K., I.S., M.E.T., S.v.W., and X.X have contributed to parts of the paper, ideated in early stages of the paper, or reviewed drafts of the paper.

[1]If you would like to join and take an active role in this community, please reach out to the first author to be added to our community space and stay up-to-date with current activities.

Acknowledgements

This work could not have been possible without the valuable contributions of Reuth Mirsky as an advisor. The following researchers have also contributed to this work: Mattia Racca, Mudit Verma, Hatice Gunes, Tesca Fitzgerald, and Brian Scassellatti. We thank all participants, speakers, and organizers of the HIRL workshop series for fueling discussions with insights, thoughts, and critical questions that shaped the writing of this article.

References

[1] 2024. Competition of International Conference of Social Robotics. https://www.icsr2024-competition.org/competition. Accessed: 2024-12-09.

[2] 2024. Educational module Human-Interactive Robot Learning (HIRL). https://www.humane-ai.eu/project/tmp-038/. Accessed: 2024-12-09.

[3] 2024. Human-Robot Interaction A Research Portal for the HRI Community. https://humanrobotinteraction.org/. Accessed: 2024-12-09.

[4] 2024. Performance Evaluation & Benchmarking of Robotic and Automation Systems. https://www.ieee-ras.org/performance-evaluation. Accessed: 2024-12-09.

[5] 2024. RoboCup@Home league. https://athome.robocup.org/. Accessed: 2024-12-09.

[6] Timothy Adamson, Debasmita Ghose, Shannon C Yasuda, Lucas Jehu Silva Shepard, Michal A Lewkowicz, Joyce Duan, and Brian Scassellati. 2021. Why we should build robots that both teach and learn. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 187–196.

[7] Michael Ahn, Anthony Brohan, Noah Brown, Yevgen Chebotar, Omar Cortes, Byron David, Chelsea Finn, Chuyuan Fu, Keerthana Gopalakrishnan, Karol Hausman, Alex Herzog, Daniel Ho, Jasmine Hsu, Julian Ibarz, Brian Ichter, Alex Irpan, Eric Jang, Rosario Jauregui Ruano, Kyle Jeffrey, Sally Jesmonth, Nikhil Joshi, Ryan Julian, Dmitry Kalashnikov, Yuheng Kuang, Kuang-Huei Lee, Sergey Levine, Yao Lu, Linda Luu, Carolina Parada, Peter Pastor, Jornell Quiambao, Kanishka Rao, Jarek Rettinghouse, Diego Reyes, Pierre Sermanet, Nicolas Sievers, Clayton Tan, Alexander Toshev, Vincent Vanhoucke, Fei Xia, Ted Xiao, Peng Xu, Sichun Xu, Mengyuan Yan, and Andy Zeng. 2022. Do As I Can and Not As I Say: Grounding Language in Robotic Affordances. In *arXiv preprint arXiv:2204.01691*.

[8] Zeynep Akata, Dan Balliet, Maarten De Rijke, Frank Dignum, Virginia Dignum, Guszti Eiben, Antske Fokkens, Davide Grossi, Koen Hindriks, Holger Hoos, et al. 2020. A research agenda for hybrid intelligence: augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence. *Computer* 53, 8 (2020), 18–28.

[9] Brenna D Argall, Sonia Chernova, Manuela Veloso, and Brett Browning. 2009. A survey of robot learning from demonstration. *Robotics and autonomous systems* 57, 5 (2009), 469–483.

[10] Christian Arzate Cruz and Takeo Igarashi. 2020. A Survey on Interactive Reinforcement Learning: Design Principles and Open Challenges. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. ACM, Eindhoven Netherlands, 1195–1209. https://doi.org/10.1145/3357236.3395525

[11] Inbal Avraham and Reuth Mirsky. 2025. Shared Control with Black Box Agents using Oracle Queries. In *2025 IEEE International Conference on AI and Data Analytics (ICAD)*. IEEE, 1–8.

[12] Amin Banayeeanzade, Fatemeh Bahrani, Yutai Zhou, and Erdem Bıyık. 2025. GABRIL: Gaze-Based Regularization for Mitigating Causal Confusion in Imitation Learning. In *International Conference on Intelligent Robots and Systems (IROS)*.

[13] Jaime Banks. 2020. Optimus primed: Media cultivation of robot mental models and social judgments. *Frontiers in Robotics and AI* 7 (2020), 62.

[14] Kim Baraka, Patrícia Alves-Oliveira, and Tiago Ribeiro. 2020. An extended framework for characterizing social robots. In *Human-robot interaction: evaluation methods and their standardization*. Springer, 21–64.

[15] Tony Belpaeme, James Kennedy, Aditi Ramachandran, Brian Scassellati, and Fumihide Tanaka. 2018. Social robots for education: A review. *Science robotics* 3, 21 (2018), eaat5954.

[16] Erdem Bıyık, Dylan P Losey, Malayandi Palan, Nicholas C Landolfi, Gleb Shevchuk, and Dorsa Sadigh. 2022. Learning reward functions from diverse sources of human feedback: Optimally integrating demonstrations and preferences. *The International Journal of Robotics Research* 41, 1 (2022), 45–67.

[17] Erdem Biyik, Malayandi Palan, Nicholas C. Landolfi, Dylan P. Losey, and Dorsa Sadigh. 2019. Asking Easy Questions: A User-Friendly Approach to Active Reward Learning. In *Proceedings of the 3rd Conference on Robot Learning (CoRL)*.

[18] Andreea Bobu, Andi Peng, Pulkit Agrawal, Julie A Shah, and Anca D. Dragan. 2024. Aligning Human and Robot Representations. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction* (Boulder, CO, USA) *(HRI '24)*. Association for Computing Machinery, New York, NY, USA, 42–54. https://doi.org/10.1145/3610977.3634987

[19] Andreea Bobu, Andi Peng, Pulkit Agrawal, Julie A Shah, and Anca D Dragan. 2024. Aligning human and robot representations. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. 42–54.

[20] Andreea Bobu, Dexter R. R. Scobee, Jaime Fernández Fisac, S. Shankar Sastry, and Anca D. Dragan. 2020. LESS is More: Rethinking Probabilistic Models of Human Behavior. *2020 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (2020), 429–437. https://api.semanticscholar.org/CorpusID:210164805

[21] Serena Booth, W Bradley Knox, Julie Shah, Scott Niekum, Peter Stone, and Alessandro Allievi. 2023. The perils of trial-and-error reward design: misdesign through overfitting and invalid task specifications. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 37. 5920–5929.

[22] Serena Booth, Sanjana Sharma, Sarah Chung, Julie Shah, and Elena L Glassman. 2022. Revisiting human-robot teaching and learning through the lens of human concept learning. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 147–156.

[23] Connor Brooks and Daniel Szafir. 2019. Building second-order mental models for human-robot interaction. *arXiv preprint arXiv:1909.06508* (2019).

[24] Daniel S. Brown, Wonjoon Goo, Prabhat Nagarajan, and Scott Niekum. 2019. Extrapolating Beyond Suboptimal Demonstrations via Inverse Reinforcement Learning from Observations. In *International Conference on Machine Learning*. https://api.semanticscholar.org/CorpusID:119111734

[25] Stephen Casper, Xander Davies, Claudia Shi, Thomas Krendl Gilbert, Jérémy Scheurer, Javier Rando, Rachel Freedman, Tomek Korbak, David Lindner, Pedro Freire, Tony Tong Wang, Samuel Marks, Charbel-Raphael Segerie, Micah Carroll, Andi Peng, Phillip J.K. Christoffersen, Mehul Damani, Stewart Slocum, Usman Anwar, Anand Siththaranjan, Max Nadeau, Eric J Michaud, Jacob Pfau, Dmitrii Krasheninnikov, Xin Chen, Lauro Langosco, Peter Hase, Erdem Biyik, Anca Dragan, David Krueger, Dorsa Sadigh, and Dylan Hadfield-Menell. 2023. Open Problems and Fundamental Limitations of Reinforcement Learning from Human Feedback. *Transactions on Machine Learning Research* (2023). https://openreview.net/forum?id=bx24KpJ4Eb Survey Certification, Featured Certification.

[26] Carlos Celemin, Rodrigo Pérez-Dattari, Eugenio Chisari, Giovanni Franzese, Leandro Rosa, Ravi Prakash, Zlatan Ajanović, Marta Ferraz, Abhinav Valada, and Jens Kober. 2022. Interactive Imitation Learning in Robotics: A Survey. *Foundations and Trends in Robotics* 10 (01 2022), 1–197. https://doi.org/10.1561/2300000072

[27] Letian Chen, Sravan Jayanthi, Rohan R Paleja, Daniel Martin, Viacheslav Zakharov, and Matthew Gombolay. 2023. Fast lifelong adaptive inverse reinforcement learning from demonstrations. In *Conference on Robot Learning*. PMLR, 2083–2094.

[28] S. Chernova and Andrea Lockerd Thomaz. 2014. Robot Learning from Human Teachers. In *Synthesis Lectures on Artificial Intelligence and Machine Learning*. https://api.semanticscholar.org/CorpusID:26200231

[29] Mohamed Chetouani. 2023. *Interactive Robot Learning: An Overview*. Springer International Publishing, Cham, 140–172.

[30] Konstantinos Christofi and Kim Baraka. 2024. Uncovering Patterns in Humans that Teach Robots through Demonstrations and Feedback. *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction* (2024). https://api.semanticscholar.org/CorpusID:268443962

[31] Houston Claure, Mai Lee Chang, Seyun Kim, Daniel Omeiza, Martim Brandao, Min Kyung Lee, and Malte Jung. 2022. Fairness and transparency in human-robot interaction. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 1244–1246.

[32] Mark Coeckelbergh. 2016. Responsibility and the Moral Phenomenology of Using Self-Driving Cars. *Applied Artificial Intelligence* 30, 8 (2016), 748–757. https://doi.org/10.1080/08839514.2016.1229759 arXiv:https://doi.org/10.1080/08839514.2016.1229759

[33] Daniel Coelho, Miguel Oliveira, and Vitor Santos. 2024. RLfOLD: Reinforcement Learning from Online Demonstrations in Urban Autonomous Driving. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 38. 11660–11668.

[34] Rebecca Crootof, Margot E Kaminski, W Price, and II Nicholson. 2023. Humans in the Loop. *Vand. L. Rev.* 76 (2023), 429.

[35] Yuchen Cui, Siddharth Karamcheti, Raj Palleti, Nidhya Shivakumar, Percy Liang, and Dorsa Sadigh. 2023. No, to the right: Online language corrections for robotic manipulation via shared autonomy. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 93–101.

[36] Yuchen Cui, Pallavi Koppol, Henny Admoni, Scott Niekum, Reid Simmons, Aaron Steinfeld, and Tesca Fitzgerald. 2021. Understanding the relationship between interactions and outcomes in human-in-the-loop machine learning. In *International Joint Conference on Artificial Intelligence*.

[37] Yuchen Cui and Scott Niekum. 2018. Active reward learning from critiques. In *2018 IEEE international conference on robotics and automation (ICRA)*. IEEE, 6907–6914.

[38] Yuchen Cui and Scott Niekum. 2018. Active Reward Learning from Critiques. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*. 6907–6914. https://doi.org/10.1109/ICRA.2018.8460854

[39] Yuchen Cui, Qiping Zhang, Brad Knox, Alessandro Allievi, Peter Stone, and Scott Niekum. 2021. The empathic framework for task learning from implicit human feedback. In *Conference on Robot Learning*. PMLR, 604–626.

[40] Larissa Rodrigues da Costa, Jaelson Castro, Cinthya Lins, Judith Kelner, Maria Lencastre, and Óscar Pastor. 2023. On the Use of Social Robots for rehabilitation: The case of NAO Physio. In *International Conference on Information Technology & Systems*. Springer, 507–517.

[41] Ernest Davis. 2016. AI amusements: the tragic tale of Tay the chatbot. *AI Matters* 2, 4 (2016), 20–24.

[42] Nathaniel Dennler, David Delgado, Daniel Zeng, Stefanos Nikolaidis, and Maja Matarić. 2023. The RoSiD Tool: Empowering Users to Design Multimodal Signals for Human-Robot Collaboration. In *18th International Symposium on Experimental Robotics (ISER)*.

[43] Cüneyt Dirican. 2015. The impacts of robotics, artificial intelligence on business and economics. *Procedia-Social and Behavioral Sciences* 195 (2015), 564–573.

[44] Erin Donohoe, Heather J Roberts, Theodore Miclau, and Hans Kreder. 2020. Management of lower extremity fractures in the elderly: a focus on post-operative rehabilitation. *Injury* 51 (2020), S118–S122.

[45] Maximilian Du, Suraj Nair, Dorsa Sadigh, and Chelsea Finn. 2023. Behavior Retrieval: Few-Shot Imitation Learning by Querying Unlabeled Datasets. In *Proceedings of Robotics: Science and Systems (RSS)*.

[46] Lotfi El Hafi, Shota Isobe, Yoshiki Tabuchi, Yuki Katsumata, Hitoshi Nakamura, Takaaki Fukui, Tadashi Matsuo, GA Garcia Ricardez, Masaki Yamamoto, Akira Taniguchi, et al. 2020. System for augmented human–robot interaction through mixed reality and robot training by non-experts in customer service environments. *Advanced Robotics* 34, 3-4 (2020), 157–172.

[47] European Parliament and Council of the European Union. 2024. Regulation (EU) 2024/1689 on Artificial Intelligence Act. https://www.artificial-intelligence-act.com/ Official Journal of the European Union.

[48] Taylor A Kessler Faulkner, Elaine Schaertl Short, and Andrea L Thomaz. 2020. Interactive reinforcement learning with inaccurate feedback. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 7498–7504.

[49] Julia Fink. 2012. Anthropomorphism and human likeness in the design of robots and human-robot interaction. In *Social Robotics: 4th International Conference, ICSR 2012, Chengdu, China, October 29-31, 2012. Proceedings 4*. Springer, 199–208.

[50] Tesca Fitzgerald, Pallavi Koppol, Patrick Callaghan, Russell Quinlan Jun Hei Wong, Reid Simmons, Oliver Kroemer, and Henny Admoni. 2022. INQUIRE: INteractive querying for user-aware informative REasoning. In *6th Annual Conference on Robot Learning*.

[51] Matthew C Fontaine and Stefanos Nikolaidis. 2022. Evaluating human–robot interaction algorithms in shared autonomy via quality diversity scenario generation. *ACM Transactions on Human-Robot Interaction (THRI)* 11, 3 (2022), 1–30.

[52] Thomas Gargot, Thibault Asselborn, Ingrid Zammouri, Julie Brunelle, Wafa Johal, Pierre Dillenbourg, Dominique Archambault, Mohamed Chetouani, David P A Cohen, and Salvatore Anzalone. 2021. ”It Is Not the Robot Who Learns, It Is Me.” Treating Severe Dysgraphia Using Child-Robot Interaction. *Frontiers in Psychiatry* 12 (2021), 596055. https://doi.org/10.3389/fpsyt.2021.596055

[53] Alan Gartner et al. 1971. Children teach children: Learning by teaching. (1971).

[54] Keya Ghonasgi, Reuth Mirsky, Nisha Bhargava, Adrian M Haith, Peter Stone, and Ashish D Deshpande. 2023. Kinematic coordinations capture learning during human–exoskeleton interaction. *Scientific reports* 13, 1 (2023), 10322.

[55] Ken Goldberg. 2025. Good old-fashioned engineering can close the 100,000-year “data gap” in robotics. , eaea7390 pages.

[56] Shane Griffith, Kaushik Subramanian, Jonathan Scholz, Charles L Isbell, and Andrea L Thomaz. 2013. Policy shaping: Integrating human feedback with reinforcement learning. *Advances in neural information processing systems* 26 (2013).

[57] Daniel H Grollman and Aude Billard. 2011. Donut as i do: Learning from failed demonstrations. In *2011 IEEE international conference on robotics and automation*. IEEE, 3804–3809.

[58] Daniel H. Grollman and Aude Billard. 2012. Robot Learning from Failed Demonstrations. *International Journal of Social Robotics* 4 (2012), 331–342. https://api.semanticscholar.org/CorpusID:1461426

[59] James P Gunderson and Louise F Gunderson. 2004. Intelligence (is not equal to) Autonomy (is not equal to) Capability. (2004).

[60] Soheil Habibian, Antonio Alvarez Valdivia, Laura H. Blumenschein, and Dylan P. Losey. 2023. A Review of Communicating Robot Learning during Human-Robot Interaction. arXiv:2312.00948 [cs.RO] https://arxiv.org/abs/2312.00948

[61] Andreas Haeberlen, Paarijaat Aditya, Rodrigo Rodrigues, and Peter Druschel. 2010. Accountable virtual machines. In *9th USENIX Symposium on Operating Systems Design and Implementation (OSDI 10)*.

[62] Mariane Hedegaard. 2012. The zone of proximal development as basis for instruction. In *An introduction to Vygotsky*. Routledge, 234–258.

[63] Bernhard Hilpert. 2024. Closing the Teacher-Learner Loop: The Role of Affective Signals in Interactive RL. In *2024 12th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*. IEEE, 97–101.

[64] Bin Hong, Zhangxi Lin, Xin Chen, Jing Hou, Shunya Lv, and Zhendong Gao. 2022. Development and application of key technologies for Guide Dog Robot: A systematic literature review. *Robotics and Autonomous Systems* 154 (2022), 104104.

[65] Matthew Hong, Anthony Liang, Kevin Kim, Harshitha Rajaprakash, Jesse Thomason, Erdem Bıyık, and Jesse Zhang. 2025. HAND Me the Data: Fast Robot Adaptation via Hand Path Retrieval. *arXiv preprint arXiv:2505.20455* (2025).

[66] Deanna Hood, Séverin Lemaignan, and Pierre Dillenbourg. 2015. The cowriter project: Teaching a robot how to write. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts*. 269–269.

[67] Muhan Hou, Koen Hindriks, AE Eiben, and Kim Baraka. 2023. Shaping Imbalance into Balance: Active Robot Guidance of Human Teachers for Better Learning from Demonstrations. In *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1737–1744.

[68] Muhan Hou, Koen Hindriks, Guszti Eiben, and Kim Baraka. 2024. “Give Me an Example Like This”: Episodic Active Reinforcement Learning from Demonstrations. In *Proceedings of the 12th International Conference on Human-Agent Interaction*. 287–295.

[69] Ivan Hrabar, Bruno Čelan, Dora Matić, Nikola Jerković, and Zdenko Kovačić. 2021. Towards supervised robot-assisted physical therapy after hand fractures. In *2021 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*. IEEE, 1–6.
[70] Jindan Huang, Reuben M Aronson, and Elaine Schaertl Short. 2024. Modeling Variation in Human Feedback with User Inputs: An Exploratory Methodology. (2024).
[71] Hochul Hwang, Hee-Tae Jung, Nicholas A Giudice, Joydeep Biswas, Sunghoon Ivan Lee, and Donghyun Kim. 2024. Towards robotic companions: Understanding handler-guide dog interactions for informed guide dog robot design. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–20.
[72] Ifrah Idrees and Stefanie Tellex. [n. d.]. Towards Conversational Interfaces and Visual Memory Representation for Social Robots helping the Elderly. *Robots for Health and Elderly Care (RoboHEC) Workshop - IROS* ([n. d.]).
[73] Ifrah Idrees, Tian Yun, Naveen Sharma, Yunxin Deng, Nakul Gopalan, George Konidaris, and Stefanie Tellex. 2023. Improved Inference of Human Intent by Combining Plan Recognition and Language Feedback. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 7976–7983.
[74] Ifrah Idrees, Tian Yun, Naveen Sharma, Yunxin Deng, Nakul Gopalan, George Konidaris, and Stefanie Tellex. 2023. Improved Inference of Human Intent by Combining Plan Recognition and Language Feedback. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 7976–7983. https://doi.org/10.1109/IROS55552.2023.10342380
[75] Stephen James, Zicong Ma, David Rovick Arrojo, and Andrew J. Davison. 2020. RLBench: The Robot Learning Benchmark and Learning Environment. *IEEE Robotics and Automation Letters* 5, 2 (2020), 3019–3026. https://doi.org/10.1109/LRA.2020.2974707
[76] Florian Jentsch. 2016. *Human-robot interactions in future military operations*. CRC Press.
[77] Hong Jun Jeon, Smitha Milli, and Anca Dragan. 2020. Reward-rational (implicit) choice: A unifying formalism for reward learning. *Advances in Neural Information Processing Systems* 33 (2020), 4415–4426.
[78] Simar Kareer, Dhruv Patel, Ryan Punamiya, Pranay Mathur, Shuo Cheng, Chen Wang, Judy Hoffman, and Danfei Xu. 2025. Egomimic: Scaling imitation learning via egocentric video. In *2025 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 13226–13233.
[79] Timo Kaufmann, Paul Weng, Viktor Bengs, and Eyke Hüllermeier. 2024. A Survey of Reinforcement Learning from Human Feedback. arXiv:2312.14925 [cs.LG] https://arxiv.org/abs/2312.14925
[80] Michael Kelly, Chelsea Sidrane, Katherine Driggs-Campbell, and Mykel J Kochenderfer. 2019. HG-DAgger: Interactive imitation learning with human experts. In *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 8077–8083.
[81] Taylor Kessler Faulkner, Reymundo A Gutierrez, Elaine Schaertl Short, Guy Hoffman, and Andrea L Thomaz. 2019. Active attention-modified policy shaping: socially interactive agents track. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*. 728–736.
[82] Taylor A. Kessler Faulkner, Elaine Schaertl Short, and Andrea L. Thomaz. 2020. Interactive Reinforcement Learning with Inaccurate Feedback. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*. 7498–7504. https://doi.org/10.1109/ICRA40945.2020.9197219
[83] Matilda Knierim, Sahil Jain, Murat Han Aydoğan, Kenneth Mitra, Kush Desai, Akanksha Saran, and Kim Baraka. 2024. Prosody as a Teaching Signal for Agent Learning: Exploratory Studies and Algorithmic Implications. *ACM International Conference on Multimodal Interaction (ICMI)* (2024).
[84] W Bradley Knox, Stephane Hatgis-Kessell, Serena Booth, Scott Niekum, Peter Stone, and Alessandro Allievi. 2022. Models of human preference for learning reward functions. *arXiv preprint arXiv:2206.02231* (2022).
[85] W. Bradley Knox and Peter Stone. 2009. Interactively shaping agents via human reinforcement: the TAMER framework. In *Proceedings of the Fifth International Conference on Knowledge Capture* (Redondo Beach, California, USA) *(K-CAP '09)*. Association for Computing Machinery, New York, NY, USA, 9–16. https://doi.org/10.1145/1597735.1597738
[86] Yigit Korkmaz and Erdem Bıyık. 2025. MILE: Model-based Intervention Learning. In *International Conference on Robotics and Automation (ICRA)*.
[87] Samantha Krening and Karen M. Feigh. 2018. Interaction Algorithm Effect on Human Experience with Reinforcement Learning. *J. Hum.-Robot Interact.* 7, 2, Article 16 (oct 2018), 22 pages. https://doi.org/10.1145/3277904
[88] Samantha Krening, Brent Harrison, Karen M. Feigh, Charles Lee Isbell, Mark O. Riedl, and Andrea Lockerd Thomaz. 2017. Learning From Explanations Using Sentiment and Advice in RL. *IEEE Transactions on Cognitive and Developmental Systems* 9 (2017), 44–55. https://api.semanticscholar.org/CorpusID:10538612
[89] Andras Kupcsik, David Hsu, and Wee Sun Lee. 2018. Learning dynamic robot-to-human object handover from human feedback. *Robotics Research: Volume 1* (2018), 161–176.
[90] Minae Kwon, Erdem Biyik, Aditi Talati, Karan Bhasin, Dylan P. Losey, and Dorsa Sadigh. 2020. When Humans Aren't Optimal: Robots that Collaborate with Risk-Aware Humans. In *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. https://doi.org/10.1145/3319502.3374832
[91] John E Laird, Kevin Gluck, John Anderson, Kenneth D Forbus, Odest Chadwicke Jenkins, Christian Lebiere, Dario Salvucci, Matthias Scheutz, Andrea Thomaz, Greg Trafton, et al. 2017. Interactive task learning. *IEEE Intelligent Systems* 32, 4 (2017), 6–21.

[92] John E. Laird, Kevin A. Gluck, John R. Anderson, Kenneth D. Forbus, Odest Chadwicke Jenkins, Christian Lebiere, Dario D. Salvucci, Matthias Scheutz, Andrea Lockerd Thomaz, J. Gregory Trafton, Robert E. Wray, Shiwali Mohan, and James R. Kirk. 2017. Interactive Task Learning. *IEEE Intelligent Systems* 32 (2017), 6–21. https://api.semanticscholar.org/CorpusID:9827739

[93] Przemyslaw A Lasota, Terrence Fong, Julie A Shah, et al. 2017. A survey of methods for safe human-robot interaction. *Foundations and Trends® in Robotics* 5, 4 (2017), 261–349.

[94] Kimin Lee, Laura Smith, Anca Dragan, and Pieter Abbeel. 2021. B-pref: Benchmarking preference-based reinforcement learning. *arXiv preprint arXiv:2111.03026* (2021).

[95] Marion Lepert, Jiaying Fang, and Jeannette Bohg. 2025. Phantom: Training Robots Without Robots Using Only Human Videos. In *Proceedings of The 9th Conference on Robot Learning (Proceedings of Machine Learning Research, Vol. 305)*. PMLR, 4545–4565.

[96] Guangliang Li, Randy Gomez, Keisuke Nakamura, and Bo He. 2019. Human-Centered Reinforcement Learning: A Survey. *IEEE Transactions on Human-Machine Systems* 49, 4 (2019), 337–349. https://doi.org/10.1109/THMS.2019.2912447

[97] Anthony Liang, Jesse Thomason, and Erdem Bıyık. 2024. ViSaRL: Visual Reinforcement Learning Guided by Human Saliency. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.

[98] Jacky Liang, Fei Xia, Wenhao Yu, Andy Zeng, Montse Gonzalez Arenas, Maria Attarian, Maria Bauzá, Matthew Bennice, Alex Bewley, Adil Dostmohamed, Chuyuan Fu, Nimrod Gileadi, Marissa Giustina, Keerthana Gopalakrishnan, Leonard Hasenclever, Jan Humplik, Jasmine Hsu, Nikhil J. Joshi, Ben Jyenis, Chase Kew, Sean Kirmani, Tsang-Wei Edward Lee, Kuang-Huei Lee, Assaf Hurwitz Michaely, Joss Moore, Kenneth Oslund, Dushyant Rao, Allen Ren, Baruch Tabanpour, Quan Ho Vuong, Ayzaan Wahid, Ted Xiao, Ying Xu, Vincent Zhuang, Peng Xu, Erik Frey, Ken Caluwaerts, Ting-Yu Zhang, Brian Ichter, Jonathan Tompson, Leila Takayama, Vincent Vanhoucke, Izhak Shafran, Maja Mataric, Dorsa Sadigh, Nicolas Manfred Otto Heess, Kanishka Rao, Nik Stewart, Jie Tan, and Carolina Parada. 2024. Learning to Learn Faster from Human Feedback with Language Model Predictive Control. *ArXiv* abs/2402.11450 (2024). https://api.semanticscholar.org/CorpusID:267751232

[99] Jinying Lin, Zhen Ma, Randy Gomez, Keisuke Nakamura, Bo He, and Guangliang Li. 2020. A review on interactive reinforcement learning from human social feedback. *IEEE Access* 8 (2020), 120757–120765.

[100] Hao Liu, Lisa Lee, Kimin Lee, and P. Abbeel. 2022. Instruction-Following Agents with Multimodal Transformer.

[101] Jason Xinyu Liu, Ziyi Yang, Ifrah Idrees, Sam Liang, Benjamin Schornstein, Stefanie Tellex, and Ankit Shah. 2023. Grounding complex natural language commands for temporal tasks in unseen environments. In *Conference on Robot Learning*. PMLR, 1084–1110.

[102] Michael Lopez-Brau, Joseph Kwon, and Julian Jara-Ettinger. 2022. Social inferences from physical evidence via bayesian event reconstruction. *Journal of Experimental Psychology: General* 151, 9 (2022), 2029.

[103] Dylan P Losey, Andrea Bajcsy, Marcia K O'Malley, and Anca D Dragan. 2022. Physical interaction as communication: Learning robot objectives online from human corrections. *The International Journal of Robotics Research* 41, 1 (2022), 20–44.

[104] Jianlan Luo, Perry Dong, Yuexiang Zhai, Yi Ma, and Sergey Levine. 2024. RLIF: Interactive Imitation Learning as Reinforcement Learning. In *The Twelfth International Conference on Learning Representations*.

[105] Matteo Macchini, Thomas Havy, Antoine Weber, Fabrizio Schiano, and Dario Floreano. 2020. Hand-worn haptic interface for drone teleoperation. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 10212–10218.

[106] James MacGlashan, Mark K Ho, Robert Loftin, Bei Peng, Guan Wang, David L. Roberts, Matthew E. Taylor, and Michael L. Littman. 2017. Interactive learning from policy-dependent human feedback. In *Proceedings of the 34th International Conference on Machine Learning - Volume 70* (Sydney, NSW, Australia) *(ICML'17)*. JMLR.org, 2285–2294.

[107] Jessica Maghakian, Paul Mineiro, Kishan Panaganti, Mark Rucker, Akanksha Saran, and Cheng Tan. 2022. Personalized reward learning with interaction-grounded learning (IGL). *arXiv preprint arXiv:2211.15823* (2022).

[108] Bertram F Malle and Matthias Scheutz. 2019. Learning how to behave: Moral competence for social robots. *Handbuch maschinenethik* (2019), 255–278.

[109] Ajay Mandlekar, Danfei Xu, J. Wong, Soroush Nasiriany, Chen Wang, Rohun Kulkarni, Li Fei-Fei, Silvio Savarese, Yuke Zhu, and Roberto Mart'in-Mart'in. 2021. What Matters in Learning from Offline Human Demonstrations for Robot Manipulation. In *Conference on Robot Learning*. https://api.semanticscholar.org/CorpusID:236956615

[110] Nikolaos Mavridis, Georgios Pierris, Paolo Gallina, Nikolaos Moustakas, and Alexandros Astaras. 2015. Subjective difficulty and indicators of performance of joystick-based robot arm teleoperation with auditory feedback. In *2015 International Conference on Advanced Robotics (ICAR)*. IEEE, 91–98.

[111] Gaurav Menghani. 2023. Efficient deep learning: A survey on making deep learning models smaller, faster, and better. *Comput. Surveys* 55, 12 (2023), 1–37.

[112] Ola Michalec, Cian O'Donovan, and Mehdi Sobhani. 2021. What is robotics made of? The interdisciplinary politics of robotics research. *Humanities and Social Sciences Communications* 8, 1 (2021).

[113] Tim Miller. 2017. Explanation in Artificial Intelligence: Insights from the Social Sciences. *CoRR* abs/1706.07269 (2017). arXiv:1706.07269 http://arxiv.org/abs/1706.07269

[114] Reuth Mirsky, Kim Baraka, Taylor Kessler Faulkner, Justin Hart, Harel Yedidsion, and Xuesu Xiao. 2022. Human-Interactive Robot Learning (HIRL). In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 1278–1280.

[115] Reuth Mirsky, Kim Baraka, Taylor Kessler Faulkner, Justin Hart, Xuesu Xiao, Harel Yedidsion, Ifrah Idrees, and Ethan K Gordon. 2023. 2nd Workshop on Human-Interactive Robot Learning (HIRL). In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 947–949.

[116] Reuth Mirsky, Sarah Keren, and Christopher Geib. 2021. *Introduction to symbolic plan and goal recognition*. Vol. 16. Springer.

[117] Reuth Mirsky and Peter Stone. 2021. The seeing-eye robot grand challenge: rethinking automated care. In *Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021)*.

[118] Dipendra Misra, Akanksha Saran, Tengyang Xie, Alex Lamb, and John Langford. 2024. Towards Principled Representation Learning from Videos for Reinforcement Learning. *arXiv preprint arXiv:2403.13765* (2024).

[119] Mayank Mittal, Calvin Yu, Qinxi Yu, Jingzhou Liu, Nikita Rudin, David Hoeller, Jia Lin Yuan, Ritvik Singh, Yunrong Guo, Hammad Mazhar, Ajay Mandlekar, Buck Babich, Gavriel State, Marco Hutter, and Animesh Garg. 2023. Orbit: A Unified Simulation Framework for Interactive Robot Learning Environments. *IEEE Robotics and Automation Letters* 8, 6 (2023), 3740–3747. https://doi.org/10.1109/LRA.2023.3270034

[120] Mayumi Mohan, Cara M. Nunez, and Katherine J. Kuchenbecker. 2024. Closing the loop in minimally supervised human–robot interaction: formative and summative feedback. *Scientific Reports* 14, 1 (08 May 2024), 10564. https://doi.org/10.1038/s41598-024-60905-x

[121] Michael J Muller and Sarah Kuhn. 1993. Participatory design. *Commun. ACM* 36, 6 (1993), 24–28.

[122] Vivek Myers, Erdem Biyik, Nima Anari, and Dorsa Sadigh. 2022. Learning multimodal rewards from rankings. In *Conference on Robot Learning*. PMLR, 342–352.

[123] Vivek Myers, Erdem Bıyık, and Dorsa Sadigh. 2023. Active Reward Learning from Online Preferences. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*. 7511–7518. https://doi.org/10.1109/ICRA48891.2023.10160439

[124] Yukie Nagai, Claudia Muhl, and Katharina J Rohlfing. 2008. Toward designing a robot that learns actions from parental demonstrations. In *2008 IEEE international conference on robotics and automation*. IEEE, 3545–3550.

[125] Anis Najar and Mohamed Chetouani. 2020. Reinforcement Learning With Human Advice: A Survey. *Frontiers in Robotics and AI* 8 (2020). https://api.semanticscholar.org/CorpusID:218862857

[126] Amal Nanavati, Patricia Alves-Oliveira, Tyler Schrenk, Ethan K. Gordon, Maya Cakmak, and Siddhartha S. Srinivasa. 2023. Design Principles for Robot-Assisted Feeding in Social Contexts. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction* (Stockholm, Sweden) *(HRI '23)*. Association for Computing Machinery, New York, NY, USA, 24–33. https://doi.org/10.1145/3568162.3576988

[127] Amal Nanavati, Patricia Alves-Oliveira, Tyler Schrenk, Ethan K Gordon, Maya Cakmak, and Siddhartha S Srinivasa. 2023. Design principles for robot-assisted feeding in social contexts. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 24–33.

[128] Jauwairia Nasir, Utku Norman, Wafa Johal, Jennifer K. Olsen, Sina Shahmoradi, and Pierre Dillenbourg. 2019. Robot Analytics: What Do Human-Robot Interaction Traces Tell Us About Learning?. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 1–7. https://doi.org/10.1109/RO-MAN46459.2019.8956465

[129] Heramb Nemlekar, Neel Dhanaraj, Angelos Guan, Satyandra K. Gupta, and Stefanos Nikolaidis. 2023. Transfer Learning of Human Preferences for Proactive Robot Assistance in Assembly Tasks. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction* (Stockholm, Sweden) *(HRI '23)*. Association for Computing Machinery, New York, NY, USA, 575–583. https://doi.org/10.1145/3568162.3576965

[130] Stefanos Nikolaidis, Yu Xiang Zhu, David Hsu, and Siddhartha Srinivasa. 2017. Human-robot mutual adaptation in shared autonomy. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. 294–302.

[131] Mary O'Keeffe, Paul Cullinane, John Hurley, Irene Leahy, Samantha Bunzli, Peter B O'Sullivan, and Kieran O'Sullivan. 2016. What influences patient-therapist interactions in musculoskeletal physical therapy? Qualitative systematic review and meta-synthesis. *Physical therapy* 96, 5 (2016), 609–622.

[132] Shayegan Omidshafiei, Dong-Ki Kim, Miao Liu, Gerald Tesauro, Matthew Riemer, Christopher Amato, Murray Campbell, and Jonathan P How. 2019. Learning to teach in cooperative multiagent reinforcement learning. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 33. 6128–6136.

[133] Patrizia Paci, Ilaria Tiddi, Daniel Preciado, and Kim Baraka. 2023. "Who'sa Good Robot?!" Designing Human-Robot Teaching Interactions Inspired by Dog Training. In *HHAI 2023: Augmenting Human Intellect*. IOS Press, 310–319.

[134] Carolina Parada. 2025. Gemini robotics on-device brings AI to local robotic devices. https://deepmind.google/discover/blog/gemini-robotics-on-device-brings-ai-to-local-robotic-devices/

[135] Priyam Parashar, Lindsay M Sanneman, Julie A Shah, and Henrik I Christensen. 2019. A taxonomy for characterizing modes of interactions in goal-driven, human-robot teams. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2213–2220.

[136] Andi Peng, Ilia Sucholutsky, Belinda Z. Li, Theodore R. Sumers, Thomas L. Griffiths, Jacob Andreas, and Julie A. Shah. 2024. Learning with Language-Guided State Abstractions. *ArXiv* abs/2402.18759 (2024). https://api.semanticscholar.org/CorpusID:266379261

[137] Karola Pitsch, Anna-Lisa Vollmer, Jannik Fritsch, Britta Wrede, Katharina Rohlfing, and Gerhard Sagerer. 2009. On the loop of action modification and the recipient's gaze in adult-child interaction. *Gesture and speech in interaction, Poznan, Poland* 24, 09 (2009), 2009.
[138] Dean A Pomerleau. 1988. Alvinn: An autonomous land vehicle in a neural network. *Advances in neural information processing systems* 1 (1988).
[139] Aaron Powers and Sara Kiesler. 2006. The advisor robot: tracing people's mental model from a robot's physical attributes. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. 218–225.
[140] Rob Price. 2016. Microsoft is deleting its AI chatbot's incredibly racist tweets. *Business Insider* (2016). https://web.archive.org/web/20190130071430/https://www.businessinsider.com/microsoft-deletes-racist-genocidal-tweets-from-ai-chatbot-tay-2016-3
[141] Zhifeng Qian, Mingyu You, Hongjun Zhou, Xuanhui Xu, and Bin He. 2023. Robot learning from human demonstrations with inconsistent contexts. *Robotics and Autonomous Systems* 166 (2023), 104466. https://doi.org/10.1016/j.robot.2023.104466
[142] Mattia Racca, Reuth Mirsky, Emmanuel Senft, Xuesu Xiao, Ifrah Idrees, Alap Kshirsagar, and Ravi Prakash. 2024. 3rd Workshop on Human-Interactive Robot Learning (HIRL). In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction* (Boulder, CO, USA) *(HRI '24)*. Association for Computing Machinery, New York, NY, USA, 1349–1351. https://doi.org/10.1145/3610978.3638160
[143] Karl C Roberts, W Timothy Brox, David S Jevsevar, and Kaitlyn Sevarino. 2015. Management of hip fractures in the elderly. *JAAOS-Journal of the American Academy of Orthopaedic Surgeons* 23, 2 (2015), 131–137.
[144] Katharina J Rohlfing, Philipp Cimiano, Ingrid Scharlau, Tobias Matzner, Heike M Buhl, Hendrik Buschmeier, Elena Esposito, Angela Grimminger, Barbara Hammer, Reinhold Häb-Umbach, et al. 2020. Explanation as a social practice: Toward a conceptual framework for the social design of AI systems. *IEEE Transactions on Cognitive and Developmental Systems* 13, 3 (2020), 717–728.
[145] Astrid Marieke Rosenthal-von der Pütten, David Sirkin, Anna Abrams, and Laura Platte. 2020. The Forgotten in HRI: Incidental Encounters with Robots in Public Spaces. 656–657. https://doi.org/10.1145/3371382.3374852
[146] SAE International. 2021. *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. Technical Report J3016_202104. SAE International. https://doi.org/10.4271/J3016_202104
[147] Akanksha Saran, Kush Desai, Mai Lee Chang, Rudolf Lioutikov, Andrea Thomaz, and Scott Niekum. 2022. Understanding Acoustic Patterns of Human Teachers Demonstrating Manipulation Tasks to Robots. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 9172–9179.
[148] Akanksha Saran, Srinjoy Majumdar, Elaine Schaertl Short, Andrea Thomaz, and Scott Niekum. 2018. Human gaze following for human-robot interaction. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 8615–8621.
[149] Akanksha Saran, Elaine Schaertl Short, Andrea Thomaz, and Scott Niekum. 2020. Understanding teacher gaze patterns for robot learning. In *Conference on Robot Learning*. PMLR, 1247–1258.
[150] Akanksha Saran, Ruohan Zhang, Elaine Schaertl Short, and Scott Niekum. 2021. Efficiently Guiding Imitation Learning Agents with Human Gaze. In *International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.
[151] William Saunders, Girish Sastry, Andreas Stuhlmueller, and Owain Evans. 2017. Trial without error: Towards safe reinforcement learning via human intervention. *arXiv preprint arXiv:1707.05173* (2017).
[152] Nienke Schrage-Prent, Daniel F Preciado Vanegas, and Kim Baraka. 2024. Interactive Robot Programming Inspired by Dog Training: An Exploratory Study. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. 965–969.
[153] Mariah L Schrum, Erin Hedlund-Botti, Nina Moorman, and Matthew C Gombolay. 2022. Mind meld: Personalized meta-learning for robot-centric imitation learning. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 157–165.
[154] Andrew D Selbst. 2020. Negligence and AI's human users. *BUL Rev.* 100 (2020), 1315.
[155] Mario Selvaggio, Marco Cognetti, Stefanos Nikolaidis, Serena Ivaldi, and Bruno Siciliano. 2021. Autonomy in physical human-robot interaction: A brief survey. *IEEE Robotics and Automation Letters* 6, 4 (2021), 7989–7996.
[156] Esmaeil Seraj, Kin Man Lee, Zulfiqar Zaidi, Qingyu Xiao, Zhaoxin Li, Arthur Nascimento, Sanne van Waveren, Pradyumna Tambwekar, Rohan Paleja, Devleena Das, and Matthew Gombolay. 2024. Interactive and Explainable Robot Learning: A Comprehensive Review. *Foundations and Trends® in Robotics* 12, 2-3 (2024), 75–349. https://doi.org/10.1561/2300000081
[157] Pratyusha Sharma, Balakumar Sundaralingam, Valts Blukis, Chris Paxton, Tucker Hermans, Antonio Torralba, Jacob Andreas, and Dieter Fox. 2022. Correcting robot plans with natural language feedback. In *Robotics: Science and Systems*.
[158] Christian Shelton. 2000. Balancing multiple sources of reward in reinforcement learning. *Advances in Neural Information Processing Systems* 13 (2000).
[159] Hua Shen, Tiffany Knearem, Reshmi Ghosh, Kenan Alkiek, Kundan Krishna, Yachuan Liu, Ziqiao Ma, Savvas Petridis, Yi-Hao Peng, Li Qiwei, Sushrita Rakshit, Chenglei Si, Yutong Xie, Jeffrey P. Bigham, Frank Bentley, Joyce Chai, Zachary Lipton, Qiaozhu Mei, Rada Mihalcea, Michael Terry, Diyi Yang, Meredith Ringel Morris, Paul Resnick, and David Jurgens. 2024. Towards Bidirectional Human-AI Alignment: A Systematic Review for Clarifications, Framework, and Future Directions. arXiv:2406.09264 [cs.HC] https://arxiv.org/abs/2406.09264
[160] Thomas B Sheridan. 2016. Human–robot interaction: status and challenges. *Human factors* 58, 4 (2016), 525–532.

[161] Einav Shpiro and Reuth Mirsky. 2024. Recognition and Identification of Intentional Blocking in Social Navigation. In *Proceedings of the 2024 International Symposium on Technological Advances in Human-Robot Interaction*. 101–110.

[162] Adrian Stoica. 1995. *Motion learning by robot apprentices: a fuzzy neural approach*. Ph. D. Dissertation. Victoria University of Technology.

[163] Theodore R. Sumers, Mark K. Ho, Robert D. Hawkins, and Thomas L. Griffiths. 2023. Show or tell? Exploring when (and why) teaching with language outperforms demonstration. *Cognition* 232 (2023), 105326. https://doi.org/10.1016/j.cognition.2022.105326

[164] Richard S Sutton and Andrew G Barto. 2018. *Reinforcement learning: An introduction*. MIT press.

[165] Bruce A Swett, Erin N Hahn, and Ashley J Llorens. 2021. Designing robots for the battlefield: State of the art. *Robotics, AI, and humanity: Science, ethics, and policy* (2021), 131–146.

[166] Aaquib Tabrez, Matthew B Luebbers, and Bradley Hayes. 2020. A survey of mental modeling techniques in human–robot teaming. *Current Robotics Reports* 1 (2020), 259–267.

[167] Susumu Tachi, Kiyoshi Komoriya, et al. 1984. Guide dog robot. *Autonomous mobile robots: Control, planning, and architecture* (1984), 360–367.

[168] Tadahiro Taniguchi, Lotfi El Hafi, Yoshinobu Hagiwara, Akira Taniguchi, Nobutaka Shimada, and Takanobu Nishiura. 2021. Semiotically adaptive cognition: toward the realization of remotely-operated service robots for the new normal symbiotic society. *Advanced Robotics* 35, 11 (2021), 664–674.

[169] Matthew E Taylor, Nicholas Nissen, Yuan Wang, and Neda Navidi. 2023. Improving reinforcement learning with human assistance: an argument for human subject studies with HIPPO Gym. *Neural Computing and Applications* 35, 32 (2023), 23429–23439.

[170] Matthew E Taylor, Halit Bener Suay, and Sonia Chernova. 2011. Integrating reinforcement learning with human demonstrations of varying ability. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*. 617–624.

[171] Emma Van Zoelen, Tina Mioch, Mani Tajaddini, Christian Fleiner, Stefani Tsaneva, Pietro Camin, Thiago S Gouvêa, Kim Baraka, Maaike HT De Boer, and Mark A Neerincx. 2023. Developing team design patterns for hybrid intelligence systems. In *HHAI 2023: Augmenting Human Intellect*. IOS Press, 3–16.

[172] Milos Vasic and Aude Billard. 2013. Safety issues in human-robot interactions. In *2013 ieee international conference on robotics and automation*. IEEE, 197–204.

[173] Jorn Verheggen and Kim Baraka. 2023. KRIS: A Novel Device for Kinesthetic Corrective Feedback during Robot Motion. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 5041–5047.

[174] Anna-Lisa Vollmer and Nikolas J Hemion. 2018. A user study on robot skill learning without a cost function: Optimization of dynamic movement primitives via naive user feedback. *Frontiers in Robotics and AI* 5 (2018), 77.

[175] Anna-Lisa Vollmer, Daniel Leidner, Michael Beetz, and Britta Wrede. 2023. From Interactive to Co-Constructive Task Learning. *arXiv preprint arXiv:2305.15535* (2023).

[176] Anna-Lisa Vollmer, Katrin Solveig Lohan, Kerstin Fischer, Yukie Nagai, Karola Pitsch, Jannik Fritsch, Katharina J Rohlfing, and Britta Wredek. 2009. People modify their tutoring behavior in robot-directed interaction for action learning. In *2009 IEEE 8th international conference on development and learning*. IEEE, 1–6.

[177] Anna-Lisa Vollmer, Manuel Mühlig, Jochen J Steil, Karola Pitsch, Jannik Fritsch, Katharina J Rohlfing, and Britta Wrede. 2014. Robots show us how to teach them: Feedback from robots shapes tutoring behavior during action learning. *PloS one* 9, 3 (2014), e91349.

[178] Anna-Lisa Vollmer and Lars Schillingmann. 2018. On studying human teaching behavior with robots: a review. *Review of Philosophy and Psychology* 9, 4 (2018), 863–903.

[179] Anna-Lisa Vollmer, Britta Wrede, Katharina J Rohlfing, and Pierre-Yves Oudeyer. 2016. Pragmatic frames for teaching and learning in human–robot interaction: Review and challenges. *Frontiers in neurorobotics* 10 (2016), 10.

[180] Thuy-Trang Vu, Shahram Khadivi, Mahsa Ghorbanali, Dinh Phung, and Gholamreza Haffari. 2024. Active continual learning: On balancing knowledge retention and learnability. In *Australasian Joint Conference on Artificial Intelligence*. Springer, 137–150.

[181] Weitian Wang, Rui Li, Yi Chen, Yi Sun, and Yunyi Jia. 2022. Predicting Human Intentions in Human–Robot Hand-Over Tasks Through Multimodal Learning. *IEEE Transactions on Automation Science and Engineering* 19, 3 (2022), 2339–2353. https://doi.org/10.1109/TASE.2021.3074873

[182] Xin Wang, Yudong Chen, and Wenwu Zhu. 2021. A survey on curriculum learning. *IEEE transactions on pattern analysis and machine intelligence* 44, 9 (2021), 4555–4576.

[183] David Whitney, Eric Rosen, James MacGlashan, Lawson LS Wong, and Stefanie Tellex. 2017. Reducing errors in object-fetching interactions through social feedback. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 1006–1013.

[184] Nils Wilde, Erdem Biyik, Dorsa Sadigh, and Stephen L. Smith. 2021. Learning Reward Functions from Scale Feedback. In *Proceedings of the 5th Conference on Robot Learning (CoRL)*.

[185] Katie Winkle, Séverin Lemaignan, Praminda Caleb-Solly, Paul Bremner, Ailie J Turton, and Ute Leonards. 2020. In-Situ Learning from a Domain Expert for Real World Socially Assistive Robot Deployment.. In *Robotics: Science and Systems*.

[186] Tengyang Xie, John Langford, Paul Mineiro, and Ida Momennejad. 2021. Interaction-grounded learning. In *International Conference on Machine Learning*. PMLR, 11414–11423.

[187] Tengyang Xie, Akanksha Saran, Dylan J Foster, Lekan Molu, Ida Momennejad, Nan Jiang, Paul Mineiro, and John Langford. 2022. Interaction-grounded learning with action-inclusive feedback. *Advances in Neural Information Processing Systems* 35 (2022), 12529–12541.

[188] Wang Xue-Song, Wang Rong-Rong, and Cheng Yu-Hu. 2023. Safe reinforcement learning: A survey. *Acta Automat-ica Sinica* 49, 9 (2023), 1813–1835.

[189] Yutao Yang, Jie Zhou, Junsong Li, Qianjun Pan, Bihao Zhan, Qin Chen, Xipeng Qiu, and Liang He. 2025. Reinforced Interactive Continual Learning via Real-time Noisy Human Feedback. *arXiv preprint arXiv:2505.09925* (2025).

[190] Zhaojing Yang, Miru Jun, Jeremy Tien, Stuart J. Russell, Anca Dragan, and Erdem Bıyık. 2024. Trajectory Improvement and Reward Learning from Comparative Language Feedback. In *Conference on Robot Learning (CoRL)*.

[191] Hamid Yeganeh. 2019. An analysis of emerging trends and transformations in global healthcare. *International Journal of Health Governance* 24, 2 (2019), 169–180.

[192] Hang Yu and Elaine Schaertl Short. 2021. Active Feedback Learning with Rich Feedback. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction* (Boulder, CO, USA) *(HRI '21 Companion)*. Association for Computing Machinery, New York, NY, USA, 430–433. https://doi.org/10.1145/3434074.3447207

[193] Ekim Yurtsever, Jacob Lambert, Alexander Carballo, and Kazuya Takeda. 2020. A Survey of Autonomous Driving: Common Practices and Emerging Technologies. *IEEE Access* 8 (2020), 58443–58469. https://doi.org/10.1109/access.2020.2983149

[194] Frauke Zeller and Lauren Dwyer. 2022. Systems of collaboration: challenges and solutions for interdisciplinary research in AI and social robotics. *Discover Artificial Intelligence* 2, 1 (2022), 12.

[195] Ruohan Zhang, Akanksha Saran, Bo Liu, Yifeng Zhu, Sihang Guo, Scott Niekum, Dana Ballard, and Mary Hayhoe. 2020. Human gaze assisted artificial intelligence: A review. In *IJCAI: Proceedings of the Conference*, Vol. 2020. NIH Public Access, 4951.

[196] Tianhao Zhang, Zoe McCarthy, Owen Jow, Dennis Lee, Ken Goldberg, and P. Abbeel. 2017. Deep Imitation Learning for Complex Manipulation Tasks from Virtual Reality Teleoperation. In *IEEE International Conference on Robotics and Automation*. https://api.semanticscholar.org/CorpusID:3720790

[197] Gaoyue Zhou, Victoria Dean, Mohan Kumar Srirama, Aravind Rajeswaran, Jyothish Pari, Kyle Hatch, Aryan Jain, Tianhe Yu, Pieter Abbeel, Lerrel Pinto, Chelsea Finn, and Abhinav Gupta. 2023. Train Offline, Test Online: A Real Robot Learning Benchmark. arXiv:2306.00942 [cs.RO] https://arxiv.org/abs/2306.00942

[198] Yilun Zhou, Serena Booth, Nadia Figueroa, and Julie Shah. 2022. Rocus: Robot controller understanding via sampling. In *Conference on Robot Learning*. PMLR, 850–860.