

Human Leading or Following Preferences: Effects on Human Perception of the Robot and the Human-Robot Collaboration

Ali Noormohammadi-Asl^{a,*}, Kevin Fan^b, Stephen L. Smith^b, Kerstin Dautenhahn^b

^a*Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada*

^b*Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada*

Abstract

Achieving effective and seamless human-robot collaboration requires two key outcomes: enhanced team performance and fostering a positive human perception of both the robot and the collaboration. This paper investigates the capability of the proposed task planning framework to realize these objectives by integrating human leading/following preference and performance into its task allocation and scheduling processes. We designed a collaborative scenario wherein the robot autonomously collaborates with participants. The outcomes of the user study indicate that the proactive task planning framework successfully attains the aforementioned goals. We also explore the impact of participants' leadership and followership styles on their collaboration. The results reveal intriguing relationships between these factors, which warrant further investigation in future studies.

Keywords: Human-robot collaboration, adaptive task planning, leading/following preference, team performance, perception of the robot and collaboration

1. Introduction

Human-robot collaboration provides the opportunity to exploit the complementary abilities of both the human and the robot in a collaborative task. This collaboration has the potential to offset human limitations, given that robots can be advantageous in tasks that require extensive repetitions, precision, speed, physical strength, and stamina. Collaborative robots (cobots) could also reduce the human's cognitive load of tasks. Simultaneously, it leverages human superior abilities to adapt and handle uncertainties and address situations that demand human-level intelligence and decision making. However, the presence of humans requires thoughtful consideration of human agents in the development of computational architectures for effective human-robot collaboration [1, 2].

The robots' ability to adapt to their human teammates is essential for fostering effective, fluent, and long-term collaboration, establishing a high level of human perception of the robot and the collaboration. Human preference, a focus of many recent studies, is a crucial factor that needs consideration in the design of human-robot collaborative scenarios and in robot programming [3–8]. This paper specifically delves into human preferences within the context of task planning and scheduling. The robot's

high computational and planning abilities enable it to assume a more significant and leading role in task planning. However, there are situations (e.g., changing environments) that necessitate human agents taking the leading role and planning for the team. Additionally, different human teammates may have varying leading/following preferences.

Although the robot needs to consider the human agent's leading/following preference and adapt its planning accordingly, the human agent's performance may conflict with the overall team goal and deteriorate team performance. This necessitates the robot programming to take into account the human agent's preference and performance simultaneously. However, the human agent's preference and performance are not fixed and may change during the collaboration. For example, for a challenging part of a task, a human agent with the overall leading preference would choose to follow the robot, or a human with an overall high performance may get tired at some point and thus start showing poor performance.

The core concept of the proposed framework for involving the human agent's preference and performance in the robot's programming is enabling the robot to adjust its planning according to the human agent's preference as long as the human preference does not lead to a substantial deviation from the optimal plan or compromise performance beyond acceptable limits.

Within this framework, unlike the majority of conventional approaches that concentrate strictly on one extreme of the leading/following spectrum, our approach can cover the entire range, allowing for roles ranging from solely act-

*This research was undertaken, in part, thanks to funding from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canada 150 Research Chairs Program.

*Corresponding author.

Email address: ali.asl@uwaterloo.ca

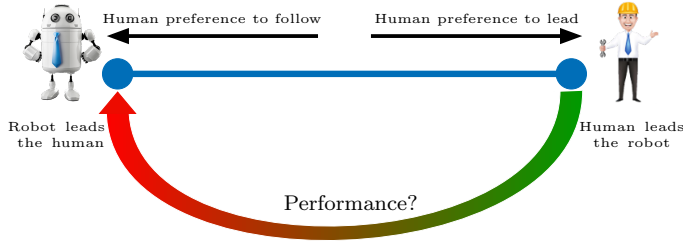


Figure 1: Encompassing the entire spectrum of leading/following roles based on human preference and performance

ing as a leader to exclusively functioning as a follower. Fig. 1 encapsulates the main idea of this framework. This framework requires the robot to continually monitor and, in real-time, gradually adapt to human preferences and performance and their changes. This framework is versatile and applicable to various collaborative scenarios, but our research specifically delves into the realms of task selection/allocation and scheduling within the context of human-robot collaboration.

The fundamental idea of task allocation problems is determining the assignments of tasks to agents in order to achieve the overall team objectives and optimize the team payoff. This has been studied extensively in both multi-robot teams [9] and Human-Robot Collaboration (HRC), which incorporates human and robot agents [10–14]. Our work centers on proactive task allocation for single-human, single-robot teams, in which the human and robot have the agency to select their own tasks and assign tasks to each other. This agency makes the problem different from offline and reactive task allocation problems, where tasks are assigned by a single agent or an external entity like a manager or central controller.

The human agent’s agency in selecting and assigning tasks allows them to implement and adjust their leading/following preference. Subsequently, the robot must continually gauge the current human preference and performance in real-time, throughout the collaboration, and adapt its planning accordingly through two-step task planning at each decision step: first, task allocation, and second, task scheduling. The robot can reassume the leading role in instances of poor human performance, even if the human agent prefers to lead the team. The concept of the robot reassuming the leading role in favor of team performance is rooted in the results of our earlier user study [15], encompassing three distinct robot strategies: prioritizing the human, prioritizing the robot, and striking a balance between both. This prior work revealed that maintaining a balanced strategy can enhance team performance without compromising the human perception of the collaboration.

Previously, we evaluated the suggested framework by applying it in a simulation environment with a simplified human decision-making model [16], and later on a physical robot, where the experimenter enacted various scenarios

of how a human agent might behave [17]. These scenarios spanned from a high leading preference to a high following preference, as well as from high performance to low performance. This work extends those studies by testing the framework’s effectiveness and applicability by recruiting participants to collaborate with the robot in a collaborative scenario, similar to the kitting task. We evaluated participants’ perceptions of the robot and collaboration as well as their collaboration strategy and decisions, which are the focus of this paper. We also analyzed the robot’s estimations and actions, which are not within the scope of this paper and will be discussed in a future publication [18].

Given the intended objective of the proposed framework – to empower the robot to adapt in real-time to the preferences and performance of its human teammate while retaining a high level of human perception regarding the robot and collaboration – we explore the following research questions:

- RQ1: To what extent can this collaboration and task planning framework improve the team performance and reduce participants’ perceived task load?
- RQ2: How do the human agent’s perception of the robot and the collaboration affect their collaboration with the robot and vice versa?
- RQ3: How does this collaboration influence the human agent’s actions and decision making?
- RQ4: How does the task difficulty influence participants’ actions and collaboration with the robot?

1.1. Contributions

The contributions of this work are fivefold:

1. Designing a collaborative scenario involving autonomous pick-and-place tasks for the robot and implementing the task planning framework along with the estimation methods for its execution.
2. Implementing dynamic task updates that take into account task states, errors made by the human agent, and subsequent actions by the robot to rectify them.
3. Studying the effects of the cobot’s adaptive task planning on participants’ perception of the robot and collaboration, e.g., trust, satisfaction, robot intelligence, and workload.
4. Studying the effect of participants’ preferences and attitude toward collaboration on the cobot’s planning.
5. Studying the effects of participants’ leadership style and initial trust in the robot on their collaboration and perception of the robot.

The earlier versions of this research were presented in conference papers [16, 17]. The paper [16] primarily concentrated on designing a robot planning framework and evaluating it in a simulation environment for a collaborative scenario with a simplified human decision-making model. In contrast, this paper extends the framework’s application to a more complex collaborative scenario, involving an actual robot collaborating with recruited participants.

The paper [17] introduced the initial implementation of this collaborative scenario and assessed the planning framework’s performance across four different scenarios conducted by the experimenter. However, the present paper unveils the final version of the experimental setup and planning framework. Importantly, it evaluates the effectiveness of the proposed planning framework for an autonomous robot collaborating individually with each of the 48 participants, involving 144 tasks. This paper focuses on evaluating participants’ perceptions of the robot, collaboration, and tasks, examining their effects on actions and collaboration. In [18], the companion paper explores the framework’s ability to enable the robot to plan and adapt to participants’ preferences and performance.

1.2. Organization

The rest of this paper is organized as follows: Section 2 reviews pertinent literature and formulates the research hypotheses. Section 3 outlines the study’s design, including both objective and subjective measures, and details the study procedure. In Section 4, we analyze the results, emphasizing participants’ perceptions of the robot, tasks, collaboration, and how these factors impact their decision-making. Section 5 concludes the paper, addressing its limitations and suggesting potential avenues for future research.

2. Related Work & Hypotheses Development

In this section, we initially conduct a review of pertinent literature focusing on the human perception of the robot and collaboration in terms of trust, self-confidence, reliance, workload, and helpfulness, and develop our hypotheses based on the insights gathered from the existing literature. We also explore the human perception of control and the various leadership and followership styles. Finally, we briefly review research concerning task allocation and adaptation to human preferences.

2.1. Human Perception of the Robot & Collaboration

2.1.1. Trust, self-confidence & reliance

Trust is a complex and multifaceted concept in human-human, human-technology, and, more specifically, in human-robot interactions (HRI). In the context of human-robot trust, contributing factors to human trust in robots can be categorized into three groups [19]:

- robot characteristics and design elements (e.g., reliability)
- human characteristics (e.g., self-confidence)
- environmental-based factors (e.g., task type)

Typically, in most human-robot collaboration scenarios, it is assumed that humans and robots share common goals, and the human agent is aware of the collaborative context of the interaction, which prevents them from considering deception. Sanders et al. mainly focussed on the first two factors mentioned above and investigated how robot performance can affect human trust [19].

Trust is a topic of interest in HRI and HRC and has been studied extensively. Some studies focus on how to model and measure trust and its dynamics and, in some cases, use it to adjust the robot’s behavior or improve human distrust or overtrust in the robot [20–22]. Some other studies, including this work, use objective or subjective measures (e.g., questionnaires) to evaluate human trust in robots before or after interaction with a robot. In our study, using Muir’s questionnaire [23], we are interested in how participants’ trust in the robot evolves during the sessions over time. We established the following hypothesis:

Hypothesis 1. *Participants’ trust in the cobot, measured before and during the collaborative tasks, will improve during the session while engaged in different tasks.*

Furthermore, it is well-established that human reliance on an autonomous system is intertwined with their trust in the system and their self-confidence (i.e., their trust in their capability to accomplish the task). Studying and modeling this relationship has been the focus of many studies. The “confidence vs. trust” hypothesis is a well-accepted concept and serves as the foundation for many proposed models [24]. It posits that a person’s reliance on an automation system is influenced by their relative trust, which refers to the difference between their trust in automation and their self-confidence. That is, if an individual’s trust in the system exceeds their self-confidence, they will opt to rely on the system; otherwise, they will rely on their own decisions and perform the task on their own. However, the findings of some studies contradict this hypothesis [25]. Authors in [26] have proposed a computational model that outperforms the “confidence vs. trust” hypothesis and signifies nuances between trust and self-confidence in predicting human reliance on automation. It is also noteworthy that in a collaborative scenario, the goal is to exploit human and robot abilities, not bypassing one by employing the other. This point may affect the full applicability of the “confidence vs. trust” hypothesis in human-robot collaboration.

We state the associated hypotheses as follows:

Hypothesis 2. *Participants’ self-confidence will improve during the collaborative tasks over time as they collaborate more with the robot.*

Hypothesis 3. *For demanding tasks, participants will trust the robot more than their own ability and will rely more on the robot.*

2.1.2. Perceived workload

There are three fundamental measures for evaluating workload: physiological (e.g., heart rate), performance-based (e.g., time), and perceptual (subjective) [27]. In this paper, we consider performance-based and subjective measures. For the subjective measure, we use NASA-TLX, which defines the workload experience in terms of “the sources of loading imposed by different tasks” [28]. We seek to study whether collaborating with the robot could positively influence participants’ perceived workload, and derive the following hypothesis.

Hypothesis 4. *Participants will experience less workload when collaborating with the cobot rather than working alone, in terms of the dimensions in the NASA task load index, including: (a) mental demand (b) physical demand (c) performance (d) effort (e) frustration. However, the temporal demand will be similar, as there was no time limit for completing tasks.*

2.1.3. Helpfulness

In [29] helpfulness is defined as “with honest intentions, trying to play a positive role with the task at hand.” In this study, we aim to investigate how humans’ expected helpfulness from the robot, which refers to their anticipation of the robot’s ability to provide effective assistance and enhance their overall collaboration experience, changes during collaboration.

Hypothesis 5. *Participants’ expected helpfulness of the cobot will increase over time as they collaborate more with the cobot.*

2.2. Being in or under control

While the concept of human-robot collaboration is a relatively recent subject, conversations surrounding human automation have endured for many years. The rise of automation systems has gradually replaced humans in various control systems, spanning from simple to complex setups. Different factors, such as human preferences and abilities, contribute to determining the autonomy levels within these systems. These can range from manual control, where human operators actively control, to semi-autonomous and fully autonomous systems, where human operators take on more supervisory roles.

The perception of control by humans is a crucial area of research, spanning psychology, social science, and human-computer (robot) interaction. In the late 1990s, researchers like Ben Shneiderman and Pattie Maes discussed human-computer interface design from distinct perspectives [30]. Shneiderman emphasized user-controlled direct manipulation, highlighting human control, predictability,

and responsibility. In contrast, Maes advocated for long-lived, proactive, adaptive software agents aligned with individual user habits and preferences. As of 2023, traditional input methods like mice and keyboards remain predominant in computer interactions, suggesting a prevailing inclination toward control and responsibility when dealing with technology. This underscores the significance of design decisions and user anticipations when incorporating intelligent technologies into human-computer interactions, comprehending users’ behaviors and control perceptions across diverse levels of agency and automation [31, 32]. When extending this concept to human-robot interaction, a preference is likely for humans to maintain control over robots.

The perception of robot control by humans can be influenced by factors such as human characteristics (e.g., locus of control), trust in robots, self-confidence, task difficulty, task criticality, human abilities, and robot abilities. In [33], the issue of being in control or under control was explored in a warehouse order-picking scenario, demonstrating higher productivity when the human leads and higher accuracy when the human follows the robot. In contrast, “6 River Systems,” a company actively engaged in collaborative robotics and warehouse automation, contends that allowing associates to have self-control over their pace leads to decreased speed and efficiency [34]. In response to this challenge, they introduced a “system-directed picking” approach, wherein robots guide associates through their tasks. Another instance in the industrial domain is the robotic picking assistive system implemented in DHL warehouses, where the robot (supplied by Locus Robotics) directs human workers to the designated pickup location [35]. In these setups, the robot is tasked with planning the tasks, while humans are responsible for picking up items.

The findings from a user study in [36], in the context of household robots, indicate that participants felt a greater sense of control when the robot operated in semi-autonomous mode. They observed a preference among participants for reduced autonomy in critical tasks (e.g., scheduling a doctor’s appointment). However, when it came to a robot transporting biscuits from the kitchen to the living room, contrary to their stated preference, participants leaned towards the autonomous robot condition. The authors attributed this shift to the perceived advantages of higher physical efficiency and comfort. It is worth noting that a more comprehensive analysis could be conducted if the researchers had also inquired about participants’ trust in the robot’s ability to move autonomously. Furthermore, in [37], the authors explore the connection between automation accuracy and task controllability, emphasizing its impact on user satisfaction. Their study underscores a strong preference for manual control over automation.

In our work, based on the collaborative and cognitively critical nature of the scenario designed (due to a penalty for mistakes), we state the next hypothesis of this paper as follows:

Hypothesis 6. *Overall, participants will prefer retaining control over the robot and being a leader.*

2.3. Leadership and Followership Styles

In this study, we are also interested in how participants' general leadership and followership styles can affect their collaboration with the robot.

Leadership Styles [38]: Leadership style pertains to the behaviors leaders demonstrate, encompassing their actions and interactions with followers. It characterizes what leaders do and how they behave in diverse circumstances. Leadership style is intricately linked to an individual's personal leadership philosophy. Commonly observed leadership styles include [38]: 1. authoritarian 2. democratic 3. laissez-faire

- **authoritarian** – involves leaders perceiving followers as requiring guidance, exerting control over their actions, emphasizing their authority, and directing communication toward themselves rather than fostering interaction among group members.
- **democratic** – involves leaders treating followers as capable collaborators, guiding rather than controlling, and fostering open, inclusive, and supportive communication. They prioritize helping followers achieve personal goals, provide guidance and suggestions without giving orders, and deliver objective praise and criticism in evaluations.
- **laissez-faire** – is a “hands-off, let it ride” approach where leaders exert minimal influence and allow followers significant freedom. They do not control or guide followers' activities and may serve as temporary or interim leaders with limited authority or interest in shaping outcomes.

It is worth emphasizing that these leadership styles are not separate categories; they can overlap and coexist. Leaders may adopt different styles depending on the situation. For example, a leader might exhibit an authoritarian approach in certain contexts while adopting a democratic style in others, showcasing a dynamic and multifaceted approach to leadership. Northouse proposed the leadership style questionnaire in [38] to identify participants' tendency to the leadership styles.

Followership Styles [39]: Kelley (1992) identified five followership styles based on two key dimensions: engagement and critical thinking. Engagement ranges from passive (waiting for direction) to active (taking the initiative to participate actively in tasks). Critical thinking spans from dependent uncritical thinking (accepting information without evaluation or questioning) to independent critical thinking (evaluating and analyzing information to identify consequences and opportunities). These five follower styles are:

- **Exemplary:** These followers show high levels of active engagement and independent critical thinking,

thinking for themselves, challenging leaders constructively, proactively supporting organizational goals, and assuming extra responsibilities.

- **Conformist:** Conformist followers are actively engaged but tend to be dependent uncritical thinkers, enthusiastically following leader directions without questioning.
- **Passive:** Passive followers have low levels of both engagement and critical thinking, needing constant direction and acting only after explicit instructions.
- **Alienated:** Alienated followers are independent critical thinkers but lack engagement. They are skeptical, may oppose management, and see themselves as mavericks.
- **Pragmatist:** Pragmatists have a moderate level of both engagement and critical thinking. They are uncommitted, prefer maintaining the status quo, and act cautiously, often waiting for crises before taking action.

Kelley also designed the questionnaire in [39] for determining followership styles.

2.4. Task Allocation and Adaptation

Task allocation in human-robot collaboration (HRC) involves assigning tasks to human and robot agents. This can be done offline, relying on prior knowledge to determine task suitability [40, 41], or online, allowing for flexibility and reactive replanning. Online approaches can either find an optimal allocation [11, 12, 42, 43] or involve human decisions, with robot agents adapting to human preferences [44–46]. Incorporating human preferences into task allocation is essential for enhancing the perception of the robot, but this aspect is often underestimated. While some studies have considered human preferences during offline planning, there is a need to assess how these preferences align with overall team performance.

In the context of task allocation, the ability of a robot to adapt is crucial for effective collaboration with humans [45, 47, 48]. Robot adaptation can be achieved through pre-programmed instruction by experts [49, 50], equipping the robot with autonomous learning and adaptation capabilities, or a combination of both [51, 52]. Learning-based adaptation approaches employ supervised or unsupervised learning algorithms [46, 53–55], with the latter involving identifying factors influencing human behavior and enriching behavioral data through expert annotation or surveys. The former enables robots to learn human preferences by observing their behavior without explicit annotation. However, many of these studies focus on the robot adapting to the human agent, while mutual adaptation, where both human and robot adjust based on each other's actions and feedback, is also vital in human-robot collaboration, as demonstrated in the studies exploring mutual adaptation between humans and robots [56].

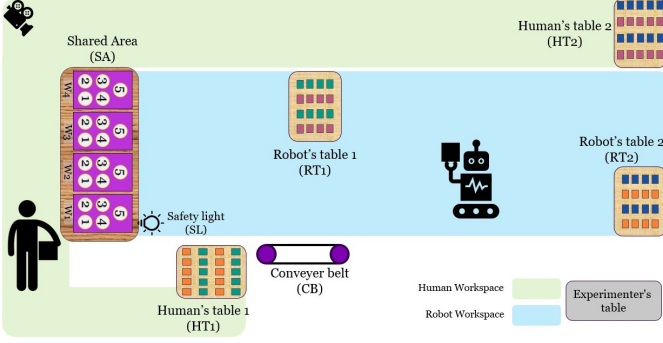


Figure 2: The layout of the experimental setup includes two tables in the robot’s workspace and two tables in the human agent’s workspace. Additionally, there is a shared area (table) where both agents need to arrange blocks. A conveyer belt beside human table 1 allows the robot to return a block to the human agent.

3. User Study: Setup & Methodology

This section provides a detailed explanation of the setup, procedure, and measures employed in this user study.

3.1. User Study Setup

The design of the user study scenario was shaped by three pivotal factors:

1. **Collaboration:** Focusing on the planning abilities of both the human and robot, leveraging the cobot’s advanced memory capacity and the human’s quicker pace,
2. **Leading/following preference:** Providing the human with the agency to adjust their role as either a leader or follower,
3. **Performance:** Introducing a task that imposes cognitive load and penalizes errors.

3.1.1. Setup

In this specific scenario, the human-cobot team collaborates to select colored blocks from tables, adhering to a specified pattern, and arrange them within a shared area with a predetermined number of spots. The designed collaborative scenario closely resembles the kitting task. The kitting task involves selecting and collecting all necessary components for assembling a particular final product [57]. Additional details about the setup are visualized in Fig.2 and Fig.3. The camera’s placement in Fig.2 corresponds to the viewpoint from which the image in Fig.3 was captured. Additionally, Fig. 2 illustrates the location of the experimenter’s table, indicating the position of both the computer and the experimenter themselves.

- **Shared area:** The shared area consists of four workspaces denoted as W_1 through W_4 , each containing five numbered spots. As per the predetermined pattern, participants are required to place a colored block in each spot, adhering to the numerical sequence, noted on each spot, from one to five.

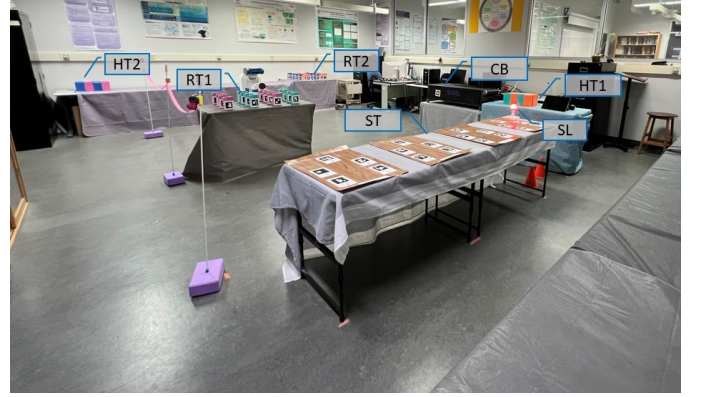


Figure 3: An overview of the experimental setting captured from the perspective of the camera illustrated in Fig. 2

- **Patterns:** The pattern depicted in Fig. 4a is an example of the pattern the human-robot team must follow when filling the spots. Participants are initially presented with a pattern (Patterns A_1 , B_1 , C_1 , and D_1) on a sheet of paper, and they are given 45 seconds for memorization. Subsequently, they return the pattern, and the experimenter provides them with a partially known version of the same pattern (Patterns A_2 , B_2 , C_2 , and D_2), again on a sheet of paper, in which certain spots contain two colors, one of which is correct. This second pattern acts as a cue to recall the initial pattern, and participants are allowed to retain it until the end of the task. Patterns A_2 , B_2 , C_2 , and D_2 feature 9, 12, 6, and 9 partially known spots, respectively, introducing varying levels of difficulty and cognitive load. The decision to incorporate a pattern memorization task was made to introduce a cognitive challenge while adhering to the limitations of a relatively short collaborative scenario. Conducting an extensive experiment that might place excessive mental and physical burdens on participants was considered impractical and challenging to obtain ethics approval. Therefore, concise tasks with durations of around 12-20 minutes were chosen to ensure participants’ sustained mental engagement.
- **Tables:** The human and robot have separate work areas, with two tables in each. In the human work area, one table is close to the shared area and contains blocks of two colors: green and orange. The other table is far from the shared area (i.e., requires travelling a longer distance) and contains blocks of two other colors: blue and pink. Likewise, the robot’s work area has two tables, one distant from the shared area, with blue and orange blocks, and one close to the shared area, including green and pink blocks. Table. 1 summarizes the distribution of the blocks. Participants were also told to pick up only one block at a time, as the cobot’s gripper only has the capacity to pick up a single block. This table and block

Table 1: Distribution of blocks with respect to the distance to the shared area

Color	Human	Robot
Green	Close	Close
Pink	Close	Far
Orange	Far	Close
Blue	Far	Far

distribution help evaluate the robot and the human’s decision-making and task planning (i.e., selecting and assigning tasks) regarding their preference, travel distance, completion time, and proximity to the optimal plan.

- **Blocks:** Four unique block colors are employed: green, blue, pink, and orange. The human’s tables are equipped with ten blocks of each color, whereas the robot’s tables contain eight blocks of each color. This ensures a total of eighteen blocks for each color, exceeding the necessary five blocks per pattern to account for possible errors. A larger quantity of blocks is placed on the human agent’s table, taking into account their swifter working pace. Each block has an ArUco marker affixed to it, aiding the robot in locating and handling the blocks within the room.
- **Conveyor Belt:** A conveyor belt has been integrated to allow the robot to return any misplaced blocks to the human. The robot places a block at one end of the conveyor belt, which then transports the block to the human’s side. This design decision accomplishes two objectives: it separates the workspaces of the human and the robot, mimicking a real industrial environment, while also enhancing safety by reducing the need for the robot and human to interact closely.
- **Light Bulb:** On the shared table, there is a red light bulb that lights up when the robot approaches to place or retrieve a block from the shared table. Participants are informed that this light serves as a signal to avoid approaching the table due to safety precautions. However, they can continue planning, navigating within their designated work area, and collecting blocks from other tables. In addition to safety concerns, the light bulb also helps regulate the participants’ pace, prompting them to wait for the robot and ensuring that the collaboration is not perceived as a competitive race.
- **Safety Measures:** Alongside the safety light bulb, we have set up safety tape and floor cones to create a well-defined physical boundary that separates the work areas of both agents. These measures are complemented by continuous vigilance from the experimenter, who is prepared to intervene as needed. Intervention options include manually controlling the

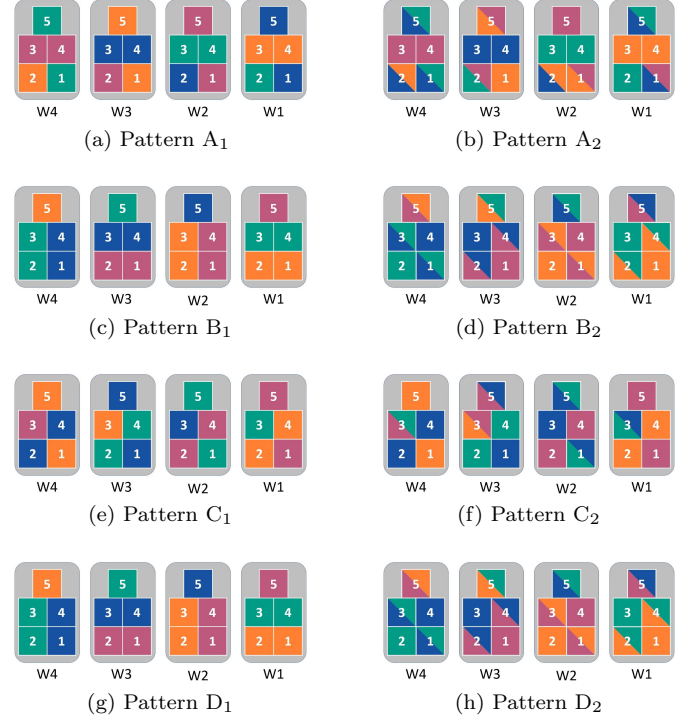


Figure 4: **a, c, e, g:** Patterns A₁, B₁, C₁, and D₁ represent the set of patterns presented on paper sheets, requiring participants to memorize them within a 45-second timeframe before returning them to the experimenter. **b, d, f, h:** Patterns A₂, B₂, C₂, and D₂ are variations with some partially known spots, serving as cues for participants throughout the collaborative task, and they are permitted to retain these patterns until task completion to aid in recalling the first pattern.

robot using a joystick or promptly activating the emergency safety stop button. During the sessions, there were instances where manual control of the robot was necessary, particularly when it came too close to the tables.

3.1.2. Tasks

Each participant is asked to complete four tasks:

- **Task 0:** In this initial task, participants engage in solo work without the robot’s involvement. For all participants, we provide pattern A and follow a consistent procedure. They are tasked with memorizing pattern A₁, printed on a sheet of paper, within a 45-second timeframe and returning it. Subsequently, we furnish them with pattern A₂, which they can retain until the task’s conclusion. It is essential to note that errors made by participants in this task do not incur penalties; rather, this task primarily serves to acquaint participants with the setup and offer them practice in block placement. This practice encompasses aspects such as adhering to the correct block order, picking up a single block at a time, and ensuring the markers on the blocks face the robot.

Table 2: The human and robot’s sets of actions

Human		Robot	
Action	Description	Action	Description
H1	Selecting a task for themselves	R1	Selecting a task for itself
H2	Assigning a task to the robot	R2	Assigning a task to the human
H3	Returning an block from the shared workspace	R3	Returning a <i>wrong</i> block from the shared workspace
H4	Performing a task assigned by the robot	R4	Performing a <i>correct</i> task assigned by the human
H5	Canceling a task assigned to the robot	R5	Canceling a task assigned to the human
H6	Rejecting a task assigned by the robot	R6	Rejecting a task assigned by the human

While focusing on task practice, we also assess participants’ accuracy, self-confidence, and their perception of workload.

- **Tasks 1, 2, 3:** The robot collaborates with the human in these tasks. We consider six distinct permutations (modes) that determine the order of presenting patterns B, C, and D (e.g., B, C, D, C, B, D). Participants are randomly allocated to one of these six modes (permutations) to ensure an equal distribution among them. Following a consistent procedure, participants are tasked with memorizing the initial pattern within a 45-second timeframe and subsequently returning it. Then, they are provided with the second pattern. To emulate real-world collaborative scenarios in which errors carry consequences, participants are informed that, upon declaring the task as completed, for each block incorrectly placed on the table, a deduction of \$1 will be applied to their total remuneration. This method of conveying disinformation, classified as a form of deception, received approval from the University of Waterloo Human Research Ethics Board.

3.1.3. Agents’ Actions

We considered a set of six distinct actions for each agent, namely the human and the robot. These actions, outlined in Table 2, are consistent for both agents, ensuring an equivalent level of agency. It is essential to recognize that the feasibility of these actions depends on the current task state, and at each decision point, certain actions may not be applicable. For instance, when the robot has not been assigned any tasks (Action H4), taking on a task delegated by the human is not a viable option. Additionally, to provide the robot with greater autonomy in adjusting its leadership role or reassuming it in the event of suboptimal human performance, we intentionally designed the rejection option (Action H6) to be less straightforward for the human. Action H6 entails the successive execution of H4 and H3 without actually doing them physically. This will be further explained in the next paragraph.

3.1.4. Human-robot communication

Both agents are required to uphold communication to keep each other informed about their forthcoming actions, as specified in Table 2. This communication is facilitated

through a purpose-built graphical user interface (GUI) installed on a tablet. Participants have the flexibility to position the tablet on a table within the room or hold it according to their preference. This GUI serves as a medium for participants to assign tasks to the robot and convey information about their intended actions. Similarly, the robot can employ the GUI to assign tasks to the participant and communicate its own planned actions. An image of the GUI is provided in Fig. 5. Notably, the GUI enforces constraints to prevent the human agent from selecting actions that would violate precedence constraints or interfere with tasks already initiated by the robot. Participants received instructions on operating the GUI and had the opportunity to practice using it before commencing the tasks. Additionally, participants were instructed to scan the marker on the blocks before placing them on the shared area, as this action enables the robot to identify the block’s ID for potential future retrieval. The GUI streamlines this process by automatically activating the tablet’s camera, allowing participants to scan the marker. To maintain brevity, the details of the GUI’s design and implementation are omitted due to space limitations.

Remark. In case participants wish to decline a task assigned by the robot, the sequence of actions involves initially performing the assigned task (H4) and subsequently engaging in the action for returning (H3). Both of these actions are executed through the GUI, without any requirement for physical execution.

3.1.5. Robot

For our experiments, we utilized the Fetch mobile manipulator robot [58]. The robot was configured to perform pick-and-place operations autonomously. While creating a fully autonomous pick-and-place system demanded a considerable amount of time, it was crucial to ensure that participants experienced a work environment resembling real industrial settings. To maintain brevity, we skip the explanation of the autonomous pick-and-place implementation on the robot.

3.2. Recruitment

We started distributing flyers and recruiting participants after securing ethics approval from the University

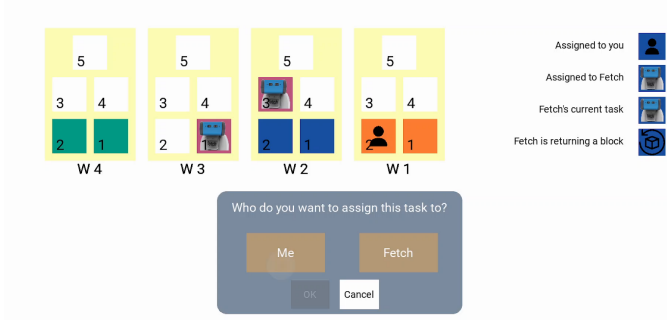


Figure 5: A screenshot of the graphical user interface (GUI) that enables participants to convey their actions to the robot and obtain information regarding the robot’s decisions and actions (refer to Table 2 for a list of actions).

of Waterloo Human Research Ethics Board. This study involves three phases for each participant.

3.3. Procedure

3.3.1. Phase 0

The first phase of our study commenced with prospective participants responding to our recruitment flyers and establishing contact with the experimenter. Following this, we sent the participants a consent form via email, along with a questionnaire aimed at gauging their background and experience in the domains of robotics and artificial intelligence.

3.3.2. Phase 1 (In-Person)

In Phase 1, we scheduled a 90-minute in-person session with each participant, where we conducted the following procedure:

Step 1: We began by welcoming participants and providing an overview of the setup, showing them their designated work area. We used presentation slides to elaborate on the specifics and informed participants that:

- The robot’s decision-making may involve occasional errors (deception).
- Upon confirming the task’s completion, There is a penalty of \$1 for the team per misplaced block (deception).

Step 2: Task 0 - To begin, we gave participants Pattern A₁, as a sheet of paper, and granted them 45 seconds to memorize it. We then collected Pattern A₁ and provided participants with Pattern A₂, a partially known version of Pattern A₁, and indicated that they could keep it until they finished the task. Participants were prompted to assess their self-confidence in successfully completing the task and subsequently initiated the retrieval and placement of the blocks. The experimenter was available to provide guidance and address any deviations from the instructions or questions. It was emphasized that the task

had no time constraint, and the completion time wouldn’t be evaluated. After successfully placing all the blocks and declaring task completion, the experimenter conducted an inspection of the blocks based on the provided pattern, counting any misplaced blocks. Importantly, participants were assured that no penalties would be applied for errors in this task. Following this, participants were requested to complete a questionnaire regarding their perceived task load.

Step 3: They watched a video of the pick-and-place of some blocks by the Fetch robot¹, and then answered the questionnaire about their trust in the robot.

Step 4: Participants familiarized themselves with the GUI and practiced its operation.

Step 5: Task 1 - Depending on the assigned mode (a permutation of patterns B, C, D), participants were provided with the corresponding pattern (e.g., B₁) and instructed to memorize it within a 45-second timeframe. Subsequently, they returned the pattern and received the second pattern, a partially known version of the first pattern (e.g., B₂). Before commencing the task, participants were asked to respond to two questions about their self-confidence and the expected helpfulness of the robot. Following this, they initiated the task, collaborating with the Fetch robot. Upon task completion, participants completed three questionnaires assessing their perceived task load, trust, and perception of the robot.

Remark. Participants initiate the task, and the robot awaits their instructions. They can assign tasks to the robot, and the robot starts working when participants allocate a task to themselves. This approach allows the robot to establish an initial understanding of participants’ preferred leading/following roles in the collaboration.

Step 6: Task 2 - This task followed the same procedure as Task 1.

Step 7: Task 3 - This task followed the same procedure as Task 1.

Step 8: Lastly, participants were asked to complete two questionnaires. The first questionnaire aimed to evaluate their performance as a team with the robot, while the second one explored their collaborative experience and utilized the shortened version of the User Experience Questionnaire (UEQ)[59, 60]. Additionally, participants were asked to rank the difficulty of the tasks (Tasks 0-3) and respond to an open-ended question: “Which abilities would you improve or add to Fetch if you were to use it in a manufacturing setting?”.

¹https://youtu.be/ahZDo0_eyJg

3.3.3. Phase 2 (Online)

In the online phase, we presented each participant with a video recording of their interaction with the robot, with a specific focus on Pattern B. These videos included synchronized content from two different camera angles, a screen recording of the graphical user interface (GUI), and the display of Patterns B₁ and B₂. One of the participant’s videos is accessible online². During the online interview, we played the video and encouraged participants to discuss their strategies, plans, and preferences, and how these evolved throughout their collaboration. We also requested them to complete two questionnaires regarding their leadership and followership styles. In line with the approved ethics application, we also informed participants about the existence of “*Deception*” elements in the study and sought their consent to use their data. To express our gratitude for their participation, participants received a \$30 gift card.

3.4. Measures

As part of our analysis aimed at assessing participants’ perception of the robot, tasks, and collaboration, we utilized a range of subjective measures, primarily derived from the items in the questionnaires they completed at various phases of the study, as well as insights from interviews. Additionally, we considered objective measures such as the number of tasks allocated to the robot by the human.

To assess participants’ trust in the robot, we employed Muir’s questionnaire [23] (refer to Table 3) using a ten-point Likert scale. We administered this questionnaire at multiple points in the study, including after they watched a video of the robot (to gauge initial trust) and following each collaboration with the robot in Tasks 1, 2, and 3. In addition, participants provided responses to two questions on a 21-point Likert scale, as shown in Table 4. These questions pertained to their self-confidence in completing each task (Tasks 0-3) and their expected helpfulness of the robot (Tasks 1-3). Participants answered these questions after reviewing the patterns but prior to commencing the task of fetching and placing the blocks. High and low values on the Likert scale represented high and low self-confidence, as well as the expected helpfulness of the robot.

Following the completion of each task, participants evaluated their workload, considering factors such as perceived mental demand, physical demand, temporal demand, performance, effort, and frustration. They did so by responding to the NASA-TLX questionnaire, as shown in Table 5, employing a scale ranging from 0 to 100. Additionally, participants completed a questionnaire designed to assess aspects of their working alliance, the traits of the robot teammate, the robot’s performance, and the collaboration experience. The first seven questions in this questionnaire,

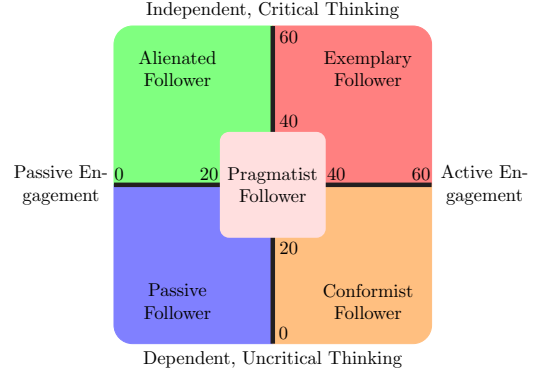


Figure 6: Kelley’s five followership styles

relevant to our study, were adapted from [61]. This questionnaire, as displayed in Table 6, used a 7-point Likert scale, presented in an agree-disagree format.

At the end of the last task (Task 4), participants were asked to complete the user experience questionnaire (UEQ). The UEQ encompasses six paired aspects, each evaluated on a 7-point semantic differential scale: obstructive/supportive, complicated/easy, inefficient/efficient, clear/confusing, boring/exciting, not interesting/interesting, conventional/inventive, and usual/leading edge. Participants also answered a 7-point Likert scale questionnaire about team fluency (i.e., efficient, synchronized collaboration with seamless coordination in shared activities [61]) and their willingness to collaborate again with fetch.

During the online phase of the experiment, we conducted interviews with participants and categorized them into four distinct groups based on their preference for leading or following in collaboration with the robot. These categories included “lead,” “collaborative-lead,” “collaborative-follow,” and “follow.” Additionally, we administered two questionnaires to assess their leadership and followership styles. The leadership style questionnaire utilized a 5-point Likert scale [38]. Scores for each leadership style (authoritarian, democratic, laissez-faire) reflect the degree of inclination toward that style, with potential scores ranging from very low (6-10) to very high (26-30), as well as intermediate levels such as low (11-15), moderate (16-20), and high (21-25). To determine participants’ followership style, we employed a seven-point Likert scale questionnaire from [39], which determines the score for independent thinking and active engagement. Based on participants’ responses, they received scores for independent thinking and active engagement, with scores ranging from 0 to 60. Subsequently, as illustrated in Fig. 6, they were categorized into one of five distinct followership styles.

4. Results & Discussion

We recruited 58 participants from our university. However, we had to omit data from 10 participants due to rea-

²<https://youtu.be/X6Rj0zwQhZ8>

Table 3: Muir’s questionnaire - measuring participants’ trust in the robot [23]

1- To what extent can the robot’s behavior be predicted from moment to moment? (predictability)
2- To what extent can you count on the robot to do its job? (dependability)
3- What degree of confidence do you have that the robot will be able to cope with similar situations in the future? (reliability over time)
4- Overall, how much do you trust the robot? (overall degree of trust)

sons such as a bug in the robot’s program and the robot’s malfunction. Consequently, our data analysis is centered on the remaining 48 participants, comprising 22 females, 24 males, and 2 individuals self-identifying as ‘others,’ with an average age of 24.02 ± 3.93 . The majority were University students (44), while 3 were postdoctoral or visiting researchers, and 1 was a staff member. Initially, we delve into participants’ perception of the tasks, the robot, and collaboration. Subsequently, we shift our focus to the analysis of participants’ actions and performance. The discussion pertaining to the robot’s actions and performance is deferred to another paper centering more explicitly on the planning framework [18].

4.1. Participants’ Perception of the Tasks, Robot, & Collaboration

This part mainly focuses on subjective measurements through questionnaires. Throughout this section, the Kruskal–Wallis H test, a nonparametric statistical test, is applied to ascertain if there are statistically significant variations among two or more groups. If a noteworthy overall difference is detected among multiple groups, the Dunn test is subsequently employed as a post hoc analysis to pinpoint specific group differences. In addition, whenever we state that the analysis is based on the tasks, it implies an examination following the chronological sequence, starting from Task 0 and progressing through Task 3.

4.1.1. Trust

We assessed participants’ trust in the robot at four stages: before collaborating with the robot, after watching a video of the robot (initial trust), and after each collaboration for Tasks 1, 2, and 3. We found a statistically significant difference among these stages ($H(3) = 13.85$, $p = .003$). Fig. 7 illustrates participants’ trust in the robot, representing the average of their responses to questions in Table 3. The figure reveals a notable difference between participants’ initial trust and trust levels after Tasks 1, 2, and 3. Furthermore, it demonstrates a gradual increase in participants’ trust over the sessions, with a significant difference observed between Task 1 and Task 3.

4.1.2. Self-confidence

Participants were also asked about their self-confidence in accomplishing the tasks before collaborating with the robot, but after seeing the pattern they needed to work on. In Task 0, participants were told they had to work alone, without the robot. To answer this question in Task 1, participants knew that the robot would work alongside them, but they had not yet worked with the robot, and it was their first experience. Fig. 8 shows the results for this question, revealing a significant difference in self-confidence across different tasks ($H(3) = 13.85$, $p = .003$). As expected, despite an increase in the average self-confidence of participants in Task 0 and Task 1, there is no significant difference between them, as both tasks involved participants working without prior experience with the Fetch robot. However, the results indicate a significant difference between Task 0 and Task 2, Task 0 and Task 3, as well as Task 1 and Task 3.

4.1.3. Helpfulness

Participants’ self-confidence improvement can be attributed to becoming more familiar with the task and finding the robot more helpful over time. However, to demonstrate that this improvement is not solely due to the former, in Tasks 1, 2, and 3, in addition to inquiring about their self-confidence, we also asked about their expectations regarding the robot’s helpfulness in completing the task. This revealed a significant difference among these responses ($H(2) = 19.76$, $p \ll .001$). As depicted in Fig. 9, participants’ expected helpfulness of the robot increased over time, with a significant difference observed between their expectations in Tasks 1 and 2, as well as Tasks 2 and 3. It’s worth noting that when participants answered this question in Task 1, they had not yet collaborated with the robot. Thus, participants perceived the robot as more helpful after the initial collaboration. However, there was no significant difference in participants’ expected helpfulness of the robot across different patterns.

4.1.4. Relative Trust

We also evaluated participants’ relative trust, i.e., the difference between participants’ trust in the robot and their self-confidence. The result shows no significant difference in participants’ relative trust between Tasks 1, 2, and 3.

4.1.5. Robot’s Performance and Traits

After each task (Tasks 1-3), participants answered the questions in Table 6. In Fig. 10, participants’ scores indicate a positive working alliance and favorable robot teammate traits. We also observe a significant improvement in how participants perceive the robot’s intelligence between Tasks 1 and 2, as well as between Tasks 1 and 3. We asked participants to rate their collaboration with the robot and its performance in terms of reliability, accuracy, and operating speed. Based on Fig. 11, participants found the

Table 4: The questions about the participants’ self-confidence and expected helpfulness of the robot. Asked after seeing the patterns but before starting the task

Q1 (Self-confidence): To what extent are you confident about your ability to successfully accomplish the task?
Q2 (Expected Helpfulness): To what extent do you think Fetch can help you improve your performance?

Table 5: NASA-TLX questionnaire asked after each task

(Mental demand) How mentally demanding was the task?
(Physical demand) How physically demanding was the task?
(Temporal demand) How hurried or rushed was the pace of the task?
(Performance) How successful were you in accomplishing what you were asked to do?
(Effort) How hard did you have to work to accomplish your level of performance?
(Frustration) How insecure, discouraged, irritated, stressed, and annoyed were you?

Table 6: A Questionnaire Assessing Working Alliance, Robot Team-mate Traits, Robot’s Performance, and Collaboration. Q1-Q7 adopted from [61]

Q1: Fetch was intelligent
Q2: Fetch was committed to the task
Q3: Fetch perceived accurately what my goals are
Q4: Fetch did not understand how I wanted to do the task.
Q5: Fetch and I worked towards mutually agreed-upon goals.
Q6: Fetch and I respected each other.
Q7: I feel that Fetch appreciates me.
Q8: Fetch and I collaborated well together.
Q9: Fetch worsened the team’s performance in terms of completion time.
Q10: Fetch improved the team’s performance in terms of accuracy.
Q11: The robot’s decisions were reliable.
Q12: The robot failed to pick up and place objects reliably.

Table 7: The robot and team’s overall performance. Q1-3 are adopted from [61]

1- Our team improved over time.
2- Our team’s fluency improved over time.
3- The robot’s performance improved over time.
4- I will be happy to collaborate again with this robot in similar future tasks

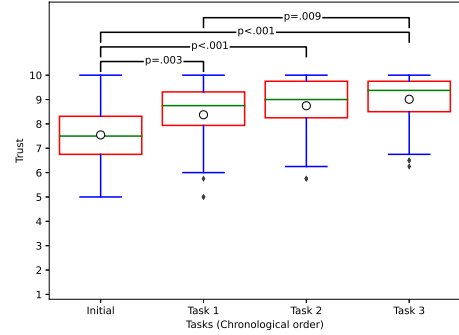


Figure 7: Participants’ trust in the robot before collaboration and after each task

robot’s decisions accurate and reliable, leading to a positive collaboration experience. It’s worth noting that Fetch was noticeably slower than the human agent in pick-and-place tasks, which is reflected in some participants’ scores. However, overall, the average of participants’ scores is low (reverse scale), indicating they believed Fetch did not significantly extend task completion time. This perception aligns with participants’ understanding of the robot’s role in the collaboration, as one participant expressed during the interview: “Mentally, it can compensate for me, but physically, I can compensate for it (Fetch).”

4.1.6. User Experience & Team Fluency

At the end of the in-person phase of the experiment, participants answered the user-experience questionnaire. Fig. 12 shows participants’ scores on this questionnaire, indicating a positive user experience. They also answered four questions in Table 7, and Fig. 13 shows their scores for these four questions. Based on it, participants perceived an improvement in their team performance and fluency and expressed a willingness to collaborate again with Fetch. Despite a high average, participants’ scores regarding the improvement in the robot’s performance cover the whole range of the scale. However, this is reasonable as the robot had the same plan for all three tasks.

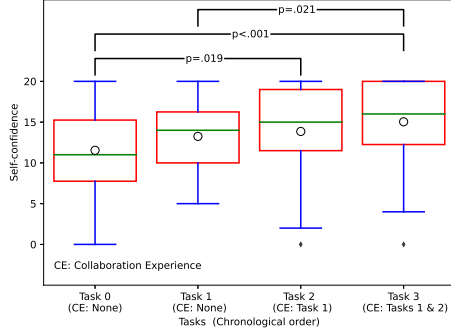


Figure 8: Participants' self-confidence in completing the tasks after seeing the patterns but before starting them based on a) the tasks

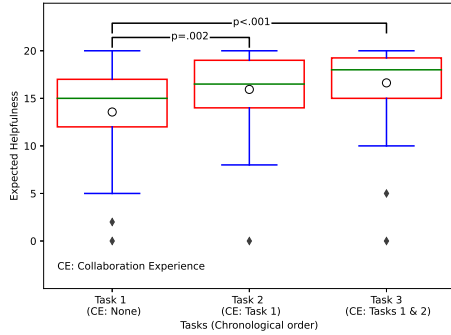


Figure 9: Participants' expected helpfulness from the robot

4.1.7. Perceived Workload

After each task (Tasks 0-3), participants answered the NASA-TLX questionnaire. Hereinafter, we discuss the results for each of the six measures in this questionnaire.

- **Mental Demand:** The mental demand for each task started from the point when participants were given the first pattern (i.e., Patterns A₁, B₁, C₁, and D₁) to memorize until the end of the task where they needed to recall the pattern with the help of the second given pattern (i.e., Patterns A₂, B₂, C₂, and D₂). We analyzed participants' scores for this question based on the tasks and the patterns. Fig. 14a shows the result for Task 0 to Task 3 (in chronological order), with a significant difference between them ($H(3) = 18.05, p < .001$). The results show a significant required mental demand for Task 0 (working alone) compared to Tasks 1, 2, and 3. This indicates that collaboration with the robot could decrease participants' perceived mental demands, supporting part of Hypothesis 4.
- **Physical Demand:** Fig. 14b illustrates the results for physical demand scores based on the chronological order of the tasks. Despite the higher mean value of physical demand for Task 0, and in contrast to Hypothesis 4, we observe no significant difference between Task 0 and Tasks 1, 2, and 3 ($H(3) = 3.35, p = .34$). The reason can be that participants placed most of the blocks on the

table, as the robot was slower than them and, therefore, did not experience a higher physical demand compared to Task 0.

- **Temporal Demand:** Participants were told several times during the experiment that there was no time limit, and we did not measure their time. However, as shown in Fig. 14c, there is a significant difference between them ($H(3) = 9.58, p = .023$), and participants perceived a higher temporal demand in Task 0 compared to Tasks 2 and 3, rejecting part of Hypothesis 4.
- **Performance:** In Tasks 1, 2, and 3, as the robot could fix participants' errors, all participants completed the tasks successfully. Conversely, in Task 0, participants made some mistakes and had no teammate to fix the errors. Therefore, as shown in Fig. 14d, participants significantly perceived themselves as more successful in completing Tasks 1, 2, and 3 compared to Task 0 ($H(3) = 52.75, p \ll .001$). This supports part of Hypothesis 4.
- **Frustration:** The results show that there is a significant difference in perceived frustration in the tasks ($H(3) = 19.25, p < .001$), specifically in Task 0 in comparison with Tasks 1, 2, and 3 (Fig. 14e). However, there is no significant difference in perceived frustration between Tasks 1-3 and Patterns B, C, and D ($H(3) = 0.27, p = .87$). This result also confirms part of Hypothesis 4, which mentions that participants experienced less frustration when collaborating with the robot.
- **Effort:** The results show a significant difference in perceived effort based on the tasks ($H(3) = 15.74, p = .001$) and patterns ($H(3) = 6.89, p = .03$). We can see in Fig. 14f that participants perceived more required effort to accomplish Task 0 in comparison with Tasks 1, 2, and 3. This confirms part of Hypothesis 6, which states collaboration can decrease perceived effort.

4.1.8. Participants' Initial Trust vs. Perception of the Robot

We assessed the correlation between participants' initial trust and their perceptions of the robot's traits, helpfulness, and collaboration (Fig. 15). The findings indicate weak correlations between participants' initial trust and their expected helpfulness of the robot, perceived robot's intelligence, understanding, and respect for them. Additionally, we observed moderate correlations between their initial trust and perceived robot's commitment to the task, understanding of goals, and collaboration.

4.1.9. Task Difficulty

Here, we investigate how task difficulty influenced participants' perception of collaboration, performance, and decision-making.

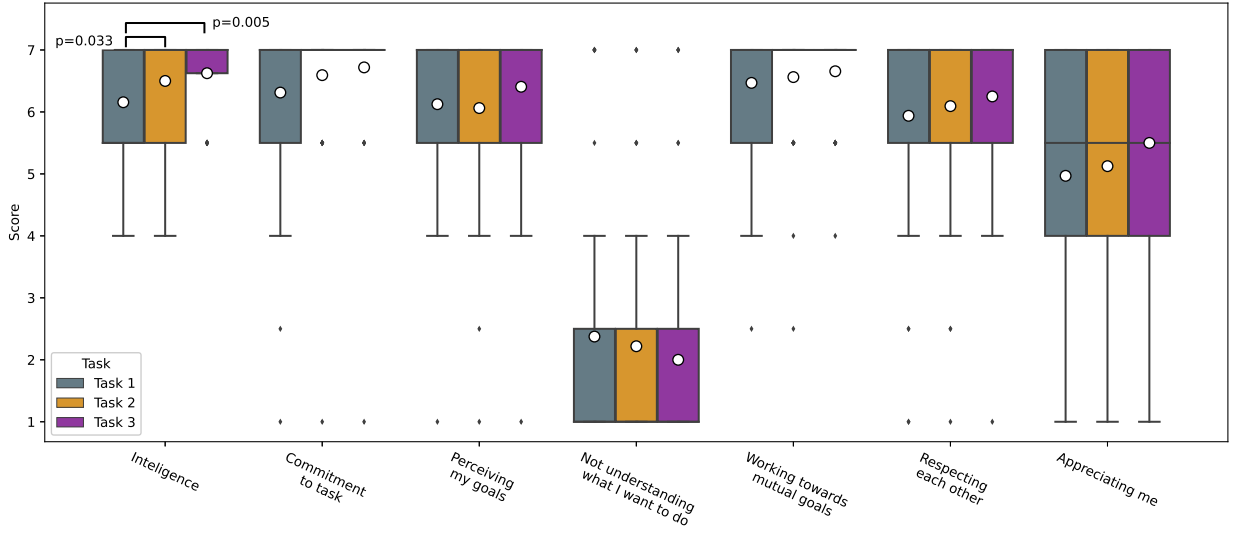


Figure 10: Participants' perception of the robot's traits and working alliance

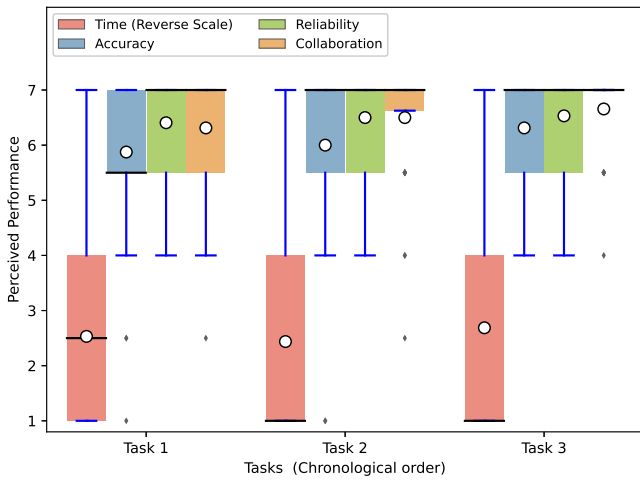


Figure 11: Participants' perception of the robot's performance and collaboration

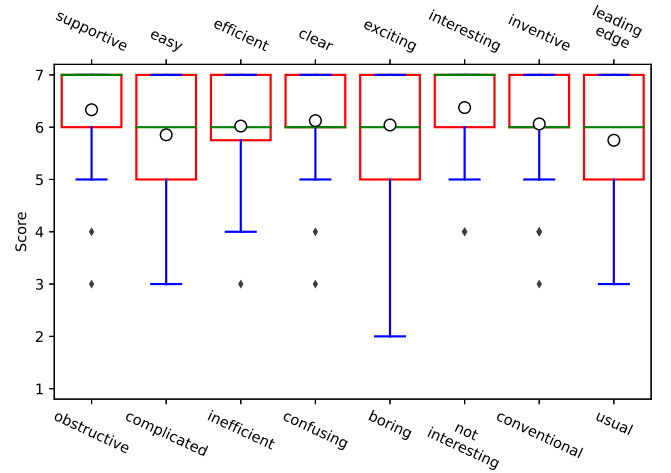


Figure 12: Scores of user experience questionnaire

- Participants Ranking:** At the end of the last task, we asked participants to rank the tasks based on their difficulty level. Fig. 16 shows their rankings based on the patterns used in the tasks. Pattern A was used only in Task 0 when they worked without the robot. Typically, the perceived difficulty of the task depends on the challenge of memorizing the first pattern within 45 seconds and then recalling it while arranging the blocks, with the second given pattern as a hint. The second patterns, Patterns A₂, B₂, C₂, and D₂, had 9, 12, 6, and 9 partially known spots, respectively. However, the number of these unknown spots is not the sole factor in determining their difficulty level; the complexity of memorizing them also plays a crucial role. This is why most participants considered Pattern D₂ with 9 unknown spots simpler than Pattern C₂ with 6 unknown spots because

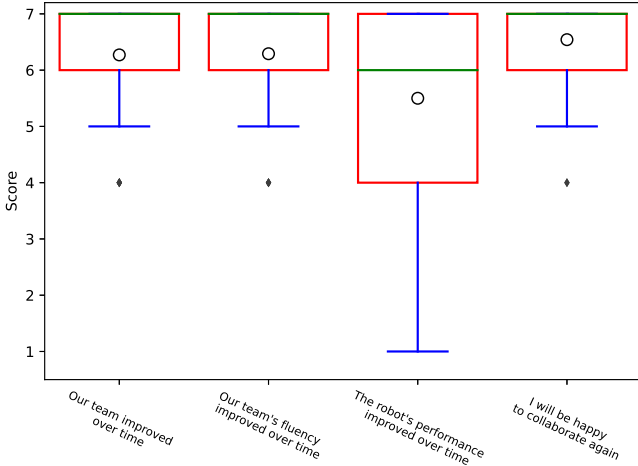


Figure 13: Team fluency and performance

it was more challenging for them to memorize Pattern C₁ compared to Pattern D₁. Participants also found Pattern B to be one of the most challenging patterns, as expected.

- **Self-confidence:** We have also analyzed participants' self-confidence based on the patterns, revealing a significant difference ($H(3) = 24.42, p \ll .001$). Pattern A was consistently used for Task 0 when participants worked alone. Fig. 17c illustrates that participants' self-confidence in accomplishing Pattern 0 (Task 0) is significantly lower than that for Patterns 2 and 3 (No significant difference observed between Patterns 0 and 1). Furthermore, we observed that participants exhibited significantly higher self-confidence in performing Pattern 3 compared to Patterns 1 and 2 (No significant difference was noted between Patterns 1 and 2).
- **Relative Trust:** A significant difference was found when evaluating participants' relative trust for the patterns ($H(2) = 12.9, p = .001$). Fig. 17d shows participants' relative trust based on the patterns, where we can observe a significant difference between Patterns B and D as well as Patterns C and D.
- **Perceived mental demand:** In Fig. 17a, as Pattern A was used only for Task 0, we only considered Patterns B, C, and D, revealing a significant difference ($H(2) = 8.05, p = .019$). The results show that participants perceived Pattern D as less mentally demanding than Patterns B and C.
- **Perceived effort:** Additionally, in Fig. 17b, participants perceived significantly higher effort for completing Pattern B compared to Patterns C and D.

Discussion: The subjective measurements support Hypothesis 1, suggesting that participants' trust in the robot grows over time during the sessions. Additionally,

according to the results, participants' self-confidence improved during the sessions as they collaborated more with the robot. This supports Hypothesis 2. Participants' answers also indicate that they expected more help from the robot during the sessions, supporting Hypothesis 5. The results also partially support Hypothesis 4. As discussed, participants perceived less workload in terms of mental demand, performance, effort, and frustration, supporting Hypothesis 4. Contrary to Hypothesis 4, participants also perceived less temporal demand when collaborating with the robot. However, participants' perceived physical demand did not significantly change when they completed the task in collaboration with the robot. *To summarize, the most important conclusion drawn from this study is that the collaboration could enhance participants' perception of the robot and collaboration and decrease their perceived workload.*

We also evaluated the participants' self-confidence, relative trust, and workload (mental demand and effort) based on the patterns. The results show that, as planned, Pattern B was more challenging for the participants than Patterns C and D. This also supports Hypothesis 3, which states that participants will trust the robot more than their own abilities for challenging tasks.

4.2. Participants' Actions and Performance

Overall preference: In the online phase, we interviewed participants and asked about their strategies and preferences for collaborating with the robot. Probing their inclination towards leading or following, we inquired whether they preferred assigning tasks to the robot or being assigned tasks. Through detailed discussions, we discerned distinct preferences, categorizing participants into four groups: lead, collaborative-lead, collaborative-follow, and follow. The groups "lead" and "follow" represent participants with completely leading or following preferences, and the groups "collaborative-lead" and "collaborative-follow" represent participants who preferred to collaborate while having a slight preference to lead or follow, respectively. Four participants fell under neither of these four groups. One of them preferred to neither lead nor follow, but do his own tasks, and he did not assign any subtask to Fetch and did all subtasks assigned by Fetch. Three other participants quickly began doing subtasks because they didn't want to forget the pattern. They mostly chose subtasks for themselves but were also willing to do subtasks assigned by Fetch. Table 8 shows the number of participants for each group. This indicates that 37 out of 48 participants had more of a leading preference, supporting Hypothesis 6.

Task difficulty: At the end of Task 0, we counted the number of misplaced blocks on the shared table. We also recorded the number of wrong actions by participants, including placing or assigning wrong blocks and returning a correct block from the shared table during the experiment. Regardless of the number of errors, 28, 12, 9, and 3 participants made at least a mistake in respectively Patterns

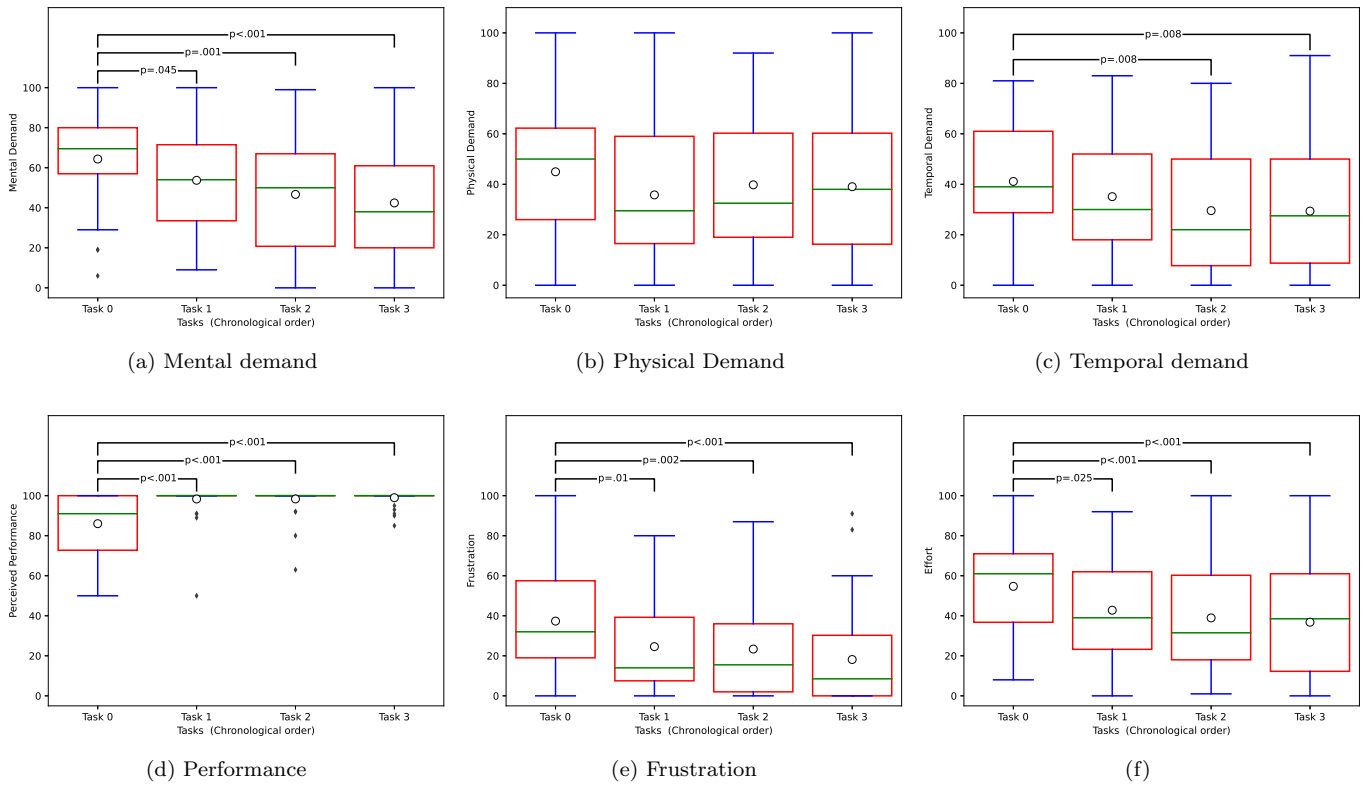


Figure 14: Participants' perceived a) mental, b) physical demand, c) temporal demand, d) performance, e) frustration, and f) effort based on the tasks

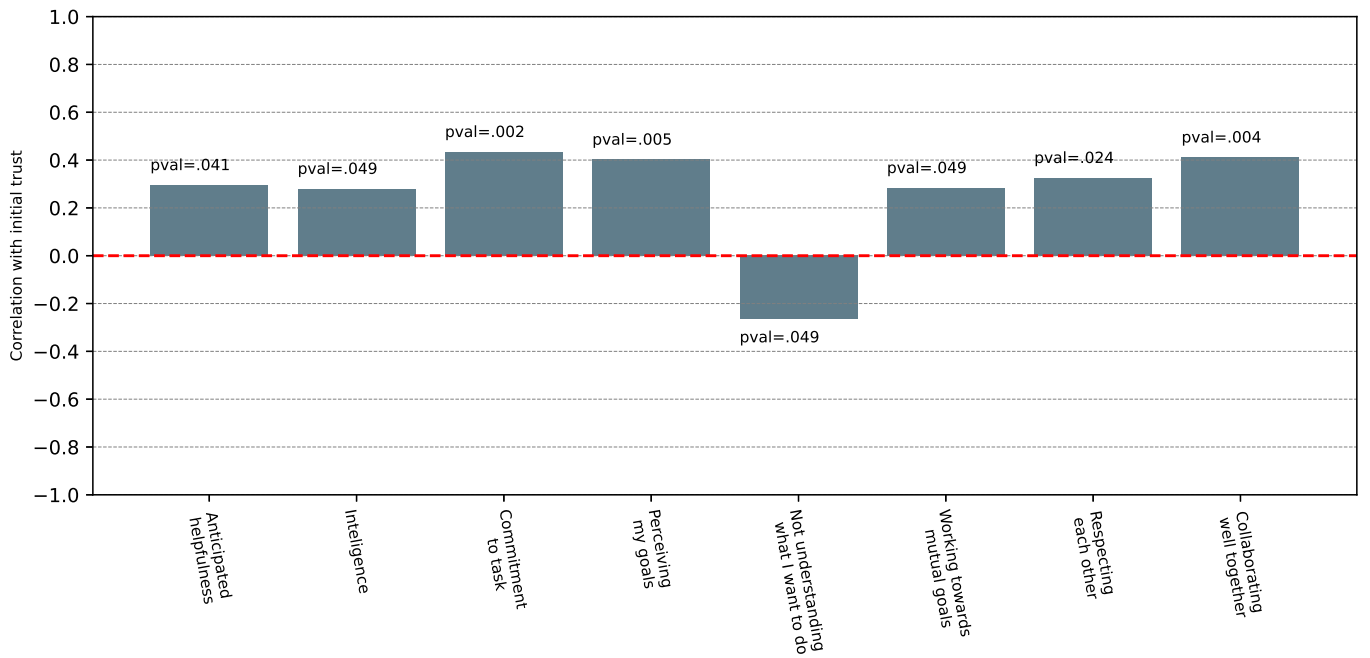


Figure 15: Correlation between participants' initial trust and perceived robot traits

Table 8: Participants’ preference based on interviews.

Preference	Lead	Collaborative-lead	Collaborative-follow	Follow	Other
# Participants	17	20	4	3	4

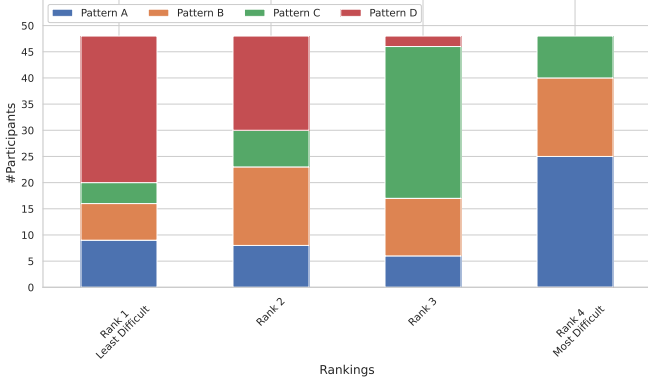


Figure 16: Ranking of task difficulty by participants

A, B, C, and D. Pattern A was used only in Task 0, with the most participants who made mistakes.

Additionally, we counted the number of subtasks that participants assigned to themselves based on the patterns. As depicted in Fig. 18, participants selected significantly fewer tasks for themselves in Pattern B compared to Patterns C and D. This indicates that they allowed the robot to assign more tasks to them. This finding aligns with their perception of task difficulty discussed earlier.

Task assignment: Additionally, we measured the number of subtasks that participants assigned to the robot, revealing a significant difference between the tasks ($H(2) = 11.56, p = .003$). Fig. 19 shows that participants preferred to assign fewer blocks to the robot, and there is a significant difference between Tasks 1 and 3.

Discussion: The results of the interviews show that most of the participants preferred to take on the leading role and have more control over the robot, which aligns with Hypothesis 6. In addition, according to the interviews, participants found the robot to be slower than themselves and preferred to handle more blocks. However, this preference may change if they were faced with a longer task requiring more physical effort. Some participants also pointed out that if the task were longer or the blocks were further apart, they would choose to assign tasks to the robot, even at the cost of completion time, to reduce their physical effort. Furthermore, the higher number of participants who made at least one mistake in Task 0 (Pattern A) compared to the other tasks indicates that the robot could assist them in improving their performance and making fewer errors. Except for Pattern A, which was only used in Task 0, participants made the most mistakes in Pattern B, a fact reflected in the subjective analysis of their perception of the task difficulty. This led them to

rely more on the robot’s decisions and help.

4.3. Leadership and Followership style

The analysis of followership styles reveals that the majority of participants (37 out of 48) are classified as “exemplary” followers, and ten of them fall into the “pragmatist” category. Only one participant exhibited traits of an “alienated” follower, and we did not encounter any instances of passive or conformist followers among the participants. This observation can likely be attributed to their shared academic background, which places a strong emphasis on critical thinking and active engagement.

It is worth highlighting that all four participants listed in Table 8, who neither assumed leadership nor followed the robot, are categorized as “pragmatist” followers. However, due to the limited number of these observations, no definitive conclusion can be drawn, and more investigation is needed.

The distribution of leadership styles is depicted in Fig. 20. The majority of participants exhibited a predominant “democratic leadership” style. Among the four participants with no clear inclination towards leading or following, three displayed a dominant “laissez-faire leadership” style. However, we cannot make conclusive statements based on this limited sample size. The sole participants characterized by an alienated followership style also demonstrated a notably dominant “laissez-faire leadership” style and exhibited a leading-collaborative approach during the collaborative interactions.

We conducted an analysis of the correlation between engagement, critical thinking, and leadership styles in relation to various participants’ factors. Intriguingly, our findings revealed a positive (negative) correlation between the number of tasks assigned by the robot to participants (the number of self-assigned tasks by participants) and their authoritarian leadership style. In other words, participants with authoritarian leadership tendencies were more inclined to accept additional subtasks assigned by the robot while selecting fewer tasks for themselves. Additionally, a moderate positive correlation was identified between participants’ initial trust in the robot and their authoritarian leadership style.

A positive, albeit nonsignificant, correlation was observed between these participants’ following preference and their perceived helpfulness of the robot in relation to their authoritarian leadership style. However, this connection warrants further investigation in future research. We also encountered a nonsignificant negative correlation between participants’ laissez-faire leadership style and the number of tasks assigned by the robot. Moreover, there

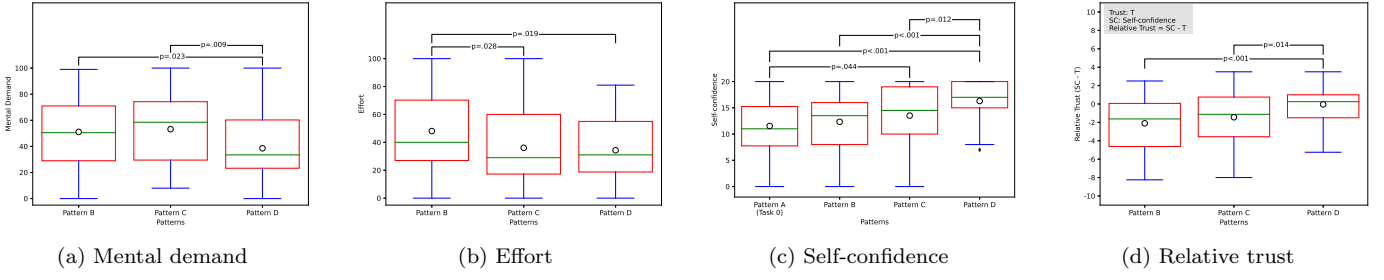


Figure 17: Participants' a and b) perceived mental demand and effort, c) self-confidence, and d) relative trust (self-confidence - trust in the robot) based on the patterns (Patterns B, C, and D)

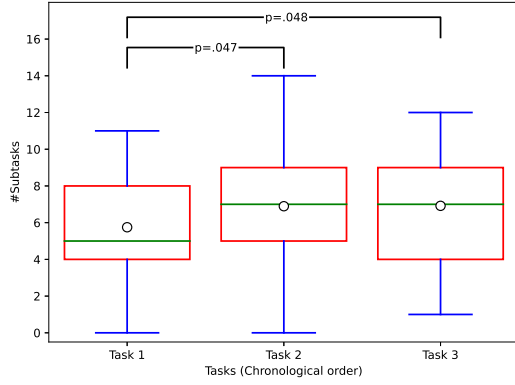


Figure 18: Number of subtasks assigned by participants to themselves in each task

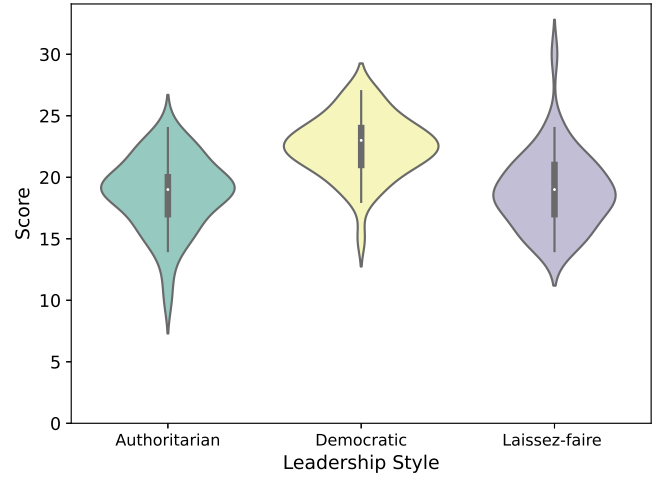


Figure 20: Caption

was a positive correlation between participants' initial trust in the robot and their level of engagement. These correlations are shown in Fig. 21.

5. Conclusion, limitations and future work

Our study explored the potential enhancement of human-robot collaboration efficiency through proactive task planning and allocation. Unlike previous literature, which often neglects either human agents' leading/following preferences or their performance, we centered on achieving a balance between human agent preference and performance, while ensuring a high level of collaboration and positive human perception of the robot.

Trust is an integral part of this collaboration. We assessed how the robot's adaptive and proactive planning can improve participants' trust in the robot. The subjective measurement of participants' trust in the robot confirmed our Hypothesis 1, showing trust enhancement as they collaborated more with the robot. Additionally, we hypothesized that participants' self-confidence would improve over collaboration with the robot, which was confirmed by analyzing their self-reported self-confidence (Hypothesis 2). Furthermore, we observed an increase in the

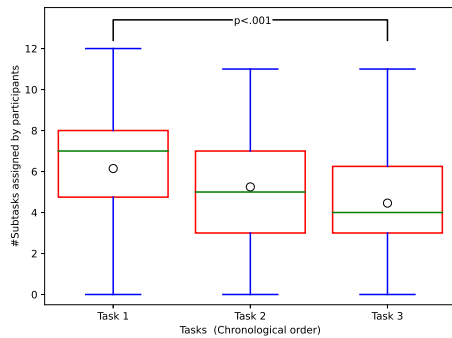


Figure 19: Number of blocks assigned by participants to Fetch in Tasks 1, 2, and 3

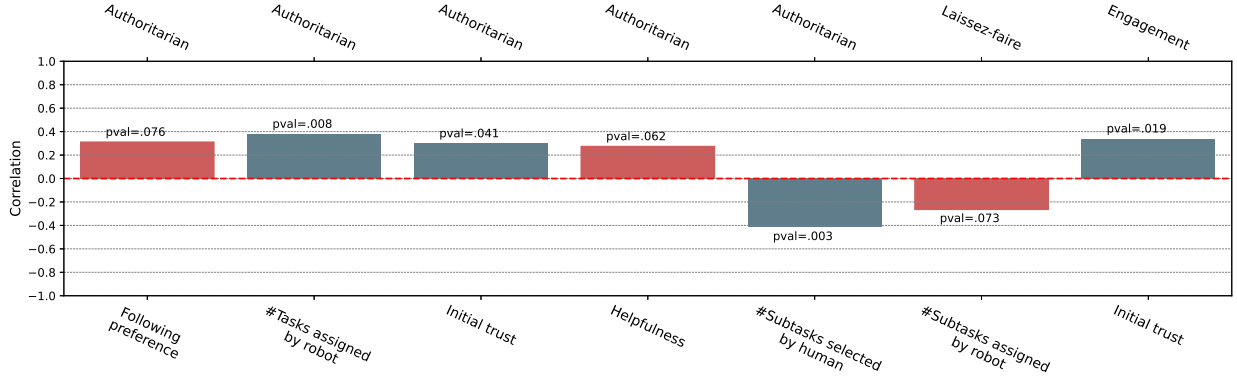


Figure 21: Correlations related to participants' leadership/follower styles

expected helpfulness of the robot by participants, supporting Hypothesis 5. According to participants' answers to the questionnaires, they found the robot to be a teammate with positive traits and a strong working alliance, with whom they were willing to collaborate again in the future.

In support of Hypothesis 3, the results indicated that for more difficult tasks, participants trusted the robot more than their own abilities, which led them to take relatively more following roles. Our proposed task planning method properly inferred this need and provided more help to participants by taking on more leading roles. The robot could also identify when participants struggled to remember the correct patterns and made errors, and accordingly, it fixed their errors and provided more help. This ability of the robot to adapt its task planning led us to hypothesize that participants would perceive, overall, a lower workload (Hypothesis 4). The results partially supported our hypothesis, and we observed an enhancement in participants' perceived mental demand, performance, effort, frustration, and unexpectedly temporal demand. However, contrary to our hypothesis, participants did not perceive a lower physical demand, likely due to the designed scenario focusing more on mental ability than physical abilities.

Based on interviews with participants, we categorized them into four groups: 'lead', 'collaborative-lead', 'collaborative-follow', and 'follow', with the majority falling into the first two categories. This supports Hypothesis 6 that participants, while having a high level of trust in the robot, would prefer to take on more of a leading role and have more control over the collaboration. This finding can guide the design of collaborative scenarios and collaborative robots.

Additionally, the results showed that participants selected fewer tasks for themselves and were allocated more subtasks in tasks that were more difficult since they preferred to follow the robot or made many errors, causing the robot to take back the leading role.

We also inquired about participants' leadership and followership styles using questionnaires. Our observations

revealed that individuals with an authoritarian leadership style exhibited greater trust in the robot. They also allowed the robot to assign more subtasks to them. Additionally, we identified a positive correlation between their initial trust and the level of engagement. Nevertheless, a more comprehensive investigation involving a diverse participant group is warranted to delve deeper into this subject.

5.1. Limitations and Future Work

This study has certain limitations concerning its design and methods. All of our participants were recruited from the University of Waterloo campus and consisted mainly of young adults. However, our ultimate goal is to study working adults in settings such as manufacturing and warehouses. It is conceivable that these two groups hold significantly different expectations and perceptions of a robot teammate. To address this, involving actual working adults in such settings could enhance the practicality of our collaborative scenarios and the robots used. Despite efforts to simulate a working environment by employing an autonomous robot for pick-and-place tasks and creating a setup with features like a conveyor belt, safety equipment, and a graphical user interface, there is room for future enhancements to make the scenarios more closely resemble realistic environments, such as warehouse automation or assembly settings. Moreover, prevalent challenges in manufacturing settings, like sudden changes or unpredictable events typically managed by humans, could be incorporated to better align the study with real-life situations.

In our current study, we assumed the infallibility of the robot's decisions. However, real-world scenarios may involve instances where either the robot or the human agents makes errors, potentially without awareness, and believe their decisions are correct, which can lead to a conflict between two agents. Additionally, our study gauged the human agent's accuracy in selecting block colors as a measure of correctness, a metric easily comprehensible to participants. While practical measures like completion time or travel distance could be introduced, these might not be as

easily understandable and measurable for human agents, potentially leading to conflicts. Numerous other avenues for exploration exist, including a comprehensive investigation into how human leadership and followership styles influence collaboration with a robot.

References

- [1] S. El Zaatari, M. Marei, W. Li, Z. Usman, Cobot programming for collaborative industrial tasks: An overview, *Robotics and Autonomous Systems* 116 (2019) 162–180. doi:<https://doi.org/10.1016/j.robot.2019.03.003>.
- [2] W. P. N. Sotirios Panagou, F. Fruggiero, A scoping review of human robot interaction research towards industry 5.0 human-centric workplaces, *International Journal of Production Research* (2023) 1–17doi:[10.1080/00207543.2023.2172473](https://doi.org/10.1080/00207543.2023.2172473).
- [3] O. Mangin, A. Roncone, B. Scassellati, How to be helpful? supportive behaviors and personalization for human-robot collaboration, *Frontiers in Robotics and AI* 8 (2022). doi:[10.3389/frobt.2021.725780](https://doi.org/10.3389/frobt.2021.725780).
- [4] H. Nemlekar, N. Dhanaraj, A. Guan, S. K. Gupta, S. Nikolaidis, 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*, 2023, p. 575–583. doi:[10.1145/3568162.3576965](https://doi.org/10.1145/3568162.3576965).
- [5] W. Wang, R. Li, Y. Chen, Z. M. Diekel, Y. Jia, Facilitating human-robot collaborative tasks by teaching-learning-collaboration from human demonstrations, *IEEE Transactions on Automation Science and Engineering* 16 (2) (2019) 640–653. doi:[10.1109/TASE.2018.2840345](https://doi.org/10.1109/TASE.2018.2840345).
- [6] E. C. Grigore, A. Roncone, O. Mangin, B. Scassellati, Preference-based assistance prediction for human-robot collaboration tasks, in: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 4441–4448. doi:[10.1109/IROS.2018.8593716](https://doi.org/10.1109/IROS.2018.8593716).
- [7] W. Wang, R. Li, Y. Chen, Y. Sun, Y. Jia, Predicting human intentions in human-robot hand-over tasks through multimodal learning, *IEEE Transactions on Automation Science and Engineering* 19 (3) (2022) 2339–2353. doi:[10.1109/TASE.2021.3074873](https://doi.org/10.1109/TASE.2021.3074873).
- [8] H. Nemlekar, R. Guan, G. Luo, S. K. Gupta, S. Nikolaidis, Towards transferring human preferences from canonical to actual assembly tasks, in: *31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 2022, pp. 1161–1167. doi:[10.1109/RO-MAN53752.2022.9900872](https://doi.org/10.1109/RO-MAN53752.2022.9900872).
- [9] A. Khamis, A. Hussein, A. Elmoggy, Multi-robot Task Allocation: A Review of the State-of-the-Art, Springer International Publishing, 2015, pp. 31–51.
- [10] C. Schmidbauer, S. Zafari, B. Hader, S. Schlund, An empirical study on workers’ preferences in human-robot task assignment in industrial assembly systems, *IEEE Transactions on Human-Machine Systems* 53 (2) (2023) 293–302. doi:[10.1109/THMS.2022.3230667](https://doi.org/10.1109/THMS.2022.3230667).
- [11] A. Pupa, W. Van Dijk, C. Brekelmans, C. Secchi, A resilient and effective task scheduling approach for industrial human-robot collaboration, *Sensors* 22 (13) (2022). doi:[10.3390/s22134901](https://doi.org/10.3390/s22134901).
- [12] Y. Cheng, L. Sun, M. Tomizuka, Human-aware robot task planning based on a hierarchical task model, *IEEE Robotics and Automation Letters* 6 (2) (2021) 1136–1143. doi:[10.1109/LRA.2021.3056370](https://doi.org/10.1109/LRA.2021.3056370).
- [13] E. Lamon, A. De Franco, L. Peternel, A. Ajoudani, A capability-aware role allocation approach to industrial assembly tasks, *IEEE Robotics and Automation Letters* 4 (4) (2019) 3378–3385. doi:[10.1109/LRA.2019.2926963](https://doi.org/10.1109/LRA.2019.2926963).
- [14] K. Darvish, B. Bruno, E. Simetti, F. Mastrogiovanni, G. Casalino, Interleaved online task planning, simulation, task allocation and motion control for flexible human-robot cooperation, in: *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2018, pp. 58–65. doi:[10.1109/ROMAN.2018.8525644](https://doi.org/10.1109/ROMAN.2018.8525644).
- [15] A. Noormohammadi, A. Dahiya, A. M. Aroyo, S. L. Smith, K. Dautenhahn, The effect of robot decision making on human perception of a robot in a collaborative task - a remote study, in: *Proceedings of the 9th International Conference on Human-Agent Interaction*, 2021, p. 423–427. doi:[10.1145/3472307.3484671](https://doi.org/10.1145/3472307.3484671).
- [16] A. Noormohammadi-Asl, A. Ayub, S. L. Smith, K. Dautenhahn, Task selection and planning in human-robot collaborative processes: To be a leader or a follower?, in: *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 2022, pp. 1244–1251. doi:[10.1109/RO-MAN53752.2022.9900770](https://doi.org/10.1109/RO-MAN53752.2022.9900770).
- [17] A. Noormohammadi-Asl, A. Ayub, S. L. Smith, K. Dautenhahn, Adapting to human preferences to lead or follow in human-robot collaboration: A system evaluation, in: *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 2023, pp. 1851–1858. doi:[10.1109/RO-MAN57019.2023.10309328](https://doi.org/10.1109/RO-MAN57019.2023.10309328).
- [18] A. Noormohammadi-Asl, S. L. Smith, K. Dautenhahn, To lead or to follow? adaptive robot task planning in human-robot collaboration (2024).
- [19] T. Sanders, K. E. Oleson, D. R. Billings, J. Y. Chen, P. A. Hancock, A model of human-robot trust: Theoretical model development, in: *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 55, SAGE Publications Sage CA: Los Angeles, CA, 2011, pp. 1432–1436.
- [20] Z. R. Khavas, S. R. Ahmadzadeh, P. Robinette, Modeling trust in human-robot interaction: A survey, in: *International conference on social robotics*, Springer, 2020, pp. 529–541.
- [21] H. Soh, Y. Xie, M. Chen, D. Hsu, Multi-task trust transfer for human-robot interaction, *The International Journal of Robotics Research* 39 (2-3) (2020) 233–249. doi:[10.1177/0278364919866905](https://doi.org/10.1177/0278364919866905).
- [22] M. Chen, S. Nikolaidis, H. Soh, D. Hsu, S. Srinivasa, Planning with trust for human-robot collaboration, in: *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 2018, p. 307–315. doi:[10.1145/3171221.3171264](https://doi.org/10.1145/3171221.3171264).
- [23] B. M. Muir, N. Moray, Trust in automation. part ii. experimental studies of trust and human intervention in a process control simulation, *Ergonomics* 39 (3) (1996) 429–460. doi:[10.1080/00140139608964474](https://doi.org/10.1080/00140139608964474).
- [24] J. D. Lee, N. Moray, Trust, self-confidence, and operators’ adaptation to automation, *International journal of human-computer studies* 40 (1) (1994) 153–184.
- [25] R. Wiczorek, J. Meyer, Effects of trust, self-confidence, and feedback on the use of decision automation, *Frontiers in Psychology* 10 (2019). doi:[10.3389/fpsyg.2019.00519](https://doi.org/10.3389/fpsyg.2019.00519).
- [26] K. J. Williams, M. S. Yuh, N. Jain, A computational model of coupled human trust and self-confidence dynamics, *ACM Transactions on Human-Robot Interaction* 12 (3) (2023). doi:[10.1145/3594715](https://doi.org/10.1145/3594715).
- [27] S. Miller, Workload measures, *National Advanced Driving Simulator*. Iowa City, United States (2001).
- [28] S. G. Hart, L. E. Staveland, Development of NASA-TLX (task load index): Results of empirical and theoretical research, in: *Advances in psychology*, Vol. 52, Elsevier, 1988, pp. 139–183.
- [29] R. G. Freedman, S. J. Levine, B. C. Williams, S. Zilberstein, Helpfulness as a key metric of human-robot collaboration, *arXiv preprint arXiv:2010.04914* (2020).
- [30] B. Shneiderman, P. Maes, Direct manipulation vs. interface agents, *interactions* 4 (6) (1997) 42–61.
- [31] G. Dove, K. Halskov, J. Forlizzi, J. Zimmerman, Ux design innovation: Challenges for working with machine learning as a design material, in: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, p. 278–288. doi:[10.1145/3025453.3025739](https://doi.org/10.1145/3025453.3025739).
- [32] J. Heer, Agency plus automation: Designing artificial intelligence into interactive systems, *Proceedings of the National Academy of Sciences* 116 (6) (2019) 1844–1850. doi:[10.1073/](https://doi.org/10.1073/)

- pnas.1807184115.
- [33] A. Pasparakis, J. De Vries, M. De Koster, In control or under control? human-robot collaboration in warehouse order picking, *Human-Robot Collaboration in Warehouse Order Picking* (March 31, 2021) (2021). doi:10.2139/ssrn.3816533.
 - [34] System-directed picking empower associates to empower associates to be more productive.
URL <https://6river.com/directed-picking/>
 - [35] How robotic picking is revolutionizing warehouse productivity.
URL <https://www.dhl.com/global-en/delivered/digitalization/locus-robotics-robotic-picking.html>
 - [36] A. Chanseau, K. Dautenhahn, K. L. Koay, M. L. Walters, G. Lakatos, M. Salem, How does peoples' perception of control depend on the criticality of a task performed by a robot, *Paladyn, Journal of Behavioral Robotics* 10 (1) (2019) 380–400.
 - [37] Q. Roy, F. Zhang, D. Vogel, Automation accuracy is good, but high controllability may be better, in: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, p. 1–8. doi:10.1145/3290605.3300750.
 - [38] P. G. Northouse, *Introduction to leadership: Concepts and practice*, Sage, 2014.
 - [39] R. E. Kelley, *The power of followership: How to create leaders people want to follow, and followers who lead themselves*, Broadway Business, 1992.
 - [40] C. Schmidbauer, S. Schlund, T. B. Ionescu, B. Hader, Adaptive task sharing in human-robot interaction in assembly, in: *2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 2020, pp. 546–550. doi:10.1109/IEEM45057.2020.9309971.
 - [41] R. Müller, M. Vette, O. Mailahn, Process-oriented task assignment for assembly processes with human-robot interaction, *Procedia CIRP* 44 (2016) 210–215, 6th CIRP Conference on Assembly Technologies and Systems (CATS). doi:<https://doi.org/10.1016/j.procir.2016.02.080>.
 - [42] M. Faroni, A. Umbrico, M. Beschi, A. Orlandini, A. Cesta, N. Pedrocchi, Optimal task and motion planning and execution for multiagent systems in dynamic environments, *IEEE Transactions on Cybernetics* (2023) 1–12doi:10.1109/TCYB.2023.3263380.
 - [43] S. Alirezazadeh, L. A. Alexandre, Dynamic task scheduling for human-robot collaboration, *IEEE Robotics and Automation Letters* 7 (4) (2022) 8699–8704. doi:10.1109/LRA.2022.3188906.
 - [44] O. C. Görür, B. Rosman, F. Sivrikaya, S. Albayrak, FABRIC: A framework for the design and evaluation of collaborative robots with extended human adaptation, *J. Hum.-Robot Interact.* 12 (3) (2023). doi:10.1145/3585276.
 - [45] M. Fiore, A. Clodic, R. Alami, *On Planning and Task Achievement Modalities for Human-Robot Collaboration*, Springer International Publishing, 2016, pp. 293–306.
 - [46] H. Nemlekar, J. Modi, S. K. Gupta, S. Nikolaidis, Two-stage clustering of human preferences for action prediction in assembly tasks, in: *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 3487–3494. doi:10.1109/ICRA48506.2021.9561649.
 - [47] A. Tausch, A. Kluge, The best task allocation process is to decide on one's own: effects of the allocation agent in human-robot interaction on perceived work characteristics and satisfaction, *Cognition, Technology & Work* 24 (1) (2022) 39–55.
 - [48] M. Gombolay, A. Bair, C. Huang, J. Shah, Computational design of mixed-initiative human-robot teaming that considers human factors: situational awareness, workload, and workflow preferences, *The International Journal of Robotics Research* 36 (5-7) (2017) 597–617. doi:10.1177/0278364916688255.
 - [49] B. Akgun, M. Cakmak, K. Jiang, A. L. Thomaz, Keyframe-based learning from demonstration: Method and evaluation, *International Journal of Social Robotics* 4 (2012) 343–355.
 - [50] J. Huang, D. Fox, M. Cakmak, Synthesizing robot manipulation programs from a single observed human demonstration, in: *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 4585–4592. doi:10.1109/IROS40897.2019.8968543.
 - [51] V. V. Unhelkar, *Effective information sharing for human-robot collaboration*, Ph.D. thesis, MIT (2020).
 - [52] A. Bajcsy, D. P. Losey, M. K. O'Malley, A. D. Dragan, Learning from physical human corrections, one feature at a time, in: *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 2018, p. 141–149. doi:10.1145/3171221.3171267.
 - [53] J. Ayoub, L. Avetisian, X. J. Yang, F. Zhou, Real-time trust prediction in conditionally automated driving using physiological measures, *IEEE Transactions on Intelligent Transportation Systems* (2023) 1–9doi:10.1109/TITS.2023.3295783.
 - [54] C. Huang, W. Luo, R. Liu, Meta preference learning for fast user adaptation in human-supervisory multi-robot deployments, in: *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021, pp. 5851–5856. doi:10.1109/IROS51168.2021.9636515.
 - [55] S. Reddy, S. Levine, A. Dragan, First contact: Unsupervised human-machine co-adaptation via mutual information maximization, in: S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, A. Oh (Eds.), *Advances in Neural Information Processing Systems*, Vol. 35, Curran Associates, Inc., 2022, pp. 31542–31556.
 - [56] S. Nikolaidis, D. Hsu, S. Srinivasa, Human-robot mutual adaptation in collaborative tasks: Models and experiments, *The International Journal of Robotics Research* 36 (5-7) (2017) 618–634. doi:10.1177/0278364917690593.
 - [57] V. Krueger, F. Rovida, B. Grossmann, R. Petrick, M. Crosby, A. Charzoule, G. Martin Garcia, S. Behnke, C. Toscano, G. Veiga, Testing the vertical and cyber-physical integration of cognitive robots in manufacturing, *Robotics and Computer-Integrated Manufacturing* 57 (2019) 213–229. doi:<https://doi.org/10.1016/j.rcim.2018.11.011>.
 - [58] M. Wise, M. Ferguson, D. King, E. Diehr, D. Dymesich, Fetch and freight: Standard platforms for service robot applications, in: *Workshop on autonomous mobile service robots*, 2016, pp. 1–6.
 - [59] B. Laugwitz, T. Held, M. Schrepp, Construction and evaluation of a user experience questionnaire, in: A. Holzinger (Ed.), *HCI and Usability for Education and Work*, Springer Berlin Heidelberg, 2008, pp. 63–76.
 - [60] M. Schrepp, J. Thomaschewski, A. Hinderks, Design and evaluation of a short version of the user experience questionnaire (ueq-s), *International Journal of Interactive Multimedia and Artificial Intelligence* 4 (6) (2017) 103–108. doi:10.9781/ijimai.2017.09.001.
 - [61] G. Hoffman, Evaluating fluency in human-robot collaboration, *IEEE Transactions on Human-Machine Systems* 49 (3) (2019) 209–218. doi:10.1109/THMS.2019.2904558.