The Effect of Robot Skill Level and Communication in Rapid, Proximate Human-Robot Collaboration

Kin Man Lee*
klee863@gatech.edu
Georgia Institute of Technology
Atlanta, Georgia, USA

Arjun Krishna*
akrishna49@gatech.edu
Georgia Institute of Technology
Atlanta, Georgia, USA

Zulfiqar Zaidi
zzaidi8@gatech.edu
Georgia Institute of Technology
Atlanta, Georgia, USA

Rohan Paleja
rpaleja3@gatech.edu
Georgia Institute of Technology
Atlanta, Georgia, USA

Letian Chen
letian.chen@gatech.edu
Georgia Institute of Technology
Atlanta, Georgia, USA

Erin Hedlund-Botti
erin.botti@gatech.edu
Georgia Institute of Technology
Atlanta, Georgia, USA

Mariah Schrum
mschrum3@gatech.edu
Georgia Institute of Technology
Atlanta, Georgia, USA

Matthew Gombolay
matthew.gombolay@cc.gatech.edu
Georgia Institute of Technology
Atlanta, Georgia, USA

ABSTRACT

As high-speed, agile robots become more commonplace, these robots will have the potential to better aid and collaborate with humans. However, due to the increased agility and functionality of these robots, close collaboration with humans can create safety concerns that alter team dynamics and degrade task performance. In this work, we aim to enable the deployment of safe and trustworthy agile robots that operate in proximity with humans. We do so by 1) Proposing a novel human-robot doubles table tennis scenario to serve as a testbed for studying agile, proximate human-robot collaboration and 2) Conducting a user-study to understand how attributes of the robot (e.g., robot competency or capacity to communicate) impact team dynamics, perceived safety, and perceived trust, and how these latent factors affect human-robot collaboration (HRC) performance. We find that robot competency significantly increases perceived trust ($p < .001$), extending skill-to-trust assessments in prior studies to agile, proximate HRC. Furthermore, interestingly, we find that when the robot vocalizes its intention to perform a task, it results in a significant decrease in team performance ($p = .037$) and perceived safety of the system ($p = .009$).

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in collaborative and social computing**; • **Computer systems organization** → Robotics.

---

*Both authors contributed equally to this research.

KEYWORDS

1 INTRODUCTION

The deployment of collaborative robots (i.e., "cobots") holds the promise of increasing productivity and enhancing safety. These robots are expected to collaborate in proximity with human workers as opposed to the caged robots typically encountered within today's manufacturing industry [67]. This paradigm shift allows for cobots to work with human workers in new ways, such as mixed-initiative teaming [29, 35, 47], to increase productivity. In mixed-initiative settings, human workers and robots are co-located in a shared workspace and must dynamically decide who should accomplish shared tasks. This online reasoning in human-robot collaboration provides a more natural and fluent interaction scheme than a strict, pre-planned allocation of work [2].

However, achieving fluency in mixed-initiative teaming is challenging as identifying precisely who should accomplish a specific task is computationally challenging for human and robot alike [55], and inferring human intent online from dynamic motion remains an open research problem requiring domain expertise for algorithm design [54]. Enabling robots to convey intent is likewise an unsolved design problem [23, 50]. Furthermore, as robots gain functionality and increased agility, close collaboration can create safety concerns that may alter the team dynamic, affecting human behavior and performance due to latent variables, such as perceived safety and trust [43, 44]. In this work, we are interested in understanding how attributes of the robot (e.g., skill-level or capacity to communicate) impact team dynamics, perceived safety, and perceived trust, and

how these latent factors affect human-robot collaboration (HRC) performance. Insight into these relationships will inform how to best design robots that work in proximity to humans.

We are specifically interested in the development of agile robotics for human-robot collaboration. Agile robots are those that are, by definition, capable of demonstrating agile behaviors in dynamic settings [21, 70]. These robots are designed to closely emulate human-level abilities across tasks such as sports [8, 70], cooking [27, 38, 64], and mobile manipulation [41, 51]. Agile robotics holds the promise of low-latency decision-making and high-speed robot maneuvers. Collaboration in proximity to humans presents a novel challenge for agile robotics, as we would like to take advantage of high-speed manipulation for effective human-robot collaboration.

Fast moving collaborative robots in proximity to humans could potentially induce fear, anxiety, and stress because the robot's intention may be unclear, the behaviors exhibited by the robot could be unpredictable, and the risk of the human coming to harm is objectively higher [3, 6, 48, 58]. Additionally, humans have less time to react, exacerbating the concerns of safety and trust [3, 11, 60, 63].

In this work, we choose the domain of table-tennis doubles for our human-subjects experiment in order to gain insights into the design of agile, collaborative robots. Here, we do not follow the typical table-tennis doubles rule, where agents must alternate hitting the ball; instead, we frame our domain as a mixed-initiative human-robot teaming task (i.e., either player can return the ball). Our modified game introduces the opportunity for players to reason about who should hit the ball, whether and how to communicate intent, and a possible physical collision if the human and robot are not in sync. Our task meets the following criterion: (1) Low-reaction time decisions from humans and robots, and (2) Shared collaboration space where either agent can successfully accomplish the task. In this environment, we evaluate how an agile robot, varying in skill level and endowed with a deliberative voice-communication channel is able to team with human end-users. We place specific emphasis on understanding how the robot's competence, voice profile, and the capacity to communicate affect perceived safety, and trust in agile, proximate human-robot collaboration. We note that we include as an independent variable the sex of the voice (i.e., male vs. female) due to (conflicting) prior work in human factors on sex-binary voices for cockpit automation indicating that the sound of female vs. male voice can have differing impacts on audibility, perceived assertiveness, and reaction time and quality [4, 24].

While we study these factors in an *ad hoc* table-tennis doubles task, the outcomes of this study are potentially applicable to settings where humans and agile-robots have to work together in a shared space. Applications include collaborative manufacturing robots, assistive home robots, athletic robots for sports and recreation, etc. Our work is the first to explore these factors in an athletic collaborative human-robot task and we hope this work inspires more research to enable the deployment of safe and trustworthy agile robots that operate in proximity with humans. We summarize our key contributions as follows:

(1) We develop a novel table tennis system for safe and dynamic HRC by extending probabilistic movement primitives [30] for motor control and incorporating a safety subsystem that uses unintrusive, camera-based human pose tracking.
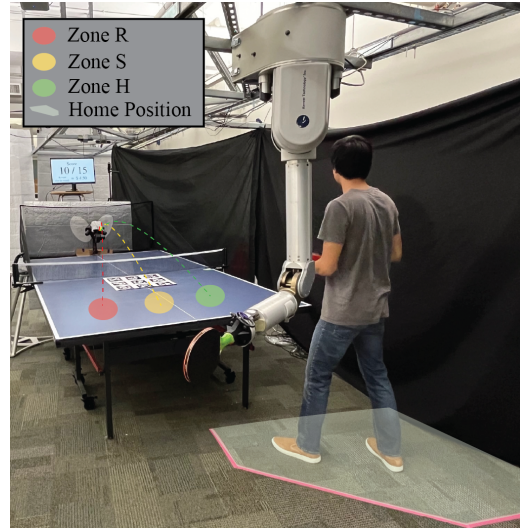

Figure 1: Setup of the collaborative table-tennis robot system.

(2) We conduct a user study and find that fast, verbal communication to convey intent reduces overall team performance ($p = .037$) and perceived safety of the robot ($p = .009$).

(3) We find that participants rarely utilized verbal signals to convey intent to the robot. Based on a small sample size of user feedback, this suggests that most felt more comfortable communicating to the robot implicitly with body motion in high-speed teaming tasks.

2 RELATED WORK

Proximity in HRC – Cobots are envisioned to be in "cage-free" environments, sharing a workspace and collaborating closely with human workers [25]. As these cage-free systems can prove detrimental due to the potential for collision, several works have attempted to create safe human-robot collaboration via algorithmic advancements [39, 45, 52], designing different manipulation techniques to stop the robot from causing bodily harm in alignment with industry guidelines [57]. Another approach has been to design compliant robots to ensure safety in proximate HRC settings [32, 59, 73]. Other work has looked into human-in-the-loop approaches that attempt to understand human intent for safe, proximate human-robot collaboration [28]. However, while such techniques can produce safe behavior and address planning within a mixed-initiative teaming paradigm, these approaches do not address the perceived safety or trust of end-users, which are important factors in achieving high-performance HRC [9]. Works in human factors have attempted to measure perceived safety through physiological signals that can indicate stress, fear, or anxiety [18, 53, 69]. Some of this work has been applied to industrial manipulators [3, 40], though none have been applied in a fast-paced HRC setting to the best of our knowledge.

Communication in HRC – Prior work displayed that verbal communication in HRC affects collaboration as well as users' perception of the robot [31, 46, 66]. Studies have shown that verbal communication between the robot and the user improves subjective and objective metrics on team performance and users prefer a speaking robot over a non-speaking one [42, 49, 61, 62]. Furthermore, [19] found
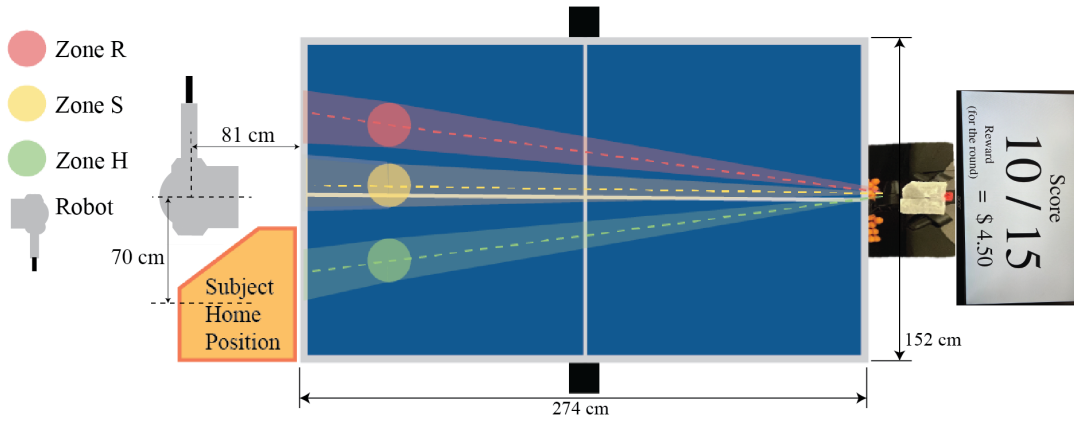
Figure 2: Top-down view of the task setup. Balls land in three zones: Robot (Zone R), Shared (Zone S), and Human (Zone H). The scoreboard on the right is updated in real-time.

that autonomous agents who exchanged information regarding their intention (i.e., deliberative communication [10]) collaborated with humans better than those sharing information in regard to the world state. It has also been suggested that a lack of truthfulness in a robot's statement can negatively affect a human's evaluation of a robot's abilities [49]. Although these studies produced interesting findings about robot communication, the corresponding tasks did not have time pressure, deviating from real-time decision-making.

In our experiment, the robot asserts its intention via a short verbal statement in a male/female voice to inform the participant that it will attempt to hit an incoming table tennis ball. The verbal communication is done in real-time by an agile robot conducting a high-speed striking maneuver. Deviating from prior literature, we additionally allow the human to communicate intent with the robot (bidirectional communication), creating a realistic, rapid, proximate HRC scenario where the human must reason about received communication, whether to communicate intent, which agent should hit the ball, and the ball trajectory, all within a short duration.

Agile Robot Manipulation – Agile robotic systems that work alongside human beings is a long-standing goal in manufacturing, where high-speed robotic maneuvers are safely conducted with humans collaborating in proximity with the robot. In this way, manufacturing is able to take advantage of a human's ability to deal with uncertainty and variability while utilizing automation to its fullest [25]. However, prior work has noted that stress, anxiety, and fear levels of a human user around a fast-moving robot with the capability of causing physical harm are high [3, 6, 11, 48, 58, 60, 63], presenting interesting challenges in producing high-performance HRC. Furthermore, as producing high-speed robotic maneuvers is a difficult challenge in itself, poor performance of the robot can adversely affect user trust [20, 33, 56, 71]. As such, to address these open problems in studying agile cobots, we study both the effects of varying skill level and communication on collaboration performance between humans and agile robot.

We utilize table tennis as a testbed for HRC, as this is a sport requiring fast reflexes and decision-making [13]. In prior work, robotic systems have been developed to play singles table tennis [8, 14–16, 30]. However, to the best of our knowledge, systems for doubles table tennis have not been explored.

3 METHODOLOGY

3.1 Task Description

The participant's task is to team-up with a 7 degree-of-freedom Barrett WAM arm to return table tennis balls launched from a Butterfly Amicus ball launcher, as shown in Figures 1 and 2. To investigate the collaborative behavior between the human and the robot, we program the launcher to project balls to three regions: Zone R (the robot zone), Zone S (the shared zone), and Zone H (the human zone). The robot has the capability to return balls landing in Zone R and Zone S, and does not attempt to return the balls landing in Zone H. The human is free to return balls in any region, but those in Zone R are more difficult as those balls are farther from the subject's starting position (displayed in Figures 1 and 2).

The task of returning the balls in Zone S is designed to be a *shared task* to evaluate the collaboration between the human and the robot. In this region, both agents (robot and participant) can comfortably attempt the task. Note, we modify the standard doubles table tennis rules for our experiment such that the two players (i.e., the participant and the robot) do not have to alternate, and the human can decide which balls to attempt. The participants must start in the "subject home position" (shown as shaded region in Figures 1 and 2) before the ball is launched to ensure consistent starting configuration and difficulty to reach Zone R.

The setup described meets the two criteria we set: 1) Low-reaction time in which the human has ≈ 0.6 seconds to make a decision on whether to attempt to return a ball or yield it to the robot. 2) Shared Human-Robot Collaboration space following the definition of mixed-initiative teaming, where the interaction strategy is flexible, dynamic, and based on agents contributing to the task that they can do best [1]. In our scenario, the subject is incentivized to rely on the robot for balls in Zone R due to the physical distance and share the Zone S with the robot. The participant score is displayed on a real-time scoreboard (Figure 2); further details are in Section 3.5.

3.2 Experiment Design

In our study, we seek to explore the impact and interplay of voice communication and robot skill level on team performance, human perceived safety, and human trust towards the robot. As such,

within our collaborative task detailed in Section 3.1, we design a study with three factors: (1) *Communication*: enabled or disabled, (2) *Robot Voice Sex*: male or female, (3) *Robot Skill*: low and high. We propose a mixed-factorial study with a 2x2x2 design, with Communication and Robot Skill as within-subject factors and Robot Voice Sex as a between-subjects factor. While a complete within-subjects design is ideal for exposing participants to all conditions, we chose to investigate robot voice sex as a between-subjects factor to prevent participants from being aware that we are investigating sex/gender bias. We detail the three independent variables below.

Communication: Drawing inspiration from how volleyball players communicate deliberative intent, we enable the robot to vocally communicate assertive audio signals such as "*Mine*" or "*Got it*" through a speaker mounted on the robot arm. The utilization of short phrases is a requirement in agile robotics as intent must be quickly conveyed due to the low-duration task. We define these Vocal, Assertive, INtention communication signals as "VAIN" signals for clarity and brevity on the form of communication used in our study. VAIN signals are provided in real-time as soon as a launched ball is detected by the vision system to indicate that the robot will attempt to return the ball. VAIN communication, when enabled, is used for all balls landing in both Zone *R* and Zone *S*.

Robot Voice Sex: We use a voice generation tool to create audio files for the phrases used in the communication factor with both a male and female tone to study whether/how the communication voice impacts the human-robot team's performance and the human's subjective perception of trust and safety. As noted, prior work has suggested that the sound of female vs. male voice can have differing impacts on audibility, perceived assertiveness, reaction time, and quality [4, 24].

Robot Skill: We have two levels of robot performance - low and high. The high performance model is comprised of a well-tuned stroke primitive (detailed in Section 3.3.2) to return balls. The low performance model replaces the well-tuned primitive with a sub-optimal primitive ($\sim$ 50% of the time), which either misses the ball or does not successfully return the ball over the net. The success rates of the robot policies are summarized in Table 1.

Prior work has shown that robot skill has a positive correlation with trust on the robot [20, 33, 56, 71]. However, most findings have come from cognitive or physical tasks in virtual environments. Additionally, the impact of a robot's skill on perceived safety in HRC is under-explored. Exploring these topics in a setting with a physical, agile robot on a fast-paced task will lead to insights that can help in designing more effective cobots.

Below, we detail our research questions under the following themes: Robot Skill and Deliberative Communication.

3.2.1 Robot Skill. As skill is a critical factor in establishing effective collaboration, we are interested in answering the following questions pertaining to this factor through our study:

(**RQ1**) How does the robot's skill level impact team performance?
(**RQ2**) How does the robot's skill level impact the perceived safety of human collaborators?
(**RQ3**) How does the robot's skill level impact human collaborators' trust on the robot?

(**RQ4**) How does the robot's skill level impact human initiative in shared regions of the task?

3.2.2 VAIN Communication. We are interested in evaluating how Vocal, Assertive, INtention (VAIN) communication (i.e., a transparent robot that always announces its intention) during robotic table tennis motions affect the dynamics of human-robot collaboration. To this end we ask the following questions:

(**RQ5**) Does VAIN communication help increase HRC team performance?
(**RQ6**) Does VAIN communication increase perceived safety of the robot in HRC?
(**RQ7**) Does communication increase people's trust on the robot in HRC?
(**RQ8**) Does the sex of the robot's voice impact the perceived safety and trust of the robot?
(**RQ9**) Does VAIN communication increase the robot's chance of taking initiative within the shared regions of task?

Stroke Type	Skill Level	Hit Rate	Success Rate
Zone *R*	high	94.36%	86.40%
Zone *R*	low	93.66%	59.86%
Zone *S*	high	86.54%	64.63%
Zone *S*	low	83.89%	30.29%

Table 1: Performance of robot strokes on balls that land in Zone *R* and Zone *S* for the different skill levels. "Hit" denotes the robot successfully hit the ball with the paddle, and "Success" represents the ball is successfully returned over the net.

3.3 System Design

Our robotic system has four main components to achieve various levels of robot skills and ensure safety, as shown in Figure 3. First, Ball Detection enables the robotic system to "see" and predict the ball trajectory in order to decide when to hit. Second, if the robot is able to strike the incoming ball (i.e., the ball must be in Zone *R* or Zone *S*), the Ball Strike Controller utilizes ProMP [30] policies to execute the corresponding pre-trained strikes for the zone. Third, Human Pose Estimation detects the 3D pose of the human. Lastly, Safe Execution determines if any human pose keypoints have entered the robot safety region, which leads to an immediate termination of the robot trajectory to avoid any possible collision between the robot and the human.

3.3.1 Ball Tracking. To accurately detect the table tennis balls, we utilize a multi-camera vision system consisting of three Stereolabs ZED2 cameras. Classical computer vision techniques, including background subtraction and color thresholding (from OpenCV [7]), are applied to identify the ball pixels within the image. With this, an estimate of the ball position in the world frame is calculated based on each camera's relative position to an AprilTag [68] on the ping pong table. All three ball poses are merged by an Extended Kalman Filter (EKF) [36, 37] to produce an accurate fused estimate of the ball. A dynamics model that factors in drag, gravity, and the coefficient of restitution of bounce is used to rollout the EKF's expected future ball trajectory, which is then utilized by the Ball Strike Controller to determine the right time to execute a stroke.
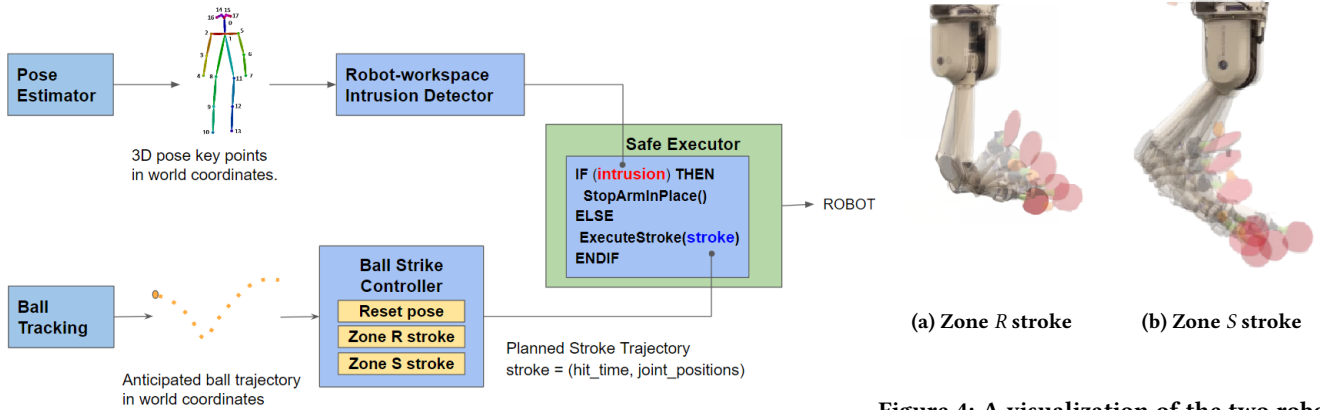
Figure 3: Overview of our collaborative table-tennis robot system.

(a) Zone R stroke   **(b) Zone S stroke**

Figure 4: A visualization of the two robot stroke primitives that hits the balls that land in Zone R and Zone S.

3.3.2 Ball Strike Controller. We trained two probabilistic movement primitives (ProMPs) [30] – one for Zone R and Zone S – using ten kinesthetic demonstrations of successful returns for each. We model the joint position trajectories of the stroke as a weighted activation over 50 radial basis functions in the phase space $[0, 1]$ and learn the weight distribution that maximizes the likelihood of the demonstrated trajectories through the Expectation-Maximization (EM) procedure outlined in Gomez-Gonzalez et al. [30]. The trained stroke primitives are visualized in Figure 4. While the learned primitive weights can be conditioned to pass through waypoints based on the ball trajectories, we execute the unconditioned mean of the learned primitive due to the possible unsafe waypoints of a conditioned ProMP. Doing so effectively ensures that the stroke performed by the robot for each zone is pre-determined, thereby keeping the robot's behavior predictable and safe for the purposes of the study.

During our experimentation, the sequence of ball launch locations is predetermined, providing the robot with access to information about which region each ball will land. To avoid the participant from obtaining information about incoming ball trajectories, for each strike, the robot starts at a common reset pose and moves to the initial pose of the primitive only once a ball is detected by the Ball Tracking component. The common reset pose was determined to be a weighted average of the initial poses of the Zone R and Zone S strokes. Once a ball is detected, the system utilizes the anticipated ball trajectory to calculate the time of impact and the approximate paddle position at the time of impact, which is used to trigger an agile robot table tennis return.

The stroke parameters, such as duration and hit phase were tuned to return as many balls in each zone as possible. These strokes serve as the *high* skill model of the robot. In order to create a *low* skill model of the robot, we introduce mistiming to the strokes 50% of the time. This offsetting during the hit phase causes the robot to poorly return the ball or completely miss. We report the performance of the robot using these two skill models over all the ball sequences considered in the experiment trial in Table 1.

3.3.3 Human Pose Estimation. We utilize a ZED2 camera for tracking human pose. The ZED software development kit provides access to a body tracking module that outputs skeleton-based human keypoints. These keypoints indicate the locations of major joints within the human body such as the head, shoulders, elbows, hands, knees, and feet. We use the 18-point 3D pose format (OpenPose [12] pose format) in our system to track the pose of participants. Each 3D keypoint consists of an X, Y, and Z position of a human joint in the camera frame. These keypoints are then transformed to world coordinates relative to an AprilTag [68] on the table, which is used in the safe execution module to avoid the robot colliding with humans.

3.3.4 Safe Execution. All control requests to the robot are routed through the safe execution module as shown in Figure 3. A fixed 3D unsafe region is created around the robot based on the reachability range of the robot strokes calculated by the Ball Strike Controller. The skeleton keypoints from the Human Pose Estimation module are passed as input into the safe execution module to detect if any part of the participant's body enters the unsafe region. If so, any robot motion in progress is immediately terminated by sending a stop command directly to the low-level controller and all further control requests are rejected. Once keypoints are no longer detected within the unsafe region (i.e., the human is outside of the robot striking area), control requests are accepted again and the robot is reset to its default joint positions. Additionally, the maximum end-effector speed of any stroke is limited to 3.5 m/s to ensure safety even if the unlikely scenario of a collision were to occur.

3.4 Metrics

3.4.1 Objective Metrics. The team performance metric is the sum of successful returns both the robot and participant achieve together in a randomized sequence of 30 balls. This randomized sequence consists of 8 balls in Zone H, 14 in Zone S, and 8 in Zone R. In order to control the impact of the participants' prior table tennis skill to the overall performance score, we measure the participant's skill level in performing the described task in a calibration phase where the participant returns a sequence of 30 balls on their own. We record for each ball, the zone in which it landed, whether it was hit, and if the ball was returned successfully. Similarly, in the main experiment, we record for each ball the landing zone, who attempted to hit it, whether it was hit, and if the return was successful.

As an additional dependent variable, we create a bidirectional communication channel by allowing the subject to speak the phrases ("Mine", "Got it") to convey their intent of returning the ball. Here, we observe if robot skill level or robot communication incentivized the human to communicate their intent. Note we explicitly state that participant communication towards the robot is optional and not required for the robot to stop its stroke. The robot always yields given participant communication, responding with either "*Okay*" or "*Alright*". Due to the latency of speech recognition and the limited timing allowed in our task, we choose a Wizard-of-Oz approach [17] to provide a timely response to the participant's declaration of intent - a button is pressed by the experimenter to trigger immediate stopping of the robot movement and provide voice responses.

3.4.2 Surveys. We conduct a pre-experiment questionnaire that includes self-reported demographics - age, gender, and dominant hand in playing table-tennis, mini-IPIP[22] to measure the big-five personality factors, and negative attitude towards robots (NARS)[65] to measure the subject's predisposition towards robot interactions. Each of these serve as potential confounds within our analysis. After each within-subject condition (described in depth within Section 3.5), the subject is asked to fill out a questionnaire assessing their interaction with the robot. This questionnaire includes the Godspeed questionnaire [5], measuring anthropomorphism, likeability, perceived intelligence, and perceived safety, and includes the trust in automated systems questionnaire [34] to measure perceived trust.

3.5 Study Procedure

To assign participants across all conditions uniformly, we sample sequences in blocks of four samples with an even split of robot voice sex in a pseudo-random fashion. The sequence order of conditions is fully randomized. The same seed is used for consecutive blocks of samples where the last 11 samples are manually assigned to ensure the gender of participants are balanced across robot voice sex.

For each participant, we start by obtaining informed consent. In the consent form, we note that our experiment may have deception approved by our IRB protocol. This hidden deception helps incentivize participants to be more involved and competitive in the task by deducting 10 cents from the study compensation for each ball missed irrespective of whether it was the human or the robot. We further utilize a digital scoreboard showing the current score alongside the loss in compensation in the current experiment block (Figure 2) to emphasize the importance of human-robot collaboration. Note, participants were debriefed about the deception at the end of our study and were paid the full amount of compensation.

Upon completion of informed consent, we give the participant a pre-experiment survey to assess self-reported demographics, personality factors, and negative attitude towards robots. Once participants have completed the survey, we perform skill calibration for a sequence of 30 balls to access the subject's performance on the task without the robot. The subject is allowed a two-minute warm up period before the calibration task begins. After the calibration phase, we move into the main study. We begin by providing a description of the collaborative study and show a demonstration over a sequence of 6 balls with the high-skill model of the robot. The demonstration provides the participant with an idea of the robot capability and show that it is safe to step into the robot's

configuration space to intercept a ball. We additionally demonstrate the high-skill model of the robot with communication enabled and show how both the robot and human can declare intent, and we inform the participant that communicating with the robot is optional. The participant is allowed a warm-up phase, actively collaborating with the robot under no communication. This allows the participant to get used to the robot's presence and minimize the learning effect in the first sequence. During this warm-up, the ball sequence order is fixed and evenly distributed across all zones.

For each condition of the 2x2 within-subject block design, we describe whether the robot has communication enabled and stress that the performance is independent of previous trials. Information about robot performance is kept secret from the participant. Each participant will experience a total of four trials varying across Robot Skill and Communication. The human-robot team will play a 30 ball sequence for each condition, after which the subject is asked to evaluate the interaction with the Godspeed [5] and Trust in Automated Systems questionnaires [34]. Upon completion of the four trials, we debrief the subject about the experiment, detail the deception/concealment, and obtain consent to use their data.

4 RESULTS

We recruited 46 participants for this in-person IRB-approved study. Each participant was compensated $20 for completing our study. Four participants were excluded from our analysis as we experienced issues with the cable tension of the robot arm, resulting in visibly degraded robot performance during the experiments. In our results, we report on 42 participants (sex: 20 Male, 20 Female, 2 Prefer not to say) and age ($M = 25.43$, $SD = 3.5$).

For each measure of trust, perceived safety, team performance, and anthropomorphism, we create a mixed-effects multiple regression model with metrics of interest as the dependent variable. The independent variables are skill, communication, and robot-voice, and we include covariates such as gender, age, personality factors, calibration score, etc. We then applied an Analysis of Variance (ANOVA) on each model to answer our research questions. The mixed-effects model accounts for the repeated measures nature of each subject completing the 2x2 within-subject blocks. For each model, we report the transforms used, results from the Shapiro-Wilk test for normality, and the Levene's test for homoscedasticity for the independent variables in the Appendix. We measure significance as $\alpha < 0.05$. If the assumptions for normality or homoscedasticity do not hold, we report a non-parametric assessment using Friedman's test followed by an all-pairs Nemenyi test comparing the within-subject factors of robot-skill and communication, and do not consider the influence of other covariates.

4.1 Perceived Trust

Across all conditions of gendered voice, skill, and VAIN communication, we compared how participants rated their levels of trust after completing each block comprising a sequence of 30 balls. We find that when the robot performed poorly, the trust in the robot reduced significantly ($F(1, 124) = 18.687$, $p < .001$), as shown in Figure 5a. This finding concurs with many previous human-robot interaction studies evaluating trust on the robot with varying skill levels [33] and addresses (**RQ3**). However, we find that VAIN communication

(a) Effect of Robot Skill on the perceived trust of the system.

(b) Impact of VAIN Communication on the perceived safety of the system.

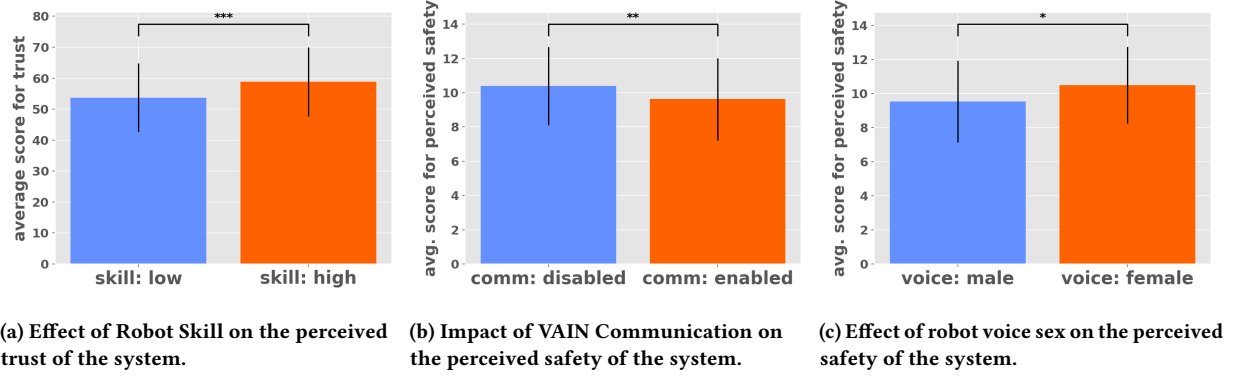(c) Effect of robot voice sex on the perceived safety of the system.

Figure 5: Effect of robot skill, VAIN communication and robot voice on the perceived trust and safety.

did not impact the level of trust on the robot ($F(1, 124) = 0.463$, $p = .497$), concluding for (**RQ7**) that communication does not improve trust on the robot.

4.2 Perceived Safety

To investigate (**RQ6**), we analyze the relationship between VAIN communication and perceived safety. We observe that this communication significantly decreased perceived safety ($F(1, 124) = 7.070$, $p = .009$), as shown in Figure 5b. We hypothesize that the constant announcement of the robot's intention induces more anxiety and agitation, which leads to lower scores within the perceived safety subscale. Additionally, we observe that the sex of the robot voice was a significant factor affecting participants' perception of the robot. On average, the robot equipped with a female voice was perceived safer than the robot with a male voice ($F(1, 38) = 7.079$, $p = .011$), as shown in Figure 5c, addressing (**RQ8**). The participant's gender also influences the perceived safety score, with male participants rating the robot as safer than female participants ($F(2, 38) = 4.709$, $p = .015$). In addressing (**RQ2**), we find that the robot skill level is marginally significant for perceived safety, with a robot with a higher skill level perceived to be safer ($F(1, 124) = 2.762$, $p = .099$).

4.3 Team Performance

For (**RQ1**), we find that a higher skill level of the robot significantly improves the team performance ($F(1, 123) = 69.25$, $p < .001$). Additionally, the skill of the human, measured by their score on the calibration task, improved team performance ($F(1, 34) = 29.929$, $p < .001$). The gender of the participant also significantly impacts team performance, with male participants having higher team performance on the task ($F(2, 34) = 6.314$, $p = 0.005$). In addressing (**RQ5**), we find that VAIN communication degrades HRC team-performance ($F(1, 123) = 4.514$, $p = .037$). We find an interaction effect between communication and skill also being significant ($F(1, 123) = 4.0266$, $p = 0.047$), indicating that communication significantly hurts team-performance when the robot's skill is low, and affects team performance minimally when robot skill is high, as visualized in Figure 6a.

To answer (**RQ4**) and (**RQ9**), we perform a non-parametric all pairs Nemenyi test, as some independent variables failed the Levene's test. The Friedman's test reported that the condition: skill + communication significantly impacts the number of shared tasks
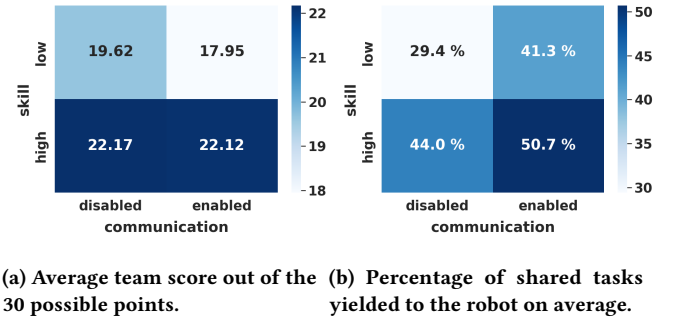

(a) Average team score out of the 30 possible points.

(b) Percentage of shared tasks yielded to the robot on average.

Figure 6: Effect of robot skill and VAIN communication on the team performance.

attempted by the robot. This metric indicates that the participant yielded these tasks to the robot ($\chi^2(3) = 18.071$, $p < .001$). Comparing the different conditions with an all pairs Nemenyi test, we find that when the communication is disabled, robot with higher skill was given more chances to play the shared task ($p = .034$). For (**RQ9**) we find that communication improves the robot's odds of playing the shared task, as depicted in Figure 6b, but the difference is not statistically significant.

Our setup additionally allowed the participant to declare their intent to the robot verbally, similar to what the robot does in the VAIN communication condition. Participants rarely made use of this channel to declare their intent, with a median of 3 declarations in a sequence of 30 balls, with 45.2% of the participants not even using the channel. We note that this metric did not correlate with other factors such as calibration score, personality factors, etc.

4.4 Anthropomorphism

We observe that the robot with high skill was rated higher in terms of anthropomorphism ($F(1, 120) = 4.762$, $p = .031$). We hypothesize this is due to the fluid nature and legible behavior of successful strokes compared to the jerky movement exhibited when the robot fails to return balls. Furthermore, we find VAIN communication was marginally significant in impacting perceived anthropomorphism ($F(1, 120) = 3.047$, $p = .083$). A participant's conscientiousness score (i.e., their propensity towards organization and attention to detail) positively impacted how much they anthropomorphized

the robot ($F(1, 35) = 9.257$, $p = .004$). Conversely, a participant's rating on the NARS subscale negatively influenced the participants perception of anthropomorphism ($F(1, 35) = 7.084$, $p = .012$). We also see that participants that are detail-oriented found the robot more anthropomorphic, while participants that are worried about robots having emotions found the robot less anthropomorphic.

When communication is enabled, the male/female voice of the robot influences the perception of anthropomorphism. The robot using a female voice is rated as more anthropomorphic than the robot with the male voice ($F(1, 120) = 4.762$, $p = .031$). The gender of the participant and its interaction with the robot voice were not significant for anthropomorphism.

4.5 Likeability

We analyze the impact of the skill and the communication factor on this dependent variable through non-parametric tests as this variable failed to meet the assumption of normality. The Friedman's test indicated that the condition: skill+communication significantly impacts likeability ($\chi^2(3) = 21.171$, $p < .001$). To further compare the different conditions, we perform an all-pairs Nemenyi test and find that when communication is disabled, higher robot skill increases likeability ($p = .043$), and with low robot skill, communication tends towards significance in improving likeability ($p = .075$).

5 DISCUSSION

Existing work has shown that verbal communication increases team performance, and that users prefer robots with vocal capabilities. Further, it has been shown that cues which signal specific robot motions (e.g., emitting a sound for a particular motion or non-verbal cues, such as eye contact, for handover tasks) improve perceived safety [3]. However, our results show that in a dynamic, mixed-initiative setting with proximate collaboration, VAIN communication negatively affects team performance and perceived safety. We hypothesize that VAIN communication overloads the subject's mental resources given the short timeframe (∼ 0.6 seconds) for task execution [26]. Within this time, participants must reason about the ball trajectory to determine who should return the ball, process received communication, and optionally communicate information. Extra information may make planning more strenuous, resulting in lower team scores on average.

Moreover, VAIN communication, while transparent, provides a signal to the human just before the robot makes a sudden and fast movement. Such explicit signals could convey to the user that the robot could be capable of overriding the person to take an unsafe action, contributing further to the ambience of physical danger and a decrease in perceived safety. Based on our current findings, we posit that VAIN communication should not be used in fast-paced human-robot teaming systems; instead, robots should be primarily defensive and avoid the appearance of conflicting with a human's autonomy. However, prior work [72] studying sound caused by robots found that louder robots are described as "scary" and "dangerous," indicating that other sound profiles of VAIN communication could potentially lead to improvement.

Additionally, we found that VAIN communication from human to the robot was rarely utilized, suggesting that participants felt more comfortable communicating to the robot implicitly with their body motion as opposed to explicitly declaring their intent with voice in fast-paced tasks. While providing general feedback, some participants noted that talking to the robot was redundant, as intercepting the ball causes the robot to stop regardless. As such, we recommend that agile collaborative systems be developed that can infer human intent from physical motion rather than relying on explicit communication. Nonetheless, it is possible other forms of voice framework and design beside VAIN communication could show benefits to using voice communication in agile collaboration. Humans in agile team sports (volleyball, tennis doubles, etc.) all incorporate some form of VAIN communication to avoid conflict, so further exploration is needed to understand how to achieve effective communication with robots.

Limitations – As indicated in Table 1, the success rate of the high robot skill level is not perfect. The number of failures vary across each trial, where the timing of these failures can alter the subject's perception of the robot's skill. While not common, some trials had a series of missed swings at the beginning of a sequence on a robot with high skill. This can lead the subject to believe the robot is poor performing, and thus receiving fewer attempts at the ball for the rest of the sequence. Additionally, our participants were largely students recruited from a university campus. The use of deception in the study also creates a high-stakes environment which can create more pronounced effects on our findings compared to a lax environment; however, we believe that our current design more accurately represents real-world HRC scenarios where failure results in negative consequences (e.g., a physical collision).

Future Work – In future work, we would like to conduct an extension to this study to identify agile robot system characteristics that lead to increased perceived safety, perceived trust, and HRC performance. Our system was designed to conservatively avoid any collision with the participants, and we plan to improve the system with tighter safety boundaries to allow closer interactions.

6 CONCLUSION

In this work, we utilize a modified human-robot teaming, table-tennis doubles task in a human-subject experiment to gain insights that can inform the design of agile, collaborative robots. Specifically, we study a mixed-initiative teaming scenario with quick decision-making from humans and robots in a shared collaboration space where either agent can successfully accomplish the task. We investigate how robot voice sex, robot skill level, and deliberative communication impact perceived safety and trust in agile, proximate human-robot collaboration. We conduct a human-subjects experiment with 42 participants and find robot skill is positively correlated with trust, extending skill-trust assessment in prior studies to agile, proximate HRC. Furthermore, interestingly, we find that VAIN communication degrades human-robot teaming performance and decreases perceived safety, indicating the need for more research enabling safe and trustworthy agile robots that operate in proximity with humans.

ACKNOWLEDGMENTS

REFERENCES

[1] James Allen, Curry I. Guinn, and E. Horvtz. 1999. Mixed-initiative interaction. *IEEE Intelligent Systems & Their Applications* 14 (1999), 14–23.

[2] Alexandre Angleraud, Amir Mehman Sefat, Metodi Netzev, and Roel Pieters. 2021. Coordinating Shared Tasks in Human-Robot Collaboration by Commands. *Frontiers in Robotics and AI* 8 (2021).

[3] T. Arai, R. Kato, and M. Fujita. 2010. Assessment of operator stress induced by robot collaboration in assembly. *CIRP Annals* 59, 1 (2010), 5–8. https://doi.org/10.1016/j.cirp.2010.03.043

[4] G Robert Arrabito. 2009. Effects of talker sex and voice style of verbal cockpit warnings on performance. *Human factors* 51, 1 (2009), 3–20.

[5] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. 2009. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics* 1, 1 (2009), 71–81.

[6] Yves Beauchamp, Terrence J. Stobbe, Kalyan Ghosh, and Daniel Imbeau. 1991. Determination of a Safe Slow Robot Motion Speed Based on the Effect of Environmental Factors. *Human Factors* 33, 4 (1991), 419–427. https://doi.org/10.1177/001872089103300404 arXiv:https://doi.org/10.1177/001872089103300404

[7] G. Bradski. 2000. The OpenCV Library. *Dr. Dobb's Journal of Software Tools* (2000).

[8] Dieter Büchler, Simon Guist, Roberto Calandra, Vincent Berenz, Bernhard Schölkopf, and Jan Peters. 2022. Learning to play table tennis from scratch using muscular robots. *IEEE Transactions on Robotics* (2022).

[9] Baptiste Busch, Jonathan Grizou, Manuel C. Lopes, and Freek Stulp. 2017. Learning Legible Motion from Human–Robot Interactions. *International Journal of Social Robotics* 9 (2017), 765–779.

[10] Abhizna Butchibabu, Christopher Sparano-Huiban, Liz Sonenberg, and Julie A. Shah. 2016. Implicit Coordination Strategies for Effective Team Communication. *Human Factors: The Journal of Human Factors and Ergonomics Society* 58 (2016), 595 – 610.

[11] John Travis Butler and Arvin Agah. 2001. Psychological Effects of Behavior Patterns of a Mobile Personal Robot. *Autonomous Robots* 10, 2 (01 Mar 2001), 185–202. https://doi.org/10.1023/A:1008986004181

[12] Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh. 2019. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2019).

[13] Carlos Castellar, Francisco Pradas, Luis Carrasco, Ana De La Torre, and José Antonio González-Jurado. 2019. Analysis of reaction time and lateral displacements in national level table tennis players: are they predictive of sport performance? *International Journal of Performance Analysis in Sport* 19, 4 (2019), 467–477. https://doi.org/10.1080/24748668.2019.1621673 arXiv:https://doi.org/10.1080/24748668.2019.1621673

[14] Letian Chen, Sravan Jayanthi, Rohan R. Paleja, Daniel Martin, Viacheslav Zakharov, and Matthew Craig Gombolay. 2022. Fast Lifelong Adaptive Inverse Reinforcement Learning from Demonstrations. *ArXiv* abs/2209.11908 (2022).

[15] Letian Chen, Rohan R. Paleja, Muyleng Ghuy, and Matthew Craig Gombolay. 2020. Joint Goal and Strategy Inference across Heterogeneous Demonstrators via Reward Network Distillation. *2020 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (2020), 659–668.

[16] Letian Chen, Rohan R. Paleja, and Matthew Craig Gombolay. 2020. Learning from Suboptimal Demonstration via Self-Supervised Reward Regression. In *Conference on Robot Learning*.

[17] Nils Dahlbäck, Arne Jönsson, and Lars Ahrenberg. 1993. Wizard of Oz studies—why and how. *Knowledge-based systems* 6, 4 (1993), 258–266.

[18] Frédéric Dehais, Emrah Akin Sisbot, Rachid Alami, and Mickaël Causse. 2011. Physiological and subjective evaluation of a human-robot object hand-over task. *Applied Ergonomics* 42, 6 (Nov. 2011), 785–791. https://doi.org/10.1016/j.apergo.2010.12.005

[19] Maaike Harbers Delft. 2012. Enhancing Team Performance through Effective Communication.

[20] Munjal Desai, Mikhail Medvedev, Marynel Vázquez, Sean McSheehy, Sofia Gadea-Omelchenko, Christian Bruggeman, Aaron Steinfeld, and Holly Yanco. 2012. Effects of changing reliability on trust of robot systems. In *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 73–80. https://doi.org/10.1145/2157689.2157702

[21] Ke Dong, Karime Pereida, Florian Shkurti, and Angela P. Schoellig. 2020. Catch the Ball: Accurate High-Speed Motions for Mobile Manipulators via Inverse Dynamics Learning. In *Proc. IROS*.

[22] M. Brent Donnellan, Frederick L. Oswald, Brendan M. Baird, and Richard E. Lucas. [n.d.]. The Mini-IPIP Scales: Tiny-yet-effective measures of the Big Five Factors of Personality. *Psychological Assessment* 18, 2 ([n.d.]), 192. https://doi.org/10.1037/1040-3590.18.2.192 Publisher: US: American Psychological Association.

[23] Anca D Dragan, Shira Bauman, Jodi Forlizzi, and Siddhartha S Srinivasa. 2015. Effects of robot motion on human-robot collaboration. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 51–58.

[24] J Edworthy, E Hellier, J Rivers, et al. 2003. The use of male or female voices in warnings systems: A question of acoustics. *Noise and Health* 6, 21 (2003), 39.

[25] Shirine el Zaatari, Mohamed Marei, Weidong Li, and Zahid Usman. 2019. Cobot programming for collaborative industrial tasks: An overview. *Robotics Auton. Syst.* 116 (2019), 162–180.

[26] Martin J. Eppler and Jeanne Mengis. 2004. The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines. *The Information Society* 20, 5 (2004), 325–344.

[27] Ema Fonseca, Inês Oliveira, Joana Lobo, Tânia Mota, José Martins, and Manuel Au-Yong-Oliveira. 2019. Kitchen Robots: The Importance and Impact of Technology on People's Quality of Life. In *New Knowledge in Information Systems and Technologies*, Álvaro Rocha, Hojjat Adeli, Luís Paulo Reis, and Sandra Costanzo (Eds.). Springer International Publishing, Cham, 186–197.

[28] Volker Gabler, Tim Stahl, Gerold Huber, Ozgur S. Oguz, and Dirk Wollherr. 2017. A game-theoretic approach for adaptive action selection in close proximity human-robot-collaboration. *2017 IEEE International Conference on Robotics and Automation (ICRA)* (2017), 2897–2903.

[29] Matthew Craig Gombolay, Anna Bair, Cindy Huang, and Julie A. Shah. 2017. 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 (2017), 597 – 617.

[30] Sebastian Gomez-Gonzalez, Gerhard Neumann, Bernhard Schölkopf, and Jan Peters. 2016. Using probabilistic movement primitives for striking movements. In *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*. IEEE, 502–508.

[31] Elena Corina Grigore, André Pereira, Ian Zhou, David Wang, and Brian Scassellati. 2016. Talk to Me: Verbal Communication Improves Perceptions of Friendship and Social Presence in Human-Robot Interaction. In *IVA*.

[32] Sami Haddadin, Alin Albu-Schäffer, Oliver Eiberger, and Gerd Hirzinger. 2010. New insights concerning intrinsic joint elasticity for safety. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, October 18-22, 2010, Taipei, Taiwan*. IEEE, 2181–2187. https://doi.org/10.1109/IROS.2010.5652037

[33] Peter A Hancock, Deborah R Billings, Kristin E Schaefer, Jessie Y C Chen, Ewart J de Visser, and Raja Parasuraman. 2011. A meta-analysis of factors affecting trust in human-robot interaction. *Hum Factors* 53, 5 (Oct. 2011), 517–527.

[34] Jiun-Yin Jian, Ann M Bisantz, and Colin G Drury. 2000. Foundations for an empirically determined scale of trust in automated systems. *International journal of cognitive ergonomics* 4, 1 (2000), 53–71.

[35] Shu Jiang and Ronald C Arkin. 2015. Mixed-initiative human-robot interaction: definition, taxonomy, and survey. In *2015 IEEE International conference on systems, man, and cybernetics*. IEEE, 954–961.

[36] Simon J Julier and Jeffrey K Uhlmann. 1997. New extension of the Kalman filter to nonlinear systems. In *Signal processing, sensor fusion, and target recognition VI*, Vol. 3068. Spie, 182–193.

[37] Rudolph Emil Kalman. 1960. A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME–Journal of Basic Engineering* 82, Series D (1960), 35–45.

[38] Petar Kormushev, Sylvain Calinon, and Darwin G. Caldwell. 2010. Robot motor skill coordination with EM-based Reinforcement Learning. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 3232–3237. https://doi.org/10.1109/IROS.2010.5649089

[39] Dana Kulić and Elizabeth A. Croft. 2007. Pre-collision safety strategies for human-robot interaction. *Autonomous Robots* 22 (2007), 149–164.

[40] Dana Kulić and Elizabeth Croft. 2007. Physiological and subjective responses to articulated robot motion. *Robotica* 25, 1 (2007), 13–27. https://doi.org/10.1017/S0263574706002955

[41] Dana Kulić, Gentiane Venture, Katsu Yamane, Emel Demircan, Ikuo Mizuuchi, and Katja Mombaur. 2016. Anthropomorphic Movement Analysis and Synthesis: A Survey of Methods and Applications. *IEEE Transactions on Robotics* 32, 4 (2016), 776–795. https://doi.org/10.1109/TRO.2016.2587744

[42] Stéphane Lallée, Katharina Hamann, Jasmin Steinwender, Felix Warneken, Uriel Martienz, Hector Barron-Gonzales, Ugo Pattacini, Ilaria Gori, Maxime Petit, Giorgio Metta, Paul Verschure, and Peter Ford Dominey. 2013. Cooperative human robot interaction systems: IV. Communication of shared plans with Naïve humans using gaze and speech. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 129–136. https://doi.org/10.1109/IROS.2013.6696343

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

[44] Przemyslaw A Lasota, Gregory F Rossano, and Julie A Shah. 2014. Toward safe close-proximity human-robot interaction with standard industrial robots. In *2014 IEEE International Conference on Automation Science and Engineering (CASE)*. IEEE, 339–344.

[45] Alessandro De Luca, Alin O. Albu-Schäffer, Sami Haddadin, and Gerd Hirzinger. 2006. Collision Detection and Safe Reaction with the DLR-III Lightweight Manipulator Arm. *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems* (2006), 1623–1630.

[46] Nikolaos Mavridis. 2015. A review of verbal and non-verbal human–robot interactive communication. *Robotics and Autonomous Systems* 63 (2015), 22–35.

https://doi.org/10.1016/j.robot.2014.09.031

[47] Lanssie Mingyue Ma, Terrence Fong, Mark J Micire, Yun Kyung Kim, and Karen Feigh. 2018. Human-robot teaming: Concepts and components for design. In *Field and service robotics*. Springer, 649–663.

[48] Vladimir Murashov, Frank Hearl, and John Howard. 2016. Working safely with robot workers: Recommendations for the new workplace. *J Occup Environ Hyg* 13, 3 (2016), D61–71.

[49] Stefanos Nikolaidis, Minae Kwon, Jodi Forlizzi, and Siddhartha Srinivasa. 2018. Planning with Verbal Communication for Human-Robot Collaboration. *J. Hum.-Robot Interact.* 7, 3, Article 22 (nov 2018), 21 pages. https://doi.org/10.1145/3203305

[50] Rohan R. Paleja, Muyleng Ghuy, Nadun R. Arachchige, Reed Jensen, and Matthew Craig Gombolay. 2022. The Utility of Explainable AI in Ad Hoc Human-Machine Teaming. *ArXiv* abs/2209.03943 (2022).

[51] Xue Bin Peng, Erwin Coumans, Tingnan Zhang, Tsang-Wei Edward Lee, Jie Tan, and Sergey Levine. 2020. Learning Agile Robotic Locomotion Skills by Imitating Animals. In *Robotics: Science and Systems*. https://doi.org/10.15607/RSS.2020.XVI.064

[52] Luka Peternel, Wansoo Kim, Jan Babič, and Arash Ajoudani. 2017. Towards ergonomic control of human-robot co-manipulation and handover. In *2017 IEEE-RAS 17th International Conference on Humanoid Robotics (Humanoids)*. IEEE, 55–60.

[53] Anita Pollak, Mateusz Paliga, Matias M. Pulopulos, Barbara Kozusznik, and Malgorzata W. Kozusznik. 2020. Stress in manual and autonomous modes of collaboration with a cobot. *Computers in Human Behavior* 112 (2020), 106469. https://doi.org/10.1016/j.chb.2020.106469

[54] H.C. Ravichandar and A. Dani. 2017. Chapter Eight - Intention Inference for Human-Robot Collaboration in Assistive Robotics. In *Human Modelling for Bio-Inspired Robotics*, Jun Ueda and Yuichi Kurita (Eds.). Academic Press, 217–249. https://doi.org/10.1016/B978-0-12-803137-7.00007-0

[55] Harish Chaandar Ravichandar, Kenneth Shaw, and S. Chernova. 2020. STRATA: unified framework for task assignments in large teams of heterogeneous agents. *Autonomous Agents and Multi-Agent Systems* 34 (2020), 1–25.

[56] Paul Robinette, Ayanna M. Howard, and Alan R. Wagner. 2017. Effect of Robot Performance on Human–Robot Trust in Time-Critical Situations. *IEEE Transactions on Human-Machine Systems* 47, 4 (2017), 425–436. https://doi.org/10.1109/THMS.2017.2648849

[57] Alexey Romanov. 2019. A review on control systems hardware and software for robots of various scale and purpose. Part 1. Industrial robotics. *Russian Technological Journal* (2019).

[58] Pericle Salvini, Diego Paez-Granados, and Aude Billard. 2022. Safety Concerns Emerging from Robots Navigating in Crowded Pedestrian Areas. *International Journal of Social Robotics* 14, 2 (01 Mar 2022), 441–462. https://doi.org/10.1007/s12369-021-00796-4

[59] Yu She, Hai-Jun Su, Deshan Meng, Siyang Song, and Junmin Wang. 2017. Design and Modeling of a Compliant Link for Inherently Safe Corobots. *Journal of Mechanisms and Robotics* 10, 1 (12 2017). https://doi.org/10.1115/1.4038530 arXiv:https://asmedigitalcollection.asme.org/mechanismsrobotics/article-pdf/10/1/011001/6403971/jmr_010_01_011001.pdf 011001.

[60] Satoru Shibata and Hikaru Inooka. 1998. Psychological evaluations of robot motions. *International Journal of Industrial Ergonomics* 21, 6 (1998), 483–494. https://doi.org/10.1016/S0169-8141(97)00004-8

[61] Aaron St. Clair and Maja Mataric. 2015. How Robot Verbal Feedback Can Improve Team Performance in Human-Robot Task Collaborations. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction* (Portland, Oregon, USA) *(HRI '15)*. Association for Computing Machinery, New York, NY, USA, 213–220. https://doi.org/10.1145/2696454.2696491

[62] Aaron St. Clair and Maja Mataric. 2015. How Robot Verbal Feedback Can Improve Team Performance in Human-Robot Task Collaborations *(HRI '15)*. Association for Computing Machinery, New York, NY, USA, 213–220. https://doi.org/10.1145/2696454.2696491

[63] Matthew Story, Phil Webb, Sarah R. Fletcher, Gilbert Tang, Cyril Jaksic, and Jon Carberry. 2022. Do Speed and Proximity Affect Human-Robot Collaboration with an Industrial Robot Arm? *International Journal of Social Robotics* 14, 4 (01 Jun 2022), 1087–1102. https://doi.org/10.1007/s12369-021-00853-y

[64] Yuta Sugiura, Daisuke Sakamoto, Anusha Withana, Masahiko Inami, and Takeo Igarashi. 2010. Cooking with Robots: Designing a Household System Working in Open Environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 2427–2430. https://doi.org/10.1145/1753326.1753693

[65] Dag Sverre Syrdal, Kerstin Dautenhahn, Kheng Lee Koay, and Michael L Walters. [n. d.]. The Negative Attitudes towards Robots Scale and Reactions to Robot Behaviour in a Live Human-Robot Interaction Study. ([n. d.]), 7.

[66] Andrea Thomaz, Guy Hoffman, and Maya Cakmak. 2016. Computational Human-Robot Interaction. *Foundations and Trends® in Robotics* 4, 2-3 (2016), 105–223. https://doi.org/10.1561/2300000049

[67] Valeria Villani, Fabio Pini, Francesco Leali, and Cristian Secchi. 2018. Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications. *Mechatronics* (2018).

[68] John Wang and Edwin Olson. 2016. AprilTag 2: Efficient and robust fiducial detection. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 4193–4198.

[69] Y. Yamada, Y. Umetani, and Y. Hirasawa. 1999. Proposal of a psychophysiological experiment system applying the reaction of human pupillary dilation to frightening robot motions. In *IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No.99CH37028)*, Vol. 2. 1052–1057 vol.2. https://doi.org/10.1109/ICSMC.1999.825408

[70] Fan Yang, Zhiwei Shi, Sixian Ye, Jiazhong Qian, Wenjie Wang, and Dong Xuan. 2022. VaRSM: Versatile Autonomous Racquet Sports Machine. In *Proc. ICCPS*. 203–214. https://doi.org/10.1109/ICCPS54341.2022.00025

[71] Sean Ye, Glen Neville, Mariah Schrum, Matthew Gombolay, Sonia Chernova, and Ayanna Howard. 2019. Human Trust After Robot Mistakes: Study of the Effects of Different Forms of Robot Communication. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 1–7. https://doi.org/10.1109/RO-MAN46459.2019.8956424

[72] Brian J. Zhang, Knut Peterson, Christopher A. Sanchez, and Naomi T. Fitter. 2021. Exploring Consequential Robot Sound: Should We Make Robots Quiet and Kawaii-et? *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (2021), 3056–3062.

[73] M. Zinn, O. Khatib, and B. Roth. 2004. A new actuation approach for human friendly robot design. In *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*, Vol. 1. 249–254 Vol.1. https://doi.org/10.1109/ROBOT.2004.1307159