

Design and Impact of Hesitation Gestures during Human-Robot Resource Conflicts

AJung Moon, Chris A.C. Parker, Elizabeth A. Croft,
H.F. Machiel Van der Loos
University of British Columbia

In collaborative tasks, people often communicate using nonverbal gestures to coordinate actions. When two people reach for the same object at the same time, they often respond to an imminent potential collision with jerky halting hand motions that we term *hesitation gestures*. Successful implementation of such communicative conflict response behaviour onto robots can be useful. In a myriad of human-robot interaction contexts involving shared spaces and objects, this behaviour can provide a fast and effective means for robots to express awareness of conflict and cede right-of-way during collaborative work with users. Our previous work suggests that when a six-degree-of-freedom (6-DOF) robot traces a simplified trajectory of recorded human hesitation gestures, these robot motions are also perceived by humans as hesitation gestures. In this work, we present a characteristic motion profile derived from the recorded human hesitation motions, called the Acceleration-based Hesitation Profile (AHP). We test its efficacy to generate communicative hesitation responses by a robot in a fast-paced human-robot interaction experiment.

Compared to traditional abrupt stopping behaviours, we did not find sufficient evidence that the AHP-based robot responses improve human perception of the robot or human-robot task completion time. However, results from our in situ experiment suggest that subjects can recognize AHP-based robot responses as hesitations and distinguish them to be different from abrupt stopping behaviours.

Keywords: Human-robot interaction, trajectory design, hesitation, nonverbal communication, reaching motions, resource conflict, collision avoidance

1. Introduction

As service robots in the workplace and personal robots in the home become increasingly commonplace, untrained users will share more spaces and resources with robots. Conflicts involving shared resources will naturally ensue between user and robot as users step away from master-slave control of robots, through consoles and teach pendants, and step into situations of direct interaction with robots as partners. The manner in which a robot collaborates and manages conflicts with people

Due to a conflict of interest with the managing editor Michael Goodrich, the editor for this paper was guest editor Magnus Egerstedt. Authors retain copyright and grant the Journal of Human-Robot Interaction right of first publication with the work simultaneously licensed under a Creative Commons Attribution License that allows others to share the work with an acknowledgement of the work's authorship and initial publication in this journal.

will affect not only the outcome of the human-robot interaction, but also the user’s perception of the robot. We believe that mechanisms to transparently communicate a robot’s internal states to users in the context of resource conflicts will facilitate positive resolution of potential conflicts and enhance the human-robot interaction experience.

In human-human interaction, people use both verbal and nonverbal communication to resolve a variety of conflicts. In this work, we focus on the resource conflict that arises when two agents reach for the same object at the same time in a non-competitive context. In human-human interaction, people often exhibit abrupt halting motions with their hands upon recognizing an imminent collision (A. Moon, Parker, Croft, & Van der Loos, 2011). These gestures convey the state of uncertainty about a person’s right or priority of access to the resource (i.e., the object), and we term them *hesitation gestures*. When generating these gestures, the trajectory of a person’s functional reaching motion smoothly transitions to the communicative halting motion of the hesitation gesture, allowing fast, seamless nonverbal negotiation with the other person.

In order for robots to safely work alongside humans while maintaining their functional usefulness in industrial environments, it is essential that robots be able to naturally and intuitively manage shared resources between humans and robots. We posit that implementing implicit communicative response mechanisms that can work in conjunction with functional motions of a robot, such as hesitation gestures, are essential to achieving this goal.

Results of our previous study (A. Moon et al., 2011) indicate that when a robot mimics the trajectories of human hesitations, untrained observers perceive it to be hesitating. In this earlier paper we also proposed a trajectory specification method, which we describe as the Acceleration-based Hesitation Profile (AHP).

In this work, we first outline a method to use AHP in a real-time human-robot interaction (HRI) context. The method allows smooth transition from a robot’s functional reaching motion to a gesture that conveys a state of hesitation to its human observer. We then verify, in a preliminary online survey that asked participants to observe and report on video recordings of robot motions, the efficacy of AHP-based robot gestures. Finally, we demonstrate AHP as a communicative collision response mechanism by implementing it in a real-time HRI experiment designed to simulate a robot-assisted assembly line scenario. Using quantitative and qualitative instruments, we investigate whether subjects can recognize AHP-based collision responses as hesitation gestures in situ, and whether the communicative response positively impacts the completion of a human-robot collaboration task and human perception of the robot.

We begin with a literature review in Section 2. In Section 3, we provide a detailed description of the AHP, its implementation on a real-time HRI system, and the methods and results of the preliminary study. We describe methods and results of the main interaction experiment in Sections 4 and 5, respectively. We close with a discussion in Section 6 and a conclusion in Section 7.

2. Related Work

Human-human collaboration typically involves people with different intentions and capabilities. Without a means to effectively communicate with each other, the collaborating partners would neither be able to establish a common ground nor interweave subplans to achieve the shared goal (Cohen & Levesque, 1991). In Bratman’s (1992) model of successful collaboration, mutual responsiveness, commitment to the joint activity, and commitment to mutual support are necessary. None of these can be established without communication between the collaborating agents. Likewise, in order for human-machine collaboration to be successful, communication mechanisms that allow the collaborating agents to interweave plans and actions and to establish mutual understanding are required (Grosz, 1996). Use of nonverbal cues is one such method for machines to communicate with humans.

Studies suggest that humans not only express their internal states and intentions, but also read others' intentions and affect via nonverbal cues (Argyle, 1994; Becchio, Sartori, Bulgheroni, & Castiello, 2008; Becchio, Sartori, & Castiello, 2010; Fox & McDaniel, 1982; Manera, Becchio, Cavallo, Sartori, & Castiello, 2011; Thorisson & Cassell, 1999). The kinematics of nonverbal cues embedded in functional motions of a person, such as reaching for an object, can influence the behaviour of their observer in an interpersonal context (Becchio et al., 2008, 2010). In addition, studies in psychology have demonstrated that humans not only infer information from motions of other biological beings as well as moving objects (Dittrich & Lea, 1994; Heider & Simmel, 1944; Ju & Takayama, 2009; Tomasello, Carpenter, Call, Behne, & Moll, 2005), but also infer information from, and are influenced by, nonverbal behaviours of computers and computer-based avatars (Reeves & Nass, 1996). Unlike computers, robots can manoeuvre in and manipulate the physical environment they share with their users. Previous research in HRI has demonstrated that human response to and perception of a robot are highly correlated to the kinematics of the robot's motions. Kim, Kwak, and Kim (2008) demonstrated that people ascribe different personalities to the robot when the trajectory parameters of the robot's gestures, including velocity and frequency, are varied. Riek et al. (2010) found that humans respond quicker to abrupt and front-oriented robot gestures than smooth or side-oriented gestures. Using a facial robot (iCat, Philips Research, Eindhoven, the Netherlands) and a two-degree-of-freedom (2-DOF) mobile robot (Roomba, iRobot, Massachusetts, USA), Saerbeck and Bartneck (2010) demonstrated that acceleration and curvature of robot motions, rather than the type of robot used, have a significant impact in conveying different affect to human observers. In Kulić and Croft's (2006) study, human subjects watched an articulated 6-DOF robot (CRS A460, Burlington, ON, Canada) perform a series of pick-and-place and reach-retract motions using different types of trajectories. Results of this study indicate that a human observer's affective response (as recorded by physiological sensors) to the robot changes significantly based on the type of trajectory used to control the robot's motion even when the trajectories functionally obtain the same result (Kulić & Croft, 2006). Findings from the abovementioned studies highlight the feasibility of nonverbal cues as human-robot communication mechanism, and echo the importance of robot motion quality in eliciting different human responses.

In addition, unlike many of the high-DOF robots used in human-robot collaboration studies, most robots in industrial settings are not equipped to display facial gestures representing the robot's state. Often, it is also impractical for a robot to have a face (Bethel & Murphy, 2006); moreover, in industry, it is necessary that the worker pay attention to the task at hand, i.e., the workpiece and potentially the robot's hand or gripper, rather than the robot's body or face (if present). Nonetheless, recent literature suggests that humans, when interacting with robots, naturally expect robots to follow social conventions even if they are non-facial and non-anthropomorphic (Fincannon, Barnes, Murphy, & Riddle, 2004). Abovementioned studies, such as that of Kulić and Croft's (2006) and Saerbeck and Bartneck's (2010) suggest that motions of even non-facial, non-anthropomorphic robots can be designed to be communicative and expressive. Based on these literature we hypothesize that motions of an articulated, non-facial robot can be designed to express hesitation to and be recognized as hesitation by its human observers in an in situ human-robot collaboration context. Verifying this will allow us to use our designed gestures as short, quick, and intuitive nonverbal means to communicate with users in myriad of human-robot resource conflict scenarios.

Prior to our previous work (A. Moon et al., 2011), studies in hesitation have been limited to its embodiment in human verbal communication and delays (such as pauses in speech or action) (Doob, 1990; Klapp, Kelly, & Netick, 1987). In the HRI domain, Bartneck, van der Hoek, Mubin, and Al Mahmud (2007) used human hesitation as a measure of cognitive dilemma in turning off a robot when instructed to do so. In another study, Kazuaki, Motoyuki, and Natsuki (2010)

programmed hesitations on a robot as the duration of time it takes for a robot (AIBO, Sony, Japan) to initiate actions after a human demonstrates to the robot how to shake hands with a person. The results of their study indicate that the management of delays in the robot's response helps improve people's experience of teaching a robot. We contribute to this body of work by investigating our empirically-driven design of hesitation as motion trajectories (AHP) proposed in our previous study (A. Moon et al., 2011).

Studies also suggest that the appearance and the behaviour of machines can affect human perception of the machines as well as human behaviour while interacting with the machines, thereby impacting the effectiveness of the human-machine teamwork (Burgoon et al., 2002; Hinds, Roberts, & Jones, 2004; Y. Moon & Nass, 1996). In an experiment involving a collaborative part retrieval task, Hinds et al. (2004) found that people exhibit more reliance and attribute more credit to their robotic partner when it has a more human-like appearance. Goetz, Kiesler, and Powers (2003) investigated the impact that a humanoid's social cues have on human acceptance of the robot as a partner. In their Wizard-of-Oz experiment, the robot's demeanour (playful vs. serious) and the nature of the cooperative task (playful vs. serious) were varied. The results suggest that the subject's compliance with the robot increases when the robot displays a demeanour that matches the nature of the task.

While the HRI in research by Hinds et al. (2004) and Goetz et al. (2003) was verbal, a number of studies have demonstrated the utility of using nonverbal gestures in conjunction with verbal communications for human-robot collaboration tasks (Breazeal, Kidd, Thomaz, Hoffman, & Berlin, 2005; Holroyd, Rich, Sidner, & Ponsler, 2011; Huang & Thomaz, 2011; Salem, Rohlfing, Kopp, & Joubin, 2011). Holroyd et al. (2011) implemented a set of policies for selecting a set of nonverbal gestures that should accompany the robot's speech in order to effectively communicate with its human partner. They demonstrated the effectiveness of their verbal/nonverbal management system in the collaborative solving of a tangram puzzle. The positive results from this study indicate that more natural management of robot gestures improves user perception of the robot and helps establish a sense of mutual understanding with the robot. Huang and Thomaz (2011) also employed nonverbal cues to supplement verbal communication with a robot. They found that such verbal communication, together with supplemental nonverbal gestures, is an effective means to acknowledge establishment of joint attention between human and robot. This approach improved human understanding of the intended robot behaviour and human-robot task performance (Huang & Thomaz, 2011). Breazeal et al. (2005) conducted an experiment with an expressive 65-DOF robot, Leo, which used shoulder shrug gestures and facial expressions to convey its state of uncertainty to a human collaborator in a joint activity. The results from this study also provide strong evidence that combined use of nonverbal gestures and speech to display a robot's behaviour state can be more effective in improving task performance than using speech alone as the only communication modality.

Even though the value of gesturing has been demonstrated in certain collaborative scenarios, designing nonverbal gestures that improve in situ human-robot collaboration has been found to be a nontrivial task. In another experiment (Reed et al., 2006), where two people worked as a haptically-linked dyad to rotate a disk to a target location collaboratively, it was found that even without verbal communication, people quickly negotiate each other's role within the team using only haptic cues. This study also demonstrated that, in comparison to completing the task alone, there is a significant increase in performance when people worked together as a team. However, when the study was repeated with human-robot dyads, human subjects did not take on a specific role within the collaborative task nor did the dyad yield an improved task performance (Reed, Patton, & Peshkin, 2007). The authors suggest that these negative results may be due to the lack of subtle haptic negotiations in the human-robot dyad condition. While these studies demonstrate the power of nonverbal communication in human-human collaboration, they also point out the

challenge of designing and exploring subtle features of communicative gestures and negotiation mechanisms to improve human-robot collaboration systems. Hence, it is uncertain whether the empirically-designed hesitation gestures we proposed in our earlier work (A. Moon et al., 2011) will improve human-robot team performance.

While the findings in studies conducted by Holroyd et al. (2011), Huang and Thomaz (2011), and Breazeal et al. (2005) emphasize the power of nonverbal communication in human-robot collaboration, unlike Reed and colleagues' experiment (Reed et al., 2007), the nature of the collaborative tasks involved implied turn-taking rules between the human and robot that may not be present in many potential human-robot collaboration scenarios. The subject's role in Breazeal et al. (2005) was to supervise and instruct the robot to perform a manipulation task while the robot waited for the subject's instruction before performing the task. These roles were reversed in research by Holroyd et al. (2011). If fixed rules exist on turn-taking or right of way and both humans and robots follow these rules perfectly, resource conflicts, such as reaching for the same object at the same time, would not occur. However, when such rules are not in place, or if at least one of the collaborating agents is not aware of or does not comply with predefined rules, transparent communication of each agent's intentions and behaviour states (dominant, submissive, collaborative, pesky) becomes even more essential for navigating the interaction. Our work contributes to the HRI body of literature by exploring the effects of nonverbal communication in collaborative scenarios without such predefined or implied hierarchy and turn-taking rules. Hence, the nature of experiments presented in this work is distinguished from these previous studies in that it features a *lack of* predefined turn-taking rules.

In summary, focusing on nonverbal gestures in HRI contexts where resource conflicts do occur and building on our previous work, we hypothesize that robot hesitation manifested as kinesic gestures based on AHP will be recognized as hesitation by human observers in situ, will elicit positive human perception of the robotic partner, and will result in improved human-robot team performance.

3. Acceleration-Based Hesitation Profile (AHP)

Our previous study (A. Moon et al., 2011) indicated that observers of a 6-DOF manipulator mimicking wrist trajectories of human hesitation gestures perceived the robot to be hesitating. This empirical result suggests that despite kinematic and dynamic differences, a robotic manipulator can display the communicative features of hesitation gestures.

Two types of hesitation gestures were identified within the context of reaching tasks. Our investigation focused on the most prevalent type of hesitation, the retraction (labelled R-type). In this type of hesitation, the subject's hand halts while reaching for a shared target object and retracts back to the starting position without continuing to reach, thereby yielding access to the object. Key features recorded from R-type human hesitation gestures then formed the basis for an end-effector trajectory specification. This specification, proposed in (A. Moon et al., 2011), is the Acceleration-based Hesitation Profile (AHP). In this section, we extend our discussion of AHP by detailing its implementation for a real-time HRI system (Section 3.1). We summarize preliminary verification of the efficacy of AHP in Section 3.2.

3.1 AHP Design

A humanlike minimum jerk motion in its principal axis of motion can be modelled as a quintic trajectory (Flash & Hogan, 1985). The following equations are quintic Hermite expressions of humanlike, uninterrupted reaching and returning motions from a starting position to a target position

and back:

$$x_{reach}(\tau_r) = (1 - 10\tau_r^3 + 15\tau_r^4 - 6\tau_r^5)x_{home} + (0.5\tau_r^3 - \tau_r^4 + 0.5\tau_r^5)\ddot{x}_{targ} \\ + (-4\tau_r^3 + 7\tau_r^4 - 3\tau_r^5)\dot{x}_{targ} + (10\tau_r^3 - 15\tau_r^4 + 6\tau_r^5)x_{targ} \quad (1)$$

$$x_{retract}(\tau_r) = (1 - 10\tau_r^3 + 15\tau_r^4 - 6\tau_r^5)x_{targ} + (\tau_r - 6\tau_r^3 + 8\tau_r^4 - 3\tau_r^5)\dot{x}_{targ} \\ + (0.5\tau_r^2 - 1.5\tau_r^3 + 1.5\tau_r^4 - 0.5\tau_r^5)\ddot{x}_{targ} \\ + (10\tau_r^3 - 15\tau_r^4 + 6\tau_r^5)x_{home} \quad (2)$$

Here, $x_{reach}(\tau_r)$ describes the quintic trajectory of a reaching motion that travels from x_{home} , the starting location of the reach, to x_{targ} , the target location of the reach; τ_r represents time, t , that is normalised by the total span of the quintic motion, t_f , ($\tau_r = t/t_f$). Analogously, $x_{retract}(\tau_r)$ describes a quintic trajectory of a retracting motion that travels from x_{targ} to x_{home} . Figure 1a illustrates the position and acceleration profiles of these motions.

The recorded human trajectories of R-type hesitation gestures have a set of relative peak acceleration magnitudes and temporal values different from that of a successful reach-retract trajectory (A. Moon et al., 2011). The relationship between the R-type peak acceleration magnitudes and their temporal values forms the basis of the acceleration-based hesitation profile (AHP). The relationship can be expressed as a set of ratios C_1 , C_2 , B_1 , and B_2 :

$$a_2 = C_1 a_1, \quad t_2 - t_1 = B_1 t_1 \quad (3)$$

$$a_3 = C_2 a_1, \quad t_3 - t_2 = B_2 t_1 \quad (4)$$

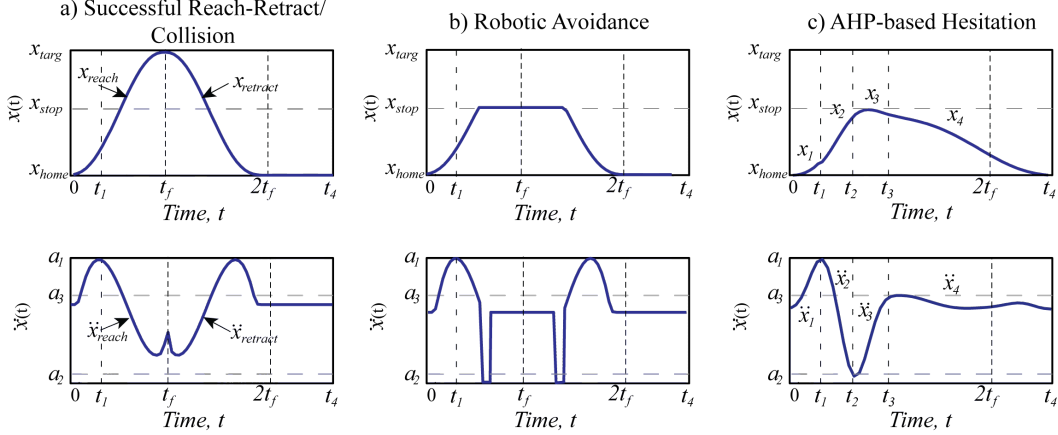


Figure 1. The position and acceleration profile of reference trajectories in the principal axis, x , is generated for the preliminary study. All four types of motions share the same values of a_1 and t_1 , the launch acceleration, and its temporal location. a) Successful reach-retract and collision motions are comprised of two quintic splines, x_{reach} followed by $x_{retract}$, each of which span total duration of t_f ; b) Robotic avoidance motions are designed to mimic emergency manoeuvres. This is achieved by flattening the two-quintic trajectories designed for the successful reach-retract motions at a desired stopping location, x_{stop} ; c) AHP-based responses consist of four splines, x_1 , x_2 , x_3 and x_4 . The variables t_1 and a_1 , t_2 and a_2 , and t_3 and a_3 respectively represent the temporal location and amplitude of acceleration at launch, braking, and yielding phases of the motion. The ratios between these variables in acceleration space comprise the key parameters of AHP. Vertical dashed lines outline the transition points for generating AHP-based trajectories with splines.

Here, a_1 , a_2 and a_3 represent the first (launch), second (braking), and third (yielding) acceleration extrema. Variables t_1 , t_2 , and t_3 , respectively represent the temporal locations of the extrema. The values used for the constants are $C_1 = -1.40$, $C_2 = 0.24$, $B_1 = 0.89$, and $B_2 = 1.14$ and are derived from recorded R-type human hesitation data collected from our previous study (A. Moon et al., 2011). Figure 1 demonstrates the motion profile differences between successful reach-retract motions and that of AHP-based motions in the position and acceleration domain.

Ratios C_1 and B_1 describe relative magnitudes between the launch acceleration and the braking acceleration, and express abruptness of the halting behaviour. Ratios C_2 and B_2 describe relative magnitudes between the launch acceleration and the start-of-return acceleration while controlling how smoothly the halting motion leads to the return phase of the motion.

To generate a reference trajectory consistent with AHP, four splines can be fit through five critical points of the AHP located at the start and end of the trajectory ($t = 0$ and t_4 , with zero accelerations) and at the temporal locations of the three acceleration extrema (t_1 , t_2 , and t_3). The first three splines (\ddot{x}_1 , \ddot{x}_2 and \ddot{x}_3) can be generated as Hermite cubic splines in acceleration space, where the initial and final values of jerk can be specified as zero for all critical points. Substituting Equations (3) and (4) into the three cubic acceleration splines and integrating them twice yields corresponding quintic position splines (x_1 , x_2 , and x_3) that are consistent with the AHP and expressed in normalised time as:

$$x_1(\tau_1) = a_1 \tau_1^4 t_1^2 \left(-\frac{\tau_1}{10} + \frac{1}{4} \right) + x_{home} \quad (5)$$

$$x_2(\tau_2) = a_1 (\tau_2 B_1 t_1)^2 \left(\frac{\tau_2^3 (1 + C_1)}{10} - \frac{\tau_2^2 (1 + C_1)}{4} + \frac{1}{2} \right) + \dot{x}_1(1) \tau_2 + x_1(1) \quad (6)$$

$$x_3(\tau_3) = a_1 (\tau_3 B_2 t_1)^2 \left(\frac{-\tau_3^3 (C_1 + C_2)}{10} + \frac{\tau_3^2 (C_1 + C_2)}{4} + \frac{C_1}{2} \right) + \dot{x}_2(1) \tau_3 + x_2(1) \quad (7)$$

$$\tau_1 = \frac{t}{t_1 - 0}, \tau_2 = \frac{t - t_1}{t_2 - t_1} = \frac{t - t_1}{B_1 t_1}, \tau_3 = \frac{t - t_2}{t_3 - t_2} = \frac{t - t_2}{B_2 t_1}, \quad (8)$$

The fourth spline, connecting the final values of $x_3(\tau_3)$ to come to a smooth stop at the final position, x_{home} , can be designed as a Hermite quintic spline in position:

$$x_4(\tau_4) = (1 - 10\tau_4^3 + 15\tau_4^4 - 6\tau_4^5)x_3(1) + (\tau_4 - 6\tau_4^3 + 8\tau_4^4 - 3\tau_4^5)\dot{x}_3(1) + \left(\frac{1}{2}\tau_4^2 - \frac{3}{2}\tau_4^3 + \frac{3}{2}\tau_4^4 - \frac{1}{2}\tau_4^5 \right)\ddot{x}_3(1) + (10\tau_4^3 - 15\tau_4^4 + 6\tau_4^5)x_{home} \quad (9)$$

Here, $\tau_4 = (t - t_3)/(t_4 - t_3)$. With this approach, an AHP-based motion is specified from only two parameters: a_1 , the launch acceleration, and its related temporal parameter, t_1 . When implemented as part of a collision response mechanism, values of a_1 and t_1 can be computed by differentiating Equation (1); the first root of $\ddot{x}_{reach}(\tau_r) = 0$ yields the value of τ_r that corresponds to t_1 ($t_1 = \tau_r t_f$). Substituting τ_r into the second derivative of Equation (1), $\ddot{x}_{reach}(\tau_r)$, yields a_1 at t_1 .

Secondary axis motion, $z(t)$, can supplement the primary axis motion $x(t)$ to generate a humanlike, two-dimensional path towards the target object while maintaining constant target-space configuration. In our experiment we implemented a second-order function to achieve this (variables a, b, c , and d are constants selected to fit the robot's range of motion within the experimental workspace): $z(x) = -a(x + c)^2 + b(x + c) + d$.

A reactive system can be built using the AHP spline specification with the significant advantage that the robot controller can decide whether or not to hesitate, even after starting to reach towards an object.

In a real-time HRI, the robot can be programmed to start reaching for the object using Equation (1), effectively replacing Equation (5) until the launch acceleration has been reached at time t_1 . At this point, the reference trajectory of the robot can smoothly switch from Equation (1) to Equation (6), since the position, velocity, and acceleration are then matched. However, this advantage comes with a limitation: the decision to hesitate has to be made before t_1 has been reached, such that the reference trajectory switch can take place at t_1 . When a collision becomes imminent after t_1 has passed, the robot would continue to follow Equation (1) and cause a collision. To account for this issue, the real-time implementation of AHP in our main experiment uses AHP in conjunction with an abrupt emergency stop to avoid collisions with the subjects. Our implementation of AHP-based trajectory generation is open source and made available online with a simulation package (Goto & Moon, 2012).

3.2 Preliminary Verification of AHP

To investigate whether AHP-based motions can be used to generate communicative and anthropomorphic hesitation gestures, we conducted an online survey-based experiment. In this preliminary study, 58 participants watched randomly ordered video recordings of the experimenter and a robot engaged in a series of reach-retract tasks.

A 6-DOF robot (A460, CRS, Burlington, ON, Canada) followed pre-generated Cartesian reference trajectories of twelve motions. Four types of robot motions — AHP-based hesitation motion, robotic avoidance motion, successful reach-retract motion, and collision — were generated using three different values of launch accelerations. These acceleration values, $\{9.5, 16.5, \text{ and } 23.5\}$ m/s², correspond to the recorded human motions from research in A. Moon et al. (2011). Since not all 6-DOF are necessary to generate the motions, the robot used only 4-DOF to create planar motions towards the target object.

Meanwhile, an experimenter stood facing the robot and enacted a series of reach-retract motions towards a target object in the middle of the scene as though sharing the object and triggering different behaviours of the robot. In the collision motions, the experimenter timed her reach such that her hand physically came in contact with the robot.

AHP-based hesitation motions were generated using the abovementioned implementation method. Successful reach-retract motions were comprised of two connected quintic splines, Equation (1) followed by Equation (2). Reference trajectories for collision motions were identical to the successful reach-retract motions. Robotic avoidance manoeuvres were designed to mimic abrupt emergency stop behaviours. To achieve this, the peak positions of the successful reach-retract motions were modified such that the robot stops at the same distance away from the target, x_{stop} , as it would for an analogous AHP-based trajectory, and then it retracts. Figure 1 shows an example of the generated reference trajectories.

Each unlabelled video contained only one of the twelve robot motions. Figure 2 shows a screenshot of the videos shown to the participants.

Using four Likert scale questions, two of which are distractor questions, we collected participant responses to *perceived hesitation* (1. Not hesitant - 5. Hesitant) and *anthropomorphism* (1. Machinelike - 5. Humanlike) of the robot motions in the videos. Data from subjects who did not complete the entire survey were dropped from our analysis. A significant result seen in a repeated-measures ANOVA, followed by results from a post-hoc analysis using the Bonferroni approach, supported the proposal that AHP-based motions are perceived to convey hesitation more than robotic avoidance motions ($F(2.49, 104.48) = 132.83, p < .0001$, post-hoc: $p = .02$). Note

that the reported results have been corrected using the Greenhouse-Geisser method in order to correct for violation of the sphericity assumption ($X^2(5) = .73, p < .05, \epsilon = .83$). Similarly, AHP-based motions conveyed hesitation more often than successful motions ($p < .001$) and collisions ($p < .001$). An overview of these responses is shown in Figure 3.

We also found significant differences in the *anthropomorphism* score across the four types of robot motions ($F(2.32, 97.54) = 12.45, p < .0001$). This result has also been corrected using Greenhouse-Geisser method in order to account for violation of the sphericity assumption ($X^2(5) = 0.63, p < .01, \epsilon = 0.77$). A post-hoc analysis using the Bonferroni approach suggests that successful motions, collisions, and AHP-based motions are all significantly more human-like than those of the robotic avoidance motions ($p < .001$ for all comparisons against robotic avoidance motions). Figure 3b graphically summarises these results. We found no significant difference in the hesitation or anthropomorphism scores across the three levels of launch acceleration.

The preliminary study provides empirical support that the AHP can be used to generate robot motions that are perceived to be anthropomorphic and hesitant to human observers. However, the main task for participants of this study was to watch videos of HRI without participating in the interaction. Hence, this result does not determine whether AHP-based motions will be perceived as hesitations when subjects are interacting with the robot in situ and focused on completing their own tasks rather than attentively observing the robot.

4. Experimental Methodology

Based on the positive results of the preliminary study, we conducted a within-subjects human-robot interaction experiment to investigate the efficacy and the impact of AHP-based conflict responses.

In this experiment, a 7-DOF robot (WAMTM, Barrett Technologies, Cambridge, MA, USA) shared a resource with a human subject. Rather than observing robot motions from video recordings, the subject in this study participated as the human interacting with the robot in a fast-paced, interactive environment.

When, by chance, the two agents reached for the shared resource at the same time, the robot responded in one of the following three ways:

Blind (B) Response: Ignore the resource conflict and continue reaching for the resource.

Robotic Avoidance (RA) Response: Respond by triggering an immediate stop.

Hesitation (H) Response: Respond with hesitation using an AHP-based trajectory.

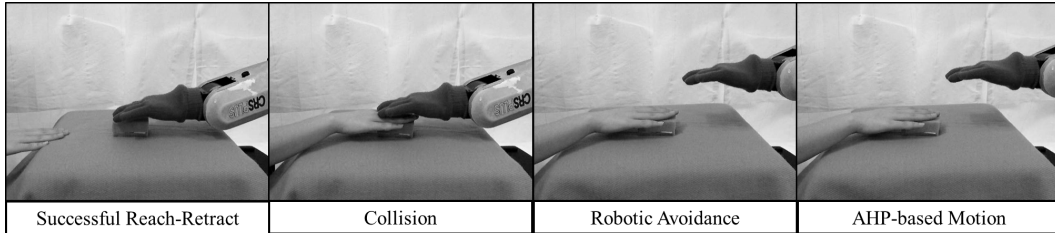


Figure 2. Screenshot of the videos shown to online participants. Each of these videos showed the experimenter and the robot reaching for the shared object in the centre of the workspace. The robot was programmed to stop at the same location for both the robotic avoidance and AHP-based motions.

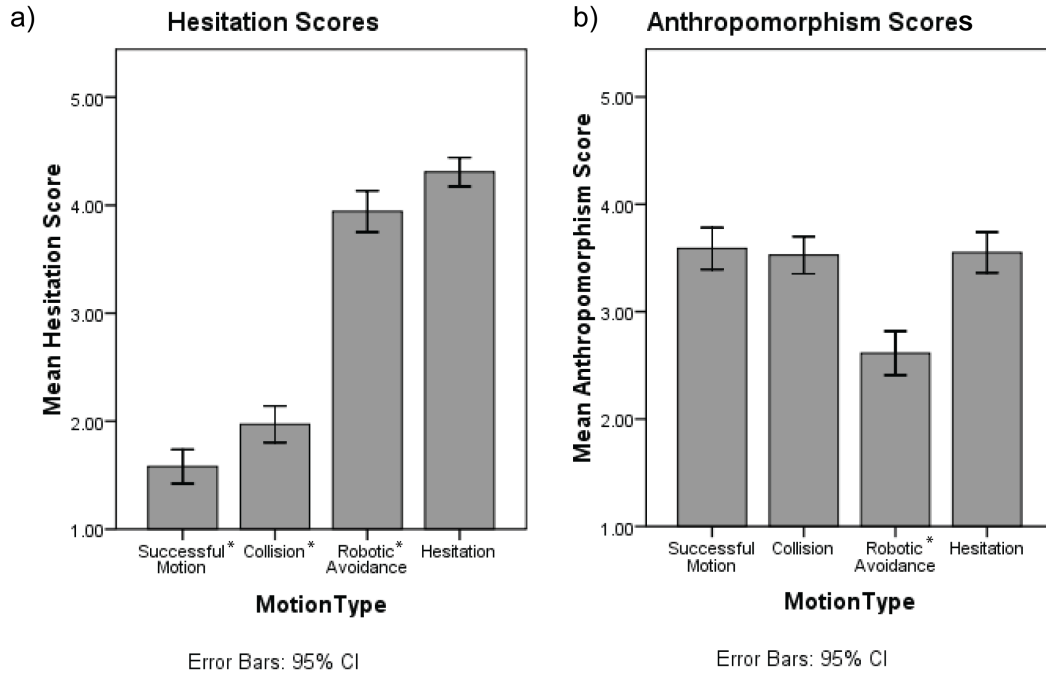


Figure 3. Preliminary study results. a) Overview of hesitation scores from a five-point Likert scale question (1. Not hesitant - 5. Hesitant). Results demonstrate significantly higher hesitation scores for AHP-based hesitation motions; b) Overview of anthropomorphism scores from the five-point Likert scale question (1. Machinelike - 5. Humanlike). Results demonstrate that AHP-based hesitation motions are perceived to be more anthropomorphic than robotic avoidance motions. The error bars show 95% confidence intervals. Conditions having significantly different scores than that of the AHP-based hesitation motions are marked with asterisks.

We hypothesized that subjects would identify AHP-based motions as hesitations even when the motions are observed in situ. In addition, even if the AHP-based motions were accurately recognized by subjects as hesitations, it was uncertain whether the added anthropomorphic feature of a conflict response would positively affect the human-robot team task. Based on numerous works that demonstrate the positive impact nonverbal human-robot communication has on efficiency and user perception, we hypothesized that untrained human users would perceive a robot more positively and complete a human-robot team task faster in the H condition than in both the B and RA conditions.

This experiment used a within-subjects design in order to account for individual differences in performing the task — both in terms of the methods used to complete the task as well as the pace of the task — and to effectively investigate our first hypothesis that the same untrained subject can distinguish collision response behaviours in the H condition from the RA condition. We outline the experimental procedure and task in Sections 4.1 and 4.2, respectively, and describe the quantitative measurement instruments used in this study in Section 4.3. We describe post-experiment interview questions and a post-hoc gesture identification experiment in Section 4.4.

4.1 Experimental Procedure

This study was approved by the University of British Columbia Behavioural Research Ethics Board and all subjects signed an informed consent form. Subjects were made aware that the task might result in physical contact with the robot, covered in soft safety padding.

The experiment included a pre-experiment questionnaire, a training session, and six trials of interaction with the robot — each followed by a questionnaire, a post-experiment interview, and a gesture identification experiment following the interview. This process is graphically summarised in Figure 4.

In total, 33 subjects (female: 13, male: 20) were recruited by posting a call for volunteers across the university campus and on the authors' lab website. The pre-experiment questionnaire was used to collect demographic information from the subjects. The age of the participants ranged from 20 to 52 (M: 26.83, SD: 7.24), and they were mostly unfamiliar with robots in general (M: 1.42, SD: .58, from a five-point Likert scale measure). All of the subjects were right handed. Only two subjects were familiar with the Barrett WAMTM robot. These two subjects indicated they had seen the robot at an exhibit or on a lab tour.

All subjects encountered each of the three conditions once in the first session (a set of three trials) and once again in the second session. The order of the trials within each session was randomised. In order to eliminate unintended visual cues or distractions, curtains surrounded the experimental area, including between the subject and the experimenter. Each trial took one to three minutes. The robot was introduced to the subjects as “the robot” or “it” rather than as “he” or “she”.

4.2 Experimental Task

The experimental task was designed to mimic an assembly line scenario. The subject's task was to pair a marble and a shape object according to provided examples and then place the pair (“finished product”) into an appropriate finished products bin. As shown in the workspace setup depicted in Figure 5, marbles were provided to the subjects in the marbles bin, which was located in the centre of the workspace and shared with the robot. The robot's task was to “inspect” the marbles bin fifteen times, by moving back and forth between the bin and its starting position, while the subject was pairing all the marbles provided. The human-robot team task was considered complete when both the subject and the robot finished their respective tasks; if one finished earlier than the other, then the subject or the robot waited for the other to finish. The subjects were told that the goal of the task was to minimize the team task completion time. To mitigate possible training effects, the example

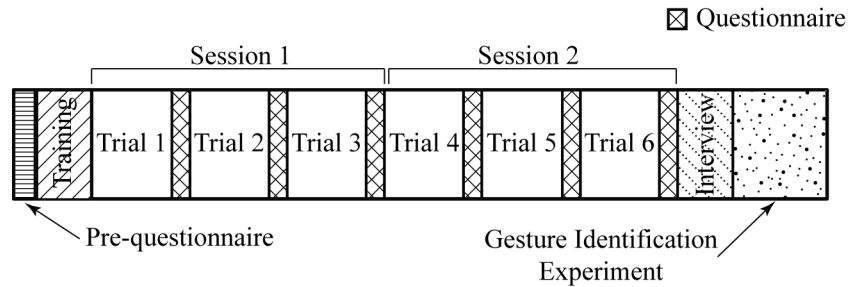


Figure 4. Overview of the experimental procedure. All subjects completed the pre-questionnaire, followed by a training session dedicated to familiarizing the subjects with the experimental task but not allowing any interaction with the robot. Two sessions of three trials were conducted in which the subjects completed the human-robot team task. The three conditions were randomized for each session.

marble-shape pairs that corresponded to each finished products bin changed at the beginning of every trial in random order.

The marbles bin contained twenty marbles of two different colours. The shapes bin placed in front of the subject contained small objects of various shapes (heart, circle, triangle, and rectangle); colours (blue, red, purple, yellow, and pink); and sizes (large, medium, and small). In pairing a marble to its respective shape, the shape object could be of any colour and size as long as the shape matched that of the respective example pairs.

During the trials, subjects wore a ring attached to a cable potentiometer (SP1-50, Celesco, Chatsworth, CA, USA) on their dominant hand at all times. They were instructed to use only the instrumented dominant hand to pick each marble from the marbles bin. This allowed us to monitor the approximate extension of the subject's hand during the trials. The subject used the non-dominant hand or both hands to handle the shape objects. The following rules were explained to the subject: any instance of collision with the robot resulted in a penalty score, and pairing mistakes made during the task were not to be corrected by the subject but were counted at the end of the trial and added to the penalty score. Counts of collisions and mistakes are reported in the results section. After details and rules of the experimental task were explained, all subjects practised the pairing task by themselves during the training session before interacting with the robot in the first trial.

4.2.1 Robot Motions and Collision Detection The robot was programmed to complete its “inspection” task by moving between the starting position and the marbles bin in a fixed task-space configuration. Instead of completing the task at a fixed speed, the robot monitored the speed of the subject's reaching motions and copied this speed in its own motions. Since the robot is capable of generating speeds that, albeit still safe due to padding and other safety mechanics, could be

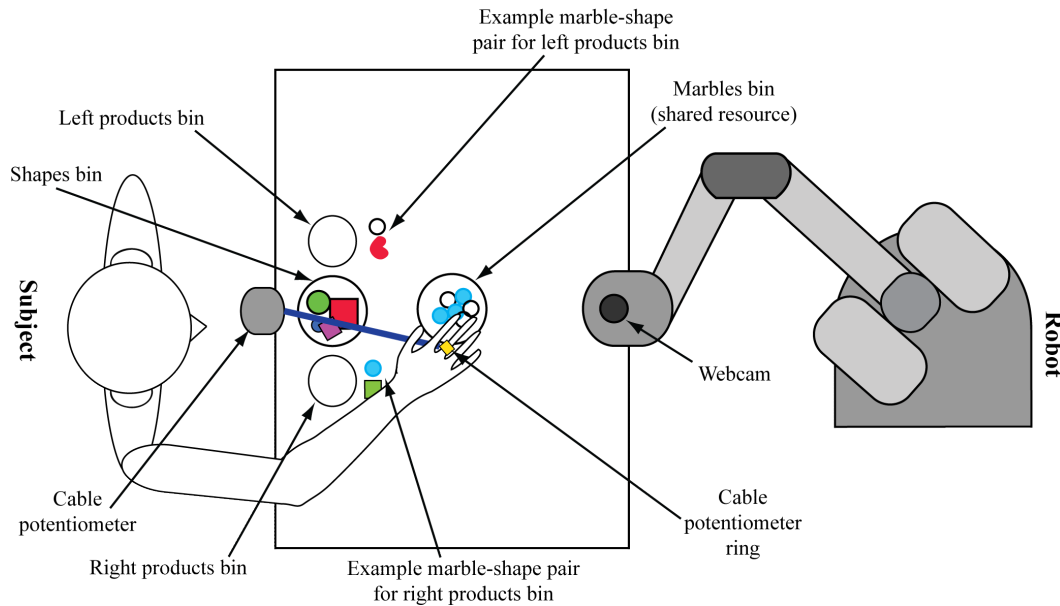


Figure 5. Top view of the experimental setup. Subjects sat across from the robot. The subjects' task was to pick up marbles from the marbles bin one at a time and assemble it with a shape from the shapes bin according to the example marble-shape pairs. The robots task was to inspect the marbles bin. The robot is shown in its initial, ready-to-reach position.

threatening to subjects, matching speeds with the subject enabled each subject to set the pace of the interaction at a comfortable level. This also allowed the a_1 and t_1 parameters of AHP-based motion trajectories to be randomised by the natural variances in each subject's motion. The subject was always the first one to move in all trials.

The robot used the subject-generated cable potentiometer readings to detect occurrence of potential collisions. In the B condition, the robot continued to reach for the marbles bin despite detecting a potential collision. This resulted in collisions and near-collision situations. In the H condition, the robot followed an AHP-based trajectory using the method described in Section 3. Immediately upon returning to its initial location, it attempted to reach for the shared resource again. In the RA condition, the robot responded by abruptly stopping at the point of collision detection and retracting back to its starting position. Similar to the H condition, the robot immediately re-attempted to access the resource. Figure 6 illustrates the flow of interaction for the three conditions.

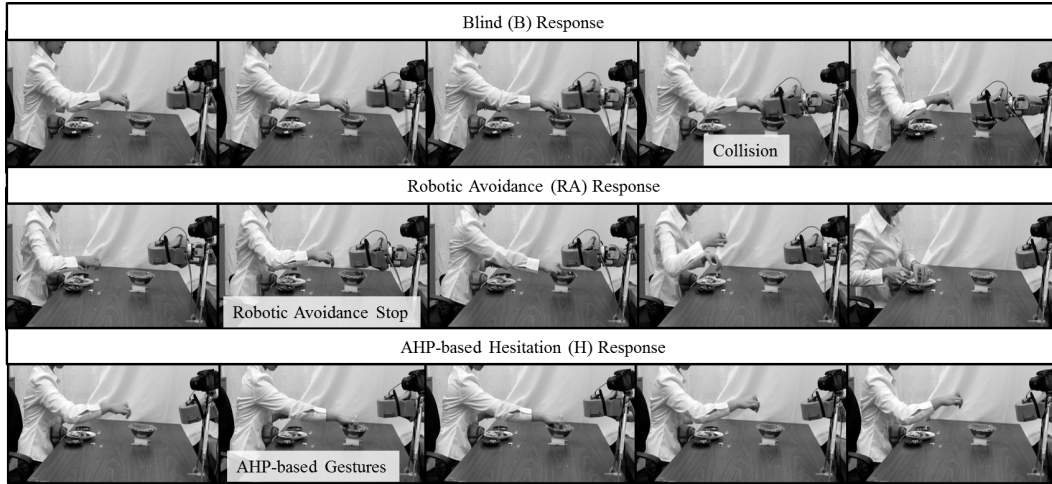


Figure 6. Sequences of video captures for the three response types. In both RA and H responses, the subject continues to reach for the shared resource and successfully picks up a marble without any collision with the robot. The position and acceleration profile of reference trajectories for each of the three response types is outlined in Figure 1.

4.3 User Perception and Task-Related Measurements

At the end of each trial, the subjects were brought away from the experimental area and completed a questionnaire. The questionnaire was comprised of select Likert scale questions from the Godspeed survey (Bartneck, Kulić, Croft, & Zoghbi, 2009) and Moon and Nass' (1996) questionnaire, which is used widely in human-computer teamwork studies (e.g., Burgoon et al., 2002; Burgoon et al., 2000). Questions from the Godspeed survey provided five key user perception measures of the robot: *animacy*, *anthropomorphism*, *perceived safety*, *perceived intelligence*, and *likeability*. Selections from the Moon and Nass' questionnaire were used for three teammate perception measures: *dominance*, *usefulness*, and *emotional satisfaction*.

For each trial, five different task-related measures were also collected. *Human task completion time* was estimated as the time between the experimenter's 'Go' signal and the point when the subject completed the pairing task. *Robot task completion time* was measured as the duration between

the start of the experiment program and the completion of the robot’s last retraction motion. The experimenter’s ‘Go’ signal coincided with the start of the experimental program. *Human-robot task completion time* was taken as the longer of the human and robot task completion times. The number of *mistakes* in a trial was measured by counting the number of misplaced marbles or shapes. A *collision* count for each trial was taken from the video recording of the experiment.

4.4 Post-Experiment Interview and the Gesture Identification Experiment

After the six trials, the experimenter conducted a post-experiment interview. Subjects were first asked to identify which one of the six trials they preferred the most. They were also asked whether and when they felt nervous or uncomfortable interacting with the robot. Qualitative responses to these questions were used to confirm the quantitative questionnaire responses.

Next, the experimenter explained the three different response modes of the robot to the subjects: B response was described as the robot not responding to potential collisions at all; RA response was described as the robot stopping abruptly upon encountering a potential collision; and H was described the robot hesitating in response to potential collisions.

Subjects were asked to comment whether they recognized the differences between the RA and H responses, and to identify in which trials they saw H responses. These questions were asked after the six trials rather than between trials, such that the subjects were not led to intentionally focus on the robot’s collision responses during the trials. Since responses to this question were from memory (rather than in situ), the possibility of bias necessitated a post-hoc experiment.

Five of the nineteen subjects who claimed to have recognized differences between the RA and H responses participated in a short additional experiment. In this gesture identification experiment, the robot was programmed to continuously reach for the marbles bin with one-second rests between attempts. Instead of completing the marble-shape pairing task, the subjects were asked to intentionally interrupt the robot by reaching for the marbles bin and then to verbally label which of the two responses (H or RA) the robot exhibited in its stopping behaviour. Each subject triggered at least five instances of collision response behaviours.

Due to the implementation limitation outlined in Section 3.1, the robot responded with an H response when a potential collision was detected before the reaching motion arrived at the launch acceleration at t_1 , and responded with a RA response if detected after t_1 . The subjects were not aware of this fact. While this approach allowed the order of the two types of responses to be naturally randomised based on the subject’s triggering of the response, this resulted in the number of triggered H and RA responses to be unbalanced; a total of 15 H and 37 RA responses were triggered.

5. Results

Due to technical problems ($N = 7$), subject failure to follow instructions ($N = 1$), and insufficient occurrence of resource conflicts ($N = 1$), data from only 24 subjects (female: 12, male: 12) were analysed and are reported here.

Subject responses from the post-experiment interview and the gesture identification experiment were used to test whether the AHP-based motions were recognized by the subjects in situ. Using the three response types — Blind (B), Robotic Avoidance (RA), and Hesitation (H) — as factors, we employed a repeated-measures ANOVA to analyse the subject perception measurements from the questionnaire and investigated whether AHP-based responses elicited a more positive perception of the robot by the subjects. A repeated-measures ANOVA was also conducted on the task-related measures to test whether the AHP-based motions positively affected the human-robot task completion in our experimental domain. To account for the within-subjects nature of the study, we also analysed the human perception and task-related measurements with the two sessions as a factor and checked for significant training effects. A significance level of $\alpha = 0.05$ was used for

all inferential statistics. Significance level adjustments for all post-hoc analyses were made using the Bonferroni approach. Measures violating the sphericity assumption were corrected with the Greenhouse-Geisser approach. Our analyses are presented in the order of the hypotheses.

5.1 Can Humans Recognize AHP-Based Motions as Hesitations in Situ?

Two coders coded the verbal responses to the final interview question, which asked the subjects whether they noticed the difference between H and RA responses. An inter-rater reliability analysis for this question yields a Cohen's Kappa of 0.75 ($p < .001$) and is considered a substantial level of consistency. A majority of the subjects (79%) said they noticed the difference between the two stopping motions and were able to correctly identify the trials that contained H responses. Considering that a minimum of one resource conflict occurred in each trial, and the fact that subjects were primarily attending to their own tasks, we consider this to be a high level of recognition.

The results of the gesture identification experiment also support our hypothesis that AHP-based trajectories are perceived as hesitations in situ. Of the 52 collision responses triggered, the subjects identified 87% correctly; only six RA responses were falsely identified as hesitations (H). One H response was identified as an RA response.

5.2 Do Humans Perceive AHP-Based Responses More Positively?

Empirical findings for the eight perception measures collected from our questionnaire yield an overarching test of our hypothesis that a robot is more positively perceived in the H than the RA and B conditions. We also provide results from the qualitative responses obtained from the post-experiment interview to support our quantitative findings.

Of the eight human perception measures collected, all but the *perceived intelligence* measure yielded an internal reliability score above 0.70. Hence, we exclude this measure from our discussion. The number of items we used to collect the perception measures was reduced or modified from methods used by Y. Moon and Nass (1996) and Bartneck et al. (2009). Nonetheless, internal reliability scores from our data show similar values as the original measures. Table 1 outlines the measurements, the items used for each measurement, and their reliability scores. We found no significant gender effects from the eight perception measures.

While all but the *dominance* measure show a tendency to increase from the first to the second session, this increase between sessions is significant only for the *perceived safety* ($F(1, 22) = 6.46, p < .05$) and *animacy* measures ($F(1, 22) = 5.63, p < .05$). We found no significant interactions between response types and sessions. Across the three response types, all but *usefulness* and *emotional satisfaction* measures show statistically significant score differences. Figure 7 graphically summarises the mean and standard error of the scores having significance.

Our data suggest that the subjects perceived the robot to be significantly less dominant in the H ($p < .001$) and RA ($p < .001$) response conditions than in the B condition ($F(1.42, 31.16) = 40.92, p < .001$).

Subjects also perceived the robot to be significantly more likeable in the H ($p < .0001$) and RA ($p < .01$) conditions than the B condition ($F(2, 44) = 18.36, p < .0001$), while the *likeability* scores of the H and RA conditions are not significantly different from each other ($p = .50$). This is consistent with subjects' interview responses to the question on which trial they preferred the most. Two individuals coded the interview responses with a high level of inter-rater reliability (Cohen's Kappa of 0.88, $p < .001$). Trials with the H condition were the most preferred (42%), followed by those with the RA (37%) and B (21%) conditions.

Responses to the perceived safety measure show a trend that the robot tended to be perceived safer in the H ($p = 0.10$) and RA ($p = 0.07$) conditions than in the B condition ($F(2, 44) = 4.03, p < .05$); however, these results are not statistically significant. The perceived safety of the

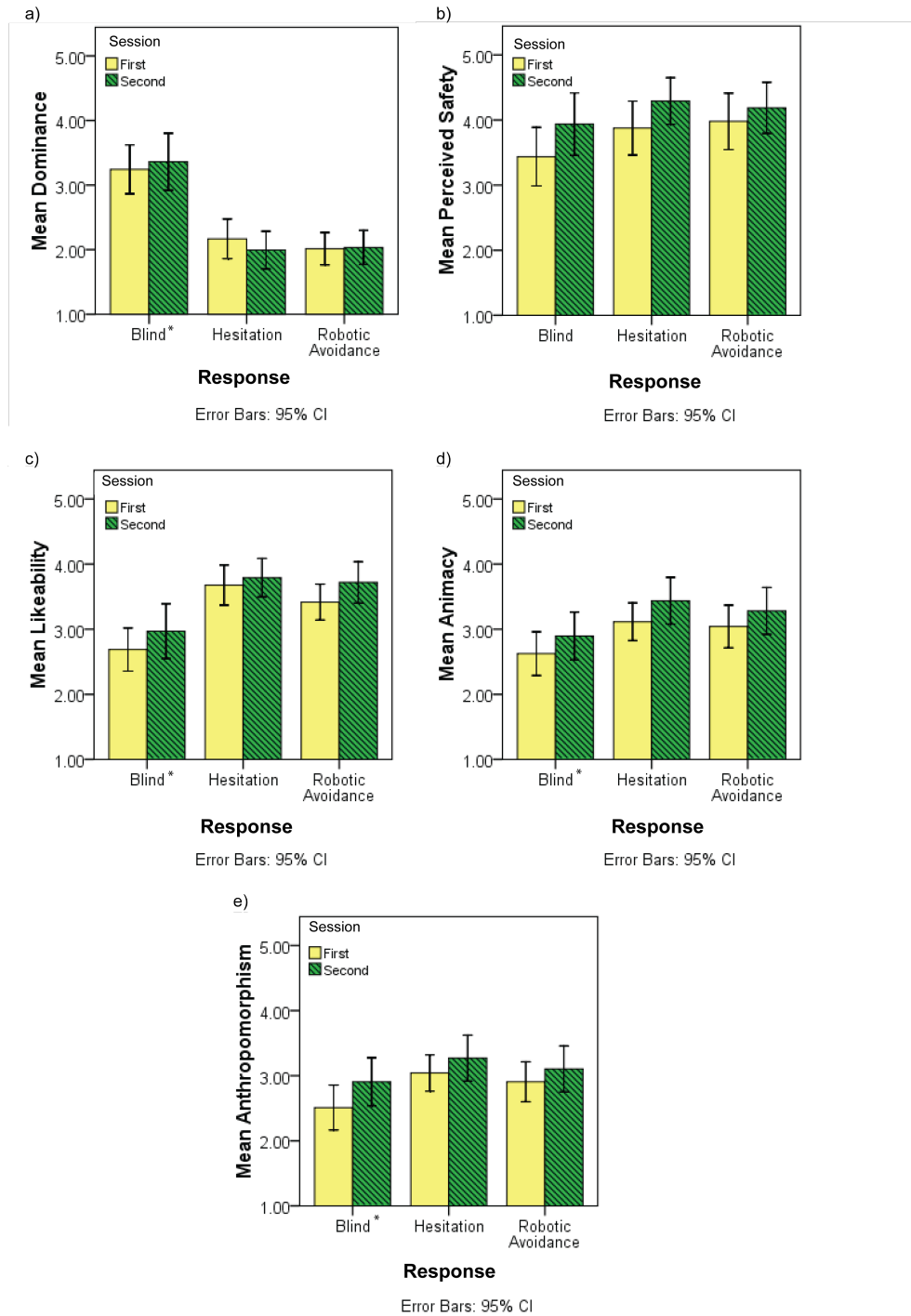


Figure 7. Overview of user perception measures having significant differences. The B condition is perceived to be more dominant, less likeable, animate, and anthropomorphic than the H condition. The RA condition is not perceived to be significantly more animate or anthropomorphic than the B condition. None of these perception measures show significant difference between the H and RA conditions. Conditions having significantly different scores than that of the H condition are marked with asterisks.

robot in the H and RA conditions were not found to be statistically different when compared against each other ($p = 1.00$). Qualitative responses to the second interview question support this finding. Over half of the subjects (58%) expressed that they felt uncomfortable or nervous interacting with the robot in at least one trial. Two coders coded the subjects' responses with a substantial level of inter-rater reliability (Cohen's Kappa of 0.75, $p < .001$). A majority of these subjects (57%) attributed this to having nearly collided with the robot during the B condition trials.

The questionnaire responses suggest that the robot's collision response behaviour affects its *perceived animacy* ($F(1.44, 31.66) = 4.96, p < .05$). The robot was perceived to be more animate in the H condition than in the B condition ($p < .05$). Significant differences also exist for the *anthropomorphism* measure ($F(1.44, 31.66) = 4.96, p < .05$). The robot was perceived to be more anthropomorphic in the H condition than in the B condition ($p < .05$). In contrast, the robot was not perceived to be more animate or anthropomorphic in the RA condition as compared to the B condition ($p = .17$ and $p = .28$ respectively). There are no significant differences in *perceived animacy* and *anthropomorphism* between the H and RA conditions ($p = .85$ and $p = .49$ respectively).

Summary. Although subjects perceived the robot to be more animate, anthropomorphic, safe, likeable, and less dominant in the H than the B condition, scores from these measures do not show significant improvements when compared to the RA condition. Hence, our results do not provide strong evidence for the hypothesis that a robot responding with AHP-based motions improves human perception of the robot.

Table 1: Internal reliabilities of the eight self-reported measures are presented here. Only the measures with a Cronbach's alpha greater or equal to 0.70 are analyzed and reported. All but *perceived intelligence* meet this requirement. Cronbach's alpha values for *dominance*, *usefulness*, and *emotional satisfaction* from the work of Y. Moon & Nass (1996) were 0.89, 0.80 and 0.86, respectively. Multiple alpha values are reported in Bartneck et al.'s (2009) work from cited studies and their range is as follows: *anthropomorphism* (0.86 to 0.93), *animacy* (0.70 to 0.76), *likeability* (0.84 to 0.92), and *perceived safety* (0.91).

Measures	Cronbach's alpha	Items
Dominance	0.88	Aggressive, Assertive, Competitive, Dominant, Forceful, Independent
Usefulness	0.79	Efficient, Helpful, Reliable, Useful
Emotional Satisfaction	0.84	How much did you like this robot? How much did you like working with this robot? Boring (reverse scale), Enjoyable, Engaging
Perceived Safety	0.91	Anxious, Agitated
Likeability	0.87	Like, Kind, Pleasant, Friendly
Animacy	0.78	Apathetic, Artificial, Mechanical, Stagnant
Anthropomorphism	0.81	Artificial, Fake, Machinelike, Moving Elegantly
Perceived Intelligence	0.54	Incompetent (reverse scale), Intelligent

5.3 Does Hesitation Elicit Improved Task Completion?

To test our hypothesis that a human-robot team will complete a given task faster when a robot responds to conflicts using AHP-based motions, we employed a repeated-measures ANOVA to analyse the *human*, *robot*, and *team task completion times* with response types as a factor.

Measured completion times for both *human-robot team task* ($F(2, 46) = 76.88, p < .0001$) and *robot task* ($F(2, 46) = 103.46, p < .0001$) are significantly shorter in the B condition as compared to the RA ($p < .001$) and H ($p < .001$) conditions. However, while a few *collisions* occurred in the B condition ($M = 1.06, SD = 1.17$), no *collision* occurred in the H and RA conditions by design ($X^2(2, N = 48) = 60.00, p < .001$). We did not obtain enough statistical power to report on significant differences in the number of mistakes made between the H and RA conditions. Table 2 summarizes the number of mistakes by condition. The two task completion times are not significantly different between the H and RA conditions ($p = .16$ and $p = .38$ for human and robot task completion time respectively, with an adjusted $\alpha = 0.017$). Figure 8 graphically illustrates these results.

Faster *team* and *robot task completion times* are observed during the second session of trials compared to the first ($F(1, 23) = 8.51, p < .01$) and ($F(1, 23) = 11.64, p < .01$) for the team and robot task completion times respectively. Across the three task completion times, we found no significant gender effects.

Summary. Results fail to support our hypothesis that AHP-based collision responses result in improved human-robot task completion time. However, they do suggest that the added communicative feature in collision responses does not significantly hinder completion time of a human-robot task compared to RA responses.

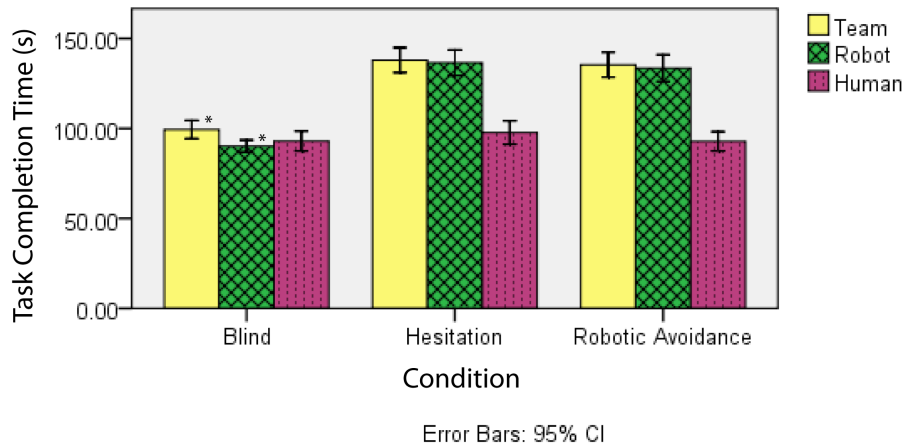


Figure 8. Overview of the task completion times. Both the subject and the robot completed their individual and team tasks faster in the B condition than in the H and RA conditions. However, significantly more number of collisions occurred in the B condition. The human task completion time is unaffected by the collision response behaviour of the robot. Task completion times that are significantly different from that of the H condition are marked with asterisks.

Table 2: The number of mistakes made in each condition is summarised by condition. Each cell contains the number of subjects who made mistakes. The total number of mistakes reported on the bottom row is the sum of the mistakes made by all subjects.

	Blind	Hesitation	Robotic Avoidance
1 Mistake	4	3	4
2 Mistakes	1	0	1
3 Mistakes	1	0	0
Total Number of Mistakes	9	3	6

6. Discussion

In our previous work (A. Moon et al., 2011), a robot traced the recorded wrist trajectories of human hesitation motions and demonstrated that anthropomorphic hesitation gestures of an articulated 6-DOF robot can be robustly recognized. This provided a strong indication that such simplified replications of human wrist trajectories are sufficient to generate visually apparent anthropomorphic hesitation gestures. The AHP was proposed as an empirically derived trajectory specification for generating hesitation gestures on a robot. In this work, we extended our earlier work by outlining an implementation method and testing the efficacy of AHP-based motions as a communicative collision response mechanism.

The findings of our study clearly demonstrate that non-expert human subjects are able to relate the proposed AHP-based trajectories to hesitation and as more anthropomorphic than the robotic avoidance behaviour. As demonstrated in the results of our main experiment, human subjects also recognized the AHP-based motions as hesitations in situ. Subjects also accurately distinguished them as different from abrupt, robotic avoidance behaviours. The fact that subjects were able to identify nuances from the brief motions of a robotic manipulator in situ gives us great confidence that anthropomorphic gestures can be useful in fast-paced human-robot collaboration.

Our approach allows a robotic system to dynamically generate trajectories it can use to exhibit communicative hesitation gestures in a variety of resource conflict scenarios involving reaching motions. Since the blind (B), hesitation (H), and robotic avoidance (RA) responses to potential collisions are planar Cartesian motions (xz -plane), both the 6-DOF and 7-DOF robots used in our studies did not need to use all of their DOF in generating these responses – the robot in the preliminary study used only 4-DOF for the planar Cartesian motions. As such, our work provides strong empirical evidence that communicative and anthropomorphic features of a gesture, and in particular the AHP-based hesitation gestures, can be superimposed onto the motions of typical 6-DOF industrial robots.

However, the impact of AHP-based hesitations in human-robot collaboration requires further investigation. While the H, and not the RA, condition was perceived significantly more anthropomorphic, animate, and likeable than the B condition, and echoes our findings from the preliminary study, this evidence is not sufficient to support our hypothesis that AHP-based collision responses improve the user’s perception of the robot. The lack of significant differences in the H and RA conditions in the main experiment is contrary to other evidence that supports AHP-based motions as the more anthropomorphic behaviour than robotic avoidance motions. In the preliminary study, where subjects’ main task was to visually observe robot motions via the online survey instrument, AHP-based motions were perceived to be significantly more anthropomorphic than robotic avoidance motions. In addition, AHP-based motions were robustly distinguished as different from robotic avoidance motions during the in situ Gesture Identification experiment.

One possible explanation for this unexpected finding is that the number of AHP-based and robotic avoidance responses triggered in each of the sessions ($M=1.06$, $SD=1.17$) was not sufficient to demonstrate the impact nuanced gesture differences have on human perception of the robot. It is also possible that, in contrast to a context where subjects are solely focused on observing robot behaviours (e.g., the preliminary online survey and the Gesture Identification experiment), subjects pay less attention to the details of the robot's motion and are therefore less likely to notice subtle differences in a context where subjects are focused on their main task (i.e., sorting). In contrast to the B condition where the robot's motions impact the subject's primary task of sorting by colliding with the subject or blocking the subject's access to the marbles bin, the robot's responses in both the H and RA conditions yield access of the marbles to the subjects, allowing subjects to continue their task relatively unaffected regardless of whether they pay attention to the robot's motions or not.

While the B condition yielded significantly faster task completion time compared to the H and RA conditions, our results also did not show a difference between task completion times for the H and RA conditions.

This lack of differences between the H and RA conditions may also be attributed to the small number of near collisions occurred during the experiment. It is possible that a significantly more number of near collisions may need to be triggered for each trial in order for the impact of nuanced gestures to be observable in task completion time.

On reflection, there are several plausible explanations for why the B condition yielded the best task completion time. First, the incentive to avoid collisions was not sufficiently strong — this is evidenced by subjects' comments from the interview stating that they found the collisions entertaining (as opposed to dangerous or frustrating). This underlines the challenge of creating experiments that reflect potential real-life human-robot collaboration scenarios. We believe that creating a perception of possible danger while not undermining the safety of subjects would yield a result more reflective of reality.

Second, by nature of the task, subjects' task performance was not directly affected by performance of the robot; in other words, subjects did not require the robot to succeed in its "inspection" task in order to continue their task, although frequently interrupting the robot's inspection tasks would result in longer human-robot task completion time. Hence, there was no tangible incentive that stopped subjects from ignoring the robot and performing their tasks that, in turn, penalised the performance of the robot in both the H and RA conditions (which were programmed to avoid collisions). Furthermore, the subjects were aware that they were being timed, potentially motivating subjects to finish their own tasks as quickly as possible, regardless of the robot's behaviours and despite being instructed to finish the human-robot team task as fast as possible. Thus, it should not surprise us that human performance remained unaffected across conditions.

Indeed, the B condition generated the best team and robot task completion times. However, this result was accompanied by larger counts of collisions compared to zero collisions in the H and RA conditions. Considering industrial applications in which correcting for mistakes or occurrence of collisions can negatively affect completion times of a task and the quality of the finished product, we feel it is important to consider these performance measures in human-robot team collaboration.

6.1 Study Limitations

As outlined in Section 3.1, real-time triggering of AHP-based responses need to take place before the reaching motion arrives at the launch acceleration at t_1 . It is possible, however, to extend the allowable period of hesitation decision-making. Since the quintic reach trajectories, Equations (1) and (6), share the same acceleration (a_1 at t_1) upon which they both start to decelerate, it is possible to use an interpolation function to smoothly transition from Equation (1) to Equation (6) at some t

after t_1 has been reached ($t_1 + \delta < t < t_2$). However, the acceptable bound for the value of δ is unknown. Hence, this technique requires further investigation and testing.

As mentioned in previous sections, the AHP describes characteristic trajectory features of only one type of hesitation gesture. Hence, the findings of our study are limited to the type of hesitation gesture that always result in yielding to the user without persistent pursuit for the resource.

In addition, we did not calculate a performance score that encompasses the team's task completion time, mistakes, and collisions, since the amount of penalty applied to each task-related measure can affect the final performance score. However, if we had a fair means to calculate such aggregate penalty scores, more complete and informative comparison of H responses' impact on human-robot team task would be available.

7. Conclusions and Future Work

Our main contribution to the field of HRI is in advancing nonverbal, communicative management of shared resources between humans and robots that involve interactions without predefined turn-taking rules.

An implementation method of AHP was introduced, and its communicative efficacy was tested. We presented strong evidence that trajectories generated using our AHP-based approach are perceived to convey hesitation when compared to robotic avoidance motions. Findings from the preliminary and main experiment support this to be true whether the motion is observed via a video recording, or in situ while engaged in a collaborative task with the robot. The quantitative results from the main study indicate that the robot is considered less dominant and more animate when it hesitates or abruptly stops than when it does not respond to subjects at all. We did not find significant evidence of improved task completion time or user perception of the robot in AHP-based responses in comparison to robotic avoidance motions (emergency manoeuvres).

The experimental system in our study only employed a robot's display of hesitations and did not have any means of responding to the subject's display of hesitations. We conjecture that by implementing a system that robustly identifies high-speed human hand gestures and their meanings, we will be able to explore socially acceptable yielding and mediation behaviours of a robot in human-robot resource conflict scenarios. With the hesitation trajectory specification that has been empirically tested in our study, these socially appropriate behaviours could be programmed into trajectory parameters of cooperative robotic devices. We believe our design of hesitation gestures has built a framework for studying such bidirectional human-robot nonverbal communication and negotiation.

Our study also shows that, while use of a "blind" robot that does not respond to human motions yields faster team completion times for a collaborative task than a robot that hesitates or abruptly stops, this was at a cost of numerous collisions. Accounting for the number of collisions and the fact that the completion time of the human's task was unaffected by the addition of hesitation gestures, we are optimistic that any human-robot collaboration system embodying hesitation gestures will produce a more positive overall outcome.

With the empirically-tested AHP-based design of hesitation gestures, we believe this work provides an important stepping stone for developing robotic systems that can respond to human-robot conflicts using communicative nonverbal gestures.

Acknowledgements

We are grateful to the participants of the experiments and to the survey respondents who volunteered for this study. Financial support for this research was provided by the Natural Sciences and Engineering Research Council of Canada, the Canada Foundation for Innovation, and the UBC Institute for Computing, Information and Cognitive Systems.

References

- Argyle, M. (1994). *The psychology of interpersonal behaviour* (5th ed.). London, UK: Penguin.
- Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, 1(1), 71–81. doi:10.1007/s12369-008-0001-3
- Bartneck, C., van der Hoek, M., Mubin, O., & Al Mahmud, A. (2007). “Daisy, daisy, give me your answer do!”. New York, NY: ACM Press. doi:10.1145/1228716.1228746
- Becchio, C., Sartori, L., Bulgheroni, M., & Castiello, U. (2008). Both your intention and mine are reflected in the kinematics of my reach-to-grasp movement. *Cognition*, 106(2), 894–912. doi:10.1016/j.cognition.2007.05.004
- Becchio, C., Sartori, L., & Castiello, U. (2010). Toward you: The social side of actions. *Current Directions in Psychological Science*, 19(3), 183–188. doi:10.1177/0963721410370131
- Bethel, C. L., & Murphy, R. R. (2006). Affective expression in appearance constrained robots. In *International Conference on Human-Robot Interaction* (p. 327). New York, NY: ACM Press. doi:10.1145/1121241.1121299
- Bratman, M. (1992). Shared cooperative activity. *The Philosophical Review*, 101(2), 327–341.
- Breazeal, C., Kidd, C., Thomaz, A., Hoffman, G., & Berlin, M. (2005). Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In *International Conference on Intelligent Robots and Systems* (pp. 383–388). IEEE/RSJ. doi:10.1109/IROS.2005.1545011
- Burgoon, J. K., Bonito, J. a., Ramirez, A., Dunbar, N. E., Kam, K., & Fischer, J. (2002). Testing the interactivity principle: Effects of mediation, propinquity, and verbal and nonverbal modalities in interpersonal interaction. *Journal of Communication*, 52(3), 657–677. doi:10.1111/j.1460-2466.2002.tb02567.x
- Cohen, P. R., & Levesque, H. J. (1991). Teamwork. *Noûs*, 25(4), 487–512.
- Dittrich, W. H., & Lea, S. E. (1994). Visual perception of intentional motion. *Perception*, 23(3), 253–68.
- Doob, L. W. (1990). *Hesitation: Impulsivity and reflection*. Westport, CT: Greenwood Press.
- Fincannon, T., Barnes, L., Murphy, R., & Riddle, D. (2004). Evidence of the need for social intelligence in rescue robots. In *International Conference on Intelligent Robots and Systems* (pp. 1089–1095). IEEE/RSJ. doi:10.1109/IROS.2004.1389542
- Flash, T., & Hogan, N. (1985). The coordination of arm movements: Mathematical model. *Journal of Neuroscience*, 5(7), 1688–1703.
- Fox, R., & McDaniel, C. (1982). The perception of biological motion by human infants. *Science*, 218(4571), 486–487. doi:10.1126/science.7123249
- Goetz, J., Kiesler, S., & Powers, A. (2003). Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *International Workshop on Robot and Human Interactive Communication* (pp. 55–60). IEEE. doi:10.1109/ROMAN.2003.1251796
- Goto, A., & Moon, A. (2012). *Hesitation - ROS Wiki*. Retrieved from <http://www.ros.org/wiki/hesitation>
- Grosz, B. J. (1996). Collaborative Systems. *AI Magazine*, 17(2), 67–85.
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, 57(2), pp. 243-259. Retrieved from <http://www.jstor.org/stable/1416950>
- Hinds, P., Roberts, T., & Jones, H. (2004). Whose job is it anyway? A study of human-robot interaction in a collaborative task. *Human-Computer Interaction*, 19(1), 151–181. doi:10.1207/s15327051hci19012_7
- Holroyd, A., Rich, C., Sidner, C. L., & Ponsler, B. (2011). Generating connection events for human-robot collaboration. In *International Symposium on Robot and Human Interactive Communication* (pp. 241–246). IEEE. doi:10.1109/ROMAN.2011.6005245
- Huang, C.-M., & Thomaz, A. L. (2011). Effects of responding to, initiating and ensuring joint attention in human-robot interaction. In *International Symposium on Robot and Human Interactive Communication* (pp. 65–71). IEEE. doi:10.1109/ROMAN.2011.6005230
- Ju, W., & Takayama, L. (2009). Approachability: How people interpret automatic door movement as gesture. *International Journal of Design*, 3(2), 1-10.
- Kazuaki, T., Motoyuki, O., & Natsuki, O. (2010). The hesitation of a robot: A delay in its motion increases learning efficiency and impresses humans as teachable. In *International Conference on Human-Robot Interaction* (Vol. 8821007, pp. 189–190). Osaka, Japan: IEEE. doi:10.1109/HRI.2010.5453200

- Kim, H., Kwak, S. S., & Kim, M. (2008). Personality design of sociable robots by control of gesture design factors. In *International Symposium on Robot and Human Interactive Communication* (pp. 494–499). Munich: IEEE. doi:10.1109/ROMAN.2008.4600715
- Klapp, S. T., Kelly, P. A., & Netick, A. (1987). Hesitations in continuous tracking induced by a concurrent discrete task. *Human Factors*, 29(3), 327–337.
- Kulić, D., & Croft, E. (2006). Physiological and subjective responses to articulated robot motion. *Robotica*, 25(01), 13. doi:10.1017/S0263574706002955
- Manera, V., Becchio, C., Cavallo, A., Sartori, L., & Castiello, U. (2011). Cooperation or competition? Discriminating between social intentions by observing prehensile movements. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, 211(3-4), 547–56. doi:10.1007/s00221-011-2649-4
- Moon, A., Parker, C. A. C., Croft, E. A., & Van der Loos, H. F. M. (2011). Did you see it hesitate? - Empirically grounded design of hesitation trajectories for collaborative robots. In *International Conference on Intelligent Robots and Systems* (pp. 1994–1999). San Francisco, CA: IEEE/RSJ. doi:10.1109/IROS.2011.6048228
- Moon, Y., & Nass, C. (1996). How “real” are computer personalities?: Psychological responses to personality types in human-computer interaction. *Communication Research*, 23(6), 651–674. doi:10.1177/009365096023006002
- Reed, K., Patton, J., & Peshkin, M. (2007). Replicating Human-Human Physical Interaction. In *International conference on robotics and automation* (pp. 3615–3620). Rome, Italy: IEEE. doi:10.1109/ROBOT.2007.364032
- Reed, K., Peshkin, M., Hartmann, M. J., Grabowewy, M., Patton, J., & Vishton, P. M. (2006, May). Haptically linked dyads: Are two motor-control systems better than one? *Psychological Science*, 17(5), 365–6. doi:10.1111/j.1467-9280.2006.01712.x
- Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge, England: Cambridge University Press.
- Riek, L. D., Rabinowitch, T., Bremner, P., Pipe, A. G., Fraser, M., & Robinson, P. (2010). Cooperative gestures: Effective signaling for humanoid robots. In *International Conference on Human-Robot Interaction* (pp. 61–68). Osaka, Japan: ACM/IEEE. doi:10.1109/HRI.2010.5453266
- Saerbeck, M., & Bartneck, C. (2010). Perception of affect elicited by robot motion. In *International Conference on Human-Robot Interaction* (pp. 53–60). New York, NY: ACM/IEEE. doi:10.1145/1734454.1734473
- Salem, M., Rohlfing, K., Kopp, S., & Joubin, F. (2011). A friendly gesture: Investigating the effect of multimodal robot behavior in human-robot interaction. In *International Symposium on Robot and Human Interactive Communication* (pp. 247–252). IEEE. doi:10.1109/ROMAN.2011.6005285
- Thorisson, K. R., & Cassell, J. (1999). The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence*, 13(4-5), 519–538. doi:10.1080/088395199117360
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(5), 675–91; discussion 691–735. doi:10.1017/S0140525X05000129

Authors’ names and contact information: AJung Moon, Department of Mechanical Engineering, University of British Columbia, Vancouver, Canada. Email: ajung@ieee.org. Chris A.C. Parker, Department of Mechanical Engineering, University of British Columbia, Vancouver, Canada. Email: parkerca@ieee.org. Elizabeth A. Croft, Department of Mechanical Engineering, University of British Columbia, Vancouver, Canada. Email: ecroft@mech.ubc.ca. H.F. Machiel Van der Loos, Department of Mechanical Engineering, University of British Columbia, Vancouver, Canada. Email: vdl@mech.ubc.ca.