

A Scoping Review and Experimental Study on Reinforcement Learning from Human Feedback for Human-Robot Collaboration

Alexandra Coroiu^{a,*}, Andrea Vogt^a, Viktor Werbilo^b, Andreas Poppele^a, Johann Christensen^b and Sven Hallerbach^b

^aDLR Institute for AI Safety and Security, Wilhelm-Runge-Straße 10, 89081, Ulm, Germany

^bDLR Institute for AI Safety and Security, Rathausallee 12, 53757, Sankt Augustin, Germany

ARTICLE INFO

Keywords:

Human-Robot Collaboration
Human-in-the-Loop
Reinforcement Learning from Human Feedback
AI Safety
Human-Robot Proxemics

Abstract

Human-Robot Collaboration (HRC) can facilitate mass customisation in Industry 4.0, with Reinforcement Learning from Human Feedback (RLHF) representing a promising approach for developing safe AI-based robots. Practical challenges remain regarding safety during AI development, human feedback quality, and bidirectional human-robot adaptation. We conducted a scoping review of RLHF in HRC systems, mapping methods that address these challenges. Following PRISMA guidelines, we screened 199 records and included 20 peer-reviewed publications (2020-2025) spanning multiple HRC domains. To our knowledge, this is the first review focused on the bidirectional, closed-loop design of RLHF. Our review found multiple feedback modalities enabling data collection in various feedback formats. Collected data can be integrated at different stages of AI training, resulting in a multi-step development process. Pilot experiments are commonly used to evaluate HRC systems based on both human and robot metrics. To empirically test a key gap identified in the review, we conducted a between-subjects VR experiment comparing system- and user-initiated feedback on robot proxemic behaviour for safe navigation. Using Bayesian models, we analysed the relation between the collected feedback and safety metrics: psychological safety (post-experiment questionnaire) and physical safety (inverse time-to-collision). Results show that user-initiated feedback captures perceived safety better than system-initiated feedback, indicating that feedback timing directly affects feedback quality. Our review and experiment findings show that RLHF relies on appropriate feedback methods to ensure AI safety in HRC, and future RLHF research should prioritise realistic HRC experiments evaluating the effects of feedback collection methods on relevant human and robot metrics.

1. Introduction

1.1. Background

Industry 4.0 has seen a surge of increasingly autonomous robots working alongside humans across various sectors, particularly in manufacturing and logistics applications. Industrial environments are transitioning from traditional isolated workcells towards Human-Robot Collaboration (HRC), where humans and robots work in shared spaces. These HRC systems combine the efficiency of automation with human flexibility, enabling robots to operate in dynamic environments [26, 48]. Such collaborative robots can be deployed for a wide range of industrial tasks: mobile robots enable navigation and transport on the work floor, and robotic arms enable variable pick-and-place operations for manipulation and assembly [28, 4]. By rapidly automating a wide range of repetitive tasks, HRC systems facilitate working with smaller, customised product batches, and allow humans to focus on higher-level decision-making. This combination of efficiency and adaptability can help address the growing need for mass customization [12, 4], offering a competitive advantage over traditional automation.

1.2. Motivation

However, without barriers to separate humans from robots, safety risks pose a significant challenge for the implementation of HRC systems in industrial settings [12]. Specifically, collisions and non-functional contacts between humans and robots represent sources of injury in the shared workspace [14, 43]. This risk is particularly critical

*Corresponding author

 alexandra.coroiu@dlr.de (A. Coroiu)

ORCID(s): 0009-0005-6643-7306 (A. Coroiu); 0009-0008-9203-2312 (V. Werbilo); 0000-0001-9871-122X (J. Christensen); 0009-0007-9163-0567 (S. Hallerbach)

for industrial robots capable of generating high forces. Safety strategies focus on preventing hazardous contact (pre-collision strategies) and limiting the impact of the contact (post-collision strategies) [18, 48]. A series of ISO standards apply to HRC and provide guidelines for the development of safe collaborative robots. ISO 10218-2 [22] prescribes safety requirements based on the interaction characteristics between the human and the robot. Specifically, collaborative robots require inherent safety features, such as power and force limiting. The main practical challenge in developing safety methods for HRC concerns ensuring ISO compliance while preserving high system performance [43]. Therefore, balancing safety and performance becomes an optimisation problem.

Besides immediate hazards like collisions, human factors represent an important latent safety concern in HRC. Human-centred design of collaborative robots should account for physical and psychological human capabilities, not only to reduce hazard risks and strain for workers [7], but also to ensure optimal system performance [13, 34, 43]. The most notable psychological factors for HRC are trust, cognitive workload, stress, and perceived safety [27, 19]. Previous research [12] identifies cognitive factors as the “most promising research challenge” for developing safe collaborative robots. The mental processes involved in HRC (i.e., perception, memory, reasoning, motor response) [7, 3] need to be carefully considered in order to neither overload nor underload workers when collaborating with robots on shared tasks [19]. These human factors are particularly relevant for ensuring safety and performance over prolonged periods of interaction with collaborative robots.

Currently, there is an ongoing effort to integrate human factors in the development process of safe collaborative robots. The main goals are to design appropriate interaction modalities for HRC [26, 28, 18], and to develop adaptable robots that can accommodate the needs and limitations of workers [13, 26, 28, 7, 14]. Existing industrial collaborative robots typically rely only on predefined safety rules, lacking the ability to process real-time human data and dynamically adapt. Processing high volumes of human data remains an open challenge, due to the difficulty in interpreting complex multimodal data and extracting relevant information about human states. Solving this challenge requires new technical solutions capable of combining a variety of sensor data and computing safe robot actions accordingly. Therefore, the development of an adaptive collaborative robot solution relies on measurement methods developed for real-time monitoring of human factors in the real workspace [27, 19, 28], and advanced artificial intelligence (AI) methods that can improve robot understanding of its environment, including human agents [48, 3]. Additionally, these monitoring and AI methods must comply with existing legal and ethical frameworks, such as the the European AI Act [10] and General Data Protection Regulation [9].

1.3. Problem Statement

The field of HRC requires new AI engineering methods to develop safe collaborative robots that can adapt to complex human environments. Human-in-the-loop (HITL) approaches constitute one potential solution that helps ensure safety-by-design by directly integrating human input in the AI development process [48]. A promising AI method is reinforcement learning, which allows robots to learn from continuous interaction with their environment, including human coworkers [13, 12]. The use of HITL methods for reinforcement learning relies on diverse interaction modalities (e.g., vision, motion, language) to collect human data that is used as feedback for RL algorithms [37, 8]. These methods can support two main robot learning techniques: imitation learning and reinforcement learning from human feedback (RLHF) [8]. Imitation learning allows a robot to quickly learn a new task from human demonstrations (e.g., using recorded video samples or hand-guiding data), facilitating human-to-robot skill transfer. RLHF helps to iteratively refine a learnt task based on human feedback (e.g., through pairwise preferences or ratings), without the need to design a rule-based reward model. RLHF can be especially useful in complex environments where robots need to adapt to workers’ subjective evaluations of safety. Therefore, RLHF can contribute to increased HRC safety by integrating human factors into AI training.

While RLHF represents a promising approach for developing adaptive collaborative robots, there are still practical challenges that need to be overcome before it can be confidently applied in real-world HRC systems. Firstly, an important aspect of developing HRC with RLHF methods is ensuring human safety during the development of the AI model. For this purpose, simulations have been identified as a safe and flexible tool to collect human data for training and testing the AI model before the AI-based robot is deployed in physical environments [13]. However, while simulations are useful in the early stages of development, further testing is necessary for validating the AI-based robot within real-world constraints. Secondly, all RLHF approaches rely on high-quality human data, which is limited by inherent uncertainty (e.g., unpredictable behaviours and human errors), as well as the difficulty of collecting enough data. To address this, human digital twins can be used in simulation to model human behaviour for training the AI-based robot [13, 12]. Nevertheless, collecting real human data remains essential for customising and validating the safety and

performance of these methods in HRC applications. Lastly, RLHF approaches are characterised by the bidirectional adaptation between the human and the robot. As the robot adapts to human feedback, the human responds to the robot. This creates a non-stationary learning environment, where iterative feedback loops lead to reciprocal learning [35]. Therefore, in order to ensure optimal HRC system safety and performance, a closed-loop design is needed to account for both directions of the HITL approach: human-to-robot and robot-to-human [19]. Our work contributes towards closed-loop RLHF design with a scoping literature review on state-of-the-art methods, restricted to bidirectional studies, and a VR experiment comparing the effects of two feedback collection conditions on human factors and feedback quality.

2. Literature Review

2.1. Research Questions

As industry shifts towards higher levels of autonomy and mass customisation, robots face increasing performance demands both in terms of efficiency and flexibility. HRC systems can help meet these performance needs, but also introduce new safety concerns in the workspace. While classic automation uses rigid constraints to ensure safety, collaborative robots should dynamically adapt to human workers to balance the safety-performance trade-off (see Figure 1). HITL approaches represent a promising solution for developing such adaptive collaborative robots. In

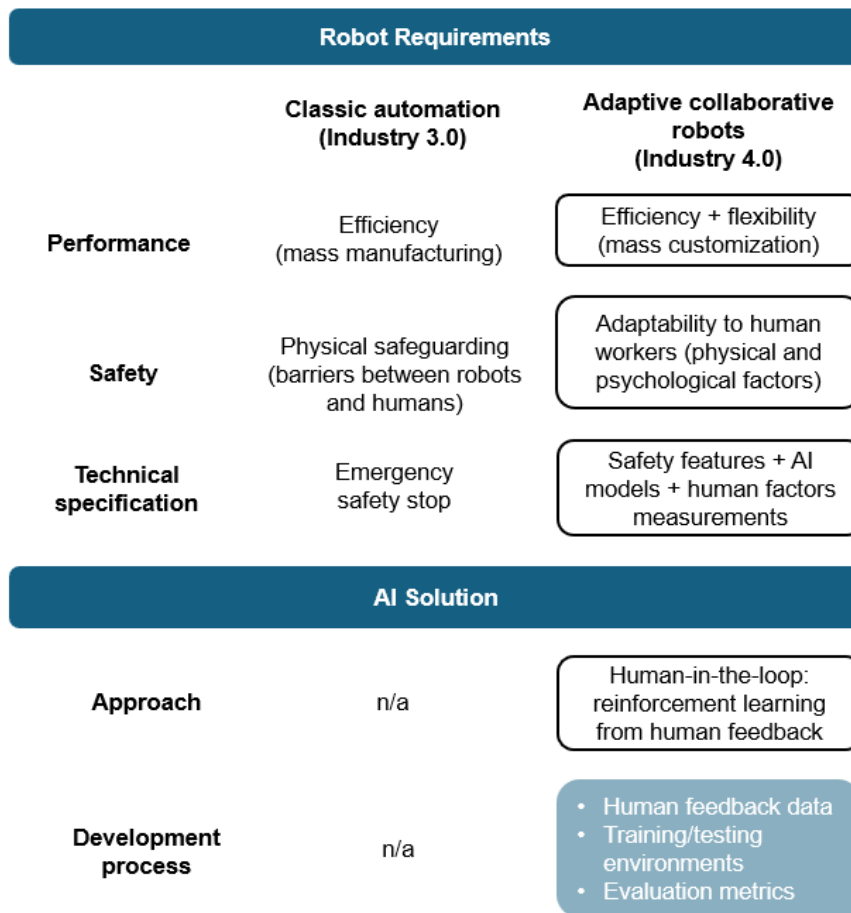


Figure 1: Requirements for adaptive collaborative robots compared to classic automation, and RLHF as a potential solution

particular, RLHF methods provide a promising paradigm to directly integrate human data into robot learning. However, the current development of RLHF methods for collaborative robots is still limited by methodological challenges regarding safety during AI development, the complexity of human feedback data, and the effects of human-robot co-adaptation. Addressing these challenges is crucial for developing RLHF applications that can be safely deployed in

real-world HRC systems. Therefore, this scoping literature review aims to investigate the state-of-the-art methods used for RLHF development in HRC systems, focusing on solutions that account for the bidirectional adaptation between the human and AI-based robot. Accordingly, we structured this review around the following research questions (RQs):

- **RQ1:** What HRC systems employ RLHF models for robot task learning?
- **RQ2:** What feedback methods are used to collect human data for RLHF models in HRC systems?
- **RQ3:** What AI development methods are used for training RLHF models in HRC systems?
- **RQ4:** What metrics are used for evaluating the effects between the human and the RLHF-trained robot in HRC systems?

2.2. Methods

To conduct this scoping literature review we searched for relevant publications using the following search string: (robot* OR cobot*) AND ("human in the loop" OR "human feedback" OR "user feedback" OR "human reward" OR "user reward" OR "human preference*" OR "user preference*") AND ("reinforcement learning" OR rl OR "preference learning" OR rlhf). We queried the Scopus database for journal articles and conference papers published between 2020 and 2025, reflecting the emergence of RLHF methods in robotics. The search resulted in 199 records. We screened the title and abstract of these records to select reports that match our research scope (see Figure 2). Each report was screened by two reviewers independently. We used the following relevance criteria:

- RLHF is used for learning a robot specific task (e.g., navigation, manipulation); excluding general AI tasks (e.g., language processing, visual processing).
- Humans directly participate in robot learning, beyond initial AI algorithm development (e.g., not limited to designing the reward model).
- The AI-based robot integrates human data as feedback for training the RLHF algorithm.
- Experiments involve real human participants; excluding the use of purely synthetic/simulated human data.
- Metrics are evaluated for both directions of the HITL: human-to-robot effects and the robot-to-human effects.

Consistent with our research questions, we recorded the following data items from each included report: application domain, specific robot task, human data measurement, feedback format, experiment environment, AI development methods, human metrics (cognitive factors and feedback behaviour), AI-based robot metrics.

2.3. Results

2.3.1. RLHF Applications

The use of HITL approaches spans across a wide range of robotic applications. RLHF is used for training robots to fulfil roles in industry [2, 42, 46], assistance and care [17, 25, 29, 32, 39, 44, 45] or emergency and disaster situations [20]. Depending on the type of robot, different tasks can be trained with RLHF. Mobile robots can be fine-tuned for navigating in busy environments where robot and human paths can easily cross [23, 31, 30, 32, 38, 42, 44]. Robotic arms can be trained to perform complex tasks that require fine movements in proximity to humans. Human feedback is then used to fine-tune the RLHF policy for robotic arm navigation in a 3D space [2, 16, 24, 46, 47, 49] or for the control of the robotic grippers [17, 25, 45]. These RLHF use cases can highly benefit from close collaboration between humans and robots to enhance capabilities for flexible tasks in uncertain environments.

2.3.2. Human Feedback Methods

There is a wide variety of proposed solutions for collecting human feedback to be used in RLHF algorithms (see Figure 3). These solutions encompass both direct and indirect feedback measurements, each requiring different modalities and different levels of data pre-processing. One of the most common feedback methods involves using a graphical user interface to collect subjective evaluations of the interaction with the robot [16, 17, 31, 47]. Language-based feedback represents another, more complex form of subjective measurement [45]. This method requires an intermediate natural language processing step, e.g., to evaluate the message or emotional valence in verbal or

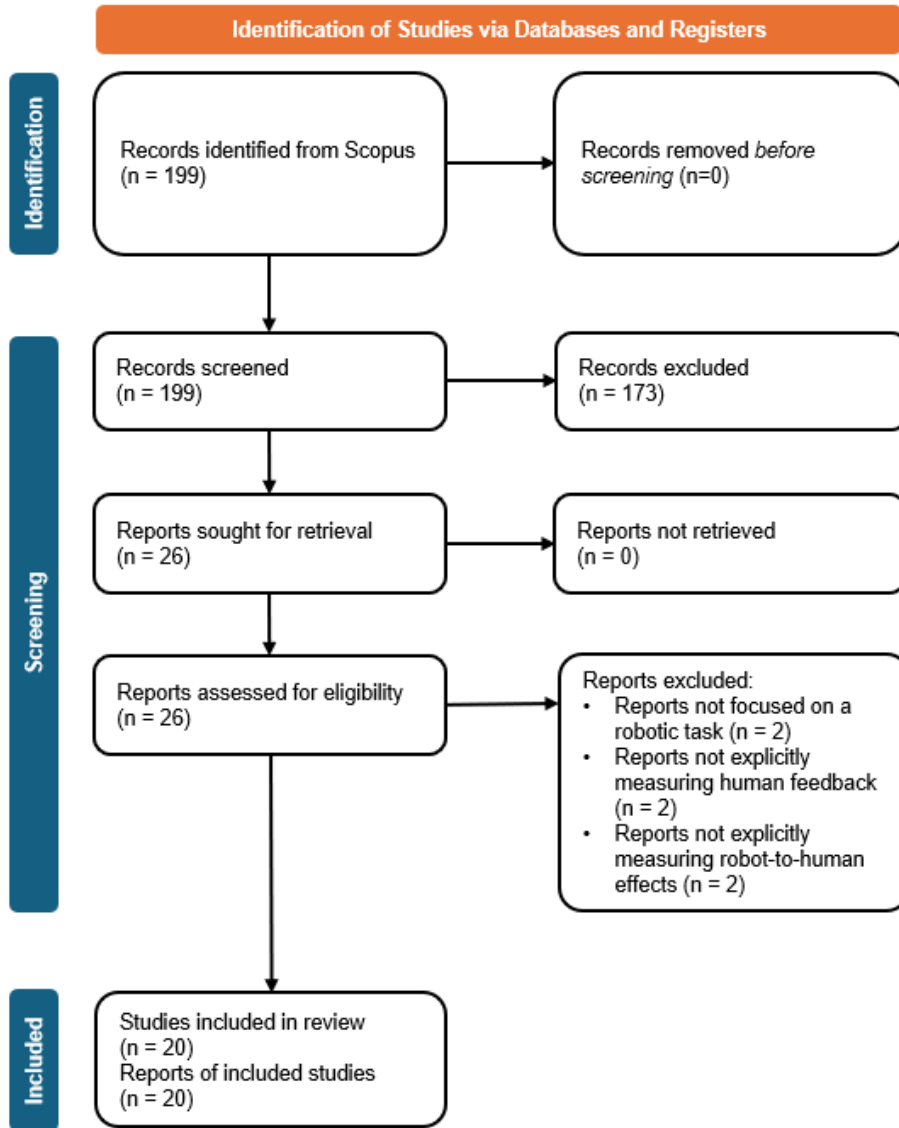


Figure 2: PRISMA flow diagram [36] of the Scopus search (2020-2025), 199 records screened, 20 included

spoken text. Indirect, objective feedback measurements also require more advanced data processing to evaluate complex input from user behaviour [46] (captured through cameras and sensors) or physiological signals like brain-computer interfaces [24]. Interpreting these signals relies on statistical or AI-based analysis to extract meaningful feedback information. Lastly, multi-modal setups can be used to collect robust feedback from combinations of various measurements [29, 32].

The types of feedback collected through the different measurement methods can vary significantly (see Figure 3). First, simple feedback formats consist of subjective preferences [5, 23, 30, 39, 44, 49, 16] and binary or scalar ratings [17, 29, 31, 32, 47, 49, 16] which can be more easily collected through GUIs. However, a key consideration in presenting preferences to users is the need to pre-select relevant pairs or groups of items for comparison [49, 5]. Second, corrections can be used to provide negative feedback for undesirable robot behaviours either as simple binary input or as a scalar value as a measure of intensity [25, 42, 45, 20]. Lastly, feedback from demonstrations [2] represents a more complex format that requires more advanced user behaviour input (e.g., from cameras, force-feedback sensors) and data processing to compare the desired actions against the robot motion and compute feedback values from this

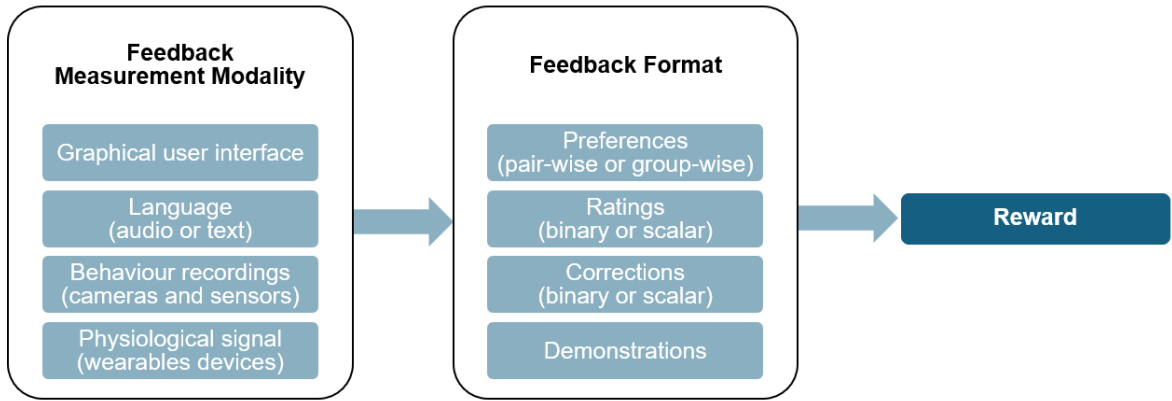


Figure 3: Feedback methods in RLHF for robotic applications: measurement modalities and formats

difference. As with the different feedback measurements, feedback formats can also be combined to acquire more robust feedback data [46, 38, 24].

Regardless of feedback measurement or format, collected data ultimately needs to be converted into a reward score that can be used in the RL algorithm. The key challenge with human feedback is the inherent uncertainty within human data. While some noise is expected, the design and implementation of feedback methods can influence the noise levels in multiple ways [31, 46, 49]. For example, [47] compares the noise levels between binary and scalar ratings. Similarly, when it comes to preferences, the specific options presented can also highly impact human feedback [5]. Furthermore, complex methods that require extensive data pre-processing make it harder to isolate the specific factors that shape the final reward score distribution. Another challenge with human feedback is collecting sufficiently large and heterogeneous datasets to create representative distributions that are not biased towards a certain type of feedback response [2, 30].

2.3.3. AI Development Methods

Developing RLHF algorithms for HRC generally requires a multi-step process to achieve the desired AI-based robot learning performance (see Figure 4). Using pre-training for the RL task before collecting human feedback [24, 39, 44, 45, 46] ensures a functioning initial policy that can then be fine-tuned and customised according to user input through subsequent learning sessions. RL algorithms can be pre-trained using trial-and-error approaches in simulation, allowing the robot to learn independently, or by incorporating synthetic human feedback to introduce human evaluation early in the learning process [20, 31, 30, 45]. These methods enable training on a large quantity of trials and data that would otherwise be difficult, if not impossible, to collect through real interaction with the environment and users. For modelling both the synthetic and the real human feedback, Bayesian methods can be employed to create distributions of the human data and account for feedback uncertainty [31, 46, 49].

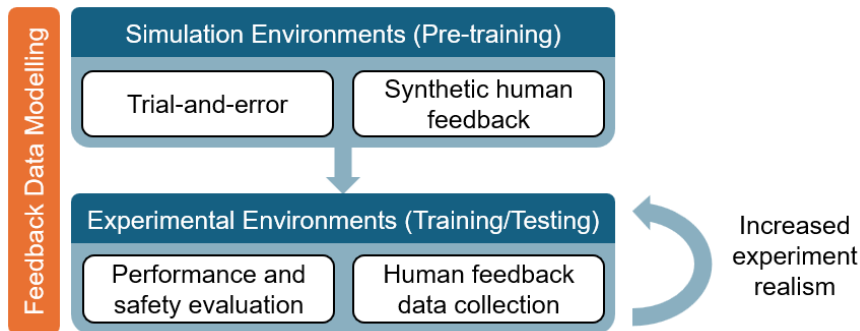


Figure 4: RLHF development methodology for robotic applications with iterative training stages

Experiments for collecting human feedback can be set up in various environments that provide different levels of realism. On the one hand, 2D or 3D computer simulations offer an easy-to-use framework for displaying AI-based robot behaviour and collecting human feedback [20, 31, 30, 38, 47, 23]. On the other hand, realistic (mock-up) environments require more space and functional physical assets [23, 32, 44, 45, 47], but offer greater realism and the opportunity to collect more representative feedback data. Additionally, an important consideration for meaningful experiments in realistic environments is the participant count and background. Pilot studies are carried out with a relatively small number of participants (< 25) [20, 24, 25, 29, 32, 42, 44, 45, 46, 49], while fewer large-scale studies exist [31, 30, 38, 39, 47]. Furthermore, although these large-scale studies include more participants, they are not representative of the intended target groups.

2.3.4. Evaluation Metrics

The human-to-robot direction within HTL concerns the effects on the AI-based robot and collaboration outcomes (see Figure 5). Within the context of RLHF for HRC applications, AI-based robot performance can be evaluated on two levels. Firstly, AI learning properties determine the suitability of the RL model for the intended task and target users. AI learning can be quantified using quality metrics such as learning efficiency [17, 20, 23, 30, 31, 44, 45], often assessed alongside adaptability [39, 32, 2, 24, 25], or explainability (especially in terms of predictability) [16]. These metrics represent AI characteristics necessary for enabling the robot to perform highly flexible tasks. Secondly, the robot's resulting behaviour directly impacts overall HRC system outcomes. While most studies focus only on robot task performance metrics (e.g., efficiency in task execution) [16, 25, 31, 42, 46, 47, 49], others specifically focus on robot actions that impact system safety (e.g., collisions in navigation) [2, 44, 17, 20].

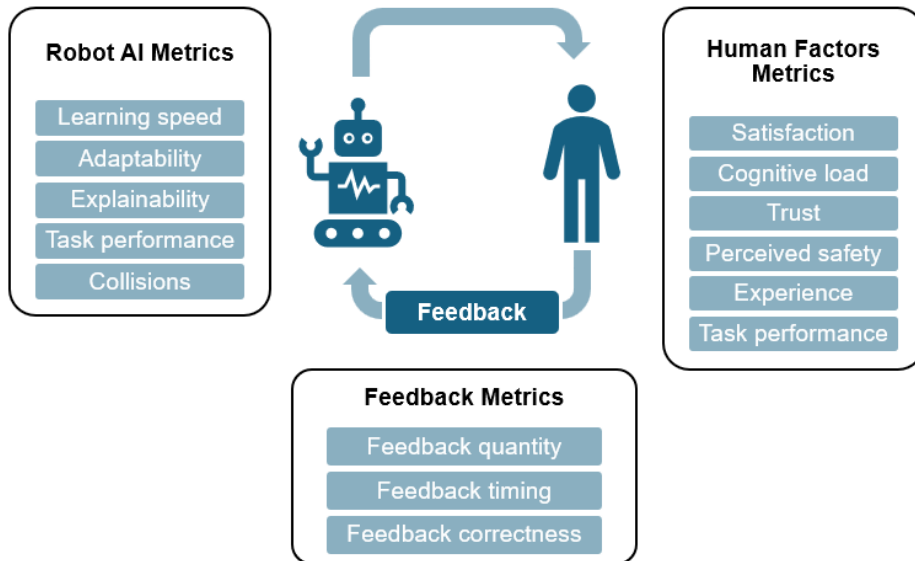


Figure 5: HRC safety and performance for the closed-loop RLHF design: AI-based robot, human and feedback metrics

The robot-to-human direction within HTL concerns the effects on human factors and collected feedback (see Figure 5). Within the context of RLHF for HRC applications, two main aspects must be considered for evaluating human factors. Firstly, psychological factors play a central role in evaluating the human-robot dynamic. The most frequently studied factor is satisfaction [5, 25, 32, 44, 46] (typically evaluated using the System Usability Scale questionnaire [6]) often assessed alongside cognitive load [2, 16, 38, 42, 49] (typically evaluated using the standard NASA TLX questionnaire [15]). Trust [39] and perceived safety [45] are only measured in one study each. Additionally, a small number of studies account for user experience and learning effects [17, 47, 39, 42]. Secondly, behavioural aspects encompass both the primary task performance [29, 47] and the provision of human feedback, which can be considered a secondary task in the context of HRC. The most studied human feedback metric is the total number of collected data points [2, 17, 20, 23, 24, 31, 42, 46], which is highly relevant for developing a robust reward function in the RL process. Additionally, it can also be relevant to evaluate feedback timing [49] or feedback correctness [24, 38].

2.4. Discussion

2.4.1. Main Findings

For RQ1, the results reveal that a wide range of HRC applications employ RLHF for robot training: from navigation AI for disaster robots to manipulation AI tasks for assistive robot arms. The common ground among all these applications is the need to operate despite high uncertainty, either because the area is unknown (e.g., in disaster sites), users are diverse (e.g., in assistive care), or tasks change frequently (e.g., mass customisation in industry). These applications often also require very close collaboration with humans: assistive robots work in close physical contact with humans, while emergency or industrial robots need to complement human efforts. Therefore, any HRC system that requires flexibility can benefit from RLHF approaches to optimise the AI-based robot using human input.

For RQ2, the results reveal a wide range of state-of-the-art technical solutions for integrating human feedback in RL algorithms. The research appears to be in an exploratory phase where various interaction tools (classic GUIs, NLP, cameras and sensors, or even BCIs) are used to collect feedback in various formats (e.g., ratings, preferences, corrections, or demonstrations). There are no universally preferred methods, as the feedback requirements can vary by application use case or AI development stage. A common issue arising across feedback methods is data uncertainty, stemming from noise and bias in human responses. Feedback strategies are often focused on maximising feedback input counts, maximising information gain from input data points, or minimising uncertainty.

For RQ3, the results show that human feedback alone cannot be directly used to produce desirable RL model performance. The AI development methodology often requires pre-training in simulation and synthetic human data to train the robot initially. Additionally, when training with real human feedback data, AI methods can benefit from Bayesian modelling to represent noisy feedback distributions. Furthermore, the collection of real human feedback currently depends on low-fidelity prototypes of the intended HRC system (computer simulations or mock-up environments), which are often used to run pilot studies. These findings suggest that the development of RLHF methods requires iterative steps to train and test the AI, from low-level simulations in the first steps to realistic environments with real users in the later steps.

For RQ4, several human and AI-based robot metrics have been identified for HRC systems. For the AI-based robot, metrics which reflect flexibility are most relevant: how well and how fast the AI model can learn to adapt to new users, tasks or environments. These AI metrics shape the overall robot behaviour, subsequently affecting HRC system outcomes. When it comes to human behaviour, the research is predominantly focused on the quantity of provided feedback, which is highly relevant within the context of RLHF. In terms of psychological factors, existing research is mainly focused on satisfaction, while the only safety-relevant metric commonly evaluated is cognitive load. This highlights the early exploratory nature of RLHF research in HRC, which currently focuses more on AI development rather than overall system safety and performance.

2.4.2. Limitations

Two main limitations can be noted. First, this scoping literature review sits at the intersection of two related, but distinct research fields: Human-Robot Interaction and AI development with HITL methods. Constructing a search string that covers relevant RLHF methods applicable to HRC systems proved to be difficult. This is an emerging field, so standard terminology is not well established yet. HRC literature focuses on human factors and robotics using vocabulary specific to the psychology and engineering domains, whereas HITL-RLHF literature focuses on AI development using technical vocabulary specific to AI model optimisation and evaluation. Our search string captures keywords specific to our research scope, but may have omitted relevant reports that use other terminology. Furthermore, we restricted our search to the Scopus database due to its multidisciplinary coverage of both HRC and AI literature, which is well suited for this emerging field, but publications indexed only in other databases may have been missed. Second, we restricted our selection criteria to studies that included both directions of the HITL approach, assuming that a closed-loop design requires specific methods to capture the bidirectional adaptation between human and AI-based robots. Nevertheless, unidirectional studies might also contain applicable feedback and AI development methods, but these methods should be validated for applicability within bidirectional studies.

2.4.3. Future Research

RLHF research for HRC can benefit from focusing solutions on target applications. When it comes to feedback formats and collection methods, tailored solutions should be developed for specific use cases [27, 28]. For example, an industrial environment with specific HRC tasks might impose various constraints on the RLHF methodology, such as environment sound and light conditions, monitoring capabilities, available interaction modalities, time and

cognitive limitations, procedural standards, resource availability. These constraints limit the viability of different feedback methods and should be considered early in the design and development of RLHF strategies. Therefore, future research should attempt to integrate and test in realistic set-ups with target user groups.

Another aspect of bringing RLHF research closer to real-world applications of HRC systems is focusing on system safety metrics. Currently, there is limited attention given to human factors that impact HRC safety: trust, perceived safety, and stress [27, 19]. These human factors should be evaluated in realistic environments, under different experimental conditions, and with varied feedback strategies. Particularly, cognitive load can be further explored, as the secondary feedback task can affect both safety and the overall outcome of the primary collaborative task. All these psychological factors are relevant not only for avoiding hazards in a shared human-robot workspace, but also for ensuring optimal HRC system performance. Furthermore, safety metrics should be evaluated in realistic experiments that best reflect the target HRC application.

Finally, the closed-loop design [19] of RLHF methods would benefit from evaluating the impact of human-robot collaboration on the collected feedback. Human feedback behaviour can be influenced by both psychological factors and real-world system properties (e.g., environmental conditions, task requirements, available feedback methods), resulting in varying levels of noise or bias. Similarly, the AI-based robot behavioural policy, defined by prior training phases (e.g., pre-training on synthetic human data), ultimately shapes the interaction with the human and the characteristics of the provided feedback. These factors should be accounted for when modelling feedback as rewards for RLHF in order to reduce uncertainty in the collected human data. A well-structured feedback strategy should capture multiple dimensions of the collaboration, including psychological and physical safety.

3. Experiment

3.1. Research Questions

RLHF applications in HRC systems rely on human data for training high-performing and safe collaborative robots. Collected feedback is characterised by both volume and quality of data. First, a high volume of feedback input is beneficial for model training, but increases effort for the human collaborator. Therefore, RLHF applications aim to have the robot learn as fast as possible with as little feedback as possible [2, 17, 49, 23, 31, 20, 24]. Feedback volume can be measured not only in terms of counts, but also as time spent per feedback query [23]. Additionally, human effort can be directly quantified as cognitive load [2, 20, 38, 49]. Second, feedback quality can be broadly defined in terms of information value for the RLHF reward model. Two key considerations that impact feedback quality are correctness [24, 38] and timing [49]. Information value per feedback point can be maximised by designing relevant feedback queries [20, 31].

The ideal feedback strategy for RLHF requires optimising the trade-off between human effort and information value obtained from feedback interactions to ensure practicality for real-world HRC systems. Binary ratings represent a fast, low-effort feedback strategy, but the lower resolution of binary signal requires more feedback queries to train the RL policy [21]. The demand for higher feedback volume could be addressed through system-initiated feedback queries at a predetermined sampling rate. In this study, we want to explore the difference between system-initiated and user-initiated binary feedback. To address both directions of a closed-loop design for HRC [19], we want to evaluate both the effects of feedback condition on the human and the effects of feedback on RLHF robot learning. Therefore, we conducted a VR experiment to answer the following research questions:

- **RQ5:** How do system-initiated versus user-initiated binary feedback strategies compare in terms of information value?
- **RQ6:** How do system-initiated versus user-initiated binary feedback strategies compare in terms of human effort?

3.2. Methods

3.2.1. Application Domain

The experiment focuses on comparing feedback conditions for RLHF methods applied to robot navigation tasks in HRC systems. Previous research [31, 32] successfully employed RLHF to integrate human proxemic zones into robot navigation and improve human factors outcomes such as perceived safety. This is particularly relevant for industrial applications where repeated proxemic violations can decrease psychological safety for workers, which in turn degrades overall HRC performance and safety [34]. Therefore, safe robot navigation in HRC systems should consider both physical and psychological safety. Furthermore, proxemic zones are not static, but dynamically change based on robot

navigation speed and interaction context [33]. RLHF solutions can more accurately capture these dynamic proxemic zones based on human input, instead of relying on AI navigation models with static proxemic rules, resulting in better safety-performance trade-offs.

3.2.2. Environment and Task

The experiment was conducted in a VR environment developed in Unreal Engine (UE) version 5.4, and displayed on an HTC Vive headset to participants. The virtual environment simulated an industrial workshop area featuring a sorting station and a conveyor belt (see Figure 6). The environment featured three humanoid robots: one co-working robot and two additional robots operating in the background. The co-working robot was responsible for delivering unsorted crates to the workstation, where the human sorts the crates by object colour using the HTC controller to interact with the virtual objects. The co-working robot is then responsible for transporting the sorted crates to the conveyor belt and storing empty crates away. As the co-working robot navigates the workshop area, its path crosses with the other two robots. Robot navigation was controlled with the UE navigation AI, and collisions were avoided using safety stops when trespassing the 1m radius around other actors.

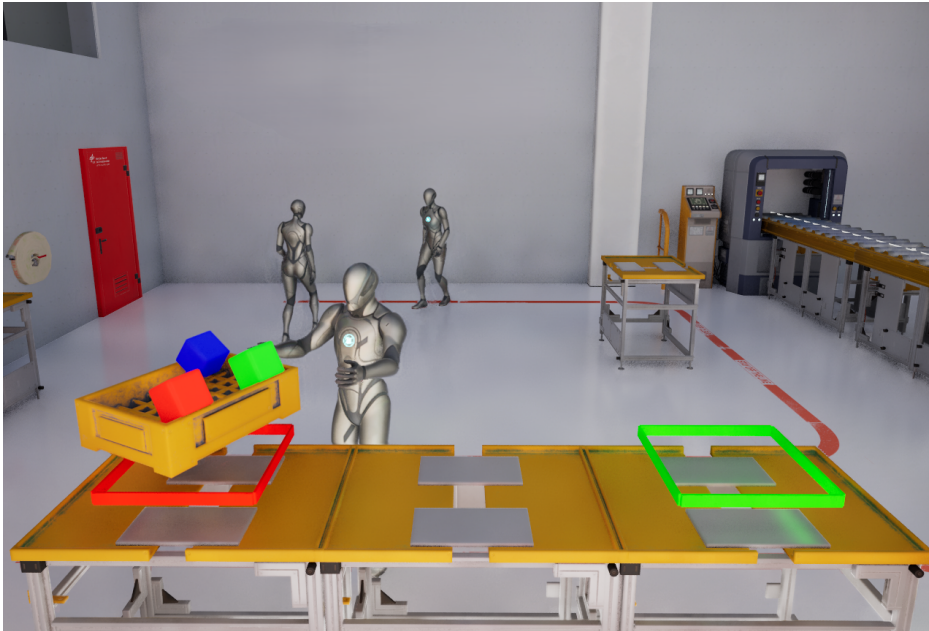


Figure 6: First-person perspective of the VR environment, including the human workstation, co-working robot and background robots

The human has to provide feedback on the proxemic behaviour of the co-working robot. A pop-up feedback interface displays the following question: “Did the robot keep sufficient distance to other actors?” (see Figure 7). The human provides binary feedback by choosing between the thumbs-up and thumbs-down buttons. Providing feedback represents a secondary task to the primary task of sorting crates. The two feedback conditions are designed to facilitate data collection while maintaining a smooth workflow. In the user-initiated condition, the participant is free to open the feedback pop-up interface whenever they want with the click of a button on the controller. In contrast, in the system-initiated condition, the participant is prompted to provide feedback for every third crate transported to the conveyor belt.

3.2.3. Participants and Procedures

Given the exploratory nature of this experiment, we used a convenience sampling strategy. Participants were recruited internally at the two institute locations in Germany. Written informed consent was obtained from all participants, including explanations on data protection and potential VR side effects. We collected demographic data and VR familiarity (6-item questionnaire) in terms of three dimensions: VR expertise, VR frequency, and VR confidence. A total of 48 participants completed the experiment, but four were excluded for not providing any feedback

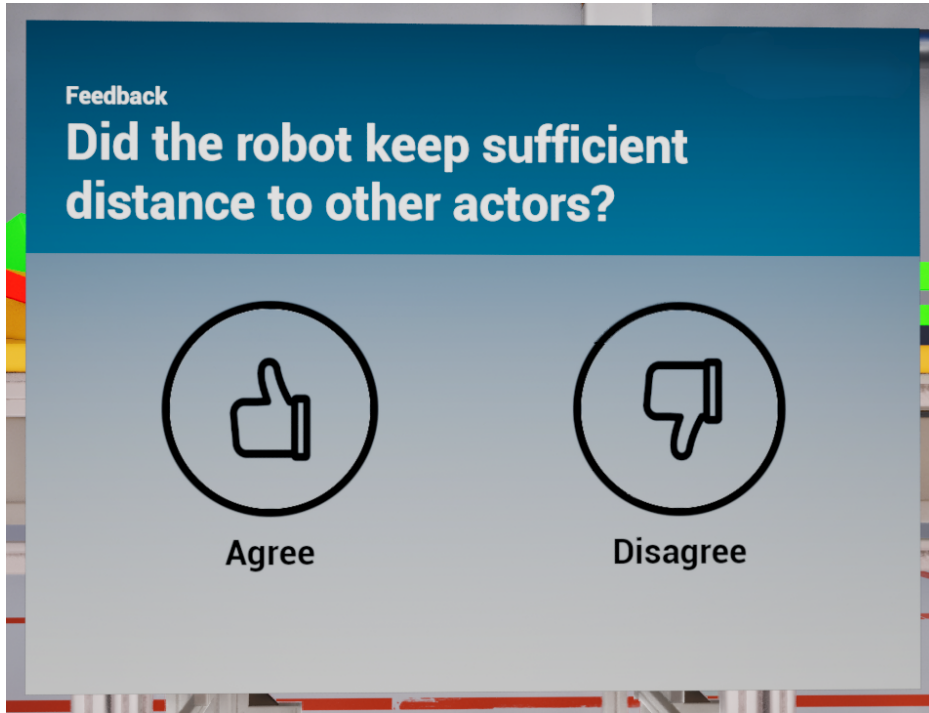


Figure 7: Feedback pop-up prompt displaying the binary feedback buttons

during the VR session. As a result, 44 participants (*Male* = 30, *Female* = 13, *NonBinary* = 1), aged between 25 and 63 years ($M = 33.3$, $SD = 8.8$) were included in the study. Most participants reported moderate VR expertise ($M = 3.07$, $SD = 1.29$) and confidence ($M = 2.92$, $SD = 0.97$), despite infrequent use ($M = 1.55$, $SD = 1.09$).

The study employed a between-subjects design with two feedback conditions: user-initiated and system-initiated. Participants were randomly split between the two experimental conditions. The final distribution of participants between the user-initiated ($N = 23$) and the system-initiated ($N = 21$) feedback conditions was balanced. To compare the two experimental conditions, we collected data on three safety-relevant psychological factors [27, 19]: perceived safety (8 item questionnaire [1]), cognitive load (NASA TLX [15]), and sense of control (3 item questionnaire [41]). The experiment followed a standardised procedure:

1. Pre-survey: Participants completed the demographic data and VR familiarity questionnaires.
2. VR introduction: Participants received an introduction to the VR equipment, environment and tasks.
3. VR session: Participants performed the task in VR. The session is capped at 15 minutes.
4. Post-survey: Participants completed perceived safety, cognitive load and sense of control questionnaires.

3.2.4. Metrics and Analysis

Human effort and feedback information value were quantified based on the survey results and data logs from UE collected during the VR sessions. First, human effort is defined by two metrics: the average duration (seconds) spent on a feedback prompt in the feedback interface, and normalised cognitive load computed from the post-survey data (in line with [2, 49, 20]). Second, feedback information value was defined by two dimensions: psychological and physical safety. Psychological safety was quantified by the two normalised human factors metrics computed from survey data: perceived safety and sense of control (in line with [31]). Physical safety was defined by a proxemic threat score quantified as inverse time-to-collision ($iTTC = TTC^{-1}$) computed from data logs [40]. We computed threat scores between the co-working robot and the human, and between the co-working robot and the other two robots in the environment. The maximum threat values were extracted from the five-second window preceding each time the feedback interface was opened and feedback was provided.

Bayesian models [11] were fitted to compare the two feedback conditions (see Appendix A). Given the nested data structure, we distinguish participant-level data from feedback-level data in our analysis to account for inter-individual

variability and avoid pseudo-replication. First, to estimate posterior differences in human effort between the feedback conditions, Bayesian models were constructed for each metric: a Gaussian model for cognitive load (participant-level), and a log-gamma model for average feedback duration (feedback-level). Second, to estimate posterior differences in information value between the feedback conditions, Bayesian models were constructed for each safety dimension: one model for psychological safety and one for physical safety. Maximum threat to the human and to the other agents were used as predictors in a Bernoulli model with feedback-level data. The psychological safety metrics were not strongly correlated ($|r| = 0.41$) and could be used as individual predictors in a binomial model with participant-level data.

3.3. Results

The system-initiated feedback led to higher feedback volume ($M = 9.22$, $SD = 0.76$) than user-initiated feedback ($M = 5.24$, $SD = 2.4$). Despite more frequent interaction with the feedback interface in the system-initiated condition, there were no credible differences in cognitive load ($M = -0.15$, 95% CrI $[-0.78, 0.48]$) or feedback duration ($M = -0.10$, 95% CrI $[-0.17, 0.39]$) between conditions. Therefore, the system-initiated condition provides more feedback data without additional human effort. Feedback valence was generally positive across both conditions, with system-initiated data showing particularly high values ($M = 0.93$, $SD = 0.13$), and user-initiated values being moderately high as well ($M = 0.78$, $SD = 0.28$). However, there are credible differences in how the two feedback conditions capture psychological safety. User-initiated feedback reflects perceived safety far better than system-initiated feedback ($M = 1.0$, 95% CrI $[0.13, 1.9]$) (see Figure 8). This credible interval excludes zero, providing strong evidence of the effect, with model convergence confirmed ($\hat{R} \approx 1.00$). Additionally, there is also a weak difference in how feedback encodes the maximum proxemic threat posed to the human ($M = -0.16$, 95% CrI $[-0.01, 0.35]$). No interaction effects were identified for sense of control or the maximum proxemic threat posed to the other robots. Full parameter estimates for the psychological and physical safety models are provided in Appendix A (Tables A.1 and A.2).

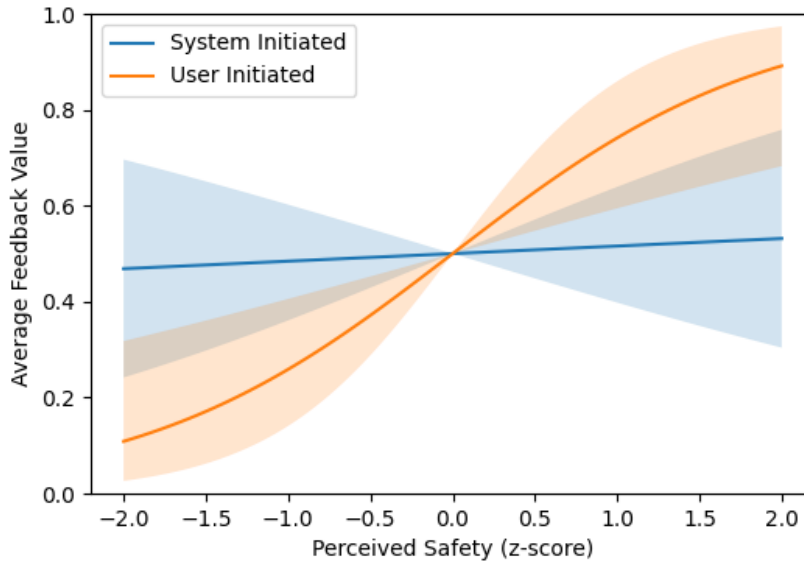


Figure 8: Feedback information value for perceived safety for user- and system-initiated conditions, Bayesian posterior sigmoids (mean and 95% CrI)

3.4. Discussion

The results indicate that user-initiated feedback provides greater information value than system-initiated feedback. This finding builds on previous research [32] demonstrating that real-time user feedback improves robot social behaviour. Additionally, it extends previous research [44] showing better outcomes with feedback-efficient preference learning. Despite a lower volume of feedback data points, the user-initiated condition captures more meaningful information about the psychological safety for the human co-worker. Capturing latent safety constructs is highly

valuable for RLHF, because it allows the encoding of not only physical state variables, but also subjective human perception [31, 23]. Additionally, the physical safety data suggests that user-initiated feedback is more sensitive to safety violations. Therefore, these findings show that user-initiated feedback compensates for the lower data volume by providing more timely and efficient feedback.

The collected feedback for the system-initiated condition was predominantly positive. This suggests that the system-initiated condition fails to capture both physical safety violations and psychological safety concerns. It appears that although human co-workers form a negative impression of perceived safety, this cannot be easily captured without timely feedback action. When prompted with the feedback interface, participants might respond positively to recent events, even though their overall safety perception is affected by proxemic violations happening outside of the immediate time window. However, two main limitations to the design of this feedback condition must be noted. First, the timing of the system-initiated prompt might influence the participants towards a positive response because it overlaps with task completion. Second, the phrasing of the feedback question might not prompt a strong negative response to proxemic violations. Because the safety stop prevents collisions, the distance between actors is always technically “sufficient”, even when it violates proxemic expectations.

The findings from this study can directly inform the design of more realistic future experiments. Moving towards higher fidelity VR simulation or physical set-up requires new feedback interfaces. A pop-up prompt similar to what was used in UE can be transferred to a mounted display on a physical workstation. However, the interaction workflow changes as the human needs to switch attention from the primary physical task to a secondary digital feedback task. The design of the workflow can impact both human effort and the quality of collected feedback. Additionally, for system-initiated feedback, a different modality might be necessary for gaining the attention of the human in a busy industrial environment. More advanced feedback modalities can be employed by having the human give feedback directly to the robot. This would considerably increase the complexity of the human-robot collaboration loop, but could provide a better effort vs information value trade-off.

4. Conclusion

RLHF enables the development of flexible AI-based robot applications across multiple domains where close HRC is required, such as manufacturing and logistics. Our work provides two contributions to RLHF in HRC: a scoping literature review focused on the bidirectional, closed-loop design of RLHF, and a VR experiment assessing the interplay between human factors, robot behaviour, and feedback quality. The review shows that RLHF is an emerging research field, currently exploring a wide range of methods for integrating human feedback into robot training during various stages of the AI development process. Our experimental results indicate that feedback timing is a key aspect in ensuring AI safety, as human data quality depends on immediate feedback action following a perceived safety threat. Future RLHF research would benefit from conducting studies in high-fidelity realistic settings to achieve the desired safety and performance goals for HRC systems. Particular attention should be paid to ensuring that feedback methods are viable in real environments, and able to produce high-quality data for RL algorithms without compromising human factors. Ultimately, RLHF methods have the potential to provide AI-based robots with the needed adaptability to bring HRC systems into highly variable real-world applications like Industry 4.0.

References

- [1] Akalin, N., Kristoffersson, A., Loutfi, A., 2019. Evaluating the sense of safety and security in human–robot interaction with older people, in: Korn, O. (Ed.), *Social Robots: Technological, Societal and Ethical Aspects of Human-Robot Interaction*. Springer, Cham. Human–Computer Interaction Series, pp. 237–264. doi:10.1007/978-3-030-17107-0_12.
- [2] Avaei, A., van der Spaa, L., Peternel, L., Kober, J., 2023. An incremental inverse reinforcement learning approach for motion planning with separated path and velocity preferences. *Robotics* 12, 61. doi:10.3390/robotics12020061.
- [3] Benos, L., Bechar, A., Bochtis, D., 2020. Safety and ergonomics in human-robot interactive agricultural operations. *Biosystems Engineering* 200, 55–72. doi:10.1016/j.biosystemseng.2020.09.009.
- [4] Bi, Z.M., Luo, C., Miao, Z., Zhang, B., Zhang, W.J., Wang, L., 2021. Safety assurance mechanisms of collaborative robotic systems in manufacturing. *Robotics and Computer-Integrated Manufacturing* 67. doi:10.1016/j.rcim.2020.102022.
- [5] Bıyık, E., Losey, D.P., Palan, M., Landolfi, N.C., Shevchuk, G., Sadigh, D., 2022. Learning reward functions from diverse sources of human feedback: Optimally integrating demonstrations and preferences. *International Journal of Robotics Research* 41, 45–67. doi:10.1177/02783649211041652.
- [6] Brooke, J., 1996. SUS: A "quick and dirty" usability scale, in: Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, I.L. (Eds.), *Usability Evaluation in Industry*. Taylor & Francis, London, pp. 189–194.

- [7] Cardoso, A., Colim, A., Bicho, E., Braga, A.C., Menozzi, M., Arezes, P., 2021. Ergonomics and human factors as a requirement to implement safer collaborative robotic workstations: A literature review. *Safety* 7. doi:10.3390/safety7040071.
- [8] Chen, H., Li, S., Fan, J., Duan, A., Yang, C., Navarro-Alarcon, D., Zheng, P., 2025. Human-in-the-loop robot learning for smart manufacturing: A human-centric perspective. *IEEE Transactions on Automation Science and Engineering* 22. doi:10.1109/TASE.2025.3528051.
- [9] European Parliament and Council of the European Union, 2016. Regulation (EU) No 2016/679 ... General Data Protection Regulation. Regulation 32016R0679. European Union. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679>.
- [10] European Parliament and Council of the European Union, 2024. Regulation (EU) 2024/... of the European Parliament and of the Council laying down harmonised rules on artificial intelligence. Regulation 32024R798. European Union. Brussels, Belgium. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R798>.
- [11] Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2013. *Bayesian Data Analysis*. 3rd ed., CRC Press.
- [12] Giallanza, A., La Scalia, G., Micale, R., La Fata, C.M., 2024. Occupational health and safety issues in human-robot collaboration: State of the art and open challenges. *Safety Science* 169. doi:10.1016/j.ssci.2023.106313.
- [13] Gonzalez-Santocildes, A., Vazquez, J.I., Eguluz, A., 2024. Enhancing robot behavior with eeg, reinforcement learning and beyond: A review of techniques in collaborative robotics. *Applied Sciences* 14. doi:10.3390/app14146345.
- [14] Gualtieri, L., Rauch, E., Vidoni, R., 2021. Emerging research fields in safety and ergonomics in industrial collaborative robotics: A systematic literature review. *Robotics and Computer-Integrated Manufacturing* 67. doi:10.1016/j.rcim.2020.101998.
- [15] Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. Technical Report NAS-98-003. NASA Ames Research Center. Moffett Field, CA.
- [16] Hindemith, L., Bruns, O., Noller, A.M., Hemion, N., Schneider, S., Vollmer, A.L., 2023. Interactive robot task learning: Human teaching proficiency with different feedback approaches. *IEEE Transactions on Cognitive and Developmental Systems* 15, 1938–1947. doi:10.1109/TCDS.2022.3186270.
- [17] Hiranaka, A., Hwang, M., Lee, S., Wang, C., Fei-Fei, L., Wu, J., Zhang, R., 2023. Primitive skill-based robot learning from human evaluative feedback, in: *IEEE International Conference on Intelligent Robots and Systems*, pp. 7817–7824. doi:10.1109/IRoS55552.2023.10341912.
- [18] Hjorth, S., Chrysostomou, D., 2022. Human–robot collaboration in industrial environments: A literature review on non-destructive disassembly. *Robotics and Computer-Integrated Manufacturing* 73. doi:10.1016/j.rcim.2021.102208.
- [19] Hopko, S., Wang, J., Mehta, R., 2022. Human factors considerations and metrics in shared space human-robot collaboration: A systematic review. *Frontiers in Robotics and AI* 9. doi:10.3389/frobt.2022.799522.
- [20] Huang, C., Luo, W., Liu, R., 2021. Meta preference learning for fast user adaptation in human-supervisory multi-robot deployments, in: *IEEE International Conference on Intelligent Robots and Systems*, pp. 5851–5856. doi:10.1109/IRoS51168.2021.9636515.
- [21] Hwang, M., Weihs, L., Park, C., Lee, K., Kembhavi, A., Ehsani, K., 2024. Promptable behaviors: Personalizing multi-objective rewards from human preferences, in: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 16216–16226. doi:10.1109/CVPR52733.2024.01535.
- [22] International Organization for Standardization, 2025. *Robotics – Safety requirements – Part 2: Industrial robot applications and robot cells*. Technical Report ISO 10218-2. International Organization for Standardization. Geneva, Switzerland. URL: <https://www.iso.org/standard/73934.html>.
- [23] Kazantzidis, I., Norman, T.J., Du, Y., Freeman, C.T., 2022. How to train your agent: Active learning from human preferences and justifications in safety-critical environments, in: *International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pp. 1654–1656. doi:10.5555/3535850.3536066.
- [24] Kim, S.K., Kirchner, E.A., Schloßmüller, L., Kirchner, F., 2020. Errors in human-robot interactions and their effects on robot learning. *Frontiers in Robotics and AI* 7, 558531. doi:10.3389/frobt.2020.558531.
- [25] Kumar Shastha, T., Kyrarini, M., Gräser, A., 2020. Application of reinforcement learning to a robotic drinking assistant. *Robotics* 9, 1. doi:10.3390/ROBOTICS9010001.
- [26] Liu, Y., Caldwell, G., Rittenbruch, M., Belek Fialho Teixeira, M., Burden, A., Guertler, M., 2024. What affects human decision making in human-robot collaboration?: A scoping review. *Robotics* 13. doi:10.3390/robotics13020030.
- [27] Lorenzini, M., Lagomarsino, M., Fortini, L., Gholami, S., Ajoudani, A., 2023. Ergonomic human-robot collaboration in industry: A review. *Frontiers in Robotics and AI* 9. doi:10.3389/frobt.2022.813907.
- [28] Lu, L., Xie, Z., Wang, H., Li, L., Xu, X., 2022. Mental stress and safety awareness during human-robot collaboration - review. *Applied Ergonomics* 105. doi:10.1016/j.apergo.2022.103832.
- [29] Maroto-Gómez, M., Malfaz, M., Castillo, J.C., Castro-González, Á., Salichs, M.Á., 2024. Personalizing activity selection in assistive social robots from explicit and implicit user feedback. *International Journal of Social Robotics* doi:10.1007/s12369-024-01124-2.
- [30] Marta, D., Holk, S., Pek, C., Tumova, J., Leite, I., 2023. Variquery: Vae segment-based active learning for query selection in preference-based reinforcement learning, in: *IEEE International Conference on Intelligent Robots and Systems*, pp. 7878–7885. doi:10.1109/IRoS55552.2023.10341795.
- [31] Marta, D., Pek, C., Melsion, G.I., Tumova, J., Leite, I., 2022. Human-feedback shield synthesis for perceived safety in deep reinforcement learning. *IEEE Robotics and Automation Letters* 7, 406–413. doi:10.1109/LRA.2021.3128237.
- [32] McQuillin, E., Churamani, N., Gunes, H., 2022. Learning socially appropriate robo-waiter behaviours through real-time user feedback, in: *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 541–550. doi:10.1109/HRI53351.2022.9889395.
- [33] Neggers, M.M., Cuijpers, R.H., Ruijten, P.A., Jsselsteijn, W.A., 2022. The effect of robot speed on comfortable passing distances. *Frontiers in Robotics and AI* 9, 915972. doi:10.3389/frobt.2022.915972.
- [34] Neumann, W.P., Winkelhaus, S., Grosse, E.H., Glock, C.H., 2021. Industry 4.0 and the human factor – a systems framework and analysis methodology for successful development. *International Journal of Production Economics* 233. doi:10.1016/j.ijpe.2020.107992.
- [35] Nixdorf, S., Zhang, M., Grosse, E.H., Ansari, F., 2026. Reciprocal learning in human–machine collaboration: a systematic literature review and implications for production and logistics. *International Journal of Production Research* 64, 1561–1585. doi:10.1080/00207543.2025.

2568177.

- [36] Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGeehan, L., ... Moher, D., 2021. The prisma 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* 372, n71. doi:10.1136/bmj.n71.
- [37] Puerta-Beldarrain, M., Gómez-Carmona, O., Sánchez-Corcuera, R., Casado-Mansilla, D., López-de Ipiña, D., Chen, L., 2025. A multifaceted vision of the human-ai collaboration: A comprehensive review. *IEEE Access* 13, 29375–29405. doi:10.1109/ACCESS.2025.3536095.
- [38] Raza, S.A., Williams, M.A., 2019. Human feedback as action assignment in interactive reinforcement learning. *ACM Transactions on Autonomous and Adaptive Systems* 14, 1–24. doi:10.1145/3404197.
- [39] Schneider, S., Kummert, F., 2021. Comparing robot and human guided personalization: Adaptive exercise robots are perceived as more competent and trustworthy. *International Journal of Social Robotics* 13, 169–185. doi:10.1007/s12369-020-00629-w.
- [40] Singamaneni, P.T., Favier, A., Alami, R., 2023. Towards benchmarking human-aware social robot navigation: A new perspective and metrics, in: 2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), pp. 914–921. doi:10.1109/RO-MAN57019.2023.10309398.
- [41] Strube, M.J., Werner, C., 1984. Personal space claims as a function of interpersonal threat: The mediating role of need for control. *Journal of Nonverbal Behavior* 8, 195–209. doi:10.1007/BF00987291.
- [42] Tahara, K., Niitsuma, M., 2025. Feedback-driven adaptive task estimation and human error handling in human-robot collaboration, in: IEEE/SICE International Symposium on System Integration (SII), pp. 891–896. doi:10.1109/SII59315.2025.10870972.
- [43] Vicentini, F., 2020. Collaborative robotics: A survey. *Journal of Mechanical Design* 143. doi:10.1115/1.4046238.
- [44] Wang, R., Wang, W., Min, B.C., 2022. Feedback-efficient active preference learning for socially aware robot navigation, in: IEEE International Conference on Intelligent Robots and Systems (IROS), pp. 11336–11343. doi:10.1109/IR0S47612.2022.9981616.
- [45] Wang, R., Zhao, D., Yuan, Z., Obi, I., Min, B.C., 2025. Prefelm: Enhancing preference-based reinforcement learning with crowdsourced large language models. *IEEE Robotics and Automation Letters* 10, 2486–2493. doi:10.1109/LRA.2025.3528663.
- [46] Yousefi, E., Chen, M., Sharf, I., 2025. Shared autonomy policy fine-tuning and alignment for robotic tasks. *International Journal of Robotics Research* 0. doi:10.1177/02783649241312699.
- [47] Yu, H., Aronson, R.M., Allen, K.H., Short, E.S., 2023. From “thumbs up” to “10 out of 10”: Reconsidering scalar feedback in interactive reinforcement learning, in: IEEE International Conference on Intelligent Robots and Systems, pp. 4121–4128. doi:10.1109/IR0S55552.2023.10342458.
- [48] Zacharaki, A., Kostavelis, I., Gasteratos, A., Dokas, I., 2020. Safety bounds in human robot interaction: A survey. *Safety Science* 127. doi:10.1016/j.ssci.2020.104667.
- [49] Zhang, R., Bansal, D., Hao, Y., Hiranaka, A., Gao, J., Wang, C., Martín-Martín, R., Fei-Fei, L., Wu, J., 2023. A dual representation framework for robot learning with human guidance, in: Conference on Robot Learning (PMLR), pp. 738–750.

A. Bayesian Statistics

Bambi specification for Gaussian model on cognitive load:

```
cognitive_load = bmb.Model(
    "NASATLX ~ condition",
    part_df,
    family = "gaussian"
)
```

Bambi specification for gamma model on feedback duration:

```
feedback_duration = bmb.Model(
    "fb_dur ~ condition + (1|pid)",
    df,
    family = "gamma",
    link = "log"
)
```

Bambi specification for binomial model on feedback value for psychological safety:

```
feedback_info_psychological = bmb.Model(
    "p(fb_sum, fb_count) ~ "
    "(PERCEIVED_SAFETY + SENSE_OF_CONTROL) * condition",
    df,
    family="binomial"
)
```

Bambi specification for Bernoulli model on feedback value for physical safety:

```
feedback_info_physical = bmb.Model(
    "fb_val ~ (log_max_threat_pr + max_threat_ar) * condition + (1|pid)",
    df,
    family="bernoulli",
)
```

Table A.1

Feedback on physical safety, parameter estimates of the Bernoulli model

Parameter	Mean	SD	95% EtiLB	95% EtiUB	R-hat
condition[UIF]	-0.21	0.81	-1.8	1.3	1.00
max_threat_human	-0.114	0.083	-0.29	0.039	1.00
max_threat_robots	-1.03	1.41	-3.7	1.7	1.00
max_threat_human:condition[UIF]	0.16	0.091	-0.01	0.35	1.00
max_threat_robots:condition[UIF]	-1.12	1.47	-4	1.8	1.00

Note: UIF: user-initiated feedback (reference = system-initiated feedback); Eti: equal-tailed interval; LB = lower bound, UB = upper bound.

Table A.2

Feedback on psychological safety, parameter estimates of the binomial model

Parameter	Mean	SD	95% EtiLB	95% EtiUB	R-hat
condition[UIF]	-1.191	0.372	-1.9	-0.47	1.00
PERCEIVED_SAFETY	0.066	0.252	-0.41	0.57	1.00
SENSE_OF_CONTROL	0.357	0.303	-0.24	0.94	1.00
PERCEIVED_SAFETY:condition[UIF]	1	0.439	0.13	1.9	1.00
SENSE_OF_CONTROL:condition[UIF]	-0.34	0.404	-1.1	0.46	1.00

Note: UIF: user-initiated feedback (reference = system-initiated feedback); Eti: equal-tailed interval; LB = lower bound, UB = upper bound.