

# Chapter 14

## Developing Robot Assistants with Communicative Cues for Safe, Fluent HRI

Justin W. Hart, Sara Sheikholeslami, Brian Gleeson, Elizabeth Croft, Karon MacLean, Frank P. Ferrie, Clément Gosselin and Denis Laurandau

### 14.1 Introduction

The Collaborative Advanced Robotics and Intelligent Systems (CARIS) laboratory at the University of British Columbia has focused on the development of autonomous

---

Justin W. Hart: Work completed while affiliated with the Department of Mechanical Engineering at University of British Columbia.

---

J. W. Hart (✉)

Department of Computer Science, University of Texas at Austin, Austin, USA  
e-mail: justinhart@utexas.edu

J. W. Hart · S. Sheikholeslami · E. Croft

Department of Mechanical Engineering, University of British Columbia,  
Vancouver, Canada  
e-mail: sara.sheikholeslami@alumni.ubc.ca

E. Croft

e-mail: elizabeth.croft@ubc.ca

B. Gleeson · K. MacLean

Department of Computer Science, University of British Columbia, Vancouver, Canada  
e-mail: brian.gleeson@gmail.com

K. MacLean

e-mail: maclea@cs.ubc.ca

F. P. Ferrie

Department of Electrical and Computer Engineering, McGill University,  
Montreal, Canada  
e-mail: ferrie@cim.mcgill.ca

C. Gosselin

Department of Mechanical Engineering, Laval University, Quebec City, Canada  
e-mail: clement.gosselin@gmc.ulaval.ca

D. Laurandau

Department of Electrical Engineering, Laval University, Quebec City, Canada  
e-mail: denis.laurandau@gel.ulaval.ca

© The Author(s) 2018

H. A. Abbass et al. (eds.), *Foundations of Trusted Autonomy*, Studies in Systems, Decision and Control 117, [https://doi.org/10.1007/978-3-319-64816-3\\_14](https://doi.org/10.1007/978-3-319-64816-3_14)

robot assistants which work alongside human workers in manufacturing contexts. A robot assistant is a device that is intended to aid and support the activities of a human worker, rather than working entirely autonomously - as a welding or painting robot in a work cell at a car factory would, or entirely under human control - as a teleoperated robot would. The creation of such robot assistants presents a number of challenges to the scientists and engineers designing them. A traditional robotic arm operating in a factory work cell is physically separated from human workers by a thick glass wall secured by steel beams that are often marked with bright orange paint. This creates an environment where the robot's operation is safe, because it is away from the people who must be kept safe from it, and where it is capable of performing its tasks with minimal input because they are repetitive and not subject to variations such as those introduced by typical human behavior. Many of the technological barriers to move these robots past the physical barriers of the work cell involve establishing a clear cycle of back-and-forth communication between the robots that would work alongside humans and their human collaborators. It is important for workers collaborating with robots to know what mode each robot is in and what actions they may take.

The purpose of a robot assistant is to aid and support the actions of a human worker. Robot assistants, by necessity, are not constrained to the confines of the work cell. For a robot assistant to be effective, it is important that the worker not be overly burdened by the task of controlling it, allowing the worker to concentrate on their portion of the shared task. It is also important that the robot assistant and human workers be able to safely operate in close proximity. Contributing to efforts to accomplish these goals, the CARIS laboratory focuses on the identification and exploitation of natural communicative cues, explicit or non-explicit, which can be implemented into the interfaces of robotic systems. These are human behaviors which communicate information to other people. Explicit cues are intentional behaviors performed with the purpose of communication, such as when one performs a hand gesture [6, 8, 9, 21]. Non-explicit cues are unintentionally performed, but broadcast intentions or other important information, such as when a person looks toward where they are about to hand an object to another person [1, 17, 23]. We posit that naturalistic communicative cues allow users to quickly benefit from collaboration with robots with minimal effort and training, and that through such communicative cues, it is possible to develop interfaces that are safe, transparent, natural, and predictable. To achieve this goal, we use a three phase design process. In the first phase, studies are performed in which two people collaborate on a shared goal. Recordings of these interactions are annotated in order to identify communicative cues used between the study participants in performing their tasks. In the second, these cues are described or mathematically modeled. In the third phase, they are implemented in robotic systems and studied in the context of a human-robot interaction.

Due to the close physical proximity between human workers and robot assistants, physical communicative cues and physical human-robot interaction play an important role in this research. An example of an interaction involving naturalistic communicative cues and physical human-robot interaction is human-robot handovers. The CARIS Laboratory has invested significant effort in the development of natural,

fluent human-robot handovers [2–4, 9, 10, 17]. Handover behaviors include not only visual cues [1, 2, 9, 10, 17, 23], but also important force and haptic cues [3, 4]. Other physical interaction studies, which have been carried out in conjunction with the Sensory Perception & Interaction (SPIN) Laboratory at UBC, include interactions in which participants tap and push on the robot to guide its behavior [5]. Our collaborators at Laval University developed an elegant backdrive mechanism for the robot assistant (discussed in Sect. 14.2), which allows users to interactively pose the device during interaction by pushing on the robot itself.

When people work together, their efforts become coordinated and fluently mesh with each other. Many prototype collaborative robot systems work by monitoring progress toward task completion, which can create a stop-and-go style interaction. Hoffman and Breazeal [11] propose a system in which a robot performs predictive task selection based on confidence based on estimates of the validity and risk of each task selection in order to improve interaction fluency. Moon et al. [17] demonstrate that a the performance of a brief gaze motion can improve the speed and perceived quality of robot-to-human handovers. A study by Hart et al. [10] found that study participants receiving an object via handover reach toward the location of a handover prior to the completion of the handover motion, indicating the importance of short-term predictions in human handover behavior. In current work, we are investigating systems which make short-term motion cue based predictions of human behavior in order to fluently and predictively act on these cues, rather than acting on completed motion gestures or task state. The intention of this work is to not only improve the efficiency with which tasks are carried out, but also the fluency of the interaction.

In order to assure the relevance of this work to manufacturing, studies at the CARIS Laboratory are grounded in real-world scenarios. Partnering with manufacturing domain experts through industrial partners provides the us and our collaborators with important insights into the applicability of our research to actual manufacturing scenarios, practices and operations. Recently the CARIS Laboratory completed a three year project that was carried out with academic partners at the Sensory Perception & Interaction (SPIN) Laboratory at UBC, the Artificial Perception Laboratory (APL) at McGill University, and the Computer Vision and Digital Systems Laboratory and Robotics Laboratory at Laval University. Importantly, General Motors of Canada served as our industrial partner in this endeavor. This project, called the Collaborative Human-Focused Assistive Robotics for Manufacturing, or CHARM project, investigated the development of robot assistants in supporting workers performing assembly tasks in an automotive plant. Partnering with General Motors provided us with manufacturing domain experts who could guide our vision based on industry best practices and provide insights based on real-world experience in manufacturing.

This chapter presents an overview of work in the CARIS Laboratory towards the construction of robot assistants. Section 14.2 provides an overview of the CHARM project highlighting the interdisciplinary nature of work in autonomous human-robot interaction and describe the robot-assistant developed for this project. Section 14.3 describes the methodology by which we study communicative cues, from observing

them in human-human interactions to description and modeling to experiments in human-robot interaction. Section 14.4 presents experimental findings that were made using this methodology in projects related to CHARM conducted at UBC. Section 14.5 describes current and future directions in which we are taking this research and conclude.

## **14.2 CHARM - Collaborative Human-Focused Assistive Robotics for Manufacturing**

The construction of robot assistants is a large undertaking involving work in several disciplines from the design of the robot itself, to sensing and perception algorithms, to the human-robot interaction design describing how the robot should behave when performing its task. To complete the construction of a state-of-the-art working system, we formed a group of laboratories specializing in these various disciplines and structured work on the robot assistant in such a way that it allowed each lab to independently pursue projects that highlight each group's expertise, while working towards the common goal of constructing an integrated robot assistant system. Insurance of progress towards a common goal was established through several avenues. At the start of the project, the plan of work was divided into a set of three research streams and three research thrusts, with teams of investigators assigned to areas which highlighted the strengths of their laboratories. Doing so allowed a large degree of freedom in the selection and pursuit of individual research projects, while carefully describing the relationships of these streams and thrusts to each other assured that the overarching research efforts of the group remained unified. Coordination of research efforts was maintained through regular teleconferences and inter-site visits by research personnel. These meetings culminated in annual integration exercises referred to as Plugfests, in which each team's developments over the past year were integrated onto a common robot assistant platform. For each Plugfest, the team agreed upon a shared collaborative assembly task to be performed by a dyad comprising the robot assistant and a human collaborator that would be implemented as a unified team spanning the research teams from each institution. This shared task would highlight the research of each group over the past year, and would be used in order to evaluate overall progress towards the shared goal of developing a robot assistant for use in a manufacturing operation. Extensive evaluation and discussion with manufacturing domain experts from General Motors provided the academic partners with the knowledge and expertise to understand and direct the impact and applicability of the work in their laboratories work in real-world usage.

### ***14.2.1 The Robot Assistant, Its Task, and Its Components***

Operating on a common platform and regular integration enabled us to make concrete progress toward our goal by constructing a real robot assistant that works on

a real-world problem. The development of autonomous robot assistants will require advancements in more than one contributing area. Without better sensing, robots will be unable to detect, track, and model all of the objects required. Without better control, they will not be able to physically interact with people in ways that users will expect them to. Without advances in human-robot interaction, they will neither understand nor be understood by their human collaborators. It was important for partners on the CHARM project to be chosen from a diverse fields representative of the challenges of the project, given the autonomy to pursue the research necessary for advancement, and be coordinated enough to remain focused on their shared goal. Keeping each other apprised of our progress through monthly teleconferences and regular email exchanges fostered an environment of collaboration between each subteam's individual efforts. As each Plugfest approached, the group would begin to focus on ways in which software could be integrated in onto the common platform. Yearly integration into the shared robot-assistant provided a real-world system to act as the testbed for these integration efforts.

#### **14.2.1.1 Car Door Assembly**

Car door assembly was chosen as the shared task for the human operator and robot assistant to collaborate on because it is presently one of the most complex and labor-intensive operations in vehicle assembly. The development of technology in this area could potentially both improve the effectiveness of workers in assembly and relieve them of portions of the assembly task that are ready to be automated. Through a series of off-site meetings, teleconferences, and site visits to vehicle manufacturing facilities, two objectives for car door assembly were identified:

- To reduce error rates and improve manufacturing quality
- To maintain or improve worker safety.

A plan was formulated to improve manufacturing quality by interactively passing parts from the robot to the worker. The robot assistant would present parts to the worker assembling a car door in the sequence in which they should be inserted, and tools in the sequence corresponding to the parts of the assembly presently worked on. This would ensure that the worker attaches the correct parts to the car door at the correct times, even as different models come down the line or as the assembly process changes. An improvement in quality would be accomplished through this process by reducing worker error rates, preventing the need to redo work that was performed incorrectly and reducing the error rate in finished products. By the end of the CHARM project, parts were presented to a worker in the test environment in an interactive fashion using an elegant handover controller which was developed by the CARIS Laboratory, with scenarios that could be either programmed into a State Machine Controller (SMC) or reasoned about automatically using a planner developed by CARIS for CHARM.

Worker safety was identified as an important concern not because present practices are unsafe, but because of two factors. The first is that safety is a priority in the both the



**Fig. 14.1** The robot-assistant supporting the activities of a worker performing a simulated car door assembly task

laboratory and manufacturing environments. None of the stakeholders in this project would want to be part of a project that put the safety of study participants, factory workers, or themselves at risk. The second is that the concept of having workers working directly alongside robots is still an evolving one, in which safety standards are currently emerging. Progress on this is proceeding at a pace that observes both an appropriate level of caution and an acknowledgement that we must embrace physical human-robot interaction in order to achieve our ultimate goal of having robots which work directly alongside humans.

#### 14.2.1.2 Robot Assistant Hardware

The robot assistant, also referred to as the Intelligent Assist Device, Fig. 14.1, in its final form consists of a robotic arm (Kuka LWR-4) with a dextrous gripper (Robotiq gripper) mounted to an overhead gantry. These devices operate on a single controller, and redundancy resolution is implemented for the entire IAD system. The IAD can be treated as a single, integrated device, rather than as its individual components. The system implements variable stiffness in all actuators, providing compliance along the robot's entire kinematic chain. This compliance contributes to worker safety by allowing the system to softly collide and interact with obstacles, and serves as an interface to the system, allowing the entire unified device to be backdriven by pushes and shoves against the robot itself.

Sensing is performed using Microsoft Kinect and PrimeSense RGBD cameras. The Kinect cameras surround the work cell in which the human and robot assistant collaborate on their shared assembly task. Their point clouds are merged using custom software, providing a global view of the scene. The PrimeSense camera is physically smaller, and focuses on a narrower field of view. It is mounted to the Kuka LWR

robotic arm, allowing the system to focus on objects to be manipulated, and providing a directional sensor which can be used for tasks such as object identification and manipulation. One benefit of such a system is that it is able to perform non-contact sensing for uninstrumented human workers and parts. No markers are required to be placed on any worker in the work cell or any part that is to be manipulated using current techniques.

### 14.2.1.3 Robot Assistant Software

The system presented at Plugfest III represents the most complete version of the robot assistant software developed for the CHARM project. It features many important components which echo familiar components in modern robotics design. The entire system operates over a high speed computer network with nodes responsible for sensing, situational awareness, control, and planning. Nodes operating on the robot assistant communicate via a communications protocol called DirectLink. One feature of DirectLink is that it is designed to present the most recent data provided by sensors, rather than the entire log of historical data, in order to provide fast responses by the robot system and avoid latency which may be introduced by processing backlogged data. The entire history of data communicated to the system can be retrieved from the Situational Awareness Database (SADB) [22], which contains both raw sensor data and perceptual data which has been processed via systems such as the device's computer vision system. It also contains information such as commands issued to the robot, either via its State Machine Controller (SMC) or planner. The SADB can be quickly accessed via a NoSQL interface, and incurs very low latencies due to mirroring across nodes. The ability to query representations in this fashion is useful for software such as planners, because this prevents them from needing to maintain this representation internally.

Worker poses and world state are measured using the aforementioned group of Kinect and PrimeSense cameras. The system is capable of merging the multiple perspectives of the Kinect cameras into a unified point cloud which is used to reliably track the worker's motion in the workspace using a skeleton track representation [20]. It is also able to track multiple objects in the work space in this fashion. As such, progress on the shared task can be measured through actions such as mounting components onto the car door. These data are stored into the SADB, where they can be processed by either the SMC or the planner. The planner developed for CHARM uses a standard representation in Planning Domain Definition Language (PDDL) [14], which is augmented with measured times for the completion of various tasks. It can choose courses of action based on this representation in order to adapt to problems which may arise, such as a worker discarding a faulty part. Additionally, it is able to use its recorded timing data in order to plan optimal timings for task execution and to begin robot motion trajectories [9, 10]. As will be discussed later, this contributes to our present work on interaction fluency.

### 14.2.2 CHARM Streams and Thrusts

Dividing CHARM into a set of complimentary streams and thrusts provided a framework within which participating investigators and laboratories could bring their best work to the project, focus on making the biggest contribution that they could to the advancement of the projects goals, and ultimately be assured that their work fit into the project’s scope and made a meaningful contribution. CHARM was structured around three interconnecting research streams that link three research thrusts representing the main research directions of the project. These are shown in Fig. 14.2. The three streams that connected the development of robot assistants across the three research thrusts of CHARM are:

- Stream 1 Define the Human-Robot Interaction with the robotic assistant.
- Stream 2 Develop relevant Situational Awareness and Data Representation.
- Stream 3 Coordinate developments through an Integration Framework.

The ordering of the streams in Fig. 14.2, represents the focus on the human as central figure in the interaction with the robot assistant, and the understanding and management of situational awareness and data representation as key to supporting this interaction. An effective and strong integration framework for communication,

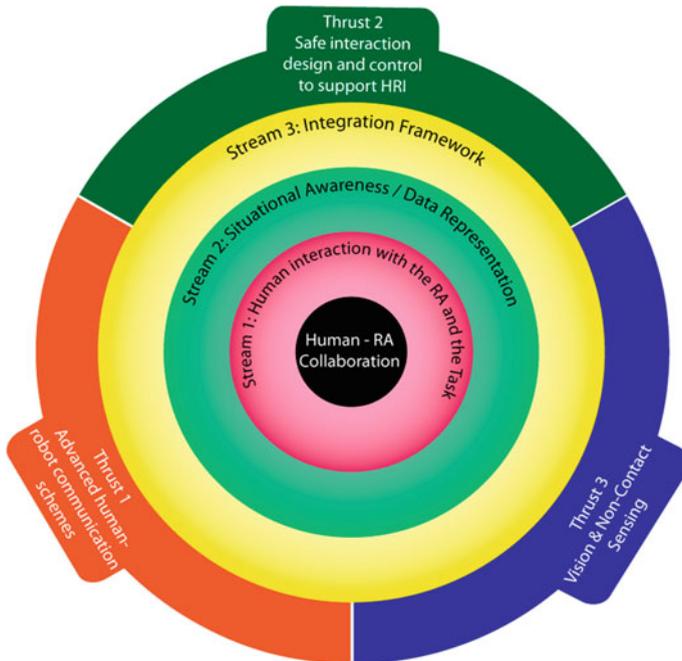


Fig. 14.2 Streams and thrusts of the CHARM project

control and perception is crucial to the successful development of a robot assistant that can effectively support a human worker at their task.

While providing support to the research thrusts, Stream 1 and Stream 2 also represent research activities that inform and are implemented within the three research thrusts, while Stream 3 is a research support activity that ties the research thrust developments together into a working prototype system.

The three research thrusts that drive the structure of Fig. 14.2 are:

Thrust 1 Advanced human-robot communication and cooperation schemes.

Thrust 2 Safe interaction design and control to support HRI.

Thrust 3 Vision and Non-Contact Sensing.

To each stream and thrust, domain experts were assigned, providing a structured exchange of knowledge and a focal point for specific needs of the project. Organizing the project in this fashion assured that responsible parties could be reached to discuss every relevant aspect of the project and that individual investigators knew who to contact, while also assuring that these investigators had the autonomy to pursue state-of-the-art work and bring it to bear on our shared task. Each stream and thrust made unique and necessary contributions to CHARM.

#### **14.2.2.1 Stream 1**

Stream 1 designed the interaction between the robot assistant and the human operator. Key systems were developed to allow human operators to interact with the robot through taps and pushes on the robot itself. Systems were developed in response to studies carried out under Thrust 1, enabling the robot to perform tasks such as elegant robot-to-human handovers. Stream 1 also developed a planner [9] that allowed the robot assistant to interactively replan for Plugfest III, as part of efforts toward using timing for fluent HRI.

#### **14.2.2.2 Thrust 1**

Thrust 1 performed studies in human-robot interaction in order to develop novel communicative cues for the robot assistant system. This thrust followed the paradigm of performing human-human studies in order to identify communicative cues, modeling and describing these cues, and then deploying them on robotic systems in order to study them in the context of Human-Robot Interaction. This process is described in greater detail in Sect. 14.3. Studies performed using this process can be found in Sect. 14.4. Contributions to the robot assistant include an elegant handover controller for robot-to-human handovers, studies in gestures, and studies in collaborative lifting.

### **14.2.2.3 Stream 2**

Stream 2 developed the Situational Awareness Database (SADB) [22] and the basic world-state and sensor data representations used in the robot assistant system. Techniques for the modeling of and storage and retrieval of data are necessary in order to enable the system to autonomously reason about its environment and the shared task that it participates in.

### **14.2.2.4 Thrust 2**

Thrust 2 developed the robot assistant Intelligent Assist Device (IAD) hardware and the control algorithms that drive this system. The software which integrates the gantry, robot arm, and robot gripper under a single controller was developed under this thrust, as well as capabilities such as the compliant control of the kinematic chain and backdrive capabilities [7, 12, 13].

### **14.2.2.5 Stream 3**

Stream 3's primary focus was to integrate components from the various streams and thrusts of CHARM. As such, inter-group communication and Plugfests were of a primary concern. Stream 3 contributed the basic communication protocol, DirectLink, which is used for communications between compute nodes in the robot assistant architecture, and the State Machine Controller (SMC), which could be used for high-level control of the robot-assistant.

### **14.2.2.6 Thrust 3**

Thrust 3's contribution was in the form of vision and non-contact sensing. For earlier revisions of the system, this involved tracking objects and actors in the interaction through the use of a Vicon motion tracking system. Later, this was replaced by a network of Microsoft Kinect RGBD cameras, and a PrimeSense camera mounted to the IAD's robotic arm. Software developed under Thrust 3 allowed the system to perform recognition and tracking, as well as to track the motion and poses of human actors in the interaction [20].

## ***14.2.3 Plugfest***

Plugfests served as a focal point for the effort of integrating components developed over each year and evaluating overall progress towards collaborative robot assistant technology. Each Plugfest was preceded by about two months of planning and prepa-

ration, with inter-site visits between investigators in each stream and thrust in order to assure that their systems properly integrated. The progress reports at the end of each Plugfest provide a picture as to how the robot assistant system matured over time.

### **14.2.3.1 Plugfest I**

At the end of the first year, the group was still in a phase of defining requirements and specifications. Each team presented initial capabilities that represented a current state of the art in their respective areas. Reports were presented describing the findings of initial studies regarding capabilities that the labs wanted to contribute to the system, as well as findings from site visits to General Motors automotive manufacturing plants. As hardware was still being acquired, the robot was not in its final form.

At this juncture the team was able to produce an integrated system which gave a vision of what was to come. A worker instrumented with markers for a Vicon motion tracker interacted with the robot which was under the control of the State Machine Controller. The robot did not yet have its arm on it, and presented parts to the worker on a tray. Data were stored and queried from a preliminary version of the Situational Awareness Database.

### **14.2.3.2 Plugfest II**

For Plugfest II, the Vicon system was removed in favor of a Kinect-based markerless tracking system. The system was able to register the point clouds from the independent Kinect sensors into a single point cloud and perform person tracking, where the representation of the worker was as a blob of points. A gesture-based system was developed in which the worker would make requests to the robot and the robot could communicate back to the worker through gestures. At this time, the robot arm was integrated into the system and a handover controller was used to hand objects to the worker. Improvements were made to the State Machine Controller and DirectLink protocols improving the overall responsiveness of the system.

### **14.2.3.3 Plugfest III**

For Plugfest III human tracking had been updated to track a full skeletal model and the SADB was capable of high-performance transactions and replication across nodes. The gantry, gripper, and arm of the robot assistant were integrated into a single controller allowing compliant control and backdrive across the entire system. Gesture-based control had largely been replaced by a system which monitored the work state of the system, removing the need for some of the gesture-based commands, which had been found to slow down the worker during the interaction. An interactive

planner had been added which allowed the system to reason about the scenario using PDDL [14].

CHARM provided its stakeholders with a project that enabled us to develop an integrated robot assistant system which brought the best of current technology to bear on the problem. By building a real, integrated system, we were able to see how our contributions impacted a real-world application and how these systems interacted with each other.

### **14.3 Identifying, Modeling, and Implementing Naturalistic Communicative Cues**

The CARIS Laboratory uses a three phased method to identify, model, and implement naturalistic communicative cues in robotic systems. These phases comprise the following steps:

Phase 1 Human-Human Studies

Phase 2 Behavioral Description

Phase 3 Human-Robot Interaction Studies

To aid in describing this process, this section will use the example of recent work in hand gestures carried out in the CARIS Laboratory [21]. In this work, first, a human-human study of a non-verbal interaction between dyads of human participants performing an assembly task is carried out in order to witness and identify the hand gestures that they use in order to communicate with each other. This is Phase 1 from the three phase method. From the data collected during this study, a set of communicative hand-gestures is identified by the researchers from annotated data. This set of gestures is validated using an online study of video-recorded example gestures, ascertaining whether study participants recognize the same gestures and intentions as those identified by the experimenters. This is Phase 2, in which an accurate description of the human behavior is identified. In Phase 3, these gestures are then programmed into a Barrett Whole-Arm Manipulator (WAM) for use in a human-robot study. In a step mirroring Phase 2, the robot gestures are studied in an online study, with results reported for recognition and understanding by participants.

#### ***14.3.1 Phase 1: Human-Human Studies***

The purpose of the Human-Human study phase is to elicit behaviors on the part of humans in a collaboration so that they can be characterized and understood, and then replicated on a robotic platform.

For this work, a study of non-verbal interaction between human dyads performing a car door assembly task was performed. The door is instrumented in seven locations

with Velcro™ strips where parts can be mounted. Correspondingly, six parts that are to be mounted onto the car door are instrumented with corresponding Velcro™ strips. Study participants were provided with a picture of a completed assembly of these parts mounted onto the car door, and asked to non-verbally communicate the proper placement and orientation of these parts on the door to a confederate through the use of hand gestures. After this assembly was completed, a second picture was presented to the participants, changing the location and orientation of four of the parts on the door. At this stage, participants were asked to direct the confederate to modify the arrangement of the parts on the door as indicated in the new picture. Items were placed on a table between the study participant and confederate in order to provide easy access to both the door and the items, as in Fig. 14.3. The participants were required to perform this task in accordance with a set of provided rules, which assure that relevant communication could be performed only via hand gestures.

- Only use one hand to direct the worker.
- Only make one gesture and hold only one part at a time.
- You must wait for the worker to complete the task before making your next gesture.
- You must remain in your home position at all times.

A group of 17 participants (female: 7, male 10) between 19 and 36 years of age participated in this study.



**Fig. 14.3** Participants in a Phase 1 human-human study of non-verbal interaction

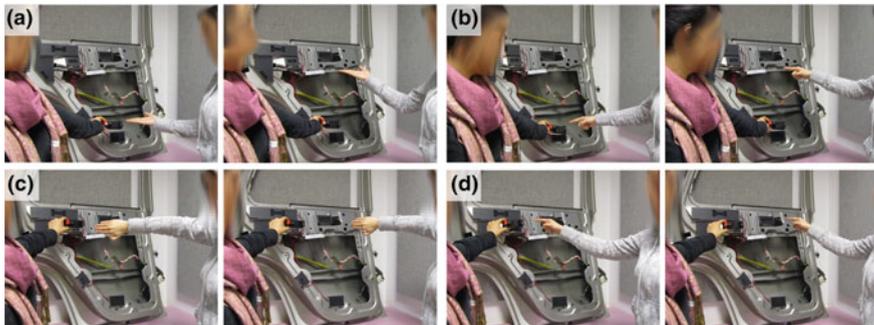
### 14.3.2 Phase 2: Behavioral Description

In the second phase, behaviors exhibited during the human-human study are characterized. In this study, gestures used by the participants were identified from annotated video data. Gestures were selected from the annotated data according to the following criteria:

- They should be understandable without trained knowledge of the gesture.
- They should be critical to task completion.
- They should be commonly used among all participants.

Based on these criteria, the experimenters identified directional gestures as a category of interest for further study, and narrowed the gestures into four categories. These were “Up,” “Down,” “Left,” and “Right.” They also identified that each of these gestures could be performed with an Open-Hand or Finger-Pointing hand pose. These gestures appear as in Fig. 14.4.

Video clips of these gestures were used in an online study of 120 participants in which participants were asked three questions. The first asks, “What do you think the worker should do with this part?” where participants were instructed to answer “I don’t know” if they did not understand the gesture. In the second they were asked to rate, “How easy was it for you to understand the meaning of this gesture (on a scale from 1 (very difficult) to 7 (very easy))?” In the third they were asked, “How certain are you of your answer to question 1 (on a scale from 1 (very uncertain) to 7 (very certain))?” At the end of this phase, survey responses to the latter two questions were found to have a high degree of internal consistency (Cronbach  $\alpha = 0.891$ ).



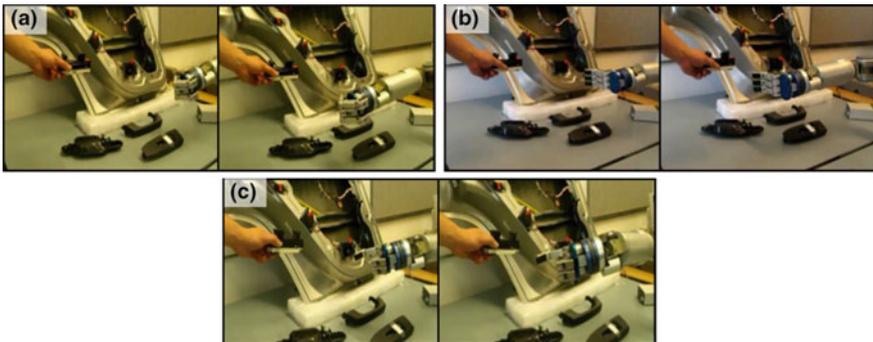
**Fig. 14.4** Directional Gestures and frequently observed accompanying hand poses identified from annotated data - “Up” and “Down” gesture with Open-Hand (a), and Finger-Pointing (b) poses, and “Left” and “Right” gesture with Open-Hand (c), and Finger-Pointing (d) poses

### 14.3.3 Phase 3: Human-Robot Interaction Studies

The purpose of Phase 3 is to attempt to replicate the identified and described communicative cues in a human-robot interaction. To do this, the study in Phase 2 was replicated with a robotic arm, a 7 Degree of Freedom (DoF) Barrett Whole-Arm Manipulator equipped with a 3-fingered BarrettHand. Gestures were programmed into the arm and presented as video clips in an online study of 100 participants. In the robotic arm case, each gesture was presented using one of three hand poses: one with an Open-Hand (OH), one with a Finger-Pointing hand pose (FP), and one with a Closed-Hand (CH), as in Fig. 14.5.

Sheikholeslami et al. [21] report recognition rates for these gestures, comparing against the human and robot conditions. Results are shown in Fig. 14.6. Their results demonstrate a similar degree of recognition and understanding of these gestures in both the human and robot conditions, and show whether they are more easily-understood for various poses of the robot's manipulator.

The purpose of this methodology is to directly identify, analyze, and implement human communicative behaviors on robotic systems in order to create human-robot interactions which are naturalistic and intuitive. The steps of performing human-human studies and characterizing the behaviors of the participants provide the data required to reproduce these behaviors, while human-robot interaction studies validate the effectiveness of the reproduced communicative cues. In the CARIS Laboratory, a central goal is to create human-robot interactions in which robots are able to interact with humans in a collaborative fashion, rather than in a manner that is more akin to direct control through a keyboard or teach-pendant interface. Exploiting naturalistic communicative behaviors is a key avenue by which we are attempting to achieve this goal.



**Fig. 14.5** “Left” and “Right” gestures implemented on the Barrett WAM with Closed-Hand (a), Open-Hand (b), and Finger-Pointing (c) hand poses implemented on the BarrettHand

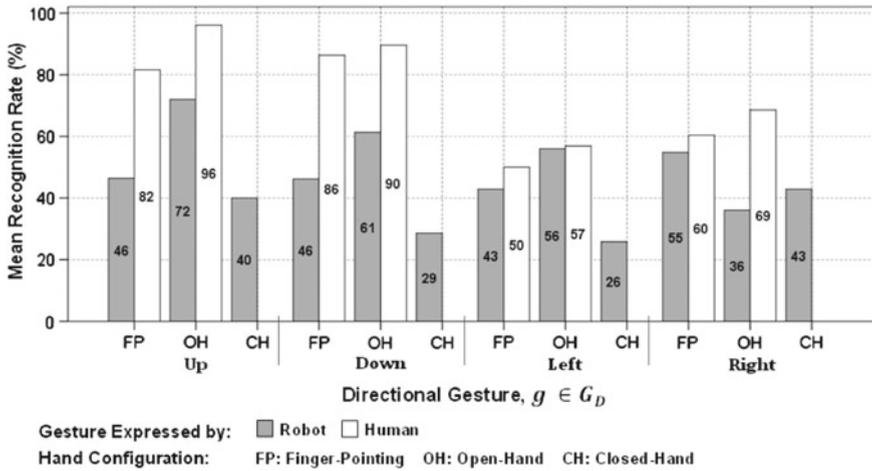


Fig. 14.6 Comparison of recognition rates for hand gestures by gesture and hand configuration between robot and human cases. OH - Open-Hand, FP - Finger-Pointing, CH - Closed-Hand

## 14.4 Communicative Cue Studies

The CARIS Laboratory has done extensive research exploring the use of communicative cues for human-robot interaction. Many of these cues are naturalistic cues; cues that emulate natural human communicative behaviors such as can be found through human-human studies and witnessed in everyday human interactions. Some of these cues are non-naturalistic, as in the case of tapping and pushing on the robot to guide its motion [5]. These cues can also be divided into explicit and non-explicit cues. An example of explicit cues would be hand gestures, as described in Sect. 14.3. An example of a non-explicit cue would be when a person inadvertently looks towards the location where they intend to hand an object over to someone else [1, 17, 23]; or the sensation of feeling another person managing the weight of an object that is handed over, allowing it to be released [3, 4].

### 14.4.1 Human-Robot Handovers

The CARIS lab has extensively studied the process of handing an object from one party to another. This is an example of an interaction that is mostly mediated by naturalistic, non-explicit cues. Important cues occurring during handovers that have been explored by CARIS include the forces acting on the object being handed over [3, 4], gaze behaviors during the handover interaction [17], motion trajectories [9, 10], and kinematic configurations [2].

To explore forces acting on an object during handovers, Chan et al. [3] constructed a baton that is instrumented with force sensing resistors (FSRs), an ATI force/torque sensor, and inertial sensors, as can be seen in Fig. 14.7. Nine pairs of participants were recruited to perform 60 handovers each in 6 different configurations. The investigators found that distinct roles for the giver and receiver of the object emerge in terms of the forces acting on the baton. The giver assumes responsible for the safety of the object by assuring that it does not fall, whereas the receiver assumes responsible for the efficiency of the handover by taking it from the giver. This can be measured as a characteristic of grip and load forces over time on behalf of the giver and receiver, with the handover concluding when the giver experiences negative load force as the receiver slightly pulls it out of their hand. This has direct implications for the design of a controller for robot-to-human handovers. Chan et al. [4] implemented such a controller on the Willow Garage PR2 robot, which mimics human behavior by regulating grip forces acting on the transferred object.

Moon et al. [17] followed up on this study by adding a gaze cue to the PR2's handover software. In a three condition, intra-participant design, 102 study participants were handed water bottles and asked which of three handovers they preferred. The robot indicated its gaze direction by tilting its head to either look down and away

**Fig. 14.7** Baton instrumented for measuring forces acting on an object during handover



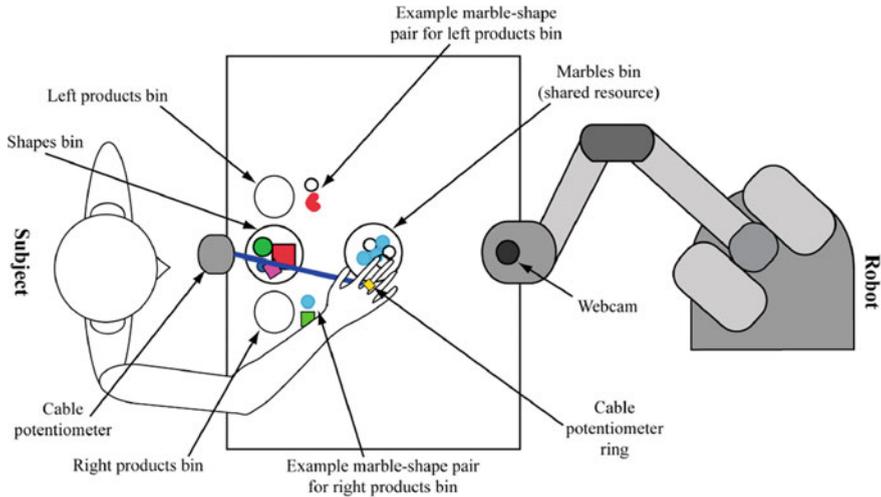
from the participant (No Gaze), look towards the water bottle (Shared Attention), or look to the water bottle and then up at the participant (Turn-Taking). The study found that participants reach for the water bottle significantly earlier in Shared Attention ( $M = 1.91s$ ,  $SD = 0.52$ ) condition over the No Gaze condition ( $M = 2.54s$ ,  $SD = 0.79$ ) ( $p < 0.005$ ). This is measured with respect to the time when the robot grasps the water bottle to be handed over and starts its trajectory toward the handover location. No significant difference was found between the Shared Attention and Turn Taking ( $M = 2.26$ ,  $SD = 0.79$ ) conditions or between the Turn Taking and No Gaze conditions. Their results also suggest that participants may prefer conditions in which the robot makes eye contact.

Additional work in conjunction with University of Tokyo went on to study how the orientation of an object interacts with how it is handed over [2]. Current work in CARIS investigates motion cues that can be exploited to detect the timing and location of handovers, for fluent human-to-robot handover behaviors [10].

### 14.4.2 *Hesitation*

Another example of a non-explicit communicative cue is motor behavior when hesitating. Moon et al. [15, 16] recorded human motion trajectories in an experiment where dyads of study participants reach for a target placed between them on a table when prompted by tones in sets of headphones. In the study, randomly-timed tones are played separately in each pair of headphones, such that only sometimes are both participants required to reach for the target at the same time. This causes participants to sometimes hesitate in accessing the shared resource placed between them, ceding access to the other participant. The study separates recorded arm motions into three categories: successful (S) - when the participant accesses the shared resource, retract (R) - when the participant retracts their hand from a trajectory directed toward the target as an act of hesitation, and pause (P) - when the participant pauses along their motion trajectory toward the target in hesitation. Accelerometer data were recorded for these motions, and a motion profile was derived from these recorded human hesitations called the Acceleration-Based Hesitation Profile (AHP). A retract-based motion profile was used to plan motion trajectories for a robot arm during a similar shared task with a human collaborator. This motion was video recorded for use in an online study. Results of the study demonstrate that study participants recognize the reproduced hesitation behavior on the part of the robot.

A follow-up study was conducted in which participants perform a shared task with the robot [16]. In this experiment, participants sit across from the robot with the task of taking marbles from a bin located in the center of the workspace shared with the robot one at a time, matching them with shapes from another bin according to a set of exemplar marble-shape pairs. The robot's task is to inspect the marble bin by moving back and forth between the bin and its starting position, see Fig. 14.8. A total of 31 participants took part in a within-participant study in which they were exposed to three conditions; Blind Response - in which the robot continues along



**Fig. 14.8** Diagram illustrating the experimental setup of a study in which human study participants interact with a robotic arm which performs hesitation behaviors during an interaction involving access to a shared resource

its trajectory, Robotic Avoidance - in which the robot arrests its motion to wait for the participant to complete their motion, and Hesitation Response - in which the robot responds with an AHP-based trajectory. Results do not show improvements in task completion or perceptions of the robot, but do demonstrate that participants recognize the hesitation behavior.

### 14.4.3 Tap and Push

Our group and our collaborators are also interested in communicative cues which are non-naturalistic, but nonetheless may be highly intuitive for human collaborators, or which are based on common human interactions but not on specific communicative cues. One example of this is a study in tap-and-push style interactions in which a human study participant taps and pushes on a robot. This study was conducted in conjunction with the Sensory Perception & INteraction (SPIN) Laboratory at UBC, by Gleeson et al. [5].

In order to study tap-and-push style interactions with a robot arm, Gleeson et al. [5] performed a series of studies in which human workers and robot assistants collaborate on an assembly task. These studies compare a set of commands based on tapping and pushing on a device to commands via a keyboard interface. In the first study, participants interact with a Phantom Omni desktop haptic device in a scripted collaborative fastener insertion task which simulates the placement and tightening of four bolts. Participants pick up bolts from open boxes, place a bolt in one of four

locations on a board, and then command the Omni to touch the bolt, simulating a tightening operation. See Fig. 14.9a. After performing its tightening operation, the Omni automatically moves to the next position. Gleeson et al. [5] found that on this task keyboard commands slightly outperform direct physical commands in the forms of taps and pushes on the robot in quantitative task performance metrics and qualitative user preference.

The second study comprises two more complex tasks which were performed as interactions with a Barrett Whole-Arm Manipulator (WAM) robotic arm. The first task is a bolt insertion task similar to that in the first study, but in which bolts are inserted in a random order that is presented to the user on notecards, and in which the robot does not automatically advance to the next bolt. See Fig. 14.9b. In the second task, participants interactively position the arm using discrete tap-and-push cues to position it over a series of cups, and then continuously guide the robot to the bottom of the cup, Fig. 14.9c. Gleeson et al. [5] found that in these more complex and less scripted tasks participants are able to more quickly complete the collaborative tasks, and that they prefer the physical interaction over the keyboard interface.

## 14.5 Current and Future Work

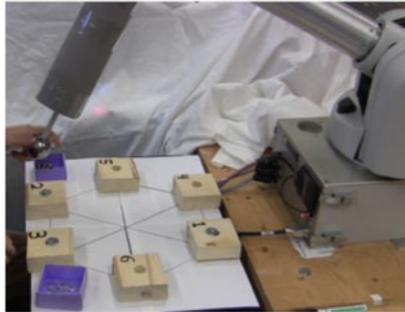
Current work in the CARIS Laboratory continues our study of communicative cues. Part of Matthew Pan's current work involves collaboration between two parties lifting an object. This expands on work by Parker and Croft [18, 19] on the development of controllers that enable robots to elegantly respond to cues based on the motion of an object that is in the process of being lifted. Another aspect of Pan's current work is the automatic detection of the intention of a person to handover an object, thus enabling a robot to identify this intention and respond by grasping the object.

Hart, Sheikholeslami, and Croft are currently working on approaches to extrapolate human motion trajectories based on hand and skeletal tracks [10]. Such an extrapolation would allow the prediction of the timing and location of the endpoint of the motion, and of what the person is attempting to do. In the context of a reach, this can inform a robot of what the person is reaching toward. Hoffman and Breazeal [11] noted that collaborative robotic systems that are based on observations of the current state of the shared task experience a stop-and-go style interaction. This is in part due to the need to respond to the current world state, rather than what is about to happen. They describe the ability to smoothly mesh the actions of collaborators as interaction fluency. By making guesses at a collaborator's intentions and extrapolations of their motions, a robot can act on these predictions. Precise predictions of the timing and location and choice of actions on a shared task can enable a robot to preemptively act on these predictions. For instance, if a person is reaching to a fastener, the robot can reach to the tool that attaches it. We have also explored optimal timing of a robot's behavior based on prior task performance on the part of both the robot and the worker.

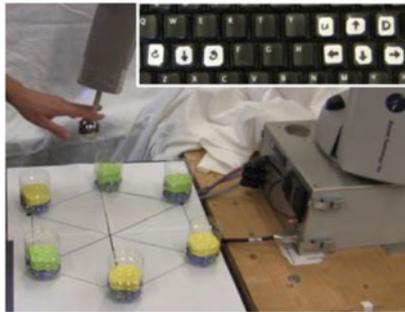
One application that we are currently pursuing is combining motion extrapolation techniques with Matthew Pan's work on handover detection, allowing us to predict



(a) Phantom Omni haptic device, interacted with via taps and pushes, or via keyboard in a bolt placement task.



(b) Barrett Whole-Arm Manipulator, interacted with via taps and pushes, or via key board in a bolt placement task.



(c) Barrett Whole-Arm Manipulator, interacted with via taps and pushes, or via keyboard in a guided manipulation task.

**Fig. 14.9** Experimental setups for studies in tap and push interactions with robotic devices

the timing and location of a handover, in addition to preemptively predicting the intention based on a predicted motion trajectory. In a study of handover motions [10], we observed that the receiver of an object being handed over begins their motion to the ultimate location of the handover prior to the arm being fully-extended to this location. We also observed that we could determine the path that the arm would follow from only a few frames of motion tracking data (as long as the track is stable). In the context of a handover motion, this can inform the timing and location at which person intends to hand over the object, thus enabling the robot to begin its trajectory in a manner similar to that of a person reaching out to accept the object. Studies on this are in progress.

The robot-assistant project itself is also exploring new directions. We are currently in the process of designing interactions around the use of advanced composite materials such as carbon fiber reinforced plastics and the construction of large components. Techniques such as gesturing to the robot will enable non-verbal communication at the longer distances required to collaborate on large composite components, whereas other naturalistic communicative cues may be combined with advanced projection mapping and augmented reality technologies to enable the robot to communicate important information back to the user. Sensing, situational awareness, and proxemics will play key roles in designing these interactions.

A closed loop of communication between a human worker and its robot collaborator is key for progress in collaborative human-robot interaction. With humans and robots working in close proximity, physical HRI techniques will become key in enabling a robot to live up to worker expectations and interpret worker intentions. For worker safety and productivity, it is crucial that robots and human collaborators are able to transparently communicate with each other and understand each other's actions and intentions. For robot assistants to act as collaborators, rather than tools directly under the control of human operators, they must behave in predictable manners and in ways that operators are able to intuitively control. Their development requires contributions from multiple disciplines in computer science, mechanical engineering, design, sensing, and control. The study of communicative cues provides a route to establishing a closed loop of transparent communication between a human worker and its robot collaborator, while behavioral predictions provide us with a route to performing this communication fluently.

**Acknowledgements** We would like to thank our collaborators on CHARM for their incredible contributions to this project. In addition to the authors of this chapter, CHARM collaborators include Ergun Caliskan, Jacques-Michel Haché, Dominic Beaulieu, AJung Moon, Marc-Antoine Lacasse, Olivier St-Martin Cormier, Andrew Phan, Brian Gleeson, Denis Ouellet, Boris Mayer St-Onge, Thierry Moszkowicz, Oleg Boulanov, Roland Y. Menassa, Stephen Hart, Muhammad Abdallah, Jim Wells, Justin Gammage and Robert Tilove. CHARM was made possible thanks to the generous support of General Motors of Canada and the Natural Sciences and Engineering Research Council of Canada under contract 140570247. Studies carried out at UBC that are reported in this chapter were approved by the University Behavioral Research Ethics Board under approval number H10-00503-A020.

## References

1. H. Admoni, A. Dragan, S.S. Srinivasa, B. Scassellati, Deliberate delays during robot-to-human handovers improve compliance with gaze communication. *Proceedings of the Ninth Annual ACM/IEEE International Conference on Human-robot Interaction (HRI '14)* (Bielefeld, Germany, 2014), pp. 49–56
2. W.P. Chan, M.K.X.J. Pan, E.A. Croft, M. Inaba, Characterization of handover orientations used by humans for efficient robot to human handovers. *Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems* (Hamburg, Germany, September 2015), pp. 1–6
3. W.P. Chan, C.A.C. Parker, H.F.M. Van der Loos, E.A. Croft, Grip forces and load forces in handovers: Implications for designing human-robot handover controllers. *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI '12)* (Boston, MA, USA, 2012), pp. 9–16
4. W.P. Chan, C.A.C. Parker, H.F.M. Van der Loos, E.A. Croft, A human-inspired object handover controller. *Int. J. Robot. Res.* **32**(8), 971–983 (2013)
5. B. Gleeson, K. Currie, K. MacLean, E. Croft, Tap and push: Assessing the value of direct physical control in human-robot collaborative tasks. *J. Hum. Robot Interact.* **4**(1), 95–113 (2015)
6. B. Gleeson, K. MacLean, A. Haddadi, E.A. Croft, J.A. Alcazar, Gestures for industry intuitive human-robot communication from human observation. *Proceedings of the Eighth Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI '13)* (Tokyo, Japan, March 2013), pp. 349–356
7. C. Gosselin, T. Laliberté, B. Mayer-St-Onge, S. Foucault, A. Lecours, V. Duchaine, N. Paradis, D. Gao, R. Menassa, A friendly beast of burden: A human-assistive robot for handling large payloads. *IEEE Robot. Autom. Mag.* **20**(4), 139–147 (2013)
8. A. Haddadi, E.A. Croft, B. Gleeson, K. MacLean, J. Alcazar, Analysis of task-based gestures in human-robot interaction. *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '13)* (Karlsruhe, Germany, May 2013), pp. 2146–2152
9. J.W. Hart, B.T. Gleeson, M.K. X.J. Pan, A. Moon, K. MacLean, E.A. Croft, Gesture, gaze, touch, and hesitation: Timing cues for collaborative work. *Proceedings of the HRI Workshop on Timing in Human-Robot Interaction (HRI '14)* (Bielefeld, Germany, March 2014)
10. J.W. Hart, S. Sheikholeslami, M.K.X.J. Pan, W.P. Chan, E.A. Croft, Predictions of human task performance and handover trajectories for human-robot interaction. *Proceedings of the HRI Workshop on Human-Robot Teaming (HRI '15)* (Portland, OR, USA, March 2015)
11. G. Hoffman, C. Breazeal, Cost-based anticipatory action selection for human-robot fluency. *IEEE Trans. Robot.* **23**(5), 952–961 (2007)
12. A. Lecours, B. Mayer-St-Onge, C. Gosselin, Variable admittance control of a four-degree-of-freedom intelligent assist device. *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)* (St. Paul, MN, USA, May 2012), pp. 3903–3908
13. A. Lecours, C. Gosselin, Computed-torque control of a four-degree-of-freedom admittance controlled intelligent assist device. *Experimental Robotics: The 13th International Symposium on Experimental Robotics*, ed. by P. Jaydev Desai, G. Dudek, O. Khatib, V. Kumar (Springer, Heidelberg, Germany, 2013), pp. 635–649
14. D. McDermott, M. Ghallab, A. Howe, C. Knoblock, A. Ram, M. Veloso, D. Weld, D. Wilkins, Pddl - the planning domain definition language. Technical Report TR-98-003, Yale Center for Computational Vision and Control, 1998
15. A.J. Moon, C.A.C. Parker, E.A. Croft, H.F.M. Van der Loos, Did you see it hesitate? Empirically grounded design of hesitation trajectories for collaborative robots. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '11)* (San Francisco, CA, USA, September 2011), pp. 1994–1999
16. A.J. Moon, C.A.C. Parker, E.A. Croft, H.F.M. Van der Loos, Design and impact of hesitation gestures during human-robot resource conflicts. *J. Hum.-Robot Interact.* **2**(3), 18–40 (2013)

17. A.J. Moon, D.M. Troniak, B. Gleeson, M.K.X.J. Pan, M. Zheng, B.A. Blumer, K. MacLean, E.A. Croft, Meet me where I'm gazing: How shared attention gaze affects human-robot handover timing. *Proceedings of the Ninth ACM/IEEE International Conference on Human-robot Interaction (HRI '14)* (New York, NY, USA, March 2014), pp. 334–341
18. C.A.C. Parker, E.A. Croft, Design & personalization of a cooperative carrying robot controller. *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '12)* (St. Paul, MN, USA, May 2012), pp. 3916–3921
19. C.A.C. Parker, E.A. Croft, Experimental investigation of human-robot cooperative carrying. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems ((IROS '11))* (San Francisco, CA, USA, September 2011), pp. 3361–3366
20. A. Phan, F.P. Ferrie, Towards 3d human posture estimation using multiple kinects despite self-contacts. *Proceedings of the 14th IAPR International Conference on Machine Vision Applications (MVA)* (Tokyo, Japan, May 2015), pp. 567–571
21. S. Sheikholeslami, A.J. Moon, E.A. Croft, Exploring the effect of robot hand configurations in directional gestures for human-robot interaction. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '15)* (Hamburg, Germany, September–October 2015), pp. 3594–3599
22. O. St-Martin Cormier, A. Phan, F.P. Ferrie, Situational awareness for manufacturing applications. *12th Conference on Computer and Robot Vision* (Halifax, Nova Scotia, Canada, 2015), pp. 320–327
23. K. Strabala, M.K. Lee, A. Dragan, J. Forlizzi, S. Srinivasa, M. Cakmak, V. Micelli. Towards seamless human-robot handovers. *J. Hum.-Robot Inter.* (2013)

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

